On the 'downward control' of extratropical diabatic circulations by eddy-induced mean zonal forces

On the "Downward Control" of Extratropical Diabatic Circulations by Eddy-Induced Mean Zonal Forces

P. H. HAYNES,* C. J. MARKS,‡ M. E. MCINTYRE,* T. G. SHEPHERD*,@ AND K. P. SHINE‡,+

*Department of Applied Mathematics and Theoretical Physics, University of Cambridge
‡Department of Atmospheric, Oceanic and Planetary Physics, University of Oxford

(Manuscript received 24 August 1989, in final form 19 April 1990)

ABSTRACT

The situation considered is that of a zonally symmetric model of the middle atmosphere subject to a given quasi-steady zonal force $\bar{\mathcal{F}}$, conceived to be the result of irreversible angular momentum transfer due to the upward propagation and breaking of Rossby and gravity waves together with any other dissipative eddy effects that may be relevant. The model's diabatic heating is assumed to have the qualitative character of a relaxation toward some radiatively determined temperature field. To the extent that the force $\bar{\mathcal{F}}$ may be regarded as given, and the extratropical angular momentum distribution is realistic, the extratropical diabatic mass flow across a given isentropic surface may be regarded as controlled exclusively by the $\bar{\mathcal{F}}$ distribution above that surface (implying control by the eddy dissipation above that surface and not, for instance, by the frequency of tropopause folding below). This "downward control" principle expresses a critical *part* of the dynamical chain of cause and effect governing the average rate at which photochemical products like ozone become available for folding into, or otherwise descending into, the extratropical troposphere. The dynamical facts expressed by the principle are also relevant, for instance, to understanding the seasonal-mean rate of upwelling of water vapor to the summer mesopause, and the interhemispheric differences in stratospheric tracer transport.

The robustness of the principle is examined when $\bar{\mathcal{F}}$ is time-dependent. For a global-scale, zonally symmetric diabatic circulation with a Brewer-Dobson-like horizontal structure given by the second zonally symmetric Hough mode, with Rossby height H_R = 13 km in an isothermal atmosphere with density scale height H = 7 km, the vertical partitioning of the unsteady part of the mass circulation caused by fluctuations in $\bar{\mathcal{F}}$ confined to a shallow layer $L_{\bar{\mathcal{F}}}$ is always at least 84% downward. It is 90% downward when the force fluctuates sinusoidally on twice the radiative relaxation timescale and 95% if five times slower. The time-dependent adjustment when $\bar{\mathcal{F}}$ is changed suddenly is elucidated, extending the work of Dickinson (1968), when the atmosphere is unbounded above and below. Above the forcing, the adjustment is characterized by decay of the meridional mass circulation cell at a rate proportional to the radiative relaxation rate τ_r^{-1} divided by $\{1 + (4H^2/H_R^2)\}$. This decay is related to the boundedness of the angular momentum that can be taken up by the finite mass of air above $L_{\bar{\mathcal{F}}}$ without causing an ever-increasing departure from thermal wind balance. Below the forcing, the meridional mass circulation cell penetrates downward at a speed $\tau_r^{-1} H_R^2/H$. For the second Hough mode, the time for downward penetration through one density scale height is about 6 days if the radiative relaxation time is 20 days, the latter being representative of the lower stratosphere. At any given altitude, a steady state is approached. The effect of a rigid lower boundary on the time-dependent adjustment is also considered. If a frictional planetary boundary layer is present then a steady state is ultimately approached everywhere, with the mass circulation extending downward from $L_{\bar{\mathcal{F}}}$ and closing via the boundary layer.

Satellite observations of temperature and ozone are used in conjunction with a radiative transfer scheme to estimate the altitudes from which the lower stratospheric diabatic vertical velocity is controlled by the effective $\bar{\mathcal{F}}$ in the real atmosphere. The data appear to indicate that about 80% of the effective control is usually exerted from below 40 km but with significant exceptions up to 70 km (in the high latitude southern hemispheric winter). The implications for numerical modelling of chemical transport are noted.

1. Introduction

The mean circulation of the middle atmosphere is one of the key factors shaping our chemical and radiative environment. In any zonally symmetric model of the mean circulation in an atmosphere whose angular momentum distribution resembles that of the Earth, the ability of tropical stratospheric air to reach extratropical latitudes and descend into the troposphere depends on the existence of an eddy-induced mean zonal force $\bar{\mathcal{F}}$ per unit mass. Such a force is needed to enable the poleward mean flow to cross surfaces of constant mean absolute angular momentum $\bar{m}$ per unit mass. This paper is concerned with the nature of the implied causal link between $\bar{\mathcal{F}}$ on the one hand, and the structure of the mean circulation on the other, where

@ Present affiliation: Department of Physics, University of Toronto.
+ Present affiliation: Department of Meteorology, University of Reading.

Corresponding author address: Dr. Peter Haynes, Dept. of Applied Mathematics and Theoretical Physics, University of Cambridge, Silver Street, Cambridge CB3 9EW, United Kingdom.

"mean" denotes a suitable measure, relevant to tracer transport, of the long-term, zonal-mean diabatic circulation (e.g., Tung 1982; WMO 1985).

The observed $\bar{m}$ distribution in the middle atmosphere is illustrated in Fig. 1 for various times of year. Because of the dominance of the Earth's rotation, extratropical regions are spanned vertically by $\bar{m}$ surfaces; and such a configuration is likely to persist throughout any plausible scenario for future change. As long as this $\bar{m}$-configuration persists, the extratropics are inaccessible to the mean circulation in the absence of a suitable force $\bar{\mathcal{F}}$. Such a force is thought to exist, in reality, as a result of the irreversible, global-scale angular momentum transport associated with dissipative eddy processes such as the breaking of gravity and Rossby waves. For a given $\bar{m}$-configuration this eddy-induced force $\bar{\mathcal{F}}$ can be regarded as controlling the strength and shape of the extratropical mean mass circulation over long timescales, neglecting temporary changes in which the $\bar{m}$-surfaces merely shift from side to side. Hence (again with a suitable definition of "mean"), $\bar{\mathcal{F}}$ can be regarded as controlling the rates at which stratospheric photochemical products are fed downward toward the extratropical troposphere. For instance, applying extra diabatic cooling to the polar stratosphere will change the temperature, but cannot cause any permanent increase in mean descent rates to the extent that $\bar{\mathcal{F}}$ and $\bar{m}$ are unchanged.

$\bar{\mathcal{F}}$ can also be regarded as holding extratropical temperatures away from their radiatively determined values. The idea of "control" by $\bar{\mathcal{F}}$ is not meant to suggest that other causal links are considered unimportant, such as the feedback of the mean state on the generation, upward propagation, and dissipation of the Rossby and gravity waves that are believed to be crucial, in turn, to determining $\bar{\mathcal{F}}$. However, one reason why thought-experiments involving manipulation of $\bar{\mathcal{F}}$ seem useful, in the present state of knowledge, is that the other causal links just mentioned are relatively ill-understood. Therefore, they need separate consideration in any attempt to understand the circulation.

All these ideas about the middle-atmospheric circulation are now widely appreciated and accepted and have been dealt with in a number of excellent reviews (e.g., Fels 1985; WMO 1985; Andrews et al. 1987). They can be traced back to pioneering work by Dickinson (1968, 1969). Less widely emphasized has been the fact, to be established below, that for given $\bar{m}$ the control of extratropical mean diabatic circulations by $\bar{\mathcal{F}}$ is exerted entirely downward when viewed on a sufficiently long timescale relative to radiative relaxation times (Haynes and McIntyre 1987; McIntyre 1987). For instance, if $\bar{\mathcal{F}}$ were confined to some layer $L_{\bar{\mathcal{F}}}$ of the middle atmosphere then the streamlines of the mean diabatic circulation would extend below, and vanish above, the layer $L_{\bar{\mathcal{F}}}$. Thus, for example, the rate at which stratospheric material becomes available for descent into the extratropical troposphere after descending across, say, the 350 K isentropic surface, is in this sense controlled exclusively by eddy dissipation

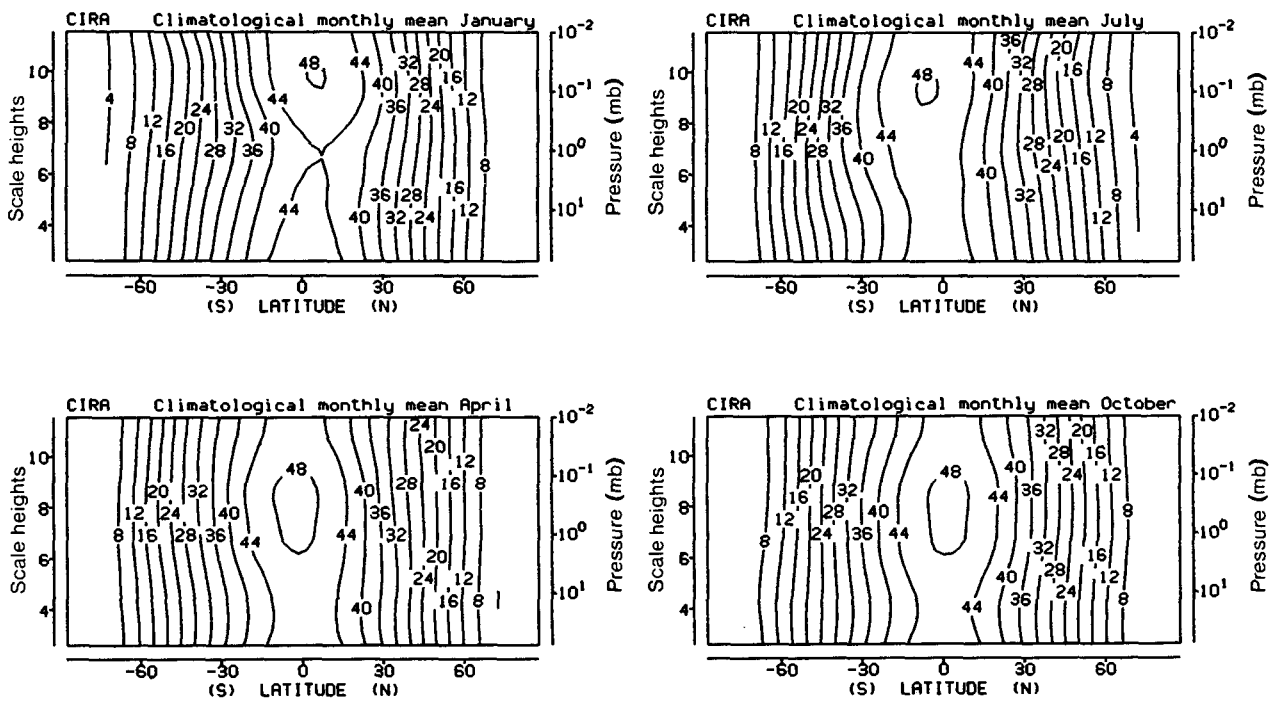

FIG. 1. Zonally averaged absolute angular momentum distribution in the middle atmosphere for January, April, July and October, in units of 6.4×10^7 m^2 s^{-1}. Calculated from the zonal wind fields in Marks (1989).

on isentropic surfaces lying *above* the 350 K surface, whatever the rate of tropopause folding, cutoff cyclone formation, or other aspects of eddy motion, below that surface. Again, the rate at which ozone becomes available for folding into, or otherwise descending into, the extratropical troposphere, is controlled from a still more restricted set of isentropic surfaces, namely those surfaces that reach into the photochemical source region for ozone. The supply of ozone depends on sustained mean diabatic descent across such surfaces, which according to conventional chemistry lie well above 350 K; the supply is clearly unaffected by branches of the circulation confined below the lowest such surface. Similar considerations evidently govern the supply of any other photochemical product from the stratosphere to the troposphere.

Because such considerations seem likely to remain an important part of our thinking about future global change assessments, it is useful, as well as interesting in its own right, to understand the physical reasons for, and range of applicability of, the dynamical principle just stated, namely, that on sufficiently long timescales and for given $\bar{m}$ the control of the mean circulation by $\bar{\mathcal{F}}$ is exerted entirely downward. Some relevant results can already be found in the literature. A special case of the principle was noticed and remarked on long ago by Dickinson (1968, p. 275) in the course of calculating the steady response to given $\bar{\mathcal{F}}$ in an idealized model that assumed constant Coriolis parameter f, and Newtonian cooling with a relaxation time τ_r varying exponentially with height. In the calculation, two exponentials exactly cancelled above a shallow forcing layer $L_{\bar{\mathcal{F}}}$ to give zero steady-state streamfunction and zero temperature change everywhere above $L_{\bar{\mathcal{F}}}$. Dickinson stated that, with more general τ_r and static stability profiles, "source points [meaning $L_{\bar{\mathcal{F}}}$] will, strictly speaking, influence points above, but . . . the same conclusions are approximately valid." However, we shall show that the downward control of the meridional mass circulation is *exactly* valid in the steady state, regardless of the detailed profiles of radiative relaxation time τ_r, static stability N^2, and zonal wind $\bar{u}$, indeed for any realistic radiation scheme that permits the steady state to be attained. Further illustrations of the general validity of the principle in a time-dependent context as steadiness is approached have recently been provided by the more realistic model calculations of Garcia (1987, Fig. 11) and of Dunkerton (1989a, Fig. 3). Moreover, as those calculations illustrate, the principle is surprisingly robust for some purposes. In particular, in the case of global circulations on the scale of the lower-stratospheric Brewer-Dobson cells, the conditions for quasi-steadiness sufficient to make the mass circulation extend predominantly downward below the forcing layer turn out not to be very stringent.

Dickinson also considered aspects of the time dependent adjustment, and one of the purposes of the present paper is to examine this more closely. In particular, we examine a) the time-dependent adjustment toward the downward-controlled steady state, b) the timescales for the different parts of the adjustment process, and c) the physical reasons why such adjustment always takes place and why it must (for any realistic basic state, and for sufficiently small $\bar{\mathcal{F}}$ that the $\bar{m}$ configuration remains realistic) be expected to lead to exact downward control in the sense discussed.

The upper branch of the change in the mass circulation will be shown to die out on a timescale $\tau_r\{1 + (4H^2/H_R^2)\}$, where H is the density scale height and H_R the Rossby height, of order fN^{-1} times the horizontal length scale, where f is the Coriolis parameter and N the buoyancy frequency. The decay of the upper circulation is understandable in terms of the finite capacity of the part of the atmosphere above $L_{\bar{\mathcal{F}}}$ to take up angular momentum while remaining in thermal wind balance. The timescale for the mean meridional circulation to turn downward can be thought of as the timescale for the upper branch of the circulation to transfer this finite amount of angular momentum upward.

Another timescale of interest is the much longer time for the middle stratosphere, say, to feel the effects of the planetary boundary layer. After the initial adjustment period, the mass circulation induced from a layer $L_{\bar{\mathcal{F}}}$ in the middle stratosphere will burrow downward at a rate proportional to the local τ_r^{-1}. In fact the downward propagation speed is $H_R^2/H\tau_r$. When the circulation reaches the Earth's surface, it is taken up by a small adjustment in the boundary-layer mass transport such that the change in surface friction just balances $\bar{\mathcal{F}}$. The accompanying small change in surface zonal wind will then work its way back upward. An assumption that this sequence of events has had time to take place is implicit in standard calculations of the zonal wind field associated with a given zonal mean temperature field and is germane, for instance, to the discussions by McIntyre and Palmer (1984) and Shine (1987) of the possible effects, on different timescales, of radiative heating upon isentropic gradients of potential vorticity at the edge of the wintertime stratospheric polar vortex.

The plan of the paper is as follows. Section 2 derives, in a slightly more general form, the formula given in McIntyre (1987, 1989b) expressing the downward control principle. The formula is a simple consequence of the assumption that a physically sensible steady state is approached and is derived in section 2 without attempting to address the nature of the adjustment mechanism. A further simplification that applies to some mesospheric problems is noted. Section 3 turns to the adjustment problem and examines the time dependence explicitly for an idealized, vertically unbounded model atmosphere, using a Laplace-transform technique as in Dickinson (1968). The results are then checked, and extended to less idealized cases, by some zonally symmetric numerical experiments (Figs. 5–9).

Graphs of the time development of the vertical mass-circulation partitioning for various horizontal scales are shown in Fig. 4. Sections 4 and 5 show how the adjustment is affected by the presence of a lower boundary. It is convenient to consider first the case of a rigid, frictionless boundary and then to incorporate a frictional planetary boundary layer. It is shown that the latter can easily take up the mass circulation with an almost imperceptible adjustment in surface zonal wind values. Section 6 gives some observational estimates of the vertical extent of the domain of influence in the real atmosphere. The implications for the vertical extent and upper boundary conditions of numerical models are noted in sections 5 and 6. In section 7, a dynamical scaling argument is presented that suggests that the nature of the adjustment to the downward-controlled steady state is not qualitatively dependent on the simplifying assumptions made in sections 3–5.

2. Steady state

The simplest derivation of the downward control principle follows from any statement of the equations that describes a steady, zonally averaged or symmetrical atmosphere. It is convenient to use the notation of the TEM (transformed Eulerian-mean) equations derived, for example, by Andrews et al. (1987), and references therein, although for the most accurate work one would use instead the mass-weighted isentropic average introduced by Tung (1982), which satisfies steady-state equations of the same mathematical form [Andrews et al. 1987, Eqs. (3.9.7a, c)], and to which, therefore, the analysis of this section applies essentially unchanged.

