Chapter 7 | The Extratropical Zonal-Mean Circulation

7.1 Introduction

In Section 5.2.1 we presented some aspects of the observed climatology of the zonal-mean circulation of the middle atmosphere, its annual cycle, and its interhemispheric variations. For example, Fig. 5.1 showed the existence, in the zonal mean, of a cold tropical tropopause, a warm stratopause, and a cold summer mesopause, while the corresponding mean zonal geostrophic winds, illustrated in Fig. 5.2, are generally westerly in winter and easterly in summer, decreasing to small values near the mesopause. In the present chapter we examine the processes that maintain the climatological zonal-mean state in the extratropics. We concentrate especially on the ways in which dynamical phenomena can lead to large departures, in certain parts of the middle atmosphere, from a hypothetical climatology determined solely by radiative and photochemical effects.

To investigate the role played by dynamical processes in producing the observed middle atmosphere circulation, it is first useful to consider what form the circulation would take in the absence of dynamical processes, other than some representation of convection, and perhaps baroclinic wave activity, in the troposphere. The temperature field associated with such a circulation can be calculated from a radiative–photochemical model of the stratosphere and mesosphere, together with a radiative–convective model of the troposphere. An example for near-solstice conditions was given in Fig. 1.2, from the time-marched calculations decribed by Fels (1985). This shows strong latitude and height variations of the resulting zonally symmetric temperature, with a maximum of about 290 K at the summer

stratopause and temperatures below 180 K throughout the middle atmosphere at the winter pole, decreasing to 130 K at the winter mesopause. The temperature field $T_r(\phi, z, t)$ say, calculated in this way will be referred to as the *radiatively determined temperature*.

A comparison of the calculated radiatively determined temperature field in Fig. 1.2 with the observed January-mean field of Fig. 5.1a reveals some overall similarities between the two, but also some striking differences. For example, the observed midlatitude summer stratopause temperature, at about 280 K, is close to T_r, but the observed north polar night is much warmer than the corresponding T_r (by about 30 K in the lower stratosphere, increasing to 100 K in the mesosphere), while the observed southern summer mesopause is much colder (by about 60 K). The July-mean zonally averaged observed temperature of Fig. 5.1c is roughly a mirror image of Fig. 5.1a, except that the southern winter polar midstratosphere, at about 180 K, is only just above the radiatively determined value.

Assuming thermal-wind balance,

$$\left(f + \frac{2u_{gr}\tan\phi}{a}\right)\frac{\partial u_{gr}}{\partial z} = -\frac{R}{aH}\frac{\partial T_r}{\partial \phi} \qquad (7.1.1)$$

[cf. Eqs. (3.4.1c) and (5.2.2)], and a suitable lower boundary condition, one may calculate the zonal gradient wind, u_{gr} say, that would be associated with the radiatively-determined temperature T_r. Figure 7.1 shows the $u_{gr}(\phi, z)$ field corresponding to the temperature field of Fig. 1.2, given that u_{gr} equals the observed climatological values at 100 mb. This "radiatively determined" gradient wind exhibits extremely strong westerlies in the winter polar night, associated with the strong latitudinal gradients of T_r there, and quite strong easterlies in the summer hemisphere. The magnitudes of the winds increase with height throughout the stratosphere and mesosphere. In contrast, the observed zonal-mean geostrophic zonal winds for January (Fig. 5.2a) show more moderate growth with height, peaking near 60 km altitude in both hemispheres and decreasing to small values near the mesopause: the observed jets "close off" in the upper mesosphere. In July (Fig. 5.2c) the Southern Hemisphere winter westerlies are stronger than their Northern Hemisphere counterparts of January, but still decrease above 60 km altitude. (Note that observed zonal-mean geostrophic zonal winds are usually small enough to be a good approximation to the observed gradient winds.)

It is possible that some of the differences between the climatological observations of Figs. 5.1 and 5.2 and Fels's time-marched radiative–photo-chemical–convective calculations of Figs. 1.2 and 7.1 may be due to deficiencies in the radiative formulation of the latter. However, by far the most important reason for these differences is the presence of dynamical processes
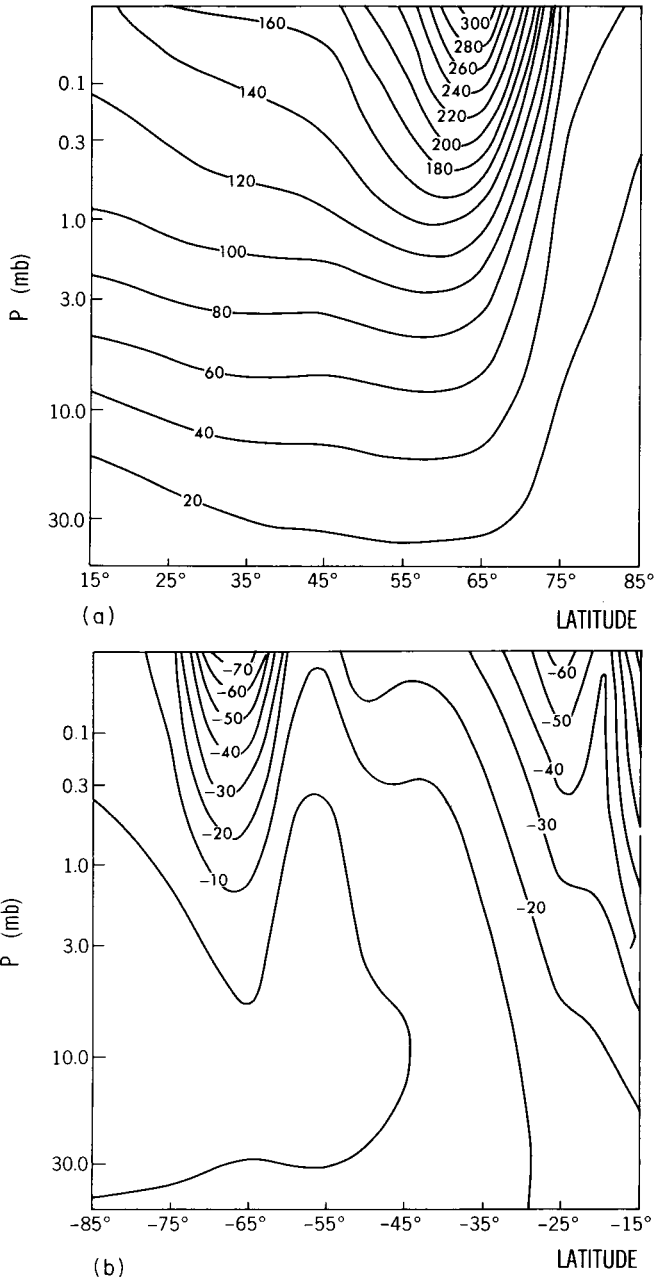
Fig. 7.1. Zonal gradient wind u_{gr} that is in thermal-wind balance with the temperature field T_r of Fig. 1.2 and equals the observed climatological zonal wind at 100 mb. (a) Northern Hemisphere (winter), (b) Southern Hemisphere (summer). (Courtesy of Dr. S. B. Fels.)

