

and Newtonian cooling considered in Section 4.5.2,  $\nabla \cdot \mathbf{F}$  is found to be negative in general, and “nonacceleration conditions” are violated.

It is also straightforward to calculate  $\mathbf{F}$  for the approximate *WKBJ* solutions of Section 4.5.4; we find

$$\mathbf{F} \approx \frac{1}{2} \rho_s k |\hat{\Psi}|^2 (0, l, \varepsilon m), \quad (4.5.33)$$

while the wave-activity density  $A = \frac{1}{2} \rho_0 \overline{q'^2} / \bar{q}_v$  [cf. Eq. (3.6.6)] is given by

$$A \approx \frac{1}{4} \rho_s (k^2 + l^2 + \varepsilon m^2)^2 |\hat{\Psi}|^2 / \bar{q}_v, \quad (4.5.34)$$

since  $q' \approx -(k^2 + l^2 + \varepsilon m^2) \psi'$  under *WKBJ* conditions, from Eq. (4.5.2). It can further be verified that

$$\mathbf{F} = (0, c_g^{(v)}, c_g^{(z)}) A \quad (4.5.35)$$

to leading order in  $\mu_w$  for these waves, using Eqs. (4.5.29), (4.5.30), (4.5.33), and (4.5.34). Thus  $\mathbf{F}$  is parallel or antiparallel to the group velocity at each point under these conditions—and therefore to the rays defined in Section 4.5.4—according as  $A > 0$  or  $A < 0$ . Note that in the present quasi-geostrophic case  $A$  has the same sign as  $\bar{q}_v$ , by Eqs. (3.6.6) or (4.5.34)—this is usually positive in the midlatitude stratosphere and mesosphere.

Following standard practice, the group velocity and the concept of a ray have been defined above only in the context of a *WKBJ* theory of almost-sinusoidal waves. However, Eq. (4.5.35) suggests that, in more general situations where the *WKBJ* approximations fail, one possible extended definition of the group velocity is

$$\mathbf{C} \equiv \mathbf{F}/A, \quad (4.5.36)$$

where  $\mathbf{F}$  and  $A$  are given by Eqs. (3.5.6) and (3.6.6), respectively, or their primitive-equation equivalents. Generalized “rays” can then be defined as curves parallel to  $\mathbf{C}$  at every point. These ideas will be used in Chapters 5 and 6 for interpreting the behavior of planetary waves in the atmosphere and in models.

## 4.6 Gravity Waves

It was mentioned in Section 4.1 that pure internal gravity waves owe their existence to buoyancy restoring forces, while inertio-gravity waves are due to the combined effects of buoyancy and Coriolis forces. The “gravity modes” identified in Figs. 4.2a,b, which include the main propagating tidally forced modes of Section 4.3, are, strictly speaking, examples of inertio-gravity waves, since they are generally affected to some extent by the rotation of the earth. In the present section we consider some rather simpler gravity

waves, restricting attention to waves of comparatively small scale (tens to hundreds of kilometers in horizontal wavelength), so that the complications of spherical geometry can be avoided.

Gravity waves of this scale appear to be common in the upper mesosphere, where they have been detected by radars and other instruments; for example, Fig. 4.15 shows some radar observations of internal gravity waves in this region. The periods of waves of this type are typically a few minutes to an hour or so, and vertical wavelengths range from 5 to 15 km. Although direct measurements are difficult, the waves are thought to have horizontal wavelengths of up to about 100–200 km, and horizontal phase speeds of up to  $80 \text{ m s}^{-1}$ . Inertio-gravity waves, with periods approaching the local “inertial period”  $2\pi f^{-1}$  and vertical wavelengths on the order of 10 km, have been detected in the upper mesosphere and also in the lower stratosphere, and these can have horizontal wavelengths of over a thousand kilometers: some observations are shown in Fig. 4.16. Internal gravity waves are also likely to be common in the lower mesosphere and the stratosphere, but observations at these levels are sparse at present.