In latitude, log-pressure coordinates (ϕ, z), the zonal momentum equation takes the form

$$\frac{\partial \bar{u}}{\partial t} + \bar{v}^* \left[\frac{1}{a\cos\phi} \frac{\partial}{\partial \phi} (\bar{u}\cos\phi) - 2\Omega \sin\phi \right] + \bar{w}^* \frac{\partial \bar{u}}{\partial z} = \bar{\mathcal{F}} \quad (2.1)$$

where $\bar{u}$ is the zonal velocity, and $\bar{v}^*$ and $\bar{w}^*$ are the latitudinal and vertical components of the mean circulation. The rotation rate and radius of the Earth are Ω and a, respectively, and t is time. $\bar{\mathcal{F}}$ represents the zonal force per unit mass due to Rossby and gravity wave breaking and other dissipative eddy processes and may be written in terms of second-order eddy quantities. The mass continuity equation may be written

$$\frac{1}{a\cos\phi} \frac{\partial}{\partial \phi} (\bar{v}^* \cos\phi) + \frac{1}{\rho_0} \frac{\partial}{\partial z} (\rho_0 \bar{w}^*) = 0 \quad (2.2)$$

where $\rho_0 = \exp(-z/H)$ and H is a constant height, defined as the scale height for the log-pressure coordinate and generally taken $\sim$7 km. In terms of a streamfunction $\psi(\phi, z)$ such that

$$\bar{v}^* = -\frac{1}{\rho_0 \cos\phi} \frac{\partial \psi}{\partial z}, \quad \bar{w}^* = \frac{1}{a\rho_0 \cos\phi} \frac{\partial \psi}{\partial \phi}, \quad (2.3a, b)$$

the steady-state version of (2.1) may be written in the form

$$\frac{\partial(\psi, \bar{m})}{\partial(\phi, z)} = \frac{\partial \psi}{\partial \phi}\frac{\partial \bar{m}}{\partial z} - \frac{\partial \psi}{\partial z}\frac{\partial \bar{m}}{\partial \phi} = \rho_0 a^2 \bar{\mathcal{F}} \cos^2\phi \quad (2.4)$$

where $\bar{m} = a\cos\phi(\bar{u} + a\Omega\cos\phi)$, the angular momentum per unit mass. This equation may be divided by $\bar{m}_\phi = \partial(\bar{m}, z)/\partial(\phi, z)$, giving $\partial(\psi, \bar{m})/\partial(\bar{m}, z)$ on the left, which is equal to $-(\partial\psi/\partial z)_{\bar{m}}$, the derivative at constant $\bar{m}$. The resulting equation may then be integrated with respect to z to give

$$\psi(\phi, z) = \int_z^\infty \left\{ \frac{\rho_0 a^2 \bar{\mathcal{F}} \cos^2\phi}{\bar{m}_\phi} \right\}_{\phi=\phi(z')} dz' \quad (2.5)$$

where the integration is along a contour, $\phi(z)$ say, of constant $\bar{m}$; i.e., along a characteristic of the hyperbolic equation (2.4). The requirement that $\rho_0\bar{w}^* \to 0$ as $z \to \infty$, and the convention $\psi \to 0$, have been used to fix the constant of integration. The typical shapes of constant $\bar{m}$ contours have already been indicated in Fig. 1. Note that dz' is the vertical projection of a segment of the $\bar{m}$ curve. Thus, again with the integration along the same curve, we deduce that

$$\bar{w}^* = \frac{1}{\rho_0 \cos\phi} \frac{\partial}{\partial \phi} \left\{ \int_z^\infty \left\{ \frac{\rho_0 a \bar{\mathcal{F}} \cos^2\phi}{\bar{m}_\phi} \right\}_{\phi=\phi(z')} dz' \right\}. \quad (2.6)$$

In the quasi-geostrophic limit, in which $|\bar{u}| \ll 2\Omega a \times |\sin\phi|$, we may write $\bar{m}_\phi \simeq -2\Omega a^2 \sin\phi\cos\phi$, $\bar{m} \simeq \bar{m}(\phi)$, and therefore

$$\bar{w}^* = -\frac{1}{a\rho_0 \cos\phi} \frac{\partial}{\partial \phi} \left\{ \int_z^\infty \left\{ \frac{\rho_0 \bar{\mathcal{F}} \cos\phi}{2\Omega \sin\phi} \right\}_{\phi=\text{const.}} dz' \right\}, \quad (2.7)$$

where the integral is now taken at constant latitude, recovering the original formula written down by McIntyre (1987). The more accurate version (2.6) takes implicit account of the fact that the $\bar{m}$ configuration differs from that of a resting atmosphere. Indeed, insofar as $\bar{\mathcal{F}}$ and $\bar{m}$ are also related through the thermodynamic and thermal wind equations, (2.6) may be regarded as a consistency condition on the steady-state circulation. A complete theory of the zonally averaged extratropical circulation would have to predict the $\bar{m}$ distribution, taking account of thermodynamic as well as mechanical forcing. Both versions (2.6) and (2.7) are most useful, however, in helping to show how small changes, or trends, in $\bar{w}^*$ are controlled dynamically in realistic circumstances for which we do not expect the $\bar{m}$ configuration to change greatly. In such circumstances, the formulae (2.6) and (2.7) state that the value of $\bar{w}^*$ at any location is directly related to

values of $\bar{\mathcal{F}}$ and its latitude gradient above that location and on the same angular momentum contour, and unrelated to values of $\bar{\mathcal{F}}$ below. In a thought-experiment in which $\bar{\mathcal{F}}$ and $\bar{m}$ are prescribed, therefore, we may say that $\bar{w}^*$ is controlled exclusively by $\bar{\mathcal{F}}$ above the location of interest.

The following points merit further comment:

1) The simplest possible theories of wave saturation for internal gravity and Rossby waves, in which particle displacements increase as $\exp(z/2H)$ until limited by some cut-off value, suggest that as z increases above the breaking level $\bar{\mathcal{F}}$ will be roughly constant in order of magnitude. The major contributions to the integrals in (2.6) and (2.7) are expected on this basis to come from a layer extending two or three scale heights above the breaking level. Section 6 will present observational evidence indicating that this simple argument is often not far wrong, but that there are exceptions that have possibly severe implications for requirements on the vertical extent of numerical models.

2) At first sight the "downwardness" seems no more than a consequence of the fact that we have chosen to apply a boundary condition on (2.5) of evanescence as $z \to \infty$ and that we expect $\int^{\infty}$ to converge. Indeed it might be argued that one could just as well apply the boundary condition as $z \to 0$ and that this would lead to a principle of "upward control." It is clear from the steadiness requirements that the two expressions for $\bar{w}^*$ will be equivalent, provided that mean conditions really are *exactly* steady. In our early thinking this "upward control" argument was rejected on the grounds that tropospheric contributions to $\bar{\mathcal{F}}$, particularly boundary-layer friction, cannot usefully be regarded as given. The time-dependent problem to be considered below provides an independent confirmation of this view and explicitly shows how easily boundary-layer friction can adjust, very slightly, as part of the response to a given middle-atmospheric $\bar{\mathcal{F}}$. As a result it becomes very clear just how ill-conditioned the converse thought-experiment, of specifying $\bar{\mathcal{F}}$ in the boundary layer and troposphere, would be.

3) In the mesospheric context mentioned earlier, $\bar{\mathcal{F}}$ may often be mainly due to vertically propagating gravity waves and may be written as $-\rho_0^{-1}\partial(\rho_0\overline{u'w'})/\partial z$. When substituted in (2.7) and after imposing the condition that $\rho_0\overline{u'w'} \to 0$ as $z \to \infty$, this gives

$$\bar{w}^* = \frac{-1}{a\cos\phi}\frac{\partial}{\partial\phi}\left\{\frac{\overline{u'w'}\cos\phi}{2\Omega\sin\phi}\right\}. \qquad (2.8)$$

This remarkably simple result states that $\bar{w}^*$ at a given level is controlled by the upward momentum flux across that level and its latitudinal derivative. The relation (2.8) has already proved useful in understanding the results from recent zonally symmetric model experiments aimed at understanding the formation of noctilucent clouds (Garcia 1989; McIntyre 1989b). For instance, it shows how the long-term, zonal-mean upwelling that cools the summer mesopause and brings water vapor up to noctilucent cloud level is not only controlled exclusively by gravity waves breaking or otherwise dissipating above the cloud level, but is independent of how far above the level in question the waves are breaking, a fact nicely illustrated by the model upwelling profiles in Fig. 5 of Garcia (1989).

4) The temperature $\bar{T}$ is related to the mass circulation ψ through the steady thermodynamic equation

$$-\frac{1}{a\rho_0\cos\phi}\frac{\partial\psi}{\partial z}\frac{\partial\bar{T}}{\partial\phi}+\frac{1}{a\rho_0\cos\phi}\frac{\partial\psi}{\partial\phi}\left(\frac{\partial\bar{T}}{\partial z}+\frac{\kappa\bar{T}}{H}\right) = \bar{\mathscr{R}}[\bar{T};\phi,z]+\bar{\mathcal{C}} \qquad (2.9)$$

[Andrews et al. 1987, Eq. (3.5.2e)], where, for each ϕ and z, $\bar{\mathscr{R}}$ is a functional of $\bar{T}(\cdot)$ expressing radiative relaxation toward some distribution T_{rad}, and $\kappa = R/c_p$, with R the gas constant and c_p the specific heat at constant pressure. $\bar{\mathcal{C}}$ is a term involving eddy heat fluxes that is small under quasi-geostrophic scaling and is found to be small in practice (see section 6). Note from (2.5) and (2.9) that if $\bar{\mathcal{C}}$ is negligible or zero and $\bar{\mathcal{F}}$ is confined to a localized region, then $\bar{\mathscr{R}}$ must vanish above that region. Since for any realistic representation of radiative processes $\bar{\mathscr{R}}$ will be a nonlocal functional of $\bar{T}$, it does not follow that $\bar{T} = T_{\text{rad}}$ immediately above the forcing layer except for a Newtonian cooling model (contrary to a statement made in McIntyre 1989a).

5) Equation (2.4) giving ψ in terms of $\bar{\mathcal{F}}$ is hyperbolic. It is to be contrasted with the elliptic equation obtained for ψ in terms of $\bar{\mathcal{F}}$ in the classical Eliassen problem [e.g., Plumb (1982, Eq. 2.9)], where the response is unsteady and radiative damping is not included. The difference between the two responses is apparent on comparison of Figs. 6(b) and 6(j), for example.

3. The time-dependent radiatively damped response of a resting atmosphere to a zonally symmetric applied force

We now consider the time-dependent adjustment problem in several ways, of which the first is an idealized case soluble analytically, followed by some less idealized cases solved numerically. The results together with a general argument given in section 7 show that a steady state is reached in circumstances of practical interest. Thus steady-state formulae like (2.6) will hold, whether expressed in the TEM, isentropic-coordinate, or other suitable formulation.

a. Mathematical development

The notation of section 2 is retained except that, following Plumb (1982), $\mu = \sin\phi$ is used as a latitudinal coordinate, and, since all quantities are independent of longitude, overbars are dropped.

The basic equations are linearized about a state of rest in which the temperature is a function of height alone, $T_0(z)$, and it is assumed that the zonal flow remains in geostrophic balance. Note that these restrictions are relaxed in the numerical experiments described later in this section. The equations therefore take the form

$$\frac{\partial u}{\partial t} - 2\Omega\mu v = \mathcal{F} \tag{3.1a}$$

$$2\Omega\mu u = -\frac{1}{a}(1-\mu^2)^{1/2}\frac{\partial \Phi}{\partial \mu} \tag{3.1b}$$

$$\frac{\partial \Phi}{\partial z} = \frac{RT}{H} \tag{3.1c}$$

$$\frac{\partial T}{\partial t} + wS = \mathscr{H}_* \tag{3.1d}$$

$$\frac{1}{a}\frac{\partial}{\partial \mu}[v(1-\mu^2)^{1/2}] + \frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0 w) = 0 \tag{3.1e}$$

where u, v, and w are eastward, northward, and upward velocities, T is the departure of the temperature from the reference profile $T_0(z)$ and Φ the departure from the corresponding reference profile of geopotential, $S = \partial T_0/\partial z + \kappa T_0/H$ is a measure of the static stability, $\mathcal{F}$ is the zonal force per unit mass, and $\mathscr{H}_*$ is the diabatic heating rate divided by c_p. [It is therefore appropriate to identify (v, w) in the zonally symmetric system (3.1) with $(\bar{v}^*, \bar{w}^*)$ in the zonally averaged system].

The variables u, v, Φ and T may be eliminated from the above equations to leave

$$\frac{4\Omega^2 a^2}{N^2}\frac{\partial}{\partial z}\left[\frac{1}{\rho_0}\frac{\partial(\rho_0 w)}{\partial z}\right] + \frac{\partial}{\partial \mu}\left[\frac{(1-\mu^2)}{\mu^2}\frac{\partial w}{\partial \mu}\right]$$
$$= \frac{1}{N^2}\frac{\partial}{\partial \mu}\left[\frac{(1-\mu^2)}{\mu^2}\frac{\partial \mathscr{H}}{\partial \mu}\right] + \frac{2\Omega a}{N^2}\frac{\partial}{\partial \mu}\left[\frac{(1-\mu^2)^{1/2}}{\mu}\frac{\partial \mathcal{F}}{\partial z}\right] \tag{3.2}$$

where $\mathscr{H} = R\mathscr{H}_*/H$ and $N^2 = RS/H$. (N will correspond to the buoyancy frequency provided that H is chosen so that there is a reasonable correspondence between z and geometric height.) As Plumb describes, there exist solutions to this equation of the form

$$w = \sum_n w_n(z, t)\Theta_n(\mu) \tag{3.3}$$

where $\Theta_n(\mu)$ is a Hough function satisfying the zero-wavenumber and zero-frequency limit of Laplace's tidal equation, provided that the forcing functions $\mathscr{H}$ and $\mathcal{F}$ may themselves be expanded as

$$\mathscr{H}(\mu, z, t) = \sum_n \mathscr{H}_n(z, t)\Theta_n(\mu) \tag{3.4a}$$

$$\mathcal{F}(\mu, z, t) = \sum_n \mathcal{F}_n(z, t)\mu B_n(\mu) \tag{3.4b}$$

where

$$\frac{d}{d\mu}[B_n(\mu)(1-\mu^2)^{1/2}] = \Theta_n(\mu). \tag{3.4c}$$

The functions $\Theta_n(\mu)$ are orthonormal on $[-1, 1]$ and therefore multiplication of (3.2) by $\Theta_n(\mu)$ and integration from $\mu = -1$ to $\mu = 1$ gives

$$\frac{\partial}{\partial z}\left[\frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0 w_n)\right] + \frac{N^2\epsilon_n}{4\Omega^2 a^2} w_n$$
$$= \frac{1}{4\Omega^2 a^2}\left[\epsilon_n \mathscr{H}_n + 2\Omega a \frac{\partial \mathcal{F}_n}{\partial z}\right]. \tag{3.5}$$

Here, ϵ_n is the nth eigenvalue of the limiting form of Laplace's tidal equation mentioned earlier. The first few values of ϵ_n have been calculated by Longuet-Higgins (1968) and are given in Table 1. (These have been verified by independent calculation. Note that the values given by Plumb (1982) are in error by 10% or so.)

Plumb (1982) noted that the form of (3.4b) appears to preclude the treatment of mechanical forcing that is non-zero at the equator. In fact, such a case causes no serious problem. Indeed one could formally derive the coefficients for such an $\mathcal{F}$ using the orthogonality relations for the B_n. The resulting solution would converge in a mean-square sense, although if evaluated at the equator, it would require a large number of terms for an accurate answer. Dunkerton (1989a) has calculated the response in w to an equatorial forcing in this way. The corresponding series for $\partial u/\partial t$, on the other hand, would not converge to the correct answer at the equator. (The problem would be equivalent to that of evaluating a discontinuous periodic function at a point of discontinuity using a Fourier series.) If $\partial u/\partial t$ were required, it would have to be calculated directly from w, and hence v, via (3.1e) and (3.1a). An alternative approach would be to divide the solution into two parts, writing

$$\frac{\partial u}{\partial t} = \frac{\partial \hat{u}}{\partial t} + \mathcal{F}, \tag{3.6a}$$

$$\frac{\partial \Phi}{\partial t} = \frac{\partial \hat{\Phi}}{\partial t} - 2a\Omega\int_0^\mu \frac{\tilde{\mu}\mathcal{F}\,d\tilde{\mu}}{(1-\tilde{\mu}^2)^{1/2}} \tag{3.6b}$$

TABLE 1. Values of the eigenvalues ϵ_n.

n	ϵ_n
1	−8.127
2	−12.54
3	−35.42
4	−44.73
5	−82.38
6	−96.62
7	−149.1
8	−168.2

and

$$\frac{\partial T}{\partial t} = \frac{\partial \hat{T}}{\partial t} - \frac{2a\Omega H}{R}\int_0^{\mu} \frac{\partial \mathcal{F}}{\partial z}\frac{\tilde{\mu}\, d\tilde{\mu}}{(1-\tilde{\mu}^2)^{1/2}}. \quad (3.6c)$$

Then $\hat{\mu}_t$ and $\hat{T}_t$ are, respectively, zonal-velocity and temperature tendencies in the problem with the right-hand side of (3.1a) equal to zero and the right-hand side of (3.1d) modified to

$$\hat{\mathscr{H}}_* = \mathscr{H}_* + \frac{2a\Omega H}{R}\int_0^{\mu} \frac{\partial \mathcal{F}}{\partial z}\frac{\tilde{\mu}\, d\tilde{\mu}}{(1-\tilde{\mu}^2)^{1/2}}, \quad (3.7)$$

which may be solved, without difficulty, using the expansions (3.4a) and (3.4b), together with the analogue of (3.3), to give the Hough coefficients for $\partial\hat{u}/\partial t$.

In the classical Eliassen problem, with $\mathscr{H}_n$ and $\mathcal{F}_n$ both prescribed and constant in time, w_n is itself constant (though the problem is inherently unsteady, since the corresponding coefficients for the zonal-velocity and temperature fields increase linearly with time). However, when $\mathcal{F}_n$ is a function of time and $\mathscr{H}_n$ a functional of the temperature field, and therefore of the time *history* of w_n, (3.5) describes the time evolution of w_n.

The boundary condition on w_n as $z \to \infty$ is set by the requirement that $\rho_0 w_n \to 0$. In this section, the atmosphere is taken to be infinitely deep and it is therefore also required that $\rho_0 w_n \to 0$ as $z \to -\infty$.