in the middle atmosphere; such processes were deliberately excluded from Fels's calculations. The extra heating or cooling that must be provided by the dynamical thermal transport is often called the "dynamical heating." Some of the dynamical processes contributing to this heating would occur in a middle atmosphere whose circulation was zonally symmetric. However, simple arguments suggest that the dynamical phenomena of greatest importance for accounting for departures from T_r,—in the extratropics at least—are associated with deviations from zonal symmetry: the "eddies" or "waves." The following sections discuss which types of wave are likely to be involved in this process, and the means by which they may force departures of $\bar{T}$ from T_r.

7.2   Some Simple Zonally Averaged Models of the Middle Atmosphere

To gain insight into the ways in which dynamical processes can lead to departures of the zonal-mean temperature from the temperature $T_r(\phi, z, t)$ of a hypothetical atmosphere controlled only by radiative, photochemical, and convective effects, it is helpful to begin by considering a hierarchy of rather simple models of the extratropical middle atmosphere. A suitable starting point is the set of quasi-geostrophic "transformed Eulerian-mean" (TEM) equations on a beta-plane (see Section 3.5). Since comparison with observed temperatures will be made, the zonal-mean temperature $\bar{T}$ will be used as a dependent variable, rather than the zonal-mean potential temperature $\bar{\theta} = \bar{T}e^{\kappa z/H}$. Then, with $\bar{J}/c_p \equiv \bar{Q}e^{-\kappa z/H}$ (see Section 3.1.1) and $N^2 \equiv H^{-1}R\theta_{0z}e^{-\kappa z/H}$ [see Eq. (3.2.13)], the quasi-geostrophic TEM set [Eqs. (3.5.5)] becomes

$$\bar{u}_t - f_0\bar{v}^* = \rho_0^{-1}\nabla \cdot \mathbf{F} + \bar{X} \equiv \bar{G}, \tag{7.2.1a}$$

$$\bar{T}_t + N^2HR^{-1}\bar{w}^* = \bar{J}/c_p, \tag{7.2.1b}$$

$$\bar{v}_y^* + \rho_0^{-1}(\rho_0\bar{w}^*)_z = 0, \tag{7.2.1c}$$

$$f_0\bar{u}_z + H^{-1}R\bar{T}_y = 0. \tag{7.2.1d}$$

It is convenient here to make a slight physical distinction from the system described by Eqs. (3.5.5) by reinterpreting the terms contributing to the zonal force per unit mass $\bar{G}$. [Note that $\bar{G}$ is not to be confused with the quantity G in Eq. (3.5.2b).] Thus $\mathbf{F}$ is now regarded as containing not only a contribution

$$\mathbf{F}_{(p)} \equiv (0, -\rho_0\overline{u'v'}, \rho_0 f_0\overline{v'\theta'}/\theta_{0z}) \tag{7.2.2a}$$

from planetary waves [cf. Eq. (3.5.6)], but also a contribution

$$\mathbf{F}_{(g)} \equiv (0, -\rho_0 \overline{u'v'}, -\rho_0 \overline{u'w'}) \tag{7.2.2b}$$

from small-scale gravity waves [cf. Eq. (3.5.3b)]. The term $\bar{X}$ now represents all further contributions to the mean zonal force per unit mass associated with gravity waves and other small-scale disturbances. The term $\bar{J}$ is the zonal-mean diabatic heating rate per unit mass, and will be assumed here to equal the zonal-mean net *radiative* heating rate: the "wave heating" term on the right of Eq. (3.5.2e) is negligible for quasi-geostrophic motions and also for gravity waves, and will be ignored here together with all other wave-induced and molecular contributions to $\bar{J}$. As in Section 3.5, $(\bar{v}^*, \bar{w}^*)$ is the residual circulation, defined by

$$\bar{v}^* \equiv \bar{v}_a - \rho_0^{-1}(\rho_0 \overline{v'\theta'} / \theta_{0z})_z, \qquad \bar{w}^* \equiv \bar{w}_a + (\overline{v'\theta'} / \theta_{0z})_y, \tag{7.2.3}$$

where $(\bar{v}_a, \bar{w}_a)$ is the Eulerian zonal-mean (ageostrophic) meridional circulation. It should be noted that $\overline{v'\theta'}/\theta_{0z}$ can alternatively be written as $\overline{v'T'}/N^2 HR^{-1}$ in Eqs. (7.2.2a) and (7.2.3). Several advantages of the set of Eqs. (7.2.1) over the conventional Eulerian-mean set were mentioned in Section 3.5.