#### 4.6.1 Pure Internal Gravity Waves

We start by considering small wave disturbances about a basic state of rest, whose frequencies are large compared to the local inertial frequency  $f = 2\Omega \sin \phi$ . We therefore neglect the effects of the earth’s rotation by putting  $f = 0$  in Eq. (4.2.1), set  $X' = Y' = J' = 0$ , and employ Cartesian coordinates  $x$  and  $y$ . We obtain

$$u'_t + \Phi'_x = 0, \quad v'_t + \Phi'_y = 0, \quad (4.6.1a,b)$$

$$u'_x + v'_y + \rho_0^{-1}(\rho_0 w')_z = 0, \quad \Phi'_{zt} + N^2 w' = 0. \quad (4.6.1c,d)$$

For simplicity we take  $N$  to be constant; we then substitute

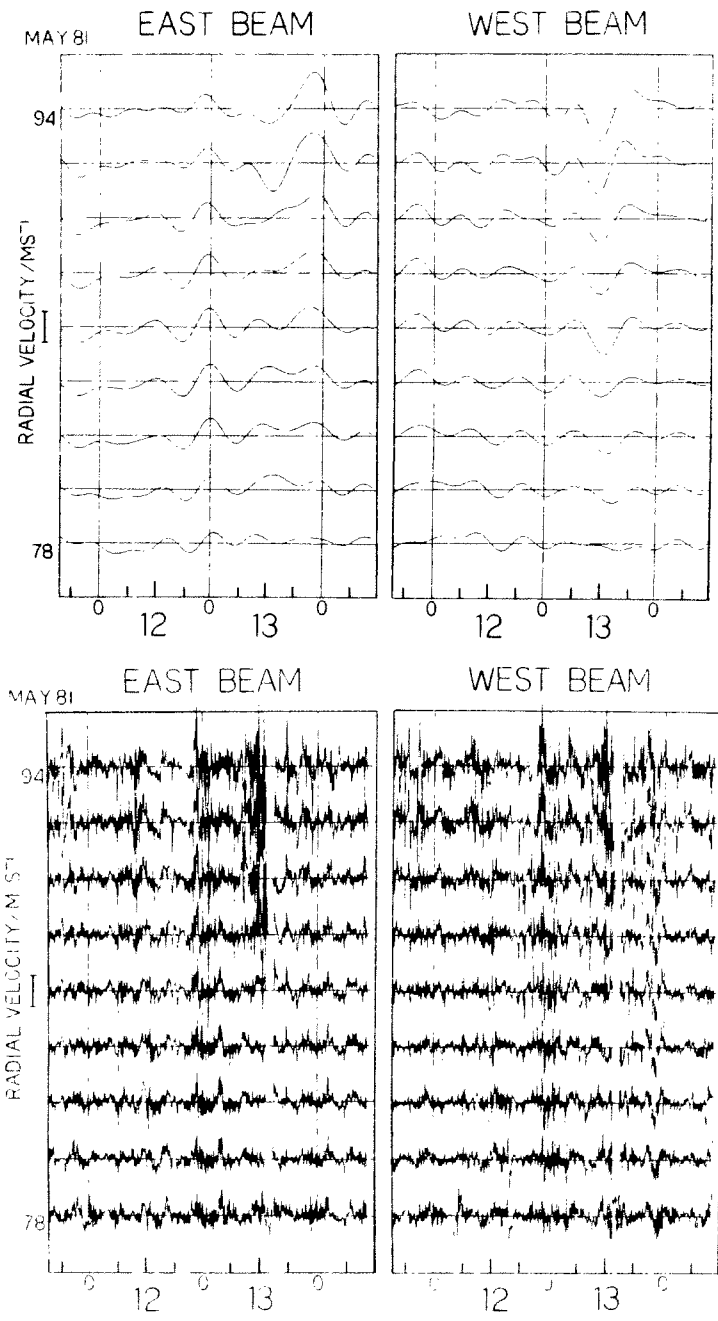
$$(u', v', w', \Phi') = e^{z/2H} \text{Re}[(\hat{u}, \hat{v}, \hat{w}, \hat{\Phi}) \exp i(kx + ly + mz - \omega t)] \quad (4.6.2)$$

into Eq. (4.6.1), where  $\hat{u}$  etc. are constant, and obtain the following equations:

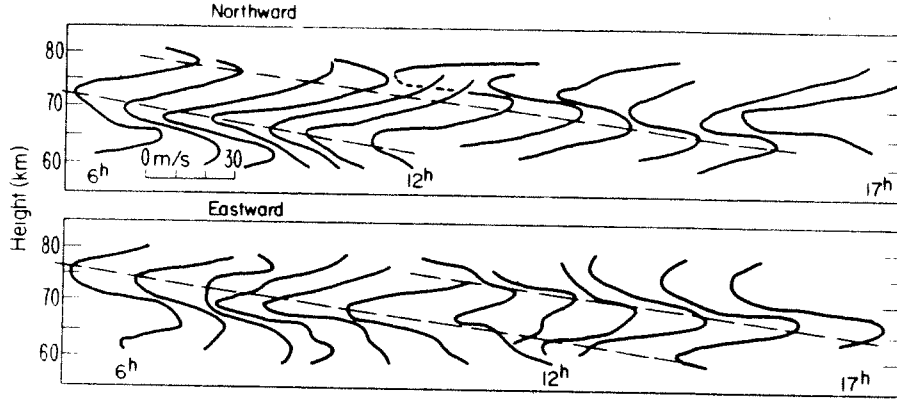
$$\hat{u} = \frac{k}{\omega} \hat{\Phi}, \quad \hat{v} = \frac{l}{\omega} \hat{\Phi}, \quad \hat{w} = -\frac{\omega}{N^2} \left( m - \frac{i}{2H} \right) \hat{\Phi}, \quad (4.6.3a,b,c)$$

and the dispersion relation

$$\omega^2 = \frac{N^2(k^2 + l^2)}{m^2 + 1/4H^2}. \quad (4.6.4)$$



**Fig. 4.15.** High-frequency radar measurements of line-of-sight velocities at heights between 78 and 94 km in the upper mesosphere measured in two directions, equally inclined at small angles to the vertical. Top panels show data filtered to include only periods longer than 8 hr and bottom panels show data filtered to include only periods from 8 min to 8 hr. The data were collected during May 11-14, 1981. [After Vincent and Reid (1983).]



**Fig. 4.16.** Time sequence of northward and eastward velocity components as measured by very-high-frequency (VHF) radar during an unusually quasi-sinusoidal inertio-gravity wave event in the mesosphere, measured October 11, 1981, in Poker Flat, Alaska (1-hr average values). Velocity scale is shown in the lower left corner of the upper panel. Dashed lines indicate approximate height of velocity extrema and show downward phase propagation. Note that the left-to-right profile placements are not precisely uniform in time. [After Balsley *et al.* (1983). American Meteorological Society.]

Note that in terms of the equivalent depth  $h$  this can be written  $\omega^2 = gh(k^2 + l^2)$ , using Eq. (4.5.19); this is consistent with the “nonrotating” limit  $\gamma^{-1/2} \rightarrow \infty$  for gravity waves given by Eq. (4.2.14a) if  $k^2 + l^2$  is replaced by its spherical equivalent  $n_G(n_G + 1)a^{-2}$ .

Using methods analogous to those of Section 4.5.3, it can be shown that particles move in elliptical orbits (with tilted axes) in vertical planes perpendicular to the horizontal vector  $(-l, k, 0)$ .

It should be noted from Eqs. (4.6.2) and (4.6.3) that

$$\begin{aligned} \overline{\rho_0 u' w'} &= \frac{1}{2} \rho_0 \operatorname{Re}(e^{z/H} \hat{u} \hat{w}^*) = \frac{1}{2} \rho_s \operatorname{Re} \left[ \frac{k}{\omega} \hat{\Phi} \left( \frac{-\omega}{N^2} \right) \left( m + \frac{i}{2H} \right) \hat{\Phi}^* \right] \\ &= -\frac{1}{2} \rho_s \frac{mk}{N^2} |\hat{\Phi}|^2 \end{aligned}$$

where an asterisk denotes the complex conjugate. Thus  $\overline{\rho_0 u' w'}$  is independent of  $z$ , so that

$$\rho_0^{-1} (\rho_0 \overline{u' w'})_z = 0,$$

which is the Eliassen–Palm theorem [Eq. (3.6.1)] for the present steady, conservative, linear case, since  $\mathbf{F} = (0, -\rho_0 \overline{v' u'}, -\rho_0 \overline{w' u'})$  here [cf. Eq. (3.5.3)], and depends only on  $z$ .