Solutions to the equations (3.1a–e) are now derived in the case where the radiative heating is represented by the Newtonian cooling approximation, with relaxation to the reference profile T_0 at a constant rate $\tau_r^{-1} = \alpha$. The heating rate is given by

$$\mathscr{H}_* = -\alpha T, \quad (3.8)$$

and substituting for T in terms of $\mathscr{H}_*$, and therefore $\mathscr{H}$, in (3.1d), gives

$$\left\{\frac{\partial}{\partial t} + \alpha\right\}\mathscr{H} = \alpha N^2 w. \quad (3.9)$$

The case where $\mathcal{F}(\mu, z, t) = \mu B_n(\mu)\mathcal{F}_n(z, t)$ is considered. The solution may then easily be generalized to any case where $\mathcal{F}$ may be written as a superposition of the form (3.4b), or (3.7) as (3.4a). A closed equation for w_n is derived by taking $[(\partial/\partial t) + \alpha]$ of (3.5) and then using (3.9) to eliminate $\mathscr{H}$, yielding

$$\left\{\frac{\partial}{\partial t} + \alpha\right\}\left[\frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0 w_n)\right)\right] + \frac{N^2\epsilon_n}{4\Omega^2 a^2}\frac{\partial w_n}{\partial t} = \frac{1}{2\Omega a}\left\{\frac{\partial}{\partial t} + \alpha\right\}\left[\frac{\partial \mathcal{F}_n}{\partial z}\right]. \quad (3.10)$$

This partial differential equation is not of a standard form although, as will be shown later, its solutions may be approximated by the solutions of an advection-diffusion equation. Therefore the solutions are investigated by examining the case where the forcing is concentrated at a particular level, say $z = 0$, so that $\mathcal{F}_n(z, t) = f_n(t)\delta(z/H)$, and noting that any $\mathcal{F}_n(z, t)$ may be constructed as a superposition of such functions. To allow analytical progress we also assume N^2 to be constant.

b. *Oscillatory forcing*

The time-harmonic case, in which $f_n(t) = \hat{f}_n \exp(i\omega t)$ and $w_n(z, t) = \hat{w}_n(z)\exp(i\omega t)$, has been studied by Garcia (1987) and Dunkerton (1989a). In this case, it follows that $\hat{w}_n$ satisfies the equation

$$(i\omega + \alpha)\frac{d}{dz}\left(\frac{1}{\rho_0}\frac{d}{dz}(\rho_0\hat{w}_n)\right) + i\omega\frac{N^2\epsilon_n}{4\Omega^2 a^2}\hat{w}_n = \frac{1}{2\Omega a}\frac{\hat{f}_n}{H}(i\omega + \alpha)\delta'\left(\frac{z}{H}\right) \quad (3.11)$$

where $\delta'(\cdot)$ is the derivative of the Dirac delta function. This equation has the solution

$$\rho_0\hat{w}_n(z, i\omega) = \frac{\Lambda_-(i\omega)}{\Lambda_-(i\omega) - \Lambda_+(i\omega)}\frac{f_n H}{2\Omega a}\exp\left\{\frac{\Lambda_+(i\omega)z}{H}\right\} \quad [z > 0], \quad (3.12a)$$

$$\rho_0\hat{w}_n(z, i\omega) = \frac{\Lambda_+(i\omega)}{\Lambda_-(i\omega) - \Lambda_+(i\omega)}\frac{f_n H}{2\Omega a}\exp\left\{\frac{\Lambda_-(i\omega)z}{H}\right\} \quad [z < 0], \quad (3.12b)$$

where

$$\Lambda_\pm(i\omega) = -\frac{1}{2} \mp \frac{1}{2}\left[1 + 4b_n\left(\frac{i\omega}{i\omega + \alpha}\right)\right]^{1/2} = -\frac{1}{2} \mp \frac{1}{2}\left(\frac{\alpha(i\omega + \hat{\alpha})}{\hat{\alpha}(i\omega + \alpha)}\right)^{1/2} \quad (3.13)$$

with

$$\hat{\alpha} = \frac{\alpha}{1 + 4b_n} \quad (3.14a)$$

and

$$b_n = -\epsilon_n\frac{N^2H^2}{4\Omega^2a^2} = \frac{H^2}{H_R{}^2}, \quad (3.14b)$$

where H_R is the Rossby height based on the horizontal length scale $a(-\epsilon_n)^{-1/2}$. (Note that ϵ_n is always negative.) The branch of the square root function is that which is positive when $\omega = 0$. The expressions (3.13) have already been derived by Garcia (1987) and Dunkerton (1989a).

The circulation pattern associated with forcing in a single Hough mode consists of n cells, corresponding to the latitudinal structure of the appropriate Hough function, with circulation of one sign above the height of the forcing and of the other sign below. The ratio,

$M(i\omega)$, of the mass circulation in each cell above the forcing to that in the cell below is given by

$$M(i\omega) = \left|\frac{\Lambda_-(i\omega)}{\Lambda_+(i\omega)}\right| \leqslant 1.$$

The variation of $M/(1+M)$, being the fraction of the circulation in the upper cell, is shown in Fig. 2 for various values of the parameter b_n. For the second Hough mode (assuming that $H = 7$ km and $N^2 = 4 \times 10^{-4}\ \mathrm{s}^{-2}$) b_n takes the value 0.29, and the maximum value of $M/(1+M)$ is 0.16, attained when $\alpha/\omega = 0$ (i.e., in the classical Eliassen problem with $\mathscr{H} = 0$). The fraction in the upper cell decreases as α/ω increases, and is less than 0.1 (for the second Hough mode) when the period of the forcing is greater than 4π multiplied by the radiative relaxation time. Note that for fixed α/ω the fraction in the upper cell is an increasing function of b_n, tending to 0.5 in the Boussinesq limit $b_n \to \infty$. Furthermore, the rate of decrease of this fraction with increasing α/ω weakens rapidly as b_n increases. It is interesting to note that for a force fluctuating at the annual frequency, with $\omega \sim (58\ \text{days})^{-1}$ and $\alpha = 0.05\ \text{days}^{-1}$, representative of the middle stratosphere, a horizontal structure corresponding to the $n = 2$ Hough mode has $M/(1+M)$ equal to about 0.08. This increases to 0.20 for the $n = 4$ mode and to 0.28 for the $n = 6$ mode.

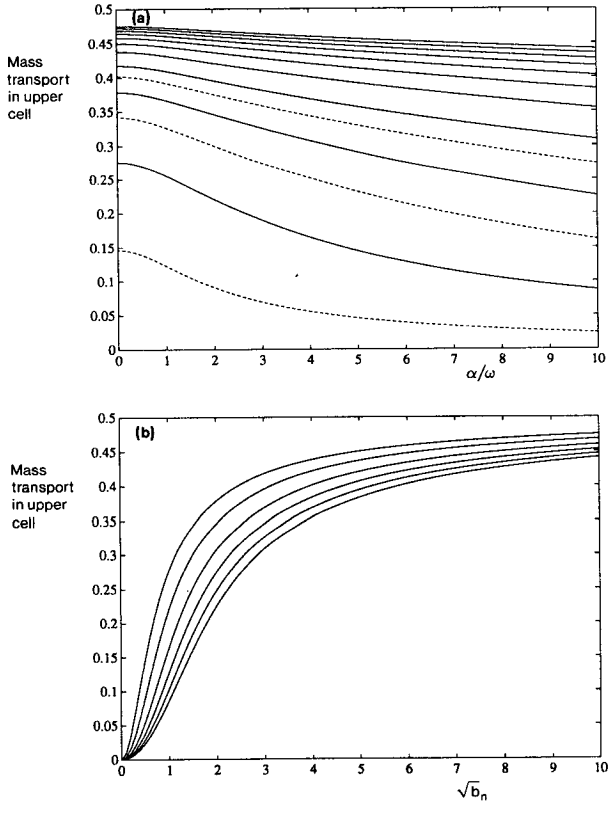

FIG. 2. (a) Fraction of mass transport in upper cell as a function of α/ω, for various values of $\sqrt{b_n}$ (0.5, 1.5, 2.5—dashed lines; 1, 2, . . . , 10—solid lines) with $\sqrt{b_n}$ increasing upward; (b) fraction of mass transport in upper cell as a function of $\sqrt{b_n}$ for various values of α/ω (0, 2, . . . , 10), with α/ω increasing downward.

c. *Switch-on forcing*

The steady-state oscillatory case gives only limited insight into the adjustment problem. Following Dickinson (1968) another simple time dependence is considered, in which the function $f_n(t) = f_n h(t)$, where f_n is a constant and $h(t)$ the Heaviside step function, zero for $t < 0$ and unity for $t > 0$. This problem may be solved by Laplace transform methods, and the transform of $w_n(z, t)$ will be written as $\tilde{w}_n(z, s)$, s being the transform variable. It is clear that the equation for $\tilde{w}_n(z, s)$ will just be (3.11), and its solution (3.12a,b), with, in each case, $i\omega$ replaced by s, $\hat{f}_n$ replaced by f_n/s, and $\hat{w}_n(z)$ replaced by $\tilde{w}_n(z, s)$.

In principle, then, the solution to (3.10) may be deduced using the inversion formula

$$\rho_0 w_n(z, t) = \frac{1}{2\pi i}\int_\Gamma \frac{\Lambda_\mp(s)}{\Lambda_-(s) - \Lambda_+(s)}\frac{f_n H}{2\Omega a} \times \exp\left\{\frac{\Lambda_\pm(s) z}{H}\right\}\frac{e^{st}}{s}\,ds \quad [z \gtrless 0], \quad (3.15)$$

where Γ is the Bromwich inversion contour. In practice this integral cannot be evaluated in closed form and analytical progress may be made only in the small- and large-t limits. It turns out that these limits give considerable insight into the behavior of the solution and results are therefore presented in some detail.

1) The small-time limit $\alpha t \ll 1$ and $|z|/H \ll (\alpha t)^{-1}$: This corresponds to Dickinson's "adiabatic" regime. In this limit, the integral above may be evaluated by considering a contour on which the quantities $\Lambda_\pm(s)$ may be approximated by their forms for large $|s|$, leading to the result

$$\rho_0 w_n(z, t) = \frac{\Lambda_\mp(\infty)}{\Lambda_-(\infty) - \Lambda_+(\infty)}\frac{f_n H}{2\Omega a}\exp\left\{\frac{\Lambda_\pm(\infty) z}{H}\right\} \times \left(1 + O(\alpha t) + O\left(\alpha t \frac{|z|}{H}\right)\right) \quad [z \gtrless 0], \quad (3.16)$$

which is just the well-known solution to the undamped problem, derived by Plumb (1982). For $|z|/H = O(1)$, the expression (3.16) is valid for times shorter than α^{-1}. However, it is interesting that for large $|z|/H$ it is valid only when $t \lesssim \alpha^{-1} H/|z|$, a severe restriction, that results from the fact that the large-$|s|$ approximation to the exponential function is not uniformly valid for large $|z|$.

2) The large-$|z|$ limit $|z|/H \gg (\alpha t) \lesssim 1$: The above-mentioned difficulty may be circumvented by again considering a contour on which $|s|$ is large but using a transformation of the integration variable that allows

a uniformly valid approximation to the exponent. This leads to the result

$$\rho_0 w_n(z,t) = \frac{\Lambda_\mp(\infty)}{\Lambda_-(\infty)-\Lambda_+(\infty)}\frac{f_n H}{2\Omega a}\exp\left\{\frac{\Lambda_\pm(\infty)z}{H}\right\} \times \exp\left\{-\alpha t\left(\frac{1+3b_n}{1+4b_n}\right)\right\} I_0\left\{\left(\frac{4\alpha t b_n|z|}{H(1+4b_n)^{1/2}}\right)^{1/2}\right\} \times\left\{1+O\left(\left(\alpha t\frac{H}{|z|}\right)^{1/2}\right)\right\} \quad [z\gtrless 0] \quad (3.17)$$

where $I_0\{\cdot\}$ is the modified Bessel function of order zero, which increases exponentially rapidly as its argument becomes large and positive. At large $|z|/H$ and fixed t, for instance, the effect of the radiative damping is therefore to increase the size of the response over its value in the adiabatic case.

3) The limit $\alpha t|z|/H \gg 1$, $\alpha t H/|z| \sim 1$: In this regime (which corresponds to Dickinson's "diffusive" regime), the integral may be evaluated using the method of steepest descents, integrating along a contour that passes through the saddle point at $s = s_o(z,t)$ where

$$t = \pm\frac{b_n\alpha z}{H(s_o+\alpha)^{3/2}((4b_n+1)s_o+\alpha)^{1/2}} \quad [z\gtrless 0]. \quad (3.18)$$

This equation for s_o is quartic and therefore soluble in closed form, but the expressions involved are cumbersome. Note, however, that the case $s_o(z,t)=0$, when

$$|z| = \alpha H b_n^{-1} t, \quad (3.19)$$

is of particular interest for two reasons. First, for any quantity which is non-zero in the steady state, the Laplace transform has a pole at $s=0$. Thus, whether $|z|$ is larger or smaller than the right-hand side of (3.19) controls whether or not this pole lies to the left or the right of the steepest-descents contour, and therefore whether the quantity in question, or its departure from the steady-state value, is exponentially small. Hence, some sort of adjustment region might be expected in the neighborhood of the location defined by (3.19). Second, for certain quantities that are zero in the steady state, like w_n for positive z, it turns out that the exponent in the expression resulting from the application of steepest descents is a maximum at given t for the value of $|z|$ given in (3.19). It follows that an approximation obtained by expanding the integrand in (3.15) about $s=0$, and retaining terms of second order in the exponent, is always good in the sense that the difference between the approximate value and the true solution is exponentially small. However, it should also be noted that the ratio of approximate and true values will not tend to unity in the appropriate limit and may well be very large or very small.

The approximate expressions for the vertical velocity, w_n, that result from this approach, are

$$w_n(z,t) \simeq \frac{f_n H}{4\Omega a\sqrt{\pi}}\left(\frac{b_n H}{(b_n+1)z}\right)^{1/2} \times\exp\left\{-\left(t-\frac{b_n z}{H\alpha}\right)^2\frac{H\alpha^2}{4b_n(b_n+1)z}\right\} \simeq \frac{f_n H}{4\Omega a\sqrt{\pi}}\left(\frac{b_n^2}{(b_n+1)\alpha t}\right)^{1/2} \times\exp\left\{-\left(z-\frac{H\alpha t}{b_n}\right)^2\frac{b_n^2}{4H^2(b_n+1)\alpha t}\right\} \quad [z>0], \quad (3.20a)$$

$$w_n(z,t) \simeq -\frac{f_n H}{4\Omega a}\exp\left(\frac{z}{H}\right) \times\operatorname{erfc}\left\{-\left(t+\frac{b_n z}{H\alpha}\right)\left[\frac{H\alpha^2}{4b_n(b_n+1)|z|}\right]^{1/2}\right\} \simeq -\frac{f_n H}{4\Omega a}\exp\left(\frac{z}{H}\right) \times\operatorname{erfc}\left\{-\left(z+\frac{H\alpha t}{b_n}\right)\left[\frac{b_n^2}{4H^2(b_n+1)\alpha t}\right]^{1/2}\right\} \quad [z<0]. \quad (3.20b)$$

The notation "erfc" is used for the complementary error function $\operatorname{erfc}(x) = 2\pi^{-1/2}\int_x^\infty e^{-y^2}dy$, which is small when its argument is large and positive, and tends to the value of two as its argument becomes more and more negative. The second expression, in each case, has been derived from the first by substituting for z in terms of t. This is allowed by the fact that, for given t, the important variation in the argument of the "erfc" or exponential function is confined to a narrow interval about the value of z defined by (3.19). Note that the leading-order expressions for w_n given in (3.20a,b) are the solutions of advection-diffusion equations. The appearance of a complementary error function in the expression for w_n for $z<0$ and an exponential function in the expression for $z>0$ is a consequence of w_n being non-zero below the forcing in the steady state but zero above the forcing. The zonal velocity u_n, on the other hand [which is expanded according to (3.4b)], is non-zero above and below the forcing in the steady state and is given approximately by the expressions

$$u_n(z,t) \approx \frac{f_n b_n}{2\alpha}\operatorname{erfc}\left\{-\left(t-\frac{b_n z}{H\alpha}\right)\left[\frac{H\alpha^2}{4b_n(b_n+1)z}\right]^{1/2}\right\} \approx \frac{f_n b_n}{2\alpha}\operatorname{erfc}\left\{\left(z-\frac{H\alpha t}{b_n}\right)\left[\frac{b_n^2}{4H^2(b_n+1)\alpha t}\right]^{1/2}\right\} \quad [z>0], \quad (3.21a)$$

$$u_n(z,t) \approx \frac{f_n b_n}{2\alpha}\exp\left(\frac{z}{H}\right) \times \operatorname{erfc}\left\{-\left(t+\frac{b_n z}{H\alpha}\right)\left[\frac{H\alpha^2}{4b_n(b_n+1)|z|}\right]^{1/2}\right\}$$

$$\approx \frac{f_n b_n}{2\alpha}\exp\left(\frac{z}{H}\right) \times \operatorname{erfc}\left\{-\left(z+\frac{H\alpha t}{b_n}\right)\left[\frac{b_n^2}{4H^2(b_n+1)\alpha t}\right]^{1/2}\right\} \quad [z<0]. \quad (3.21b)$$

Similar expressions may be derived for other dynamical variables. Again, these results are implicit in the work of Dickinson (1968), but he gives explicit solutions only for the geopotential height.

4) The limit $\alpha t \gg |z|/H \gg 1$: This regime holds once the adjustment region highlighted in the description of regime 3 has passed. Here, the steepest-descents method may not be applied naively to evaluate the integral since the stationary point of the exponent lies close to a branch point of the functions $\Lambda_\pm(s)$ at $s = -\hat{\alpha}$. However, it is clear that the behavior of the departure from the steady solution is dominated by exponential decay in time at rate $\hat{\alpha}$.

In summary, the adjustment to a sudden change in mechanical forcing comprises the instantaneous appearance of meridional circulations familiar from the classical theory, followed by the downward extension of the lower cell, and the upward translation, away from the forcing region, of the upper cell, as described by (3.20a,b). The downward extension occurs as if the vertical mass transport were a passive scalar, with a continuously applied source in the forcing region, which is advected downward with speed w_α [equivalent to the speed c in Dickinson (1968)] in fluid with molecular diffusion coefficient κ_α, where

$$w_\alpha = \frac{H\alpha}{b_n} = \frac{\alpha H_R^2}{H}, \quad (3.22a)$$

$$\kappa_\alpha = \frac{H^2\alpha(b_n+1)}{b_n^2} = \alpha H_R^2\left(1+\frac{H_R^2}{H^2}\right). \quad (3.22b)$$

The upward translation occurs as if the vertical velocity were a similar passive scalar, but advected upward with speed w_α from a dipole source. There is a related pattern to the response in the zonal velocity. Above the forcing region, the response is that associated with upward advection of the zonal velocity from a continuously applied source, and below, with downward advection of the mass-weighted velocity, again from a continuously applied source. This evolution is depicted schematically in Fig. 3.