7.2.1   Steady-State Model

We first consider a hypothetical steady-state atmosphere in which the seasonal cycle is absent; with time derivatives set to zero, Eqs. (7.2.1a,b) give

$$-f_0 \bar{v}^* = \bar{G}, \qquad N^2 HR^{-1}\bar{w}^* = \bar{J}/c_p, \tag{7.2.4a,b}$$

and substitution into Eq. (7.2.1c) yields

$$-\bar{G}_y + f_0 \rho_0^{-1}(\rho_0 \bar{J}\kappa / N^2 H)_z = 0 \tag{7.2.4c}$$

(since $\kappa = R/c_p$), showing how the diabatic heating rate $\bar{J}$ must be related to $\bar{G} \equiv \rho_0^{-1} \nabla \cdot \mathbf{F} + \bar{X}$ in this hypothetical state. [Note that Eq. (7.2.4c) is a steady-state version of the quasi-geostrophic potential vorticity equation; this can be seen by substituting Eq. (3.5.10) into Eq. (3.3.4).] If all eddy and small-scale effects are absent, so that $\bar{G}$ vanishes, then $\bar{v}^* = 0$ by Eq. (7.2.4a): the continuity equation [Eq. (7.2.1c)] then implies that $|\bar{w}^*|$ grows exponentially with z, in general. To prevent this unphysical behavior we can impose the boundary condition $\bar{w}^* = 0$ at a lower boundary $z = z_0$, say. Then $\bar{w}^* = 0$ everywhere and so $\bar{J} = 0$ everywhere, by Eq. (7.2.4b). The atmosphere is then in radiative equilibrium, under our assumption that $\bar{J}$ is the net *radiative* heating rate: thus, the temperature $\bar{T}$ must equal the the radiatively determined value $T_r(y, z)$, so that the long-wave cooling

everywhere balances the solar heating. Moreover, $\bar{u} = u_r$, where u_r is the geostrophic radiatively determined wind, satisfying

$$f_0 \frac{\partial u_r}{\partial z} + H^{-1} R \frac{\partial T_r}{\partial y} = 0 \qquad (7.2.5)$$

[see Eq. (7.2.1d), but contrast Eq. (7.1.1)], and since $\bar{v}^* = \bar{w}^* = 0$ and eddies are absent, $\bar{v}_a = \bar{w}_a = 0$ as well, by Eq. (7.2.3).

7.2.2   Annually Varying Model with No Waves

We next include time-dependence by letting the solar heating rate take on an annual variation, $\bar{J}_s(y, z, t)$, say, but still assume that eddy and small-scale effects are absent, so that $\bar{G} = 0$. Further qualitative insight can be obtained by parameterizing $\bar{J}$ in terms of $\bar{T}$. As a simple example we use the Newtonian cooling form

$$\bar{J}/c_p = -[\bar{T} - T_r(y, z, t)]/\tau_r(z), \qquad (7.2.6)$$

where $T_r(y, z, t)$ is the temperature calculated from a time-dependent radiative–photochemical model (such as that from which Fig. 7.1 was obtained) with specified solar heating $\bar{J}_s(y, z, t)$, and $\tau_r(z)$ is a radiative relaxation time. This parameterization is not expected to be quantitatively accurate for large departures of $\bar{T}$ from T_r; however, it does contain the important physical feature of relating the net heating rate to departures of $\bar{T}$ from the radiatively determined temperature T_r.

Using Eqs. (7.2.1) with $\bar{G} = 0$ and $\bar{J}/c_p$ given by Eq. (7.2.6), it can be shown that

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial}{\partial z}\left(\rho_0^{-1} \frac{\partial}{\partial z} \rho_0 \varepsilon \right) \right] \bar{T}_t + \left\{ \rho_0^{-1} \left[\frac{\rho_0 \varepsilon}{\tau_r} (\bar{T} - T_r) \right]_z \right\}_z = 0, \quad (7.2.7)$$

where $\varepsilon(z) = f_0^2/N^2(z)$, as in Eq. (3.2.16). In this equation the term in T_r (which depends on the solar heating $\bar{J}_s$) provides the *forcing*, to which $\bar{T}$ is the *response*. In general $\bar{T}$ will follow $T_r(y, z, t)$ but will be somewhat lagged in time (since the relaxation time τ_r is nonzero) and somewhat more smoothly distributed in space [because of the properties of the elliptic operator on the far left of Eq. (7.2.7)]: the zonally symmetric dynamics thus provide a kind of "inertia." Since $\bar{T}$ is not equal to T_r in general, Eq. (7.2.6) implies that the net radiative heating rate does not generally vanish; by Eqs. (7.2.1) the residual circulation $(\bar{v}^*, \bar{w}^*)$ does not vanish, either. This idealized model makes it clear that the nonvanishing of the *net* heating rate $\bar{J}$ is essentially due to the presence of the "dynamical inertia," and cannot be regarded as *imposed* by external radiative agencies. To put it another way,

although the solar heating $\bar{J}_s$ has been specified in advance in this model, the "long-wave cooling" $\bar{J} - \bar{J}_s$ must be determined as part of the solution.

Assuming vertical scales of order H, where H is the scale height, and horizontal scales of order L, where $f_0^2 L^2 \sim N^2 H^2$, we can apply a simple order-of-magnitude argument to Eq. (7.2.7) to obtain

$$\frac{\partial \bar{T}}{\partial t} \sim (\bar{T} - T_r)/\tau_r \qquad (7.2.8)$$

approximately, in the present model. (This argument does not apply near the equator, where $f_0^2 L^2 \ll N^2 H^2$: see Section 8.1 for a discussion of the relevant time-dependent balance in the tropics.) If $\Delta \bar{T}(y, z)$ is the maximum annual variation of $\bar{T}$ in the model and τ is a seasonal timescale (say 3 months), an estimate of the temperature tendency is

$$\partial \bar{T}/\partial t \sim \Delta \bar{T}/\tau. \qquad (7.2.9)$$

However, typical radiative relaxation times are usually a few days ($\tau_r \ll \tau$), so that

$$\bar{T} - T_r \sim \frac{\tau_r}{\tau} \Delta \bar{T} \ll \Delta \bar{T} \qquad (7.2.10)$$

from Eqs. (7.2.8) and (7.2.9), and thus departures of $\bar{T}$ from the annually varying radiatively determined value $T_r(y, z, t)$ are much smaller in this model than the amplitude of the annual swing $\Delta \bar{T}$. It follows that these departures are also much smaller in magnitude than the annual swing ΔT_r in the radiatively determined temperature, which can be estimated by comparison of opposite hemispheres in Fig. 7.1. The model, therefore, predicts extratropical temperatures $\bar{T}(y, z, t)$ that are always "close" to the annually varying temperatures $T_r(y, z, t)$ determined from radiative-photo-chemical-convective considerations alone.[1]