As mentioned above, the small-scale gravity waves observed in the middle atmosphere tend to have vertical wavelengths less than about 15 km, so that  $4H^2 m^2 \geq 34$  if  $H = 7$  km, and thus  $m^2 \gg 1/4H^2$ . It is therefore reasonable to neglect  $1/4H^2$  compared with  $m^2$  in Eq. (4.6.4); this is equivalent to making the “Boussinesq approximation,” and gives  $\omega^2 = N^2(k^2 + l^2)/m^2$ . The solution with positive vertical group velocity  $c_g^{(z)} \equiv \partial\omega/\partial m$  is

$$\omega = -N(k^2 + l^2)^{1/2}/m. \quad (4.6.5)$$

Numerous other properties of internal gravity waves are given in standard texts. In particular, the frequency  $\omega$  is always smaller in magnitude than the buoyancy frequency  $N$ ; indeed, under the hydrostatic approximation of Eq. (3.1.3c),  $|\omega| \ll N$ , as shown for example by Gill (1982, Section 6.14), and thus the period is much greater than  $2\pi N^{-1}$  ( $\approx 5$  min in the middle atmosphere). The same inequality implies that the horizontal wavelength  $2\pi(k^2 + l^2)^{-1/2}$  must be much larger than the vertical wavelength  $2\pi|m|^{-1}$ , by Eq. (4.6.5). Another interesting property is that if  $i/2H$  is neglected in Eq. (4.6.3c)—a fair approximation for waves of vertical wavelength less than 15 km, for which  $2H|m| \geq 5.8$ —then Eqs. (4.6.3) and (4.6.5) give  $(\hat{u}, \hat{v}, \hat{w}) \cdot (k, l, m) \approx 0$ , so that the velocity  $(u', v', w')$  is perpendicular to the wave vector  $\mathbf{k} \equiv (k, l, m)$  and thus lies in planes of constant phase,  $kx + ly + mz = \text{constant}$ . In this case parcel orbits are straight lines, also perpendicular to the wave vector.

It is often convenient to choose the horizontal axes so that  $\mathbf{k} = (k, 0, m)$  and  $l = 0$ : this is possible since no preferred horizontal direction is imposed by the motionless basic state considered here. Then  $\hat{v} = 0$  by Eq. (4.6.3b), and Eq. (4.6.5) becomes

$$\omega = -Nk/m, \quad (4.6.5')$$

given the convention  $k > 0$ , as before. The vector group velocity is then

$$\mathbf{c}_g = \left( \frac{\partial\omega}{\partial k}, 0, \frac{\partial\omega}{\partial m} \right) = \frac{N}{m^2} (-m, 0, k), \quad (4.6.6)$$

and the tangent of the angle it makes with the horizontal has magnitude

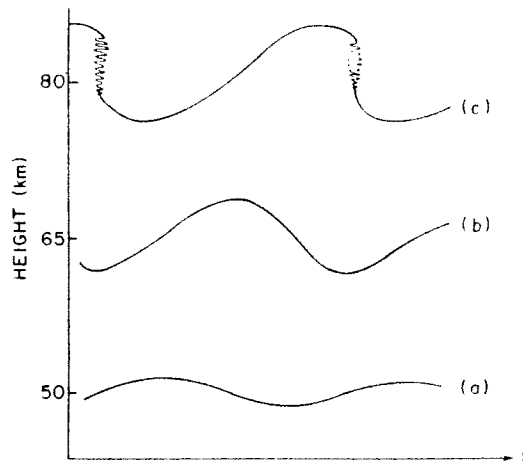
$$|c_g^{(z)}/c_g^{(x)}| = |k/m| = |\omega/N|.$$

The foregoing theory can be extended to allow for a basic flow  $[\bar{u}(z), \bar{v}(z), 0]$  and buoyancy frequency  $N(z)$  that vary with height (although in this case one cannot generally choose axes such that  $l = 0$ ). When these quantities vary only on height scales much greater than a vertical wavelength, *WKB* methods analogous to those of Appendix 4A and Section 4.5.4 can be used, and ray-tracing can be carried out. Once again critical levels, at which  $\omega = k\bar{u} + l\bar{v}$ , need special attention.