As the upper cell moves away from the forcing region its mass transport weakens, ultimately leaving all the mass transport in the lower cell. The ratio $M(t)$ of the

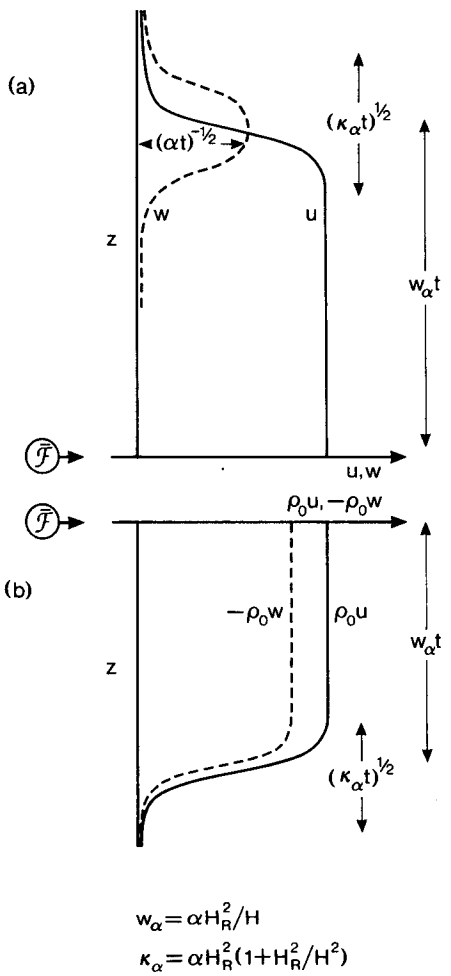

FIG. 3. Schematic diagram showing the response to a force localized at $z = 0$ and having the latitudinal structure of a single Hough mode: (a) u and w for $z > 0$; (b) $\rho_0 u$ and $\rho_0 w$ for $z < 0$.

mass transport in the upper cell to that in the lower is given by

$$M(t) = \left|\frac{w_n(0^+,t)}{w_n(0^-,t)}\right| = \frac{1-m(t;\alpha,\hat{\alpha})}{1+m(t;\alpha,\hat{\alpha})} \quad (3.23a)$$

where

$$m(t;\alpha,\hat{\alpha}) = \frac{1}{2\pi i}\left(\frac{\hat{\alpha}}{\alpha}\right)^{1/2}\int_\Gamma \frac{e^{st}(s+\alpha)^{1/2}}{s(s+\hat{\alpha})^{1/2}}\,ds. \quad (3.23b)$$

After some manipulation it may be shown that

$$m(t;\alpha,\hat{\alpha}) = \left(\frac{\hat{\alpha}}{\alpha}\right)^{1/2}\left[I_0((\alpha-\hat{\alpha})t/2)e^{-(\alpha+\hat{\alpha})t/2} + \alpha\int_0^t I_0((\alpha-\hat{\alpha})t'/2)e^{-(\alpha+\hat{\alpha})t'/2}dt'\right], \quad (3.23c)$$

where $I_0(\cdot)$ is the modified Bessel function. It follows from standard properties of Bessel functions (see

Gradshteyn and Ryzhik 1980, §6.611, Eq. 4) that $m(t; \alpha, \hat{\alpha}) \to 1$ as $\alpha t \to \infty$, the difference decaying at a rate $\hat{\alpha} = \alpha/(1 + 4b_n)$. Figure 4 shows the fraction of the transport in the upper cell, and that in the lower cell, against time, for various values of b_n. As shown earlier, for $b_n = 0.25$, about 85% of the transport is in the lower cell initially, and it may be seen from the figure that in the adjustment problem this proportion increases to 95% after two radiative time scales. A similar overall picture emerges as in the oscillatory problem. As b_n increases the initial proportion in the lower cell decreases to 50%. For fixed b_n this proportion is an increasing function of time such that the transport in the upper cell ultimately disappears. However, the number of radiative time scales required for reduction by a certain fraction is proportional to $\alpha/\hat{\alpha} = 1 + 4b_n = 1 + (4H^2/H_R^2)$ and increases as the inverse square of the horizontal scale when $H_R \ll H$. Thus, precise steadiness is not critical to the downward control principle at large scales, but it is critical at small scales, where in practical terms "large" means planetary scale and "small" means synoptic or smaller.

d. The Boussinesq limit

The limit $b_n \to \infty$ corresponds to that of a Boussinesq fluid. In this case, the circulation cells are symmetric above and below the forcing level as may be seen from the appropriate limit of (3.15), which reduces to

$$\rho_0 w_n = \pm \frac{1}{2\pi i} \int_\Gamma \frac{f_n H}{4\Omega a s} \exp\left\{ st \mp \frac{zN(-\epsilon_n)^{1/2}}{2\Omega a} \left(\frac{s}{s+\alpha}\right)^{1/2} \right\} ds \quad [z \gtrless 0]. \quad (3.24)$$

The integrand has branch points at $s = 0, -\alpha$, and a branch cut may be taken between these points along the real axis. The integration contour may then be deformed until it encloses (and follows) the branch cut. For $\alpha t \gg 1$ the dominant contribution comes from close to $s = 0$, and with a rescaling of the integration variable the integral may be approximated as

$$\rho_0 w_n = \pm \frac{f_n H}{4\Omega a \pi} \int_0^\infty \sin\left(\frac{|z| N(-\epsilon_n)^{1/2}}{2\Omega a} \frac{s^{1/2}}{(\alpha t)^{1/2}} \right) \frac{e^{-s}}{s} ds \times \left\{ 1 + O\left(\frac{|z| N(-\epsilon_n)^{1/2}}{2\Omega a (\alpha t)^{3/2}} \right) \right\} \quad [z \gtrless 0]. \quad (3.25)$$

The value of this integral may be deduced by integrating one given by Gradshteyn and Ryzhik (1980, §3.896, Eq. 4) with respect to one of the parameters. It follows that

$$\rho_0 w_n = \pm \frac{f_n H}{4\Omega a} \operatorname{erfc}\left(\frac{|z| N(-\epsilon_n)^{1/2}}{4\Omega a (\alpha t)^{1/2}} \right) \times \left\{ 1 + O\left(\frac{|z| N(-\epsilon_n)^{1/2}}{2\Omega a (\alpha t)^{3/2}} \right) \right\} \quad [z \gtrless 0]. \quad (3.26)$$

The approximation involved in this estimate breaks down for sufficiently large z, but $\rho_0 w_n$ is then exponentially small.

From the above expression, it follows that the symmetric meridional cells extend further away from the forcing level with time, the extension being purely "diffusive" rather than advective. Indeed, the estimates of the relevant velocity and diffusion coefficient made in the previous section still hold.

If the ratio H_R/H were small, rather than zero, it might be anticipated that non-Boussinesq effects would be felt after a time t, such that the extension of the cells was comparable to H, i.e., such that $H_R(\alpha t)^{1/2} \sim H$, and so $t \sim H^2/H_R^2\alpha$. This latter time scale is the time scale on which the cells become asymmetric and is the same (in this limit) as that estimated in the previous section for the decay of the upper cell.

e. Numerical experiments

The pattern of evolution just described is appealingly simple, but, of course, pictures such as those in Fig. 3 will be quantitatively relevant only a long time after the forcing has been switched on, and only if the forcing is situated many scale heights above the ground. It is therefore of interest to examine the response under less restrictive conditions. This is done through simple experiments with a numerical model, based on the University of Reading sigma-coordinate model (Hoskins and Simmons 1975), which uses a spherical harmonic representation in the horizontal. The model levels are taken to be equally spaced at intervals of about 2 km, with the highest level at about 70 km. The model fields are constrained to be zonally symmetric. Radiative damping is applied in the form of a Newtonian cooling, which restores the temperature field at a constant rate of (4 days)$^{-1}$ toward its initial distribution. A high-

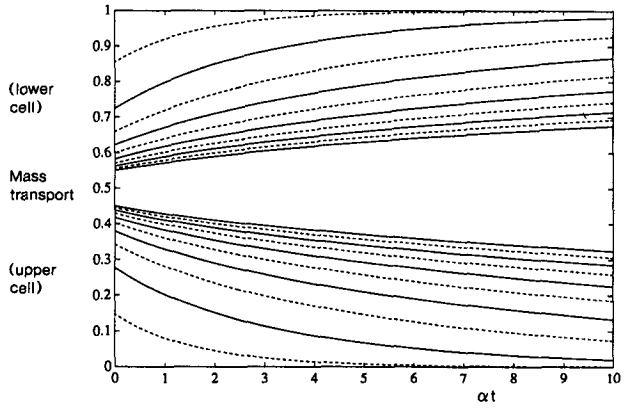

FIG. 4. Fraction of mass transport in each cell as a function of αt, for various values of $\sqrt{b_n}$ (0.5, 1.5, ..., 4.5—dashed lines; 1, 2, ..., 5—solid lines), with $\sqrt{b_n}$ decreasing outward.

order diffusivity, effective only at small horizontal scales, is applied to all dynamical fields. Further mechanical friction is confined to a boundary layer including the lowest two levels. The exact role that the boundary layer plays in the response will be elucidated in section 5. Here, it is simply remarked that the only visible effect of the boundary layer in the numerical experiments is to take up any mass circulation that penetrates to the boundary. Any "rebound" circulation is too weak to be visible in the figures shown.

In the first experiment reported here, the flow was confined to a hemispheric domain by a rigid equatorial boundary and the initial configuration was one of uniform temperature (250 K) and zero zonal flow. A body force $\mathcal{F}$, taking the form of the appropriate $n = 2$ Hough function multiplied by a Gaussian in height, centered at 50 km with a half-width of 1.6 km and having a maximum value of 3.4×10^{-6} m s^{-2}, was increased from zero at day zero. After day 5, it was held constant. This increase was made smooth in order to avoid exciting zonally symmetric gravity waves. The magnitude of the force was small enough that the linearizing assumptions needed to derive (3.1a–e) remained valid. A sequence of pictures of the zonal velocity and meridional mass flow at various times is shown in Fig. 5. The first (for comparison purposes) is from a similar experiment in which there was no radiative damping, i.e., the classical linearized Eliassen adjustment, in which the velocity tendency and the meridional circulation are independent of time (apart from the increase in forcing amplitude up to day 5). The solution of the radiatively damped adjustment problem just after a force has been instantaneously applied, but assuming that the flow remains in geostrophic balance (i.e., $t = 0_+$ in the idealized problem considered earlier in this section), is the same as that of the Eliassen problem. As expected from the results shown in Fig. 4, the upper cell is weak even at this early stage. Furthermore, the large horizontal scale also leads to a long downward decay scale of the mass circulation, and a large proportion of the transport penetrates immediately to the boundary. In the radiatively damped case, as time progresses (Figs. 5d and 5f) penetration of the lower cell increases so that after five radiative timescales (20 days) almost all the lower branch of the flow across latitude circles is confined to the frictional boundary layer. Note also the tendency for the zonal velocity change to become independent of height above the forcing, corresponding to zero temperature change and again as expected from the simple steady-state considerations of section 2.

In a second experiment, the vertical dependence and peak value of the force were kept the same but the horizontal dependence was made to be a Gaussian distribution in latitude, with a half-width of 5 degrees and centered at 45 degrees. The response to this force consisted of a number of Hough modes. The zonal velocity and meridional mass flow at various times are shown in Fig. 6. Again, the equivalent without radiative damping are shown first, in Figs. 6(a) and (b). Note that the presence of higher Hough modes, with relatively large $\Lambda_\pm(\infty)$, leads to a shorter vertical decay scale for the lower cell than that seen in 5(b). Turning now to the radiatively damped evolution, we see that after 2.5 radiative relaxation times (10 days) a large-scale meridional circulation has penetrated to the ground (as would be expected from the first experiment). However, the downward propagation of the smaller-scale parts of the circulation takes place much more slowly. At this time, some narrowing of the response in the neighborhood of the forcing is apparent. As time increases this narrowing occurs at successively lower levels, as is required if the steady solution, in which the meridional circulation is confined to the region directly below the forcing, is to be approached. After 12.5 radiative relaxation times the circulation at all levels is close to that expected in the steady state. Note that at a given level the adjustment to a steady state has the character of negative diffusion to the extent of a sharpening rather than a weakening of horizontal gradients as the character of the problem makes the transition from elliptic to hyperbolic.

Both these experiments were performed with a rigid hemispheric boundary. Without such a boundary, the initial circulation may be of global extent. The circulation seen in a repeat of the second experiment, but with a global domain, is shown in Fig. 7. The equatorial boundary is seen to have little qualitative impact on the mass flow in extratropical regions of the hemisphere in which the forcing is applied.

Finally, it is interesting to consider the response to a forcing when the basic state is not at rest. A modified version of the second experiment, with a basic zonal flow characteristic of winter conditions (Fig. 8a), was therefore carried out. Figure 8b shows the angular momentum contours. In the case studied earlier, where the background state was at rest, these would have been strictly vertical. Figure 8c shows the corresponding meridional circulation in the nonradiatively damped atmosphere, and Figures 8d–h show the evolution of the meridional circulation over a number of radiative time scales. The final figure is broadly consistent with the steady-state prediction that the circulation should be confined to regions spanned by angular-momentum contours extending downward from the region of forcing. In order to emphasize the latter point, the same set of calculations was done with a very strong background velocity profile for which the angular-momentum contours were slanted further from the vertical. To make the slant conspicuous, it was necessary to increase the maximum winds to greater than 180 m s^{-1}. The corresponding results are shown in Fig. 9.

The idea of downward control, manifested by a downward-penetrating lower circulation cell, has arisen quite naturally in the above analysis and has been verified in the numerical experiments. An estimate can

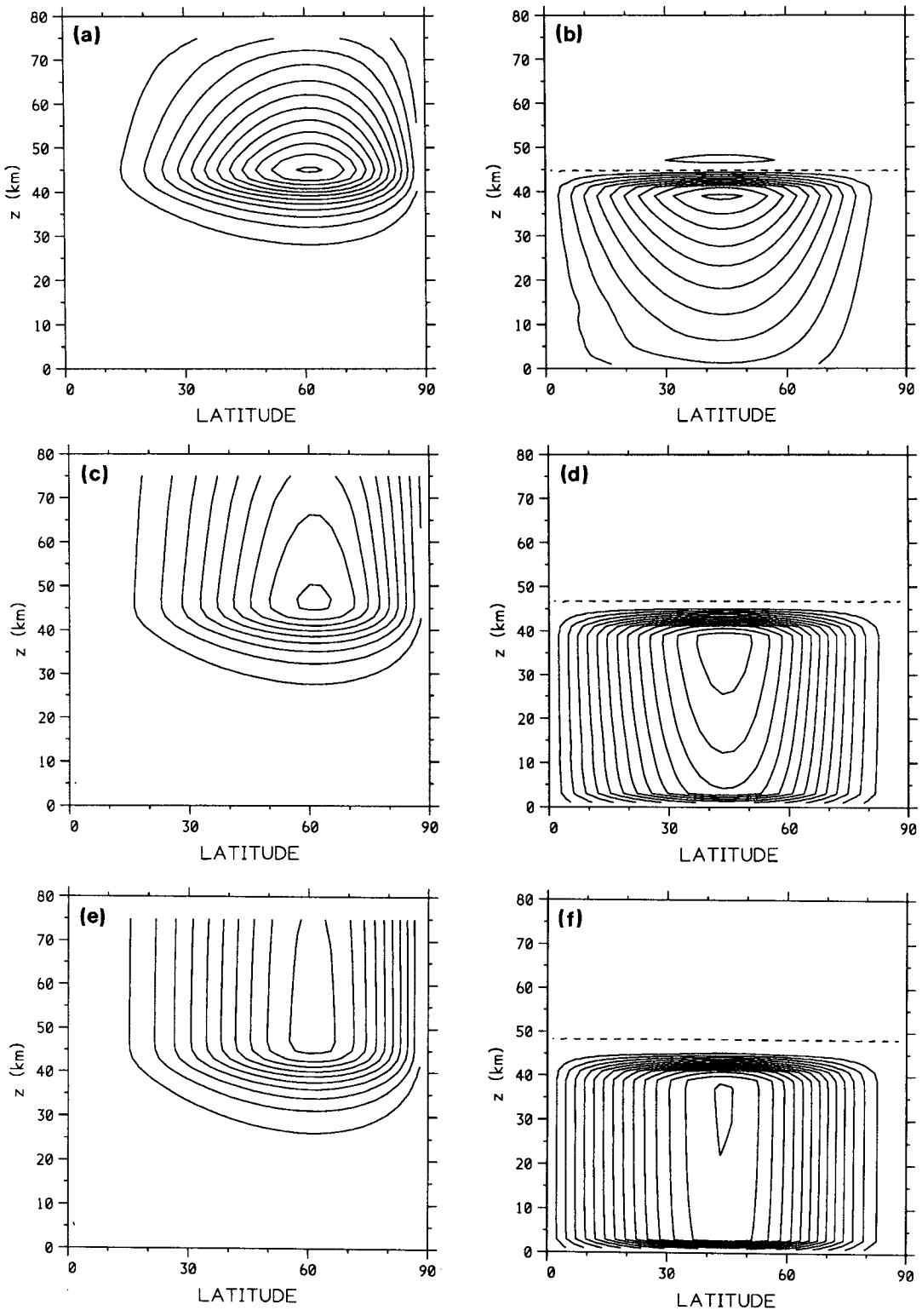

FIG. 5. Response of hemispherical atmosphere to forcing in second Hough mode, localized about 45 km. (a) Zonal wind u and (b) meridional circulation ψ for atmosphere without radiative damping: the same fields for a model atmosphere in which the temperature field is relaxed on a time scale of four days are shown in (c), (d) at 10 days and in (e) and (f) at 20 days. The contour interval for u is (a) 6×10^{-3} m s^{-1} and (c) and (e) 3×10^{-3} m s^{-1}, that for ψ is 5×10^{-7} m s^{-1} $\times$ 1000 mb in (b), (d) and (f); in (a) the magnitude corresponds to the accumulated velocity tendency at day 10.

be provided for the time lag between the change in forcing and the resulting change in the mass transport at a distance Δz below,

$$t_\alpha(\Delta z) \sim \frac{\Delta z H}{\alpha {H_R}^2}. \tag{3.27}$$

The time t_α is also that required to set up the steady state at distances Δz above and below the forcing. It is clear from (3.21a,b) that in an unbounded atmosphere the vertically integrated momentum in the downward-penetrating cell increases linearly with time, consistent with the fact that, in the model problem under consideration, there is nothing to oppose the mechanical forcing. In the real atmosphere, of course, the change in mechanical forcing will ultimately be opposed by a frictional or other additional force at some location, such as the Earth's surface. This part of the adjustment will now be examined.

4. The influence of a rigid frictionless boundary

Garcia (1987) has already investigated the effect of mechanical dissipation, represented by a constant Rayleigh friction coefficient, on the steady-state response to time-harmonic forcing, and one might consider the corresponding problem of adjustment to a sudden change in a force imposed in addition to the frictional force. However, there is little a priori evidence for processes that can be represented by strong, quasi-uniform Rayleigh friction, at least in the stratosphere. It seems more useful in the present state of knowledge to regard the entire stratospheric force as given, since it does not necessarily have a frictional character. On the other hand, we treat separately the frictional force in the planetary boundary layer, to express the fact that the boundary layer can adjust so as to allow a steady-state momentum balance. In what follows, we analyze this adjustment. First, though, it is useful to examine the role of a rigid but frictionless boundary since there are nontrivial technicalities even in that case, and the corresponding thought experiment is interesting for the way in which the angular momentum budget is satisfied.