On the other hand, an examination of the fields of T_r in Fig. 1.2 and the observed climatological $\bar{T}$ in Fig. 5.1a shows that in January at 0.01 mb (in the upper mesosphere, near 80 km altitude), $\bar{T} - T_r \approx 60$ K at 60°N and -45 K at 60°S, values comparable with that of $\Delta T_r \approx 50$ K (as obtained by subtracting T_r at 60°N from T_r at 60°S). Similarly, at 10 mb and 60°N, in the winter middle stratosphere, $\bar{T} - T_r \approx 45$ K, comparable with $\Delta T_r \approx 55$ K for that latitude and height. Clearly, the model described in the present section fails to predict these *large* departures of $\bar{T}$ from T_r in the upper

---

[1] The same conclusion holds if Eq. (7.2.6) is regarded as an order-of-magnitude estimate, rather than an equality: the argument does not depend on the precise form of the parameterization of $\bar{J}$, but only on the fact that the radiative timescale is much less than the seasonal timescale.

mesosphere and winter stratosphere. Additional effects must be included if a basic understanding of the observed annual variations of the temperature structure of these regions is to be obtained.

7.2.3   Inclusion of Wave-Forcing Effects

To obtain the next model in our hierarchy, we now suppose that $\bar{G}$ is nonzero. The set of Eqs. (7.2.1) with parameterization as in Eq. (7.2.6) yields the equation

$$\left[\frac{\partial^2}{\partial y^2} + \frac{\partial}{\partial z}\left(\rho_0^{-1}\frac{\partial}{\partial z}\rho_0\varepsilon\right)\right]\bar{T}_t + \left\{\rho_0^{-1}\left[\frac{\rho_0\varepsilon}{\tau_r}(\bar{T} - T_r)\right]_z\right\}_z + f_0 H R^{-1}\bar{G}_{yz} = 0$$

$$[1] \qquad\qquad\qquad\qquad [2] \qquad\qquad\qquad\qquad [3] \qquad (7.2.11)$$

for $\bar{T}$ and, using Eq. (7.2.5), the equation

$$\left[\frac{\partial^2}{\partial y^2} + \rho_0^{-1}\frac{\partial}{\partial z}\left(\rho_0\varepsilon\frac{\partial}{\partial z}\right)\right]\bar{u}_t + \rho_0^{-1}\left[\frac{\rho_0\varepsilon}{\tau_r}(\bar{u} - u_r)_z\right]_z - \bar{G}_{yy} = 0 \quad (7.2.12)$$

for $\bar{u}$: the latter is a form of Eq. (3.5.7). [Note that the elliptic operator acting on $\bar{u}_t$ in Eq. (7.2.12) differs slightly from that acting on $\bar{T}_t$ in Eqs. (7.2.11) and (7.2.7).]

We now suppose that the wave forcing represented by $\bar{G}$ varies on a timescale τ_w: except for rapid events like sudden warmings we can expect τ_w to be comparable to the seasonal timescale τ and much greater than the radiative timescale τ_r. If rapid wave events like sudden warmings *are* present, the application of a running time average over a period $\tau_w = O(\tau)$ will remove shorter-period fluctuations in Eq. (7.2.11), leaving only smooth variations of timescale $O(\tau_w)$. In either case a scaling argument similar to the one that was applied to Eq. (7.2.7) gives the ratio of terms in Eq. (7.2.11) as

$$[1] : [2] : [3] = \frac{\Delta\bar{T}}{\tau} : \frac{\bar{T} - T_r}{\tau_r} : f_0 L R^{-1}\Delta\bar{G} \qquad (7.2.13)$$

if $f_0^2 L^2 \sim N^2 H^2$ again (thus excluding equatorial regions once more) and $\Delta\bar{G}$ is a typical variation in $\bar{G}$ over a time τ_w. As mentioned in the previous section, it is generally found that $(T - \bar{T}_r)/\tau_r \gg \Delta T/\tau$ in the polar night and the summer upper mesosphere, so in these regions the term [1] in Eq. (7.2.11) is small, and the effects represented by $\bar{G}$ must be large enough to give a balance between terms [2] and [3]. Equivalently, the time derivatives in Eqs. (7.2.1a,b) are small, so that—according to this model—the balances expressed by Eq. (7.2.4) hold approximately, at each t, in those regions where $\bar{T}$ exhibits large departures from T_r. This is especially true near the

solstices, when time derivatives are expected to be small, in any case; for example, $\bar{T}_t$ is then smaller than the estimate of Eq. (7.2.9).

This simple model suggests that dynamical effects contributing to the mean zonal force per unit mass $\bar{G} \equiv \rho_0^{-1} \nabla \cdot \mathbf{F} + \bar{X}$ are the primary agents responsible for maintaining the large departures of $\bar{T}$ from T_r that are observed in parts of the middle atmosphere, and thus for producing the net radiative heating rates $\bar{J}$ that occur there. At the same time, these dynamical processes also drive the residual circulation, as can be seen from the continuity equation [Eq. (7.2.1c)] together with the approximate equation Eq. (7.2.4a) or more generally from Eq. (3.5.8), which can be written as

$$\left[\frac{\partial^2}{\partial y^2} + \rho_0^{-1} \frac{\partial}{\partial z} \left(\rho_0 \varepsilon \frac{\partial}{\partial z} \right) \right] f_0 \bar{v}^* + \rho_0^{-1} (\rho_0 \varepsilon \bar{G}_z)_z + \rho_0^{-1} \left(\frac{\rho_0 f_0 \kappa \bar{J}}{N^2 H} \right)_{yz} = 0$$

(7.2.14)

in the present notation. [The transformation of Eq. (7.2.3) shows that the Eulerian-mean meridional circulation is also dynamically driven.] Conversely, the model also suggests that those regions that are observed to be close to radiative equilibrium, such as parts of the midlatitude lower stratosphere and the summer stratopause, are in such a state because of the absence of any dynamical effects that can lead to significant values of $\bar{G}$ there.