## 4.6.2 A Simple Model of Breaking Gravity Waves

Owing to the presence of the  $e^{-2H}$  factor, proportional to  $\rho_0^{1/2}$ , in Eq. (4.6.2), the linear, nondissipative theory of Section 4.6.1 predicts velocity and geopotential disturbances that grow with altitude; at some height the nonlinear terms that have been neglected will become important, and the linear theory will break down.

A physical picture of this breakdown can be obtained by considering a set of material surfaces at various levels, which are undulating as an internal gravity wave propagates vertically through them; Fig. 4.17 is a schematic diagram of this situation. In the lower mesosphere, say, the material surface (a) has a gentle sinusoidal variation, as predicted by linear theory. For gravity waves of period much less than a day the effects of radiative relaxation are small, and in the absence of other diabatic processes we can use an isentrope (a surface of constant  $\theta$ ) as the material surface (a). In the middle mesosphere the material surface (b) is also sinusoidal, but of larger amplitude than (a); linear theory still holds, and (b) can also be taken as an isentrope. In the upper mesosphere, however, nonlinear effects become important, leading to the rapid and irreversible deformation of material contours such as (c), followed by turbulence, small-scale mixing,



**Fig. 4.17.** Schematic diagram illustrating the breaking of vertically propagating internal gravity waves in the mesosphere. The curves labeled (a), (b), and (c) denote material surfaces. At the level of (a) and (b) the linear nondissipative theory of Section 4.6.1 is approximately valid. At the level of (c) nonlinear effects are important, with irreversible deformation of previously wavy material surfaces, and turbulence near the wave crests, presumably followed by small-scale mixing and dissipation.

and dissipation. Isentropes are no longer material surfaces, owing to the excitation of diabatic effects.

The process described here is known as *gravity-wave breaking*, by analogy with the overturning and breaking of oceanic surface waves on a shelving beach. It also has points in common with the phenomenon of planetary-wave breaking, described in Section 5.2.3. It will tend to limit the  $e^{z/2H}$  growth of gravity wave amplitudes with height, and this has important consequences for the large-scale flow in the middle atmosphere, as will be seen below.

A simple model of this breaking process was suggested by Lindzen (1981), who considered linearized disturbances to a basic zonal flow  $\bar{u}(z)$ , and used a *WKBJ* method to generalize the solutions of Eq. (4.6.2). He then defined the *breaking level*,  $z_b$ , to be that altitude at which the isentropes first become vertical, with  $\partial\theta/\partial z = 0$ , thus implying a loss of static stability and the onset of turbulence and mixing. From Eqs. (3.2.13) and (3.4.2c)

$$\theta_z = \bar{\theta}'_z + \theta'_z = HR^{-1}e^{\kappa z/H}[N^2 + \Phi'_{zz} + \kappa H^{-1}\Phi'_z], \quad (4.6.7)$$

and since  $\Phi'$ , as calculated by linear theory, grows exponentially with  $z$  [cf. Eq. (4.6.2)], we expect that  $-(\Phi'_{zz} + \kappa H^{-1}\Phi'_z)$  will at a sufficient altitude become large enough to cancel  $N^2$  at some values of  $x$ ,  $y$ , and  $t$ , at the "crests" of the waves. At such points  $\theta_z = 0$ , and breaking occurs, in Lindzen's sense. This approach is not strictly self-consistent, since the linear solutions will break down before such a height is reached; nevertheless it should give a qualitative feel for the fully nonlinear behavior.

In the special case where  $\bar{u} = 0$  we can use the linear theory of Section 4.6.1 to illustrate Lindzen's approach. We choose axes such that  $l = 0$  and thus  $v' = 0$ , and suppose that  $\Phi' = \Phi_0 \cos k(x - ct)$  at some lower level, which can be taken as  $z = 0$  by suitable choice of  $p_s$ . Thus the solutions of Eq. (4.6.2) apply, subject to Eq. (4.6.3) with  $l = 0$ ,  $\hat{\Phi} = \Phi_0$  (real), and  $\omega = ck$ ; in particular,

$$\Phi' = \Phi_0 e^{z/2H} \cos[k(x - ct) + mz]. \quad (4.6.8)$$

Assuming once more that vertical wavelengths are small enough to ensure that  $|m| \gg 1/2H$ , we have

$$\Phi'_{zz} + \kappa H^{-1}\Phi'_z \approx -m^2\Phi' = -m^2\Phi_0 e^{z/2H} \cos[k(x - ct) + mz].$$