At the lower boundary it is conventional to approximate the true boundary condition, that the geometrical vertical velocity should vanish at the ground, by $w_n = 0$ on $z = z_0 < 0$, where z_0 is an appropriate value of the log pressure coordinate (equivalent to ignoring the surface pressure change). This approximation is seriously inaccurate when considering large-scale tropospheric dynamics, for which the change in pressure at the surface is often an important part of the response to zonally symmetric forcing (Haynes and Shepherd 1989). Even for forcing in the lower and middle stratosphere, certain quantities are quite sensitive to the lower boundary condition. Note from Fig. 5b, for instance, the proportion of the apparent mass transport that enters the boundary (and is therefore associated with surface pressure changes). This sensitivity is a consequence of the fact that below the forcing, the nondimensional vertical decay rate of mass-weighted quantities, $\Lambda_-(\infty)$, is small, at least for $n = 2$ (when it is about 0.23, implying that the decay scale of the effect is about 30 km). A corrected boundary condition, therefore, is used here. For a single Hough mode it may be written in the form [Haynes and Shepherd 1989, Eq. (2.14b)]

$$w_n + \frac{4\Omega^2 a^2 H}{RT_{00}\epsilon_n} e^{z/H} \frac{\partial}{\partial z}(e^{-z/H} w_n) = 0 \quad \text{at} \quad z = z_0 \tag{4.1}$$

where $T_{00} = T_0(z_0)$. The rigid boundary has the effect of requiring an additional term in the inversion integral (3.15), leading to a correction ${w_n}^b$ to the vertical velocity, where

$$\rho_0 {w_n}^b = \frac{-1}{2\pi i} \int_\Gamma \frac{f_n H}{2\Omega a} \frac{\Lambda_+(s)}{\Lambda_-(s) - \Lambda_+(s)} \times \left[\frac{RT_{00}\epsilon_n + 4\Omega^2 a^2 \Lambda_-(s)}{RT_{00}\epsilon_n + 4\Omega^2 a^2 \Lambda_+(s)}\right] \times \exp\left\{\frac{\Lambda_+(s)(z - z_0) + \Lambda_-(s) z_0}{H}\right\} \frac{e^{st}}{s}\, ds. \tag{4.2}$$

Similar considerations to those in the previous section show that, for sufficiently large t,

$$\begin{aligned} {w_n}^b &\simeq \frac{f_n H}{4\Omega a} \exp\left(\frac{z_0}{H}\right)\left\{1 - \frac{4\Omega^2 a^2}{RT_{00}\epsilon_n}\right\}^{-1} \times \operatorname{erfc}\left\{-\left(t - \frac{b_n z'}{H\alpha}\right)\left[\frac{H\alpha^2}{4b_n(b_n+1)z'}\right]^{1/2}\right\} \\ &\simeq \frac{f_n H}{4\Omega a} \exp\left(\frac{z_0}{H}\right)\left\{1 - \frac{4\Omega^2 a^2}{RT_{00}\epsilon_n}\right\}^{-1} \times \operatorname{erfc}\left\{\left(z' - \frac{H\alpha t}{b_n}\right)\left[\frac{{b_n}^2}{4H^2(b_n+1)\alpha t}\right]^{1/2}\right\} \end{aligned} \tag{4.3}$$

where $z' = z - 2z_0$, and that the corresponding correction to u_n, say ${u_n}^b$, is given by

$${u_n}^b \simeq \frac{f_n}{2} \exp\left(\frac{z_0}{H}\right)\left[\frac{4z' b_n(b_n+1)}{H\alpha^2}\right]^{1/2}\left\{1 - \frac{4\Omega^2 a^2}{RT_{00}\epsilon_n}\right\}^{-1} \times \left\{\mathcal{X}\left(\left(\frac{b_n z'}{H} - \alpha t\right)\left[\frac{H}{4z' b_n(b_n+1)}\right]^{1/2}\right) - \mathcal{X}\left(\frac{b_n z'}{H}\left[\frac{H}{4z' b_n(b_n+1)}\right]^{1/2}\right)\right\} \tag{4.4}$$

where $\mathcal{X}(u) \equiv \int_u^\infty \operatorname{erfc} u'\, du'$ and $\mathcal{X} \sim -2u + O(1)$ as $u \to -\infty$. Thus, for large times, this part of the solution dominates the one calculated in the previous section and leads to an asymptotic state in which w_n is independent of height and constant in time, and u_n is independent of height (and therefore has no asso-

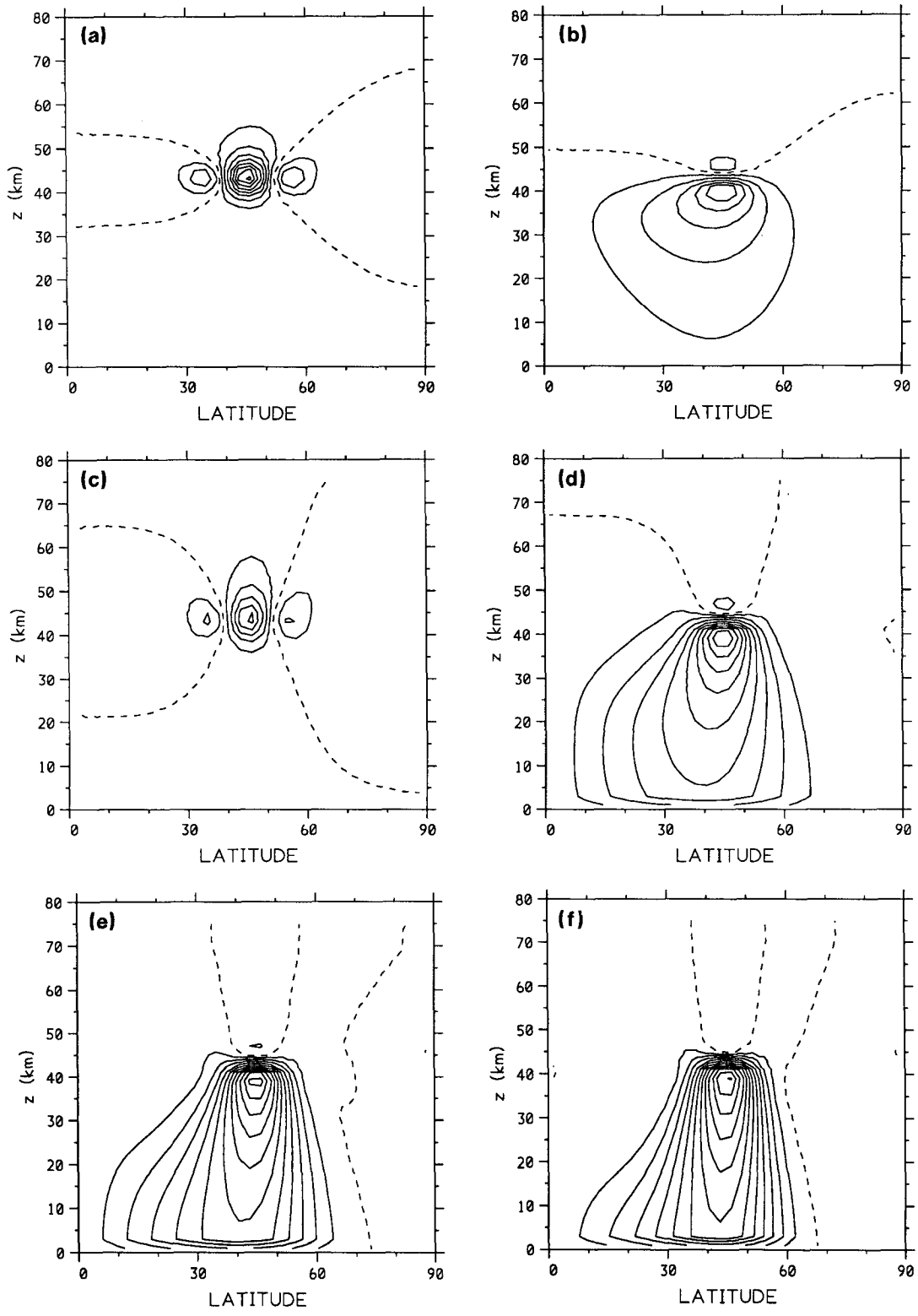

FIG. 6. Response to forcing localized both in latitude (about 45°) and height (about 45 km). (a) Zonal wind u and (b) meridional circulation ψ for atmosphere without radiative damping; (c)–(j) the same for atmosphere with four day radiative timescale after 10, 20, 30, 40, 50 and 100 days, except that the zonal wind is shown only in (c) 10 and (i) 100 days. The contour interval for u is 5×10^{-2} m s^{-1}, that for ψ is 10^{-6} m s^{-1} $\times$ 1000 mb; in (a) the magnitude corresponds to the accumulated velocity tendency at day 10.

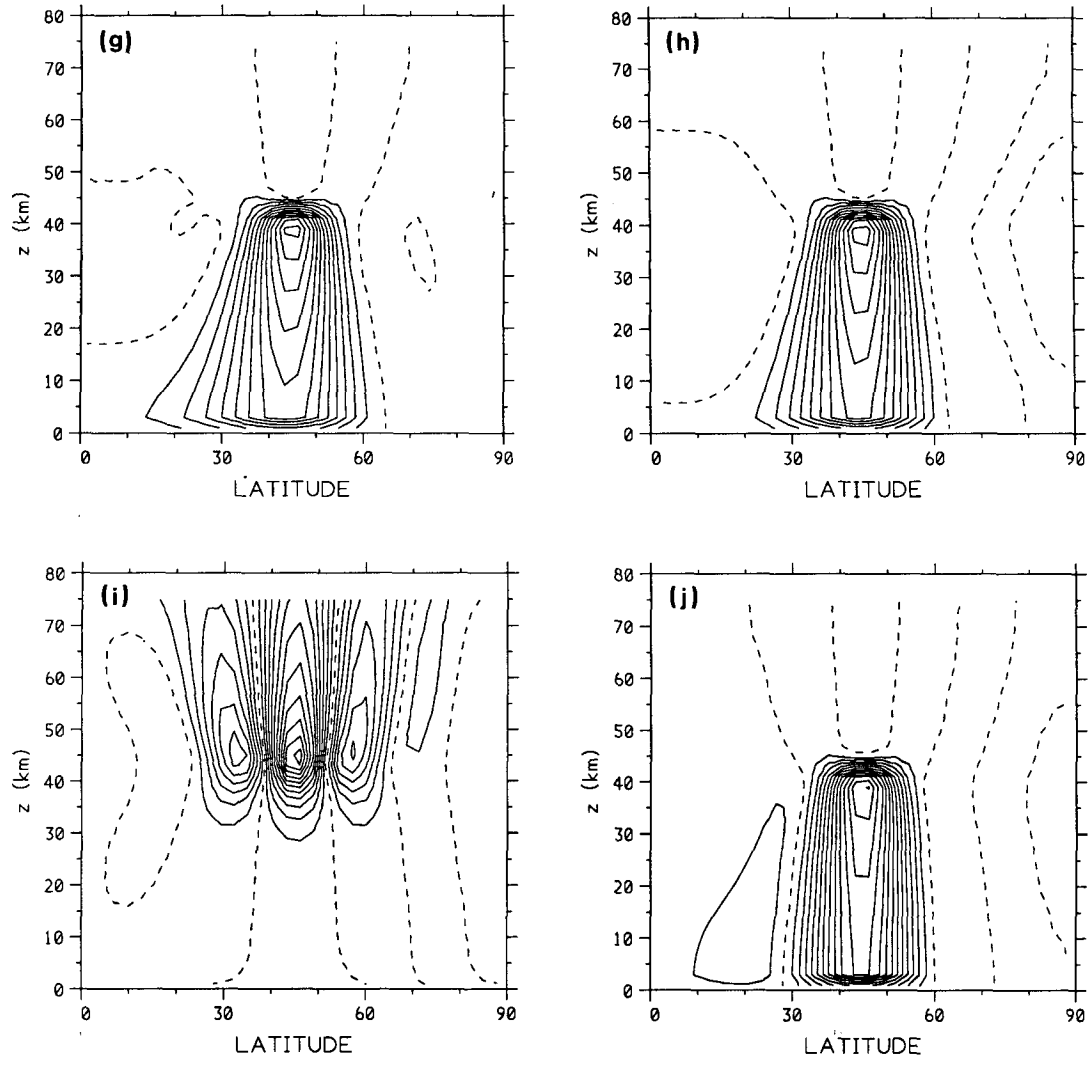

FIG. 6. (*Continued*)

ciated temperature perturbation) but increases linearly with time. The linear increase is a consequence of the fact, noted earlier, that there is nothing to balance the mechanical forcing. The presence of the rigid but frictionless boundary limits the amount of angular momentum that can be taken up below the forcing without a secular increase in zonal velocities. The application of the correct boundary condition (4.1) allows explicit consideration of changes in the surface pressure and, therefore, a latitudinal shift in mass, with a resulting partition of the excess angular momentum between relative and planetary components. It is apparent from (4.4) that a fraction $\{1 - (4\Omega^2 a^2/RT_{00}\epsilon_n)\}^{-1} = \{1 + (H_R{}^2N^2/RT_{00})\}^{-1}$ appears as an increase in the relative angular momentum. For the case of an isothermal atmosphere, this fraction may also be written as $\{1 + (\kappa H_R{}^2/H^2)\}^{-1}$. Note that this fraction is achieved once the downward mass circulation has reached the ground, and that it does not depend on the height of the forcing above the ground. As before, the adjustment zone across which flow variables change to their asymptotic behavior propagates upward at speed w_α, and thickens as if diffusion were acting with diffusivity κ_α. However, it needs propagation through only a few density scale heights before the vertically integrated rate of change of relative angular momentum becomes close to its steady state value. Figure 10 shows schematically the configuration of the adjustment in the vertical. Note that, below the upward propagating adjustment region, the vertical velocity becomes independent of height, but not the vertical mass transport, and it is the associated latitudinal mass transport that leads to a steady non-zero Coriolis torque.

5. Adjustment in the presence of a planetary boundary layer

Now suppose that mechanical dissipation is represented by an extra term on the right-hand side of (3.1a)

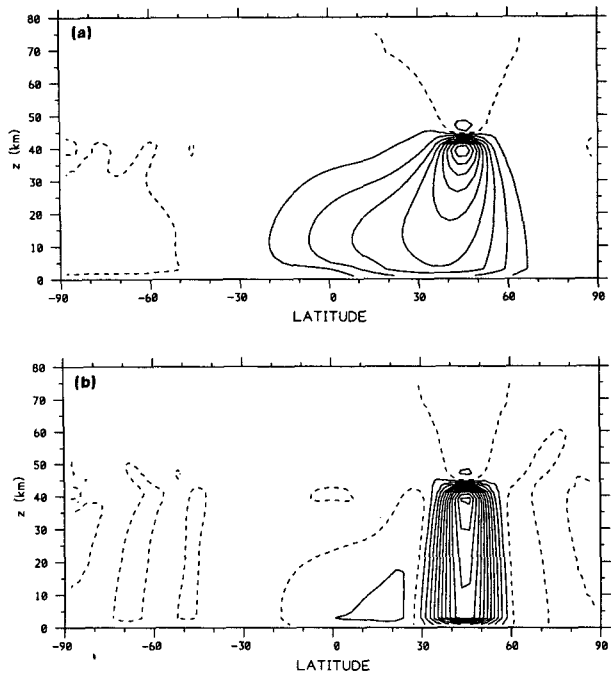

FIG. 7. As Fig. 6, but response of meridional circulation (a) after 10 days and (b) after 100 days for spherical atmosphere.

of the form $\mathcal{G}[u; \mu, z]$, a functional of u for each μ and z. Thus, the total zonal force is now $\mathcal{F} + \mathcal{G}$. In the presence of a rigid lower boundary, the steady-state form of (3.1a) may be integrated in the vertical to give

$$\int_{z_0}^{\infty} e^{-z/H}\mathcal{F}dz = -\int_{z_0}^{\infty} e^{-z/H}\mathcal{G}[u; \mu, z]dz. \quad (5.1)$$

This gives one constraint on the vertical profile of u for each μ, exactly what is required to supplement the thermal wind relation and determine the velocity field uniquely from the steady state equations. Without such information on the form of the mechanical dissipation, the velocity field remains indeterminate up to a function of μ.

It is convenient to investigate the manner in which the flow adjusts to achieve the balance (5.1) by considering the simplest possible choice of $\mathcal{G} = -\gamma\delta((z - z_0)/H)u$, representing a boundary layer at $z = z_0 < 0$. Integration across this layer, and use of the continuity relation (3.1e) in the standard way, implies that the effect of the friction may be represented by a suitably modified boundary condition on $\tilde{w}_n(s, z)$ at $z = z_0$ of the form

$$\tilde{w}_n + \left(\frac{4\Omega^2a^2H}{RT_{00}\epsilon_n} - \frac{\gamma H}{s}\right)e^{z/H}\frac{\partial}{\partial z}(e^{-z/H}\tilde{w}_n) = 0. \quad (5.2)$$

The corresponding term in the vertical velocity, associated now with both the rigid boundary and the frictional layer, is given by

$$\rho_0 w_n{}^b = -\frac{1}{2\pi i}\int_\Gamma \frac{f_nH}{2\Omega a}\frac{\Lambda_+(s)}{\Lambda_-(s) - \Lambda_+(s)} \times \left[\frac{4\Omega^2a^2s\Lambda_-(s) + RT_{00}\epsilon_n(s - \gamma\Lambda_-(s))}{4\Omega^2a^2s\Lambda_+(s) + RT_{00}\epsilon_n(s - \gamma\Lambda_+(s))}\right] \times \exp\left\{\frac{\Lambda_+(s)(z - z_0) + \Lambda_-(s)z_0}{H}\right\}\frac{e^{st}}{s}ds. \quad (5.3)$$

Note that the pole in the integrand at $s = 0$ has disappeared. The expression is closer in form to that found for the vertical mass transport above the forcing in section 3. By analogy, it is evident that a meridional cell propagates away from the boundary, leaving behind it a zonal velocity that is constant in height, and such that the steady state requirement (5.1) is approached. Asymptotic investigation of the relevant expressions for large t confirms this qualitative picture. The velocity correction, $u_n{}^b$, may be estimated as

$$u_n{}^b \simeq \frac{f_n}{2}\exp\left(\frac{z_0}{H}\right)\left\{\frac{1}{\gamma} - \frac{b_n}{\alpha}\right\} \times \operatorname{erfc}\left\{\left(z' - \frac{H\alpha t}{b_n}\right)\left[\frac{b_n{}^2}{4H^2(b_n + 1)\alpha t}\right]^{1/2}\right\}. \quad (5.4)$$

This expression is independent of the factor $4\Omega^2a^2/RT_{00}\epsilon_n$ that is an indicator of the importance of the surface pressure changes, implying (as might have been expected) that these are zero in the steady state limit.

It may be deduced that the time, τ_b, required for the velocity at the level of the forcing to achieve its steady state value is just the time required for propagation of the signal to the ground, and then upward again, and is given by

$$\tau_b = \frac{2|z_0|}{w_\alpha} = \frac{2|z_0|H}{\alpha H_R{}^2}.$$

Adjustment on this timescale is an implicit assumption in any calculation of gradient wind from a given radiative temperature field that requires the velocity distribution to be given at the lower boundary.

The qualitative picture described here has been confirmed by numerical integration of the partial differential equation (3.10), and a schematic picture of the adjustment is given in Fig. 11.