The residual circulation in the middle atmosphere can be calculated *diagnostically* using *observed* temperatures, an accurate radiative heating algorithm for $\bar{J}$, the spherical-geometry equivalents of Eqs. (7.2.4b) [or, more accurately, Eq. (7.2.1b)] and (7.2.1c), and appropriate boundary conditions. A pioneering calculation of this kind was performed by Murgatroyd and Singleton (1961), who attempted to estimate the Eulerian-mean meridional circulation $(\bar{v}, \bar{w})$ by making what is now known to be an unjustifiable approximation, namely that the eddy heating terms on the right of the Eulerian-mean thermodynamic equation [Eq. (3.3.2e)] are negligible. However, under this assumption Eq. (3.3.2e) becomes essentially isomorphic to Eq. (7.2.1b); the continuity equation [Eq. (3.3.2d)] is of the same form as Eq. (7.2.1c), and so the Murgatroyd and Singleton circulation, sometimes also called a "diabatic circulation" (see Section 9.3.1), is a close approximation to $(\bar{v}^*, \bar{w}^*)$. It is shown schematically in Fig. 7.2: generally rising motion takes place above 30 km in the summer hemisphere, with flow from the summer to the winter hemisphere in the upper stratosphere and mesosphere, and descent in the winter hemisphere. The lower stratospheric circulation is more symmetric about the equator, with rising motion at low latitudes and descent in high latitudes.

It must be emphasized that the Murgatroyd and Singleton diagnostic calculation assumes $\bar{T}(y, z, t)$ to be given, and then derives $(\bar{v}^*, \bar{w}^*)$. It does
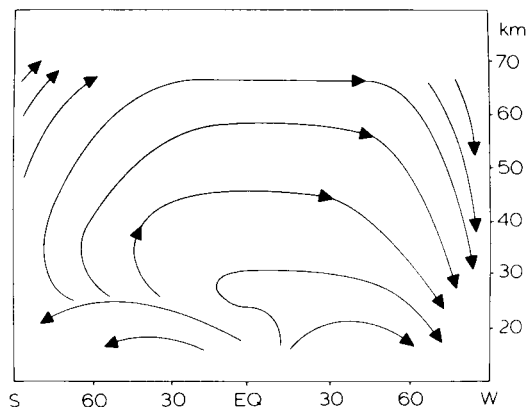
Fig. 7.2. Schematic streamlines of the solstice diabatic circulation in the middle atmosphere, as obtained from the Murgatroyd and Singleton (1961) calculation. S, summer pole; W, winter pole. [After Dunkerton (1978). American Meteorological Society.]

not require details of the eddy-forcing $\bar{G}$. A more predictive approach is to solve the complete set of mean-flow equations for $\bar{u}$, $\bar{T}$, $\bar{v}^*$, and $\bar{w}^*$, given some knowledge of $\bar{G}$. A classic study of this type was that of Leovy (1964b), who parameterized $\bar{G}$ in terms of $\bar{u}$ by assuming a linear "Rayleigh friction" drag,

$$\bar{G} = -\bar{u}/\tau_{\mathrm{m}}, \qquad (7.2.15)$$

where τ_{m} is a constant mechanical relaxation time, and used an expression of the form of Eq. (7.2.6) for $\bar{J}$. A variety of choices of τ_{r} and τ_{m} were examined: with $\tau_{\mathrm{r}} = \tau_{\mathrm{m}} \approx 15$ days, for example, several basic features of the seasonal cycle of $\bar{T}$ and $\bar{u}$ were simulated, although the details of the polar night stratosphere were not captured.

Current wave, mean-flow interaction theory leads us to expect that the parameterization of Eq. (7.2.15) is not likely to be very accurate, even with a relaxation time τ_{m} that depends on z. For some types of wave (e.g., gravity waves: see Sections 4.6.2 and 7.3), improved parameterizations are now available; for others, no satisfactory alternative parameterization has yet come to light.

A different, more qualitative, approach makes use of the result of Section 3.6 that the quantity $\nabla \cdot \mathbf{F}$ depends on wave transience, nonconservative wave effects, and wave nonlinearity. Thus, if waves are in some sense strongly transient, nonconservative, or nonlinear (or any combination of these) in a particular region, we can anticipate large local values of $\rho_0^{-1} \nabla \cdot \mathbf{F}$ there and perhaps large local departures of $\bar{T}$ from T_{r}. (Under the same conditions large values of $\bar{X}$ often tend to occur as well.) The identification

of the mechanisms responsible for local deviations of the climatological temperature fields from the radiatively determined value thus amounts to a search for wave motions that significantly violate the "nonacceleration conditions" in the relevant regions of the middle atmosphere. Likely candidates in the mesosphere and in the winter stratosphere are discussed in the next two sections.

7.3 The Upper Mesosphere

As noted in Section 7.1, a comparison of Figs. 5.1a,c with Fig. 1.2 reveals that the observed climatological temperature field $\bar{T}$ at the solstices in the upper mesosphere is much warmer than the radiatively determined value T_r in midwinter and much colder in midsummer. In accordance with the arguments of Section 7.2.3 we consider in this section what wave motions could significantly break the nonacceleration conditions in this part of the atmosphere, in a climatological sense, and thus lead to the large climatological values of $\bar{G} = \rho_0^{-1} \nabla \cdot \mathbf{F} + \bar{X}$, on the order of $100 \text{ m s}^{-1} \text{ day}^{-1}$, required to account for such discrepancies.

It is now generally believed that gravity waves provide the main part of the necessary forcing $\bar{G}$ in the upper mesosphere: as discussed in Section 4.6.2, such waves grow in amplitude as they propagate from tropospheric source regions up into the rarefied mesosphere, and "break," leading to turbulence, small-scale mixing, and dissipation. Associated with these waves there is an Eliassen–Palm flux divergence $\nabla \cdot \mathbf{F}_{(g)} \approx -\rho_0^{-1} (\overline{\rho_0 u' w'})_z$ (called $\bar{X}_1$ in Section 4.6.2), representing a zonal force per unit mass on the zonal-mean flow: the mixing induced by the waves also leads to diffusion of mean-flow properties. Parameterizations of these processes, due to Lindzen (1981), are given in Eq. (4.6.18). The net contribution to the forcing $\bar{G}$ associated in this way with gravity waves tends to drag the mean flow toward the horizontal phase speed of the waves.