The breaking level  $z_b$  is defined by  $\{\max|\Phi'_{zz} + \kappa H^{-1}\Phi'_z|\}_{z=z_b} = N^2$ , and so

$$z_b \approx 2H \ln \left| \frac{N^2}{m^2\Phi_0} \right| = 2H \ln |c^2\Phi_0^{-1}| \quad (4.6.9)$$

by Eqs. (4.6.8) and (4.6.5') with  $c = \omega/k$ . Note that  $z_b$  depends on the phase speed  $c$  and the initial geopotential amplitude  $\Phi_0$  of the wave; it increases as  $\Phi_0$  decreases because waves of small amplitude need to penetrate to

greater heights before they have grown enough to break than do waves of larger amplitude.

The next step is to describe what happens above the breaking level. The turbulence that presumably sets in leads to diffusion of heat and momentum, and a crude way of parameterizing this diffusion is to modify Eqs. (4.6.1a,d) thus:

$$u'_t + \Phi'_x = Ku'_{zz}, \quad \Phi'_{zt} + N^2 w' = K\Phi'_{zzz}, \quad z > z_b, \quad (4.6.10a,b)$$

where  $K$  is a constant diffusion coefficient; above  $z_b$  the continuity equation, Eq. (4.6.1c), still holds. On seeking solutions of the form

$$\Phi' = e^{z/2H} \operatorname{Re}[\Phi_1 \exp i(kx + m_1 z - \omega t)]$$

above  $z_b$ , chosen to match Eq. (4.6.8) at  $z_b$ , we find

$$\omega + im_1^2 K = -Nk/m_1 \quad (4.6.11)$$

by analogy with (4.6.5'), where the convention  $k > 0$  is retained. Thus if  $K$  is chosen so small that

$$m^2 K \ll |\omega|, \quad (4.6.12)$$

it follows that

$$m_1 \approx -\frac{Nk}{\omega} + \frac{iKN^3 k^3}{\omega^4} = m + \frac{iKN^3}{c^4 k}$$

and so

$$\Phi' = \Phi_0 \exp \left[ \frac{z}{2H} - \frac{KN^3}{c^4 k} (z - z_b) \right] \cos[k(x - ct) + mz], \quad z > z_b, \quad (4.6.13)$$

to satisfy continuity with Eq. (4.6.8) at  $z_b$ . The extra exponential factor involving  $K$  results from the postulated diffusive damping of the waves due to breaking. Lindzen hypothesizes that above  $z_b$  the waves are *saturated*, or just on the verge of breaking; thus  $\max|\Phi'_{zz} + \kappa H^{-1}\Phi'_z| = N^2$  for all  $z \geq z_b$ . Hence the diffusive decay must be such as to exactly balance the  $e^{z/2H}$  growth, and  $K$  must be chosen such that the coefficient of  $z$  in the exponential term in Eq. (4.6.13) vanishes, that is,

$$K = c^4 k / 2HN^3. \quad (4.6.14)$$

Note that this implies

$$m^2 K = \frac{N^2 k^2}{\omega^2} \left( \frac{c^4 k}{2HN^3} \right) = \frac{c^2 k}{2HN} \left| \frac{\omega}{2Hm} \right|$$

by Eq. (4.6.5'), and thus Eq. (4.6.12) is consistent with our previous assumption that  $|m| \gg 1/2H$ . Substitution of Eq. (4.6.14) into Eq. (4.6.13) gives

$$\Phi' = \Phi_0 e^{\tilde{\nu}/2H} \cos[k(x - ct) + mz] \quad \text{for } z \geq z_b, \quad (4.6.15a)$$

while Eqs. (4.6.10a,b) yield

$$(u', w') = (1, -k/m)c^{-1}\Phi_0 e^{\tilde{\nu}/2H} \cos[k(x - ct) + mz] \quad \text{for } z \geq z_b \quad (4.6.15b,c)$$

at leading order in the small parameter  $m^2K/\omega$ . Thus Lindzen's saturation hypothesis implies no further growth of wave amplitude above  $z_b$ ; this is roughly consistent with observed gravity-wave amplitudes in the mesosphere. Note, incidentally, that  $\max|u'| = |c|$  for  $z \geq z_b$ , from Eqs. (4.6.15b) and (4.6.9).