If the Rayleigh friction layer is not at the ground but at the level $z = z_1$, say, then the picture is more complicated. The solution $\tilde{w}_n$ is now required to satisfy (4.1), together with the condition

$$[\tilde{w}_n]_{z_1^-}^{z_1^+} - \frac{\gamma H}{s}e^{z/H}\frac{\partial}{\partial z}(e^{-z/H}\tilde{w}_n)\Big|_{z=z_1} = 0 \quad (5.5)$$

(u, Φ being assumed continuous, and the second term having equal limiting values just above and below z_1; contrast Ekman layer theory). The linear equation for $\tilde{w}_n$ may be solved as before, but the solutions are sufficiently complicated that it is difficult to make ana-

lytical progress in parameter regimes that allow new behavior. Nonetheless, a qualitative description of the solution may be attempted on the basis of what has been found already. Again one expects a period in which the circulation burrows downward toward the boundary. This part of the evolution is essentially unaffected by the presence of the friction layer, apart from a reduction in strength of the circulation cell by a fraction that is equal to $\gamma/(\gamma + b_n^{-1}\alpha)$ when $z_1 \leq 0$ and $\gamma e^{-z_1/H}/(\gamma + b_n^{-1}\alpha)$ when $z_1 > 0$. This reduction may be much smaller in the latter case, and, furthermore, appears below $z = 0$ only after a time lag $2z_1/w_\alpha$. Once the cell has reached the ground the boundary adjustment described earlier takes place, in which a region of constant acceleration extends upward at a speed w_α. Once this region has reached the friction layer at a time $(z_1 - 2z_0)/w_\alpha$, the zonal velocity there increases initially as

$$u_n \simeq \left(t - \frac{z_1 - 2z_0}{w_\alpha}\right) f_n e^{z_0/H}\left(1 - \frac{\gamma e^{-[z_1]/H}}{\gamma + b_n^{-1}\alpha}\right)$$

where $[z_1] = \max(z_1, 0)$. For the system to come into equilibrium (5.1) implies the balance $f_n = \gamma u_n(z_1) \times e^{-z_1/H}$, and the time τ_f required may be estimated as

$$\tau_f = \frac{(z_1 - 2z_0)}{w_\alpha} + \frac{e^{(z_1 - z_0)/H}}{\gamma}\left(1 - \frac{\gamma e^{-[z_1]/H}}{\gamma + b_n^{-1}\alpha}\right)^{-1}.$$

An important point to note is that the time for this balance to be achieved, and the ultimate steady state velocity, both increase exponentially with the number of scale heights that the friction layer is above the ground when this number is sufficiently large.

The case $z_1 > z_0$ may seem somewhat abstract, but might be relevant if the frictional force $\mathcal{G}$ that ultimately opposed the effect of planetary-wave induced forcing in the stratosphere, say, arose in regions of high gravity-wave breaking (e.g., at the tropopause or in the mesosphere). Certainly in some middle-atmosphere models the effect of this gravity-wave breaking is represented as a layer of increased Rayleigh-friction coefficient near the top of the model. If there were no other angular momentum sink to balance the prescribed force $\mathcal{F}$, then the mass circulation associated with $\mathcal{F}$ would have to close off in that layer. The idea of "upward control" might then seem more appropriate, but in the time-dependent problem would be achieved in the manner just outlined, leading eventually to a cell extending downward from the friction layer, which would ultimately shut off all meridional circulation below the level of the applied forcing. The steady state tracer-bearing circulation in such a model would be far from that actually observed in the real atmosphere. There would be meridional flow only within the forcing layer and, in the opposite sense, in the friction layer, with vertical flow between forcing and friction layers to close the circulation. In fact, such a case is unlikely to have arisen, because most models include a relatively large mechanical damping that is applied over most heights and latitudes. This is also a potentially serious problem, since, like the upper damping layer, the more extensive mechanical damping prevents the penetration of the steady state mass circulation to levels far below the forcing (see Garcia 1987, especially Fig. 1). It is, therefore, likely that such models significantly underestimate that part of the lower stratospheric diabatic circulation that is controlled by $\mathcal{F}$ in the middle and upper stratosphere, for example. Evidence that, in some parts of the atmosphere, this problem may indeed be serious is presented in the next section. The performance of such models might well be improved by reducing internal friction and, if necessary, including a frictional layer at the lower boundary of the model. Such a change would allow the mass circulation to penetrate downward and to close off in the boundary layer. Note incidentally that Matsuno-type models, with geopotential specified at an artificial lower boundary, allow the circulation to penetrate *across* the boundary which can therefore act as an angular momentum sink without any friction having been specified.

6. Downward control domain depths estimated for the real atmosphere

In this section, satellite observations of temperature and ozone in the middle atmosphere are used, in conjunction with a radiative transfer scheme, to calculate the diabatic circulation. From this circulation an estimate can be made of the height to which the atmosphere must be modeled to simulate reliably the diabatic vertical velocity in, for instance, the lower stratosphere.

A large number of recent studies have used satellite data to derive the diabatic or residual circulation (e.g., Solomon et al. 1986; Gille et al. 1987). The results reported here are derived from the diabatic circulation calculated by Shine (1989) and defined as the flow with meridional and vertical velocities $(\bar{v}^*, \bar{w}^*)$ that satisfy the thermodynamic and continuity equations:

$$\frac{\partial \bar{\theta}}{\partial t} + \frac{\bar{v}^*}{a}\frac{\partial \bar{\theta}}{\partial \phi} + \bar{w}^*\frac{\partial \bar{\theta}}{\partial z} = \bar{\mathscr{Q}} e^{\kappa z/H} \qquad (6.1a)$$

and

$$\frac{1}{a\cos\phi}\frac{\partial}{\partial \phi}(\bar{v}^*\cos\phi) + \frac{1}{\rho_0}\frac{\partial}{\partial z}(\rho_0 \bar{w}^*) = 0, \qquad (6.1b)$$

where θ is the potential temperature. This is the TEM residual circulation as defined by Andrews et al. [1987, Eqs. (3.5.2a–e)], apart from a term on the right-hand side of the thermodynamic equation involving eddy fluxes [cf. (2.9) above]. This term has been calculated explicitly using the three-dimensional circulation derived in Marks (1989), and is less than 0.2 K per day everywhere. The residual circulation $(\bar{v}^*, \bar{w}^*)$ is nearly

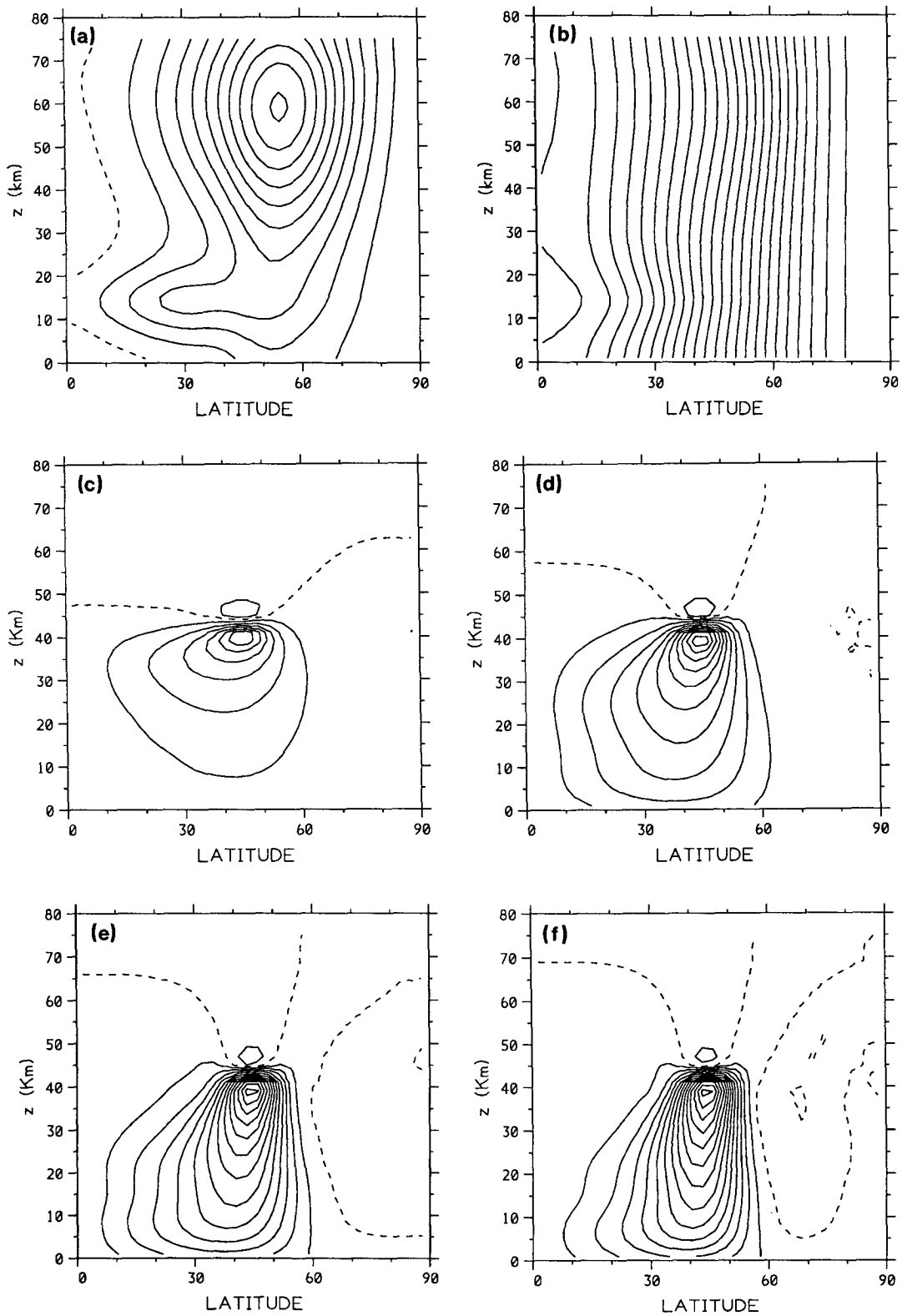

FIG. 8. Response of a nonresting hemispherical atmosphere, with a realistic background zonal flow, to force localized about 45°N and 45 km: (a) background flow; (b) associated angular momentum distribution; (c) meridional circulation in undamped case; (d)–(i) meridional circulation in radiatively damped case after 10, 20, 30, 40, 50 and 100 days. The contour interval for (a) is 10 m s^{-1}, that for (b) is 20 m $s^{-1} \times a$, and that for (c)–(i) is 10^{-6} m $s^{-1} \times 1000$ mb.

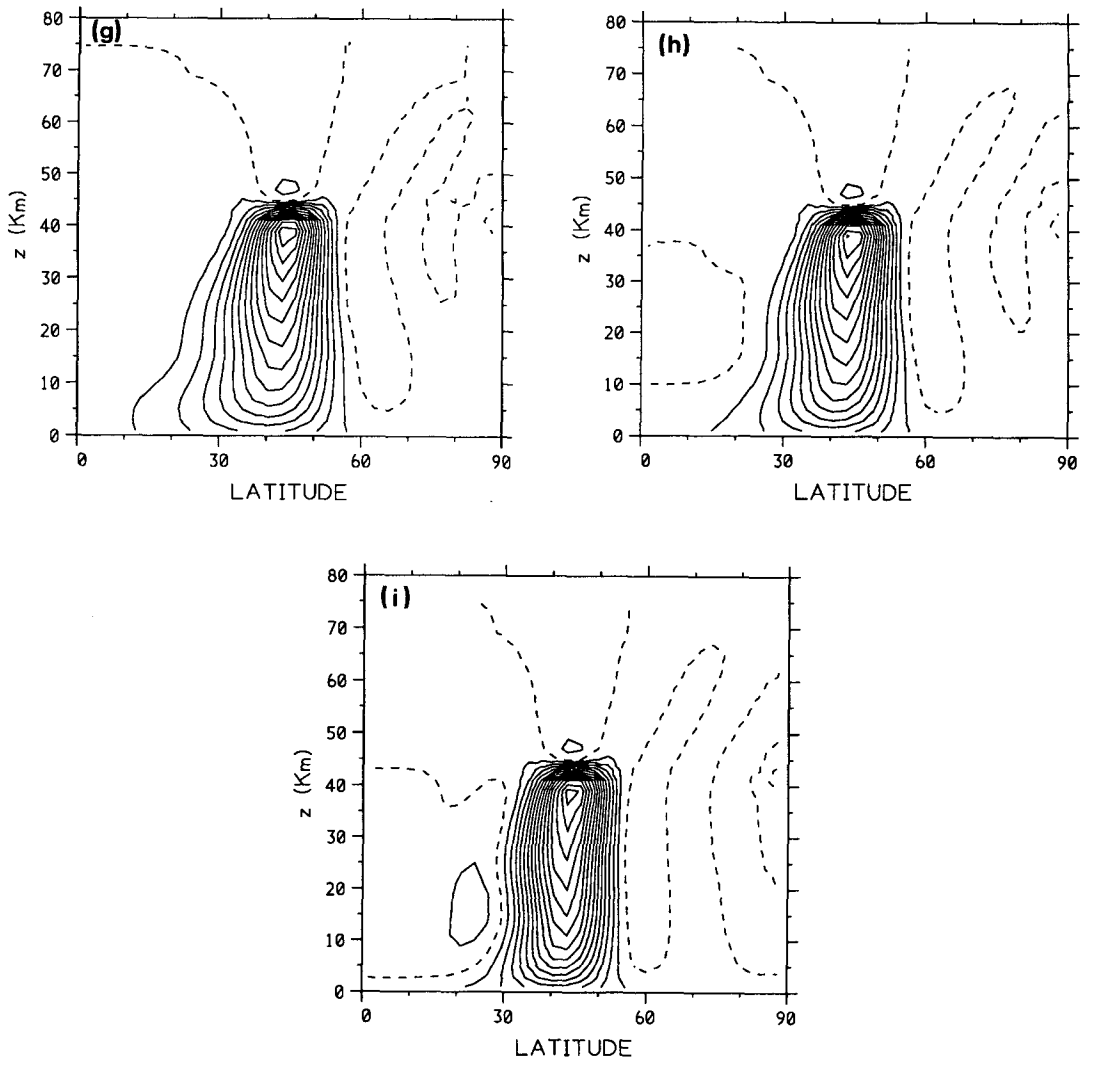

FIG. 8. (*Continued*)

equal to the diabatic circulation since $\partial\bar{\theta}/\partial t$ makes only a relatively small contribution to (6.1a) (Shine 1989). Recent satellite climatologies of temperature (from Barnett and Corney 1985) and ozone [from Keating et al. (1987) for pressures less than 20 mbar, and reprocessed SBUV data for pressures greater than 20 mbar] were used in conjunction with the radiative transfer scheme described by Shine (1987) to give the diabatic heating rate $\bar{\mathscr{H}}$. Equations (6.1a) and (6.1b) were then solved for $\bar{v}^*$ and $\bar{w}^*$. This method of diagnosing the circulation is essentially the same as that used by Murgatroyd and Singleton (1961), although it was not then known as the diabatic circulation, nor was its full significance and relationship with other aspects of the circulation appreciated until some seventeen years later (Dunkerton 1978).

The mean velocity field is now treated as if the horizontal component $\bar{v}^*$ were derived from the momentum equation, in which case the vertical component could be calculated from the continuity equation (6.1b) using the requirement that $\rho_0\bar{w}^* \to 0$ as $z \to \infty$; viz.

$$\bar{w}^*(\phi, z) = \frac{1}{a\rho_0(z)\cos\phi}\int_z^\infty \rho_0(z') \times \frac{\partial}{\partial\phi}[\bar{v}^*(\phi, z')\cos\phi]dz' = \lim_{\zeta\to\infty} W(\phi, z, \zeta) \quad (6.2a)$$

where

$$W(\phi, z, \zeta) = \frac{1}{a\rho_0(z)\cos\phi}\int_z^\zeta \rho_0(z') \times \frac{\partial}{\partial\phi}[\bar{v}^*(\phi, z')\cos\phi]dz'. \quad (6.2b)$$

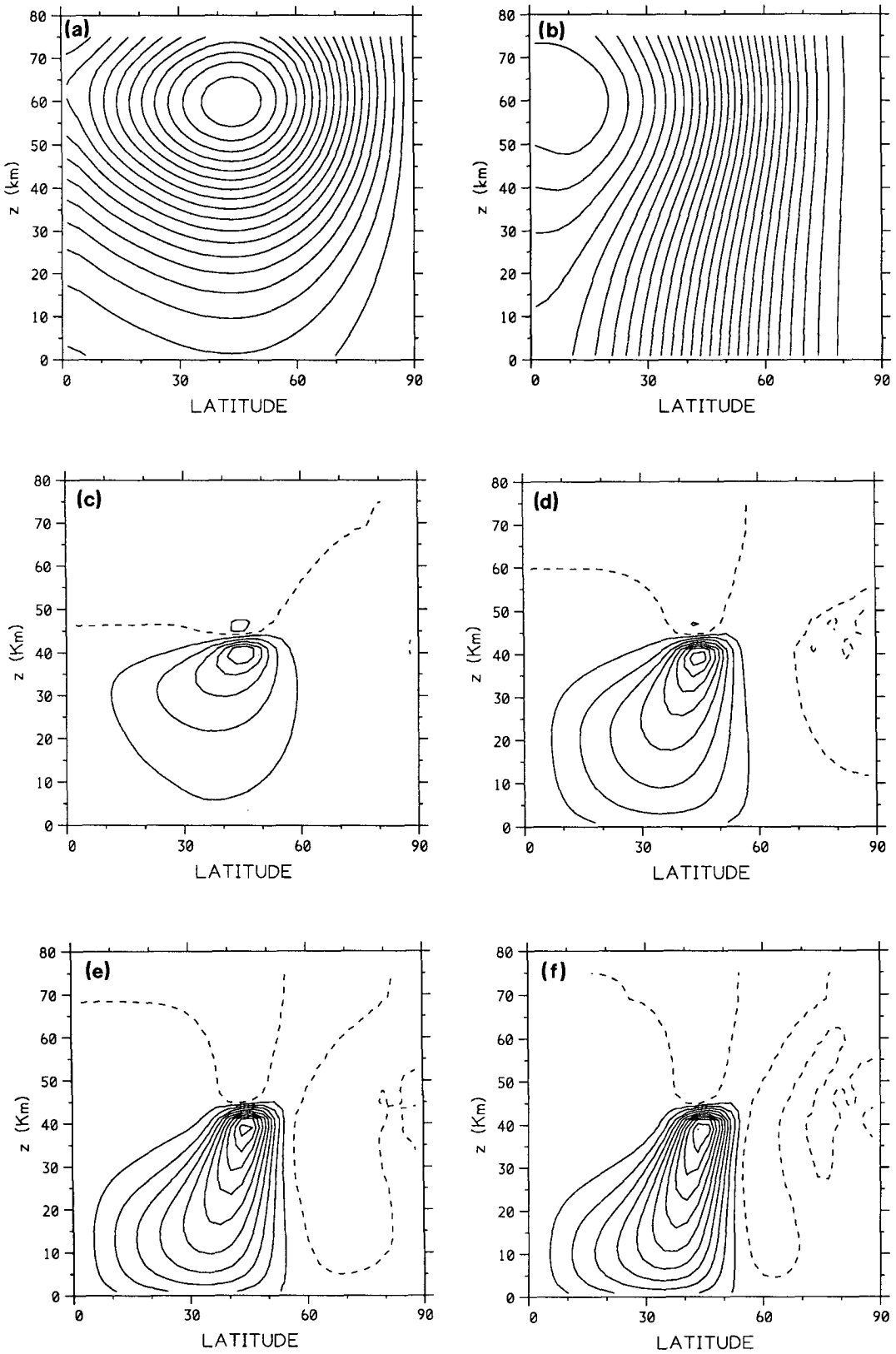

FIG. 9. As Fig. 8, but for hemispherical atmosphere with very strong background flow and angular momentum contours departing strongly from the vertical.