Since gravity waves are absorbed at or near critical levels, where the local horizontal wind speed equals their horizontal phase speed, the only gravity waves reaching the mesosphere from below are expected to have phase speeds outside the range of horizontal wind speeds occurring in the underlying stratosphere. Thus, when winter westerly winds are present in the stratosphere we can anticipate that gravity waves within a range of easterly phase speeds will occur in the mesosphere and break there. Conversely, when summer easterly winds are present in the stratosphere, gravity waves with westerly phase speeds would be expected to appear, and break, in the mesosphere. In winter, the breaking gravity waves of easterly phase speed will exert an easterly force on the westerly jet, and hence tend to

close it off; similarly, the breaking gravity waves in summer will tend to
close off the easterly jet. (An analogous "shielding effect" has also been
postulated as an explanation of the mesospheric equatorial semiannual
oscillation: see Section 8.5.2.)

A more detailed argument is as follows. Consider the solstices [see Figs.
5.2a,c]: in the winter mesosphere, above the peak of the westerly jet, the
gravity-wave contribution to $\bar{G}$ is negative, while in the summer mesosphere,
above the peak of the easterly jet, $\bar{G}$ is positive. By the approximate Eq.
(7.2.4a) we expect $\bar{v}^*$ to be directed from summer to winter hemisphere at
these altitudes: mass continuity then requires that $(\rho_0 \bar{w}^*)_z > 0$ in the winter
mesosphere and $(\rho_0 \bar{w}^*)_z < 0$ in the summer. Assuming that $\bar{w}^*$ does not
vary rapidly over a pressure scale-height H (or, alternatively, that $\bar{w}^* = 0$
at some level at or above the mesopause), this implies descent in the winter
mesosphere and ascent in the summer (see Fig. 7.2). By Eqs. (7.2.4b) and
(7.2.6), this means that $\bar{T}$ should be warmer than T_r in the winter upper
mesosphere and cooler than T_r in the summer, in accordance with the
observations. By the thermal wind relation of Eq. (7.2.1d), the reversed
latitudinal gradients in $\bar{T}$ imply reversed vertical shears in $\bar{u}$ and hence a
closing-off of the jets.

Experiments with Lindzen's parameterizations in fairly simple models of
the middle atmosphere suggest that the gravity-wave drag and diffusion
contributions to $\bar{G}$ can account satisfactorily for much of the departure of
$\bar{T}$ from T_r in the mesosphere, given reasonable values of the wave para-
meters: see Fig. 7.3. The magnitude of the computed drag in such models
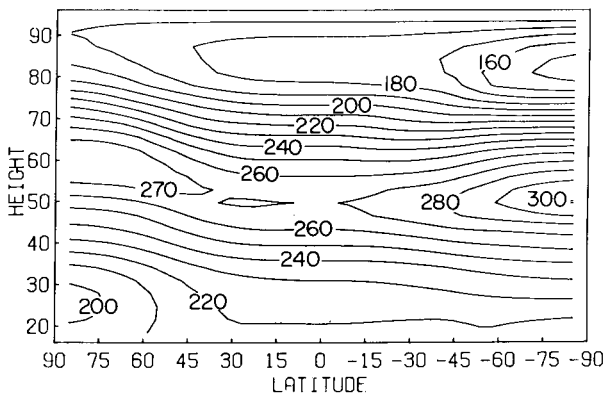

Fig. 7.3. Zonal-mean temperature (K) at the Northern-Hemisphere winter solstice, derived
from a zonally symmetric model including a parameterization of the zonal drag associated
with breaking gravity waves, but no representation of planetary waves. The reversed meridional
temperature gradient in the upper mesosphere should be compared with the observations in
Fig. 5.1a and contrasted with the radiatively determined calculations in Fig. 1.2. [After Holton
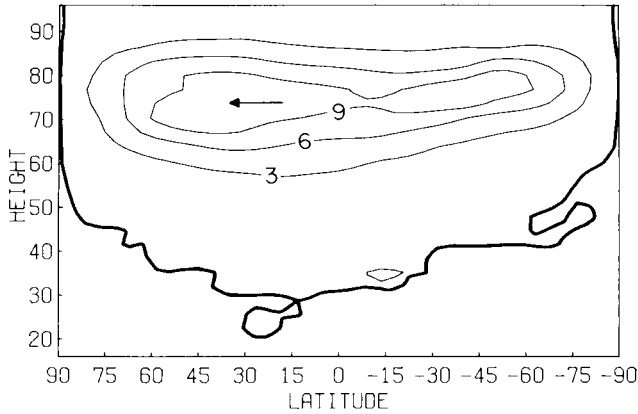(1983a). American Meteorological Society.]

Fig. 7.4. As in Fig. 7.3, but for residual mean meridional wind $\bar{v}^*$ (m s^{-1}). [After Holton (1983a). American Meteorological Society.]
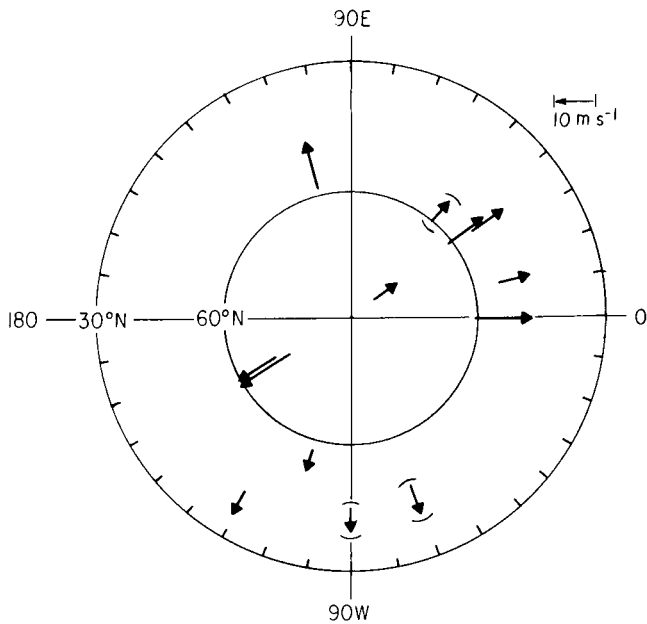

Fig. 7.5. Time-averaged north–south wind components near 90 km in June in the Northern Hemisphere, measured by various methods at different locations. Note the strong equatorward flow in all cases. [After Nastrom *et al.* (1982).]

peaks at about $100 \text{ m s}^{-1} \text{ day}^{-1}$ in the upper mesosphere and the diffusion coefficient reaches values of about $100 \text{ m}^2 \text{ s}^{-1}$. The implied value of $\bar{v}^*$ can amount to some 10 m s^{-1} near the mesopause, in rough order-of-magnitude agreement with the few radar observations of $\bar{v}$ that are available: see Figs. 7.4 and 7.5.