We now calculate the quantity

$$\bar{X}_1 \equiv -\rho_0^{-1}(\rho_0 \overline{u'w'})_z; \quad (4.6.16)$$

using Eqs. (4.6.15b,c), (4.6.9), (4.6.5'), and (4.6.14) it follows that

$$\bar{X}_1 = c^3k/2NH = \frac{N^2K}{c} \quad \text{for } z > z_b, \quad (4.6.17)$$

which is a nonzero constant in general. Thus if  $c > 0$  there is a zonal-mean vertical wave momentum flux convergence or Eliassen-Palm flux divergence above the breaking level (and vice versa if  $c < 0$ ), and this implies a contribution to the wave-induced forcing of the zonal-mean flow. Lindzen also postulates that the momentum and heat diffusion represented by  $K$  acts on the zonal-mean flow, as well as on the gravity waves.

As mentioned above, Lindzen's method is more general than that given here, in that it includes a slowly varying mean flow  $[\bar{u}(z), 0, 0]$  and nonzero meridional wave number  $l$ . The generalization of Eq. (4.6.17) is

$$\bar{X}_1 = \frac{(c - \bar{u})^3 k}{2N} \left[ \frac{1}{H} + \frac{3 d\bar{u}/dz}{c - \bar{u}} \right] = \frac{N^2K}{c - \bar{u}} \quad (4.6.18a,b)$$

when  $l = 0$  but  $\bar{u} \neq 0$ . On substituting typical wave parameters, such as those given at the beginning of Section 4.6, into these formulas, it is found that Lindzen's parameterization of the possible frictional effects due to breaking gravity waves implies a strong forcing of the zonal-mean circulation of the upper mesosphere, in accordance with observations that indicate that  $\bar{X}_1$  may be on the order of several tens of meters per second per day (see Fig. 4.18). This topic will be discussed further in Sections 7.3 and 8.5.

It is also possible to allow for the radiative damping of gravity waves in the above theory. Since damping tends to decrease the amplitude of the

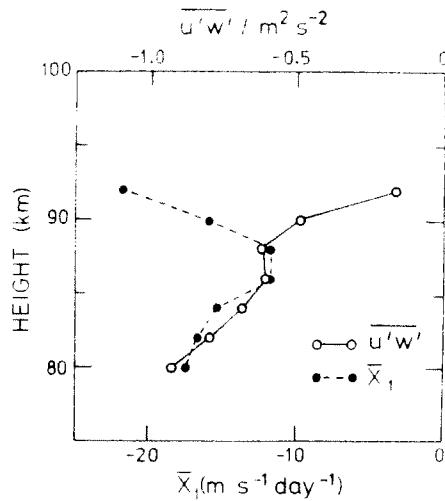


Fig. 4.18. Height profiles of  $\overline{u'w'}$  and  $\overline{X_1}$ , in the upper mesosphere, derived from double-beam radar measurements in May 1981. [After Vincent and Reid (1983).]

upward-propagating waves, the breaking level tends to be raised; indeed, waves of sufficiently small intrinsic phase speed may not break at all. However, if they do break, the radiative damping has little effect on the resulting values of  $\overline{X_1}$ .

It has been emphasized that Lindzen's model is highly simplified, being essentially based on a linear theory of monochromatic gravity waves. Much more observational and theoretical work needs to be done to investigate the validity of the model and to understand the detailed nonlinear dynamics of a complex spectrum of breaking gravity waves (for which a "breaking amplitude" is more appropriate than a "breaking level") and their effects on the mean flow.

#### 4.6.3 Inertio-Gravity Waves

On somewhat larger space and time scales (horizontal wavelengths  $\sim 1000$  km and periods of several hours) than those considered in Section 4.6.1, gravity waves will be influenced by the rotation of the earth, and the theory given there will need modification. As a simple example that avoids the complexity of the spherical geometry used in Section 4.2, we consider small disturbances to a state of rest on an " $f$ -plane," in which the Coriolis