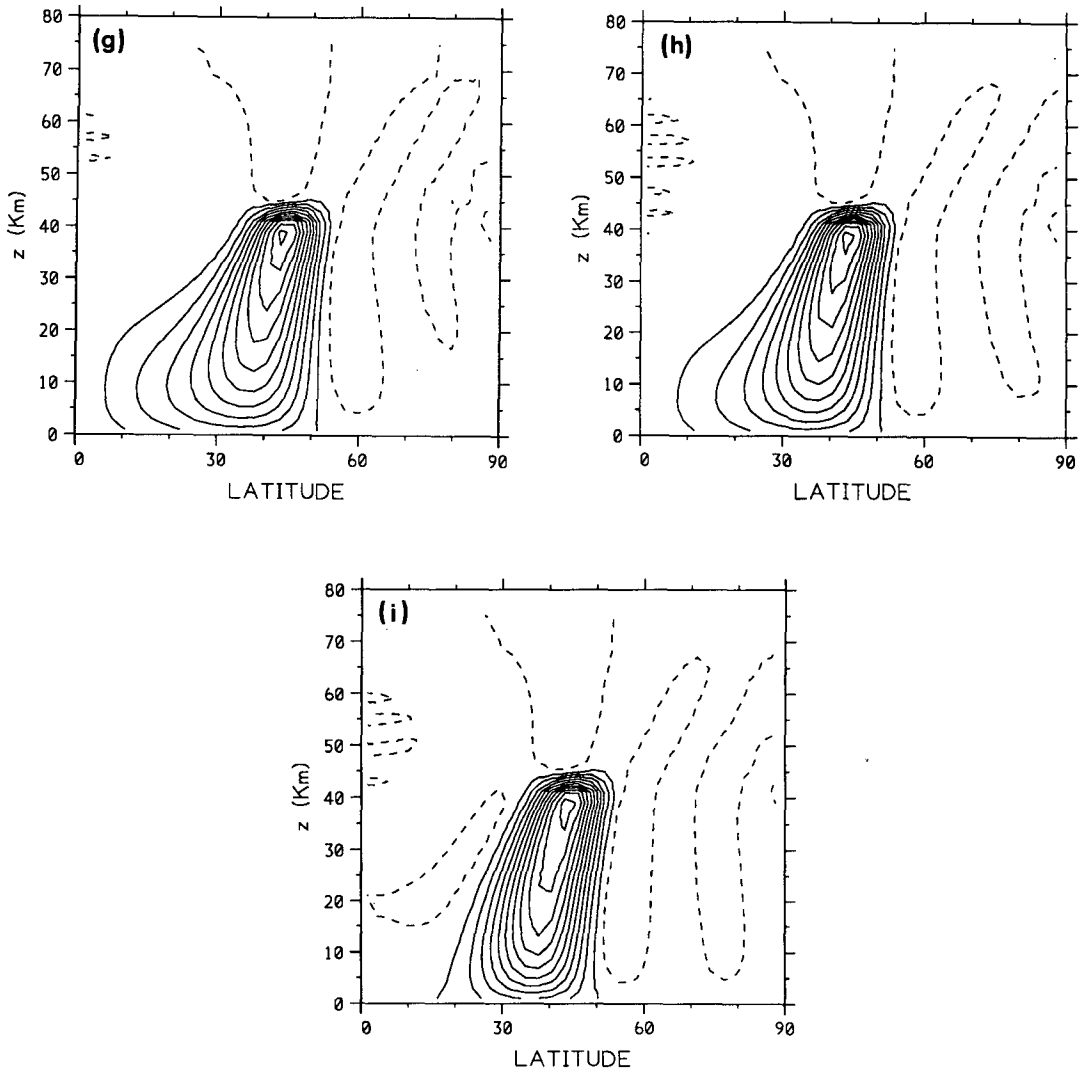

FIG. 9. (*Continued*)

Under the assumption of a steady balance in the momentum equation, (6.2a) is equivalent to (2.7) and expresses the downward control of the diabatic vertical velocity at any level by contributions from the zonal drag at all higher levels, embodied here by the meridional velocity $\bar{v}^*$ associated with the balancing Coriolis torque. The question to be addressed is therefore: How large does the upper limit ζ have to be taken in practice, in order to obtain a good approximation to the infinite integral?

The $\bar{v}^*$ fields of Shine (1989) have been used to calculate $W(\phi, z, \zeta)$ with $z = 2.608H$ (approximately 18 km) for the whole year, and for ζ up to $11.5H$. When $\zeta = 11.5H$ (approximately 80 km) expression (6.2b) reproduces almost exactly the $\bar{w}^*$ fields at $z = 2.608H$ calculated by Shine (1989). (This says no more than that $\rho_0\bar{w}^*$ at $11.5H$ is very much less than $\rho_0\bar{w}^*$ at $2.608H$.) Figure 12 shows the latitude time series of $W(\phi, 2.608H, 11.5H)$ from the CIRA climatology. Note the ascent at low latitudes and descent at high latitudes (which is particularly strong at the northern winter pole).

The convergence of $W(\phi, 2.608H, \zeta)$ to a limit as ζ increases, which is assured by the presence of the density term in the integrand of (6.2a), indicates the "top" of the downward control region. This height is presumably that beyond which general circulation models must extend in order to model lower stratospheric diabatic circulations correctly, including transport of stratospheric photochemical products toward the troposphere, and cross-tropopause transport. To fix ideas this level is taken to be the height ζ above which $W(\phi, 2.608H, \zeta)$ is within 20% of its "true" value taken as $W(\phi, 2.608H, 11.5H)$.

Figure 13 shows a latitude time plot of the height (km) of the "top" of the downward control region for

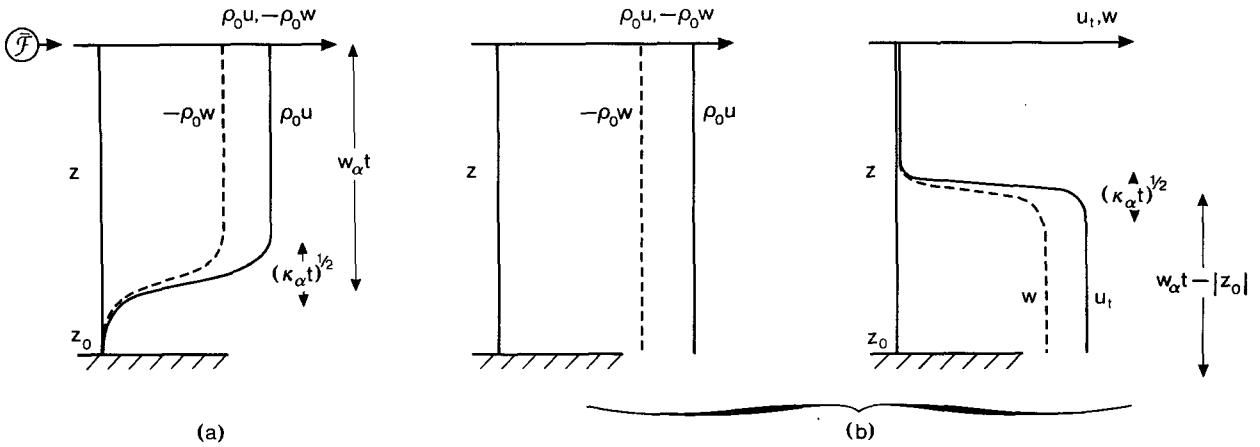

FIG. 10. As Fig. 3, but for atmosphere with a rigid boundary at z_0 (<0): (a) $t < |z_0|/w_\alpha$; (b) $t > |z_0|/w_\alpha$, showing the two components of the solution, the correction due to the boundary being on the right. Note that u_t, rather than u, is shown in the rightmost panel.

$z = 2.608H$ as diagnosed from the CIRA data. It is clear that in midlatitudes the top of the region is at 30–40 km, i.e., two or three scale heights above the level of interest, as expected from simple scale analysis (see section 2). Polewards of about 60° in June and July, however, the top of the region is above 50 km, with a maximum above 60 km in very high southern latitudes in July.

An important aspect of Shine's (1989) study concerns the problem of correcting the vertical velocities to ensure that the constraint of zero globally averaged mass flux across an isobaric surface is maintained. The circulation was shown to be significantly sensitive to the manner in which this correction is performed. The sensitivity of the top of the downward control region to the manner of this correction must also be considered here, in order to assess how robust the present results are.

Figure 14 shows the sensitivity of W at $z = 2.608H$ and $\phi = -67.5°$ to various forms of the correction. The precise details of the corrections are described in Shine (1989) but are not of crucial importance here. The important point is that the conclusions seem independent of the way in which the heating rate (and hence the diabatic circulation) is adjusted to ensure mass balance. Although the diagnosed value of W at $\phi = -67.5°$, $z = 2.608H$, $\zeta = 11.5H$ is sensitive to the way in which the correction is made (the values vary between -0.35 mm s^{-1} and -0.1 mm s^{-1}), the shape of the profiles shows that convergence is not satisfactory until $\zeta \gtrsim 60$ km for all profiles.

It is important to note from Fig. 14 that the sign of the contribution to the integral in expression (6.2) varies with height; hence, if the vertical extent of a model falls far short of adequacy, it may not be true that a vertical extension will improve the representation of,

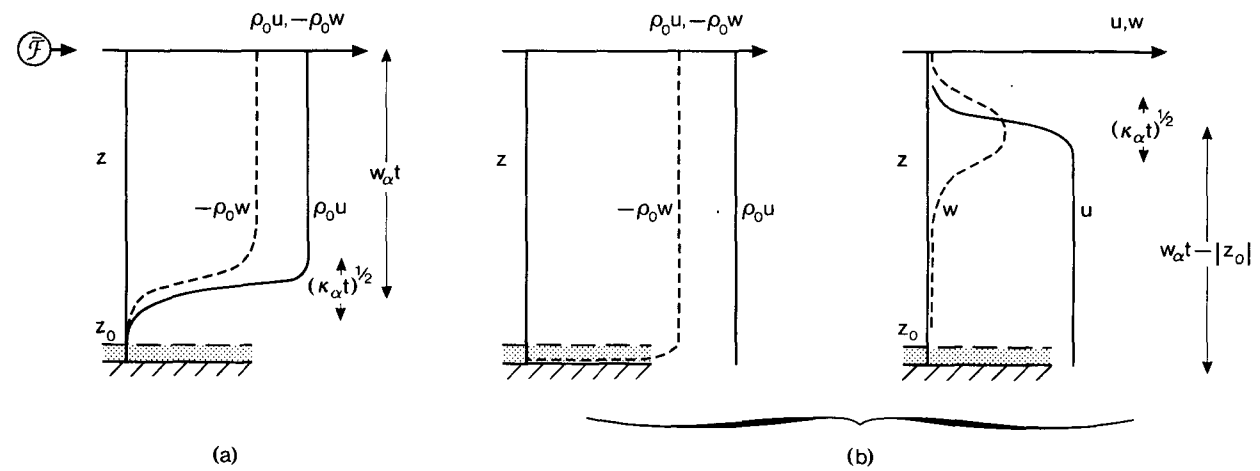

FIG. 11. As Fig. 10, but for atmosphere with a rigid boundary and a thin boundary layer at z_0 (<0), with $\rho_0 u$ or u being shown in each case.

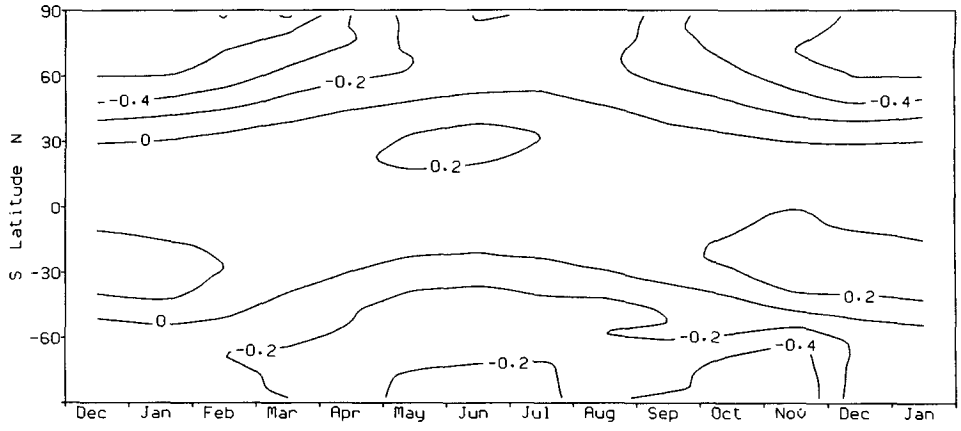

FIG. 12. Latitude time plot of $W(\phi, 2.608H, 11.5H)$, in mm s^{-1}, which approximates $\bar{w}^*$ at about 18 km.

for instance, the stratosphere–troposphere exchange rate, unless that vertical extension goes beyond altitudes that exert any influence on that representation.

It might seem somewhat surprising that, in some locations, a substantial fraction of lower-stratospheric descent rates arises from eddy-induced forces exerted well into the mesosphere. That this is likely to be the case may be seen directly from the distribution of the zonal force $\bar{\mathcal{F}}$ per unit mass estimated by Shine (1989) on inserting the mean velocities in the TEM momentum equation and is confirmed below by order-of-magnitude arguments. Shine found that $\bar{\mathcal{F}}$ was maximum in the midlatitude mesosphere; in January in the Northern Hemisphere it reached -30 m s^{-1} day^{-1}, whilst in the Southern Hemisphere in July it reached -60 m s^{-1} day^{-1}. It was argued by Shine that this interhemispheric difference is probably associated with the increased ability of gravity waves with a weak westerly phase speed (therefore, easterly relative to the mean flow) to penetrate through the less disturbed southern winter stratosphere, to be dissipated in the midmesosphere. The strong interhemispheric difference is also reflected, as it must be, in the marked asymmetry in middle mesosphere temperatures; in July, the lower mesosphere temperatures near the South Pole are more than 16 K warmer than their Northern Hemisphere counterparts in January. Shine's (1989) results also indicate that in the lower stratosphere, on the other hand, wintertime high-latitude values of $\bar{\mathcal{F}}$ were smaller in the Southern Hemisphere, entirely consistent with theoretical notions that besides differences in tropospheric forcing the larger, stronger Southern Hemisphere vortex tends to exclude planetary wave activity.

Given the sense, at the different levels, of the interhemispheric differences in $\bar{\mathcal{F}}$, it is thus clear from (6.2a) that the altitude from which substantial control of $\bar{w}^*$ in the lower stratosphere is exerted is considerably higher in the Southern Hemisphere winter than in the Northern winter.

Shine (1989) discusses the relation of his estimates

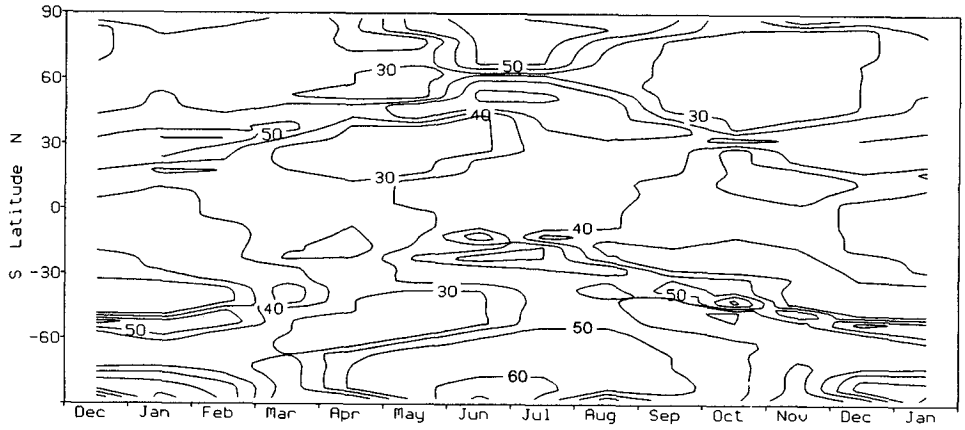

FIG. 13. Latitude time plot of the approximate altitude (in km) to which the diabatic circulation must be calculated to represent $\bar{w}^*$ at about 18 km to within 20%.

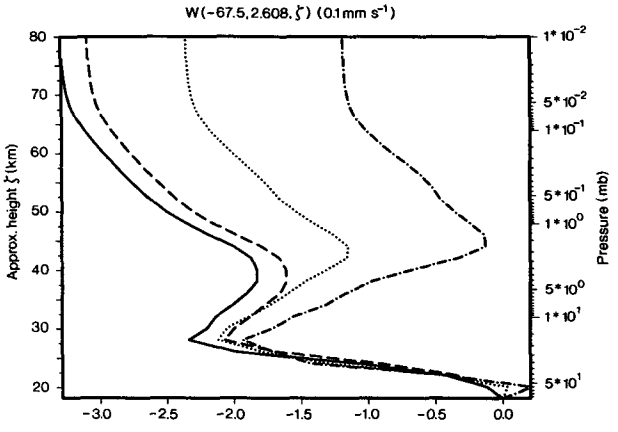

FIG. 14. Variation of the quantity $W(-67.5°, 2.608H, \zeta)$ defined in (6.2b) (in 0.1 mm s^{-1}) in July as a function of ζ, for four different methods of correcting the circulation to ensure the globally averaged mass flux across an isobaric surface is zero, as it must be by definition in a hydrostatic atmosphere. (Dotted line, correction independent of latitude; solid line, correction assuming all error in solar heating calculations; dashed line, correction entirely in Northern Hemisphere; dash-dot line, correction entirely in Southern Hemisphere). The calculations shown in Figs. 12 and 13 were done assuming a correction independent of latitude.

of $\bar{\mathcal{F}}$ to those of other authors and finds general agreement as to the order of magnitude for the mesospheric force per unit mass, for example. An estimate of around 20 m s^{-1} day^{-1} for the force per unit mass exerted at about 60 km seems conservative and consistent, at 60° latitude, with a meridional velocity of 2 m s^{-1}. Assuming a horizontal scale of 2000 km, and a vertical scale of 7 km, this implies through the continuity relation an associated vertical velocity of 7 mm s^{-1}. The vertical mass transport associated with such a velocity corresponds at 18 km to a vertical velocity of 0.02 mm s^{-1}, entirely consistent with the results shown in Fig. 14.