Other wave motions that may contribute to the maintenance of climatological-mean departures of $\bar{T}$ from T_r in the upper mesosphere are atmospheric tides (see Section 4.3), which may break in the mesosphere and could contribute to $\bar{G}$ both below and above their breaking levels. However, planetary waves are unlikely to be important, at least in the summer mesosphere, since their amplitudes are small there.

7.4   The Winter Polar Stratosphere

The other region of the middle atmosphere in which large climatological departures from the radiatively determined temperature are observed is the winter polar stratosphere. As discussed in Chapter 5, the occurrence of planetary waves is a familiar feature of the winter stratosphere, and it was Dickinson (1969) who first suggested that these waves should be regarded as the prime agents responsible for forcing departures of $\bar{T}$ from T_r there.

Dickinson's formulation was partly based on the quasi-geostrophic potential vorticity equation [Eq. (3.3.4)]; he identified the mean northward geostrophic eddy flux of quasi-geostrophic potential vorticity, $\overline{v'q'}$, as the essential wave property required for driving departures from the radiatively determined state. The relationship

$$\rho_0^{-1} \mathbf{\nabla} \cdot \mathbf{F}_{(p)} = \overline{v'q'} \qquad (7.4.1)$$

[see Eq. (3.5.10)] demonstrates that this is equivalent, for quasi-geostrophic disturbances, to the formulation of Section 7.2.

If the "breaking" transient planetary waves mentioned in Section 5.2.3 are common events in the winter stratosphere, they could well contribute to systematically large negative values of $\rho_0^{-1} \mathbf{\nabla} \cdot \mathbf{F}_{(p)}$ there. The fact that the sign is negative is made plausible by the isentropic maps of Ertel's potential vorticity P, such as those in Fig. 5.6, which suggest that breaking planetary waves effect a systematic southward eddy flux of larger, polar values of P and a systematic northward eddy flux of smaller, equatorial values of P. If the same holds for quasi-geostrophic quantities on z surfaces, then $\overline{v'q'} < 0$ and hence $\rho_0^{-1} \mathbf{\nabla} \cdot \mathbf{F}_{(p)} < 0$ by Eq. (7.4.1): as expected, these waves are violating nonacceleration conditions. [The analogous result for isentropic eddy fluxes of P is expressed by Eq. (3.9.11): see Section 7.5.]

Dissipating stationary planetary waves (cf. Section 5.2.2) are also likely to be associated with negative values of $\mathbf{\nabla} \cdot \mathbf{F}_{(p)}$, as in the simple example mentioned in Section 4.5.5. Radiative damping is one source of dissipation for such waves; it has also been suggested that breaking, small-scale waves may introduce nonconservative mechanical forcing or damping of the planetary waves, just as they do for the zonal-mean flow. However, one type of planetary wave that appears unlikely to induce significant forcing of the zonal-mean climatological flow is the global normal mode or free traveling planetary wave discussed in Sections 4.4 and 5.4. Except perhaps during their growth and decay phases, such waves do not strongly violate nonacceleration conditions.

Further research will be required to determine whether the values of $\rho_0^{-1} \mathbf{\nabla} \cdot \mathbf{F}$, or equivalently of $\overline{v'q'}$, associated with planetary waves are large enough in magnitude and of the appropriate distribution in latitude and height to account for the observed departures of $\bar{T}$ from T_r in the winter polar stratosphere. [Note incidentally that the calculation of the *response* to such a zonal force is not entirely trivial: equations like Eqs. (7.2.11), (7.2.12), and (7.2.14) must be solved subject to appropriate boundary conditions.] However, support for some of the ideas presented in this section comes from the general circulation-model experiments described in Section 11.2.1: with some versions of the "SKYHI" model, the polar night stratosphere temperature appears to be too close to T_r because the planetary waves are too weak. A similar picture emerges from experiments with the NCAR Community Climate Model, mentioned at the end of Section 11.1. Also relevant is the fact that the observed Southern-Hemisphere winter stratosphere is closer to the radiatively determined state than is the Northern-Hemisphere winter stratosphere: this may be due to the weaker southern winter planetary-wave activity mentioned above.

It should finally be mentioned that observational and model studies suggest that gravity-wave drag may also have some role to play in the winter stratosphere in contributing *directly* to the zonal force $\bar{G}$ as well as perhaps forcing or dissipating the planetary waves there.

7.5 Interpretation and Generalization

The models described in Section 7.2 are highly simplified and cannot be expected to give full quantitative agreement with observations. Their purpose is rather to provide qualitative insights into the physical mechanisms that determine the zonal-mean climatological state of the middle atmosphere. They suggest how eddy motions on various scales can keep certain parts of the middle atmosphere in a state far from that predicted by

radiative–photochemical models; further, they shed light on some of the physical properties of the eddies that may be responsible for this maintenance process.

These models show how dynamical effects are associated with a "dynamical heating" [as represented, say, by the terms on the left of Eq. (7.2.1b)] which must be balanced by a net radiative heating. In particular, they make it clear that the residual circulation cannot be regarded as a flow that is purely *driven by* an externally imposed net radiative heating rate $\bar{J}$, and independent of the eddy forcing. Indeed, in the steady-state limit described by Eqs. (7.2.4)—which is shown in Section 7.2.3 to be a reasonable first approximation for some purposes—the residual circulation and the net radiative heating are both *entirely* eddy-driven under appropriate boundary conditions, and would both vanish if the eddy-forcing $\bar{G}$ were zero. Of course, the eddy-forcing $\bar{G}$ itself is not generally independent of the mean flow structure (this fact is made explicit, for example, by the parameterizations mentioned in Section 7.3). As a result, interesting and complex feedbacks may occur in the wave, mean-flow interaction and radiative-dynamical interaction processes.

Some of the results derived above for quasi-geostrophic flows on a beta plane can be extended to flows described by the primitive equations on the sphere; it is convenient for this purpose to use the isentropic coordinates introduced in Sections 3.8 and 3.9. For example, the analogs of the steady-state model equations [Eqs. (7.2.4a,b)] and the continuity equation [Eq. (7.2.1c)] in these coordinates are

$$a^{-1}\bar{v}^*\bar{m}_\phi + \bar{Q}^*\bar{m}_\theta = \bar{G}\cos\phi, \qquad (7.5.1a)$$

$$(a\cos\phi)^{-1}(\bar{\sigma}\bar{v}^*\cos\phi)_\phi + (\bar{\sigma}\bar{Q}^*)_\theta = 0, \qquad (7.5.1b)$$

from Eqs. (3.9.7a,c), where $a\bar{m} \equiv a^2\Omega\cos^2\phi + a\bar{u}\cos\phi$ is the zonal-mean absolute angular momentum per unit mass, $Q = (J/c_p)e^{\kappa z/H}$, and $\bar{G}$ now represents the terms $\bar{X}^* + (\bar{\sigma}a\cos\phi)^{-1}\tilde{\mathbf{\nabla}}\cdot\tilde{\mathbf{F}}$ in Eq. (3.9.7a). [Note that $\overline{(\sigma'u')}_t = 0$ because the waves are assumed steady.] Equation (7.5.1b) implies the existence of a stream function $\bar{\psi}^*$ such that

$$\bar{\sigma}a(\cos\phi)\bar{Q}^* = \bar{\psi}^*_\phi, \qquad \bar{\sigma}(\cos\phi)\bar{v}^* = -\bar{\psi}^*_\theta \qquad (7.5.2a,b)$$

and substitution into Eq. (7.5.1a) then yields

$$\frac{\partial(\bar{\psi}^*, \bar{m})}{\partial(\phi, \theta)} = \bar{\sigma}a(\cos^2\phi)\bar{G}. \qquad (7.5.3)$$

If eddy and frictional effects are absent, then $\bar{G} = 0$ and Eq. (7.5.3) implies that

$$\bar{\psi}^* = \Psi(\bar{m}), \qquad (7.5.4)$$

say, for some function Ψ; that is, $\bar{\psi}^*$ is constant on surfaces of constant $\bar{m}$. The residual circulation $(\bar{v}^*, \bar{Q}^*)$ flows along surfaces of constant zonal-mean absolute angular momentum, since it must conserve angular momentum when the right-hand side of Eq. (7.5.1a) vanishes. Now if $\bar{Q}^* = 0$ everywhere along some isentrope, $\theta = \theta_0$ say (perhaps a nominal lower boundary for the stratosphere), it follows from Eq. (7.5.2a) that $\bar{\psi}^*$ is constant on θ_0 and hence, by Eq. (7.5.4), is constant on all $\bar{m}$ surfaces that intersect θ_0. Thus, by Eq. (7.5.2), $\bar{v}^*$ and $\bar{Q}^*$ both vanish everywhere in the region threaded by such surfaces. Since $\bar{Q}^* = \bar{Q}$ in the absence of eddy motions, $\bar{Q} \equiv 0$ here and the region is in radiative equilibrium, with $\bar{v}^* \equiv 0$ also, by analogy with Section 7.2.1.[2]

The results of Sections 7.2.2 and 7.2.3 for the annually varying case can similarly be extended, but once again the scaling arguments that demonstrate the primacy of the eddy forcing in the extratropical middle atmosphere fail near the equator. Note incidentally that the mean zonal momentum equation in the form of Eq. (3.9.9) shows how Dickinson's argument concerning the role of the northward eddy potential vorticity flux generalizes to the primitive equations in isentropic coordinates, with Ertel's potential vorticity P replacing the quasi-geostrophic potential vorticity q.

References

7.1. The approach of considering how dynamical processes can lead to large departures from a hypothetical radiatively determined climatology was pioneered by Dickinson (1969). More recent use of the same ideas has been made by Kurzeja (1981), Plumb (1982), and Mahlman *et al.* (1984), among others.

7.2. The use of a hierarchy of models of the kind described here was suggested by Dr. S. B. Fels (1982 personal communication).

7.3. The original suggestion that gravity waves might by responsible for most of the forcing $\bar{G}$ in the upper mesosphere was made by Houghton (1978). Some indirect evidence for the "filtering" of gravity-wave phase speeds by the mean winds through which they propagate is provided by the radar measurements of Balsley *et al.* (1983). Simple models of the middle atmosphere involving parameterizations of gravity-wave drag in the mesosphere include those of Holton (1982, 1983a), Matsuno (1982), and Schoeberl *et al.* (1983). The effects of tides are considered, for example, by Lindzen (1981) and Miyahara (1984).

---

[2] Note that this argument does not exclude the possibility of regions containing steady, closed meridional circulations, in which angular momentum is conserved, so that Eq. (7.5.4) holds, but that are not in radiative equilibrium since the streamlines do not cut any isentropes on which $\bar{Q}^* = 0$. Such flows are useful idealizations of the low-latitude tropospheric Hadley cell, but their relevance for the middle atmosphere is not at present clear.

7.4. The forcing and dissipation of planetary waves by breaking gravity waves has been investigated by Schoeberl and Strobel (1984), Holton (1984a), and Miyahara (1985). Observations suggesting that direct gravity-wave drag on the zonal-mean flow may be important in the stratosphere are presented by Hamilton (1983) and Smith and Lyjak (1985).

7.5. Questions of wave, mean-flow feedbacks and radiative-dynamical feedbacks are considered for example by Fels *et al.* (1980) and Fels (1985). Models of angular-momentum-conserving Hadley circulations are discussed by Held and Hou (1980), Held and Hoskins (1985), and Tung (1986), among others.