The above results clearly have important implications for the vertical extent necessary for a model if it is to simulate the diabatic circulation in the lower stratosphere correctly. It is important to take account not only of the convergence properties of the integral (6.2a), but also of the corollary (2.8) noted in (3) of §2. One might be tempted to conclude that, if much of the contribution to $\bar{\mathcal{F}}$ above 45 km, say, arises from vertically propagating gravity waves, then, apart from a passive "sponge" layer to absorb wave activity, it would be unnecessary to extend models above this level. However, it would be important that the correct momentum fluxes were included in such a model, so the gravity waves would have to be parameterized in some way even if they were not needed to close off the mesospheric jets, for instance. More seriously, given the horizontal inhomogeneity in the gravity-wave sources, it may well be that horizontal momentum fluxes are important above the stratopause, acting to redistribute (perhaps through locally forced Rossby waves) the effect of breaking gravity waves. Furthermore, it seems unwise, on the present evidence, to assume that latitudinally propagating Rossby waves of tropospheric origin play no substantial role in the momentum budget of the mesosphere. It therefore seems difficult to rule out the possibility that accurate modeling well into the mesosphere is necessary to obtain a good representation of the lower stratospheric diabatic circulation, and hence stratosphere–troposphere exchange rates.

7. Discussion

The regions where the dynamical ideas discussed above are relevant, particularly the principle of downward control, are those where the angular momentum ($\bar{m}$) contours shown in Fig. 1 span the atmosphere in the vertical. On the other hand, the downward control principle fails in regions not vertically spanned by angular-momentum surfaces, such as must always occur at least somewhere in the tropics. For instance there may be regions where the angular momentum is nearly constant within the tropical tropospheric Hadley circulation (Held and Hou 1980) and perhaps also in the tropical stratosphere (Dunkerton 1989b); see also Fig. 1. If such a region exists then it would allow vertical recirculation with no requirement for an eddy-induced force $\bar{\mathcal{F}}$. Also, as Dunkerton (1989b) points out, even a body force confined to high latitudes may change the tropical circulation through its remote influence on tropical angular momentum distributions. What seems clear is that for realistic angular momentum distributions the net mass flux out of the tropical regions is controlled, in the sense explained earlier, by the distribution of $\bar{\mathcal{F}}$, and thus that the distribution of $\bar{\mathcal{F}}$ may be used to give an estimate of net tropical upwelling (Holton 1990). This may well be a good estimate, but it should be cautioned that there are at least two unanswered questions about whether it is sufficient for quantifying troposphere–stratosphere exchange in the tropics. First, whatever the *net* tropical upwelling, there remains the possibility, already mentioned, of a significant recirculation within the tropics that is to a large extent independent of the extratropical $\bar{\mathcal{F}}$. The possibility that such a purely tropical recirculation might contribute significantly to the stratospheric photochemical processing of tropospheric pollutants has not, as far as we know, hitherto been considered. In reality, it would depend upon longitudinal as well as latitudinal asymmetries in radiative temperatures T_{rad}. Second, there is the question of whether the net, zonally averaged mass transport is a good indication of the net transport of a chemical tracer. There are tracers such as water vapor for which it is evident that bulk advection by the mean meridional circulation is not the only physical process playing an important role in determining the transport into the stratosphere [see, e.g., Holton (1984), and refs.].

The steady state solution was derived in section 2 without relying on any simplifying assumptions about the background state, or about the radiative heating or cooling. It seems clear from the weight of analytical and numerical modeling evidence that the time-dependent adjustment will yield this steady state for a wide range of realistic model atmospheres and large-scale $\bar{\mathcal{F}}$ distributions. An important requirement is that the $\bar{m}$ distributions remain realistic and broadly resemble those in Fig. 1. This places certain restrictions on the magnitude of $\bar{\mathcal{F}}$.

It would be desirable to have a correspondingly general mathematical proof that the steady state assumed in section 2 is ultimately attained. However, quite apart from the expected restrictions on $|\bar{\mathcal{F}}|$, it has already been shown that the downward control principle does not usefully apply on small spatial scales with $H_R \ll H$, because of the long adjustment times involved. This means that a general, rigorously provable statement would have to recognize these restrictions in some precise way. Such a statement and its proof have so far eluded us, and so, for cases more general than those considered in sections 3c and 5, we must rely on numerical experiments like those presented in section 3e (see Figs. 8 and 9), and on arguments such as the following.

For simplicity we confine attention to demonstrating the downturn of the circulation in a model atmosphere that is unbounded below. Unless otherwise stated, symbols have the same meaning as in section 2. Consider a force $\bar{\mathcal{F}}$ confined to a shallow layer $L_{\bar{\mathcal{F}}}$ of depth scale D and width scale L, centered on $z = 0$. Given that the mean meridional circulation arises in response to $\bar{\mathcal{F}}$, one expects the contribution $f\bar{v}^*$ in the zonal momentum balance not to be of larger order than $\bar{\mathcal{F}}$, and hence that the net meridional mass transport is bounded by some multiple of $\rho_0\bar{\mathcal{F}}D/f$, where f is the Coriolis parameter. This suggests that the corresponding vertical velocity $\bar{w}^*$ near $L_{\bar{\mathcal{F}}}$ is bounded by $\bar{\mathcal{F}}D/fL$. We next assume that $\bar{w}^*$ is bounded elsewhere according to

$$|\bar{w}^*| \leqslant \frac{\bar{\mathcal{F}}D}{fL}\min(1, e^{z/H}). \tag{7.1}$$

For $z < 0$ this says that $\rho_0\bar{w}^* \sim e^{-z/H}\bar{w}^*$ is bounded, as we expect from the requirement that the meridional mass transport remains bounded. For $z > 0$, (7.1) says that $\bar{w}^*$ remains bounded. This may seem a relatively strong assumption, but it is practically certain to be true, being consistent with the analytical structure of solutions representing the response of a compressible atmosphere to given mechanical and thermal forcing (Plumb 1982); for all real ω in (3.13), Re $\Lambda_+ < -1$ so $\bar{w}^*$ decays with height (though possibly only weakly) above the forcing. Note that it is essential that density scale-height effects be included, even in this qualitative argument, since downward control does not occur in a Boussinesq model (see section 3d). An estimate for the temperature change $\Delta\bar{T}$ now follows by assuming that the adiabatic heating estimated from (7.1) balances the change in radiative cooling, so that

$$|\Delta\bar{T}| \leqslant |\bar{w}^*|\,\frac{N^2H}{R}\,\tau_r \leqslant \frac{\bar{\mathcal{F}}D}{fL}\,\frac{N^2H}{R}\,\tau_r \min(1, e^{z/H}) \tag{7.2}$$

where τ_r is an upper bound on the radiative relaxation time. Note that it is not being assumed here that the radiative effects are equivalent to a Newtonian cooling, only that the radiation acts to relax the temperature to a given profile and that a maximum timescale for the relaxation exists.

From thermal wind balance it now follows that the corresponding change $\partial(\Delta\bar{u})/\partial z$ in the zonal wind shear satisfies

$$\left|\frac{\partial\Delta\bar{u}}{\partial z}\right| \leqslant \frac{\bar{\mathcal{F}}D}{f^2L^2}\,N^2\tau_r \min(1, e^{z/H}). \tag{7.3}$$

If $\Delta\bar{u}$ vanishes sufficiently far below the forcing [again plausible from Plumb's (1982) solutions, and from the evanescence of $\bar{w}^*$ and hence $\bar{v}^*$ below $L_{\bar{\mathcal{F}}}$] it then follows from (7.3) that

$$|\Delta\bar{u}| \leqslant \frac{\bar{\mathcal{F}}D}{f^2L^2}\,HN^2\tau_r e^{z/H} \quad \text{for} \quad z < 0. \tag{7.4}$$

Consider the zonal momentum integrated above some height $z_0 \leqslant 0$, say. The estimate (7.3) ensures that the integral converges as $z \to \infty$. Integrating by parts, we have

$$\begin{aligned}\left|\int_{z_0}^{\infty} e^{-z/H}\Delta\bar{u}\,dz\right| &= H\left|\Delta\bar{u}(z_0)e^{-z_0/H} + \int_{z_0}^{\infty} e^{-z/H}\frac{\partial\Delta\bar{u}}{\partial z}\,dz\right| \\ &\leqslant H(2H + |z_0|)\frac{\bar{\mathcal{F}}D}{f^2L^2}\,N^2\tau_r \quad [z_0 \leqslant 0],\end{aligned} \tag{7.5}$$

using (7.3) and (7.4), the term involving $|z_0|$ arising from the portion of the integral between z_0 and 0.

We are now in a position to confirm the general applicability of the principal results of section 3. First consider the case $|z_0| \simeq 0$. The left-hand side of (7.5) then represents the total momentum change above the forcing layer, and the mass partitioning in the classical (Eliassen) adjustment problem [see above Fig. 2, also (3.16)] shows that the momentum increases initially as $\bar{\mathcal{F}}Dt$ multiplied by a factor of order $\min(1, H^2/H_R^2)$. Assuming a simple monotonic time dependence, it follows that the circulation above $z = 0$ dies out on a timescale obtained by equating $\bar{\mathcal{F}}Dt\min(1, H^2/H_R^2)$ to the right-hand side of (7.5) with $z_0 = 0$, that is on a timescale at most

$$t \sim \frac{N^2H^2\tau_r}{f^2L^2}\max\left(1, \frac{H_R^2}{H^2}\right) \sim \tau_r \max\left(\frac{H^2}{H_R^2}, 1\right), \tag{7.6a}$$

consistent with (3.23a,c). Second, taking $z_0 < 0$ and $|z_0| \gtrsim \max(H_R, H_R^2/H)$, the same argument applied to the layer $z_0 < z < 0$ implies that the momentum of that layer (which initially increases simply as $\bar{\mathcal{F}}Dt$) will become steady on the timescale

$$t \sim \frac{N^2H|z_0|\tau_r}{f^2L^2} \sim \frac{H|z_0|\tau_r}{H_R^2}. \tag{7.6b}$$

Since the total momentum must continue to increase as $\bar{\mathcal{F}}Dt$ beyond that time, a non-negligible (mass-weighted) zonal acceleration must eventually occur below z_0, at least after the time (7.6b), consistent with a meridional cell penetrating downward at speed $f^2L^2/N^2H\tau_r = H_R^2/H\tau_r$ as described by (3.20b) and (3.22a).

This argument strongly supports the idea that the general picture of the adjustment to a steady state revealed by the calculations of section 3 does not depend on any particular assumption about the basic state, or about the radiative relaxation such as the Newtonian cooling assumption made there. The picture should hold whenever the radiative effects act to restore temperatures toward some notional equilibrium profile on a finite timescale.

It might be asked how the downward control principle for zonally symmetric mean circulations, as discussed in this paper, is related to the form of the principle noted by Haynes and McIntyre (1987) for the exact three-dimensional diabatic circulation. In a certain sense the latter form is more general, and (like the relation between $\bar{\mathcal{F}}$ and the flow across $\bar{m}$ surfaces) not dependent on an assumption of quasi-steadiness. There is of course a price for this generality. To use this form of the principle, more information needs to be given, in this case the rates of change of Rossby-Ertel potential vorticity (PV) on isentropic trajectories (i.e., more detail about the process summarized in the zonally averaged description as the advective PV flux due to Rossby-wave breaking and similar processes) together with the curl of the gravity-wave drag, and certain other terms that might be cumulatively significant even though negligible under quasi-geostrophic scaling. The usefulness or otherwise of a thought-experiment in which all this is given is far from clear at present; and it may be best to think of it for the time being as a purely diagnostic statement of the vertical velocity, based on the PV equation, and the assumption that the mass-carrying capacity of an atmospheric layer per unit altitude z goes to zero as $z \to \infty$. One might refer to this as the "weak" form of the principle. By contrast, the results in this paper give credence to a "strong" form of the principle, in which, in return for assuming zonal symmetry, a simple and definitive causal connection is made between the zonal force exerted above a particular level and the vertical mass flow across that level. The causal connection is given a clear justification using an analysis of the initial-value problem. As has been shown, for planetary-scale forcing this causal link is set up over only a few radiative timescales. One might ask whether such a "strong" form could be obtained for three-dimensional cases in which either the basic state, or the applied forcing, were zonally asymmetric. This appears unlikely, since, for example, a longitude-dependent force localized in latitude and height would generally set up a Rossby-wave field as part of the response, including an upward propagating part. The latter would break or otherwise dissipate at some level above the applied force, and thereby exert a zonal mean force at that level. A sustained diabatic circulation would, therefore, generally exist above the level of the original applied force.

Acknowledgments. We thank D. G. Andrews, T. J. Dunkerton, R. R. Garcia, J. R. Holton, B. J. Hoskins, N. A. McFarlane, A. O'Neill and R. A. Plumb for helpful comments and correspondence, and S. P. Cooper, L. I. Kurtul, and A. Rickaby for programming assistance. This work received support from the US Office of Naval Research, from the UK Meteorological Office through a Royal Society Meteorological Office Fellowship supporting PHH, and from the Natural Environment Research Council both through the UK Universities' Global Atmospheric Modelling Project and also through individual grants supporting CJM, TGS, and KPS. TGS acknowledges present support from the Natural Sciences and Engineering Research Council and the Atmospheric Environment Service of Canada.

REFERENCES

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

Barnett, J. J., and M. Corney, 1985: Middle atmosphere reference model derived from satellite data. *Handbook for MAP, 16: Atmospheric Structure and its Variations in the Region 20–120 km: Draft of a New Reference Middle Atmosphere,* K. Labitzke, J. J. Barnett and B. Edwards, Eds., 47–85.

Dickinson, R. E., 1968: On the excitation and propagation of zonal winds in an atmosphere with Newtonian cooling. *J. Atmos. Sci.,* **25,** 269–279.

——, 1969: Theory of planetary wave-zonal flow interaction. *J. Atmos. Sci.,* **26,** 73–81.

Dunkerton, T., 1978: On the mean meridional mass motions of the stratosphere and mesosphere. *J. Atmos. Sci.,* **35,** 2325–2333.

——, 1989a: Body force circulations in a compressible atmosphere: Key concepts. *Pure Appl. Geophys.,* **130,** 243–262.

——, 1989b: Nonlinear Hadley circulation driven by asymmetric differential heating. *J. Atmos. Sci.,* **46,** 956–974.

Fels, S. B., 1985: Radiative-dynamical interactions in the middle atmosphere. *Issues in Atmospheric and Oceanic Modeling,* S. Manabe, Ed. *Advances in Geophysics,* **28A,** 277–300. [Academic Press.]

Garcia, R. R., 1987: On the mean meridional circulation of the middle atmosphere. *J. Atmos. Sci.,* **44,** 3599–3609.

——, 1989: Dynamics, radiation, and photochemistry in the mesosphere: Implications for the formation of noctilucent clouds. *J. Geophys. Res.,* **94,** 14 605–14 615.

Gille, J. C., L. V. Lyjak and A. K. Smith, 1987: The global residual mean circulation in the middle atmosphere for the northern winter period. *J. Atmos. Sci.*, **44,** 1437–1452.

Gradshteyn, I. S., and I. M. Ryzhik, 1980: *Table of Integrals, Series, and Products,* Corrected and Enlarged Edition. Academic Press, 1160 pp.

Haynes, P. H., and M. E. McIntyre, 1987: On the evolution of vorticity and potential vorticity in the presence of diabatic heating and frictional or other forces. *J. Atmos. Sci.*, **44,** 828–841; also **47,** 2021–2031.

——, and T. G. Shepherd, 1989: The importance of surface pressure changes in the response of the atmosphere to zonally-symmetric thermal and mechanical forcing. *Quart. J. Roy. Meteor. Soc.*, **115,** 1181–1208.

Held, I. M., and A. Y. Hou, 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37,** 515–533.

Holton, J. R., 1984: Troposphere-stratosphere exchange of trace constituents: The water vapour puzzle, *Dynamics of the Middle Atmosphere,* Holton, J. R., T. Matsuno, Eds. Reidel, 369–385.

——, 1990: On the global exchange of mass between the stratosphere and troposphere: *J. Atmos. Sci.*, **47,** 392–395.

Hoskins, B. J., and A. J. Simmons, 1975: A multi-layer spectral model and the semi-implicit method. *Quart. J. Roy. Meteor. Soc.*, **101,** 637–655.

Keating, G. M., D. F. Young and M. C. Pitts, 1987: Ozone reference models for CIRA. *Adv. Space Res.*, **7,** 105–115.

Longuet-Higgins, M. S., 1968: The eigenfunctions of Laplace's tidal equations over a sphere. *Phil. Trans. Roy. Soc. London,* **A262,** 511–607.

Marks, C. J., 1989: Some features of the climatology of the middle atmosphere revealed by Nimbus 5 and 6. *J. Atmos. Sci.*, **46,** 2485–2508.

McIntyre, M. E., 1987: Dynamics and tracer transport in the middle atmosphere: An overview of some recent developments. *Transport Processes in the Middle Atmosphere* (Proc. NATO workshop, Erice), G. Visconti, R. R. Garcia, Eds., D. Reidel, 267–296.

——, 1989a: On the Antarctic ozone hole. *J. Atmos. Terr. Phys.*, **51,** 29–43.

——, 1989b: On dynamics and transport near the polar mesopause in summer. *J. Geophys. Res.*, **94,** 14 617–14 628.

——, and T. N. Palmer, 1984: The "surf zone" in the stratosphere. *J. Atmos. Terr. Phys.*, **46,** 825–849.

Murgatroyd, R. J., and F. Singleton, 1961: Possible meridional circulations in the stratosphere and mesosphere. *Quart. J. Roy. Meteor. Soc.*, **87,** 125–135.

Plumb, R. A., 1982: Zonally symmetric Hough modes and meridional circulations in the middle atmosphere. *J. Atmos. Sci.*, **39,** 983–991.

Shine, K. P., 1987: The middle atmosphere in the absence of dynamical heat fluxes. *Quart. J. Roy. Meteor. Soc.*, **113,** 603–633.

——, 1989: Sources and sinks of zonal momentum in the middle atmosphere diagnosed using the diabatic circulation. *Quart. J. Roy. Meteor. Soc.*, **115,** 265–292.

Solomon, S., J. T. Kiehl, R. R. Garcia and W. L. Grose, 1986: Tracer transport by the diabatic circulation deduced from satellite observations. *J. Atmos. Sci.*, **43,** 1603–1617.

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

WMO, 1985: *Atmospheric Ozone 1985:* Assessment of our understanding of the processes controlling its present distribution and change. World Meteorological Organization Global Ozone Research and Monitoring Rep. No. 16, 1095 pp. + 86 pp. refs. [Available from Global Ozone Research and Monitoring Project, World Meterological Organization, Case Postale 5, CH 1211, Geneva 20, Switzerland.]