Vacillation, Sudden Warmings and Potential Enstrophy Balance in the Stratosphere

MARK R. SCHOEBERL

Geophysical and Plasma Physics Dynamics Branch, Plasma Physics Division, Naval Research Laboratory, Washington, DC 20375

(Manuscript received 11 June 1981, in final form 6 April 1982)

ABSTRACT

The balance of potential enstrophy and its relationship to vacillation cycles and the sudden warming is studied for a β-channel model of the stratosphere. It is shown that the mean flow cannot be steady in the presence of large-amplitude quasi-geostrophic waves [$\sim$1–0.25 geopotential kilometers (gpkm)] when any dissipation is present, and the maximum wave amplitude allowed is $\sim$2 gpkm.

If wave forcing (transience plus dissipation) is artificially maintained, the mean flow decelerates slowly at first then explosively as the potential vorticity gradient of the basic state is wiped out over the channel. This process is called wave saturation. The initial phase of the explosive deceleration resembles both the observed and modeled mean flow evolution during a sudden stratospheric warming. A simple vacillation model based upon these ideas shows remarkable similarity to the results of Holton and Mass (1976) and Davies (1981).

1. Introduction

Holton and Mass (1976) and Holton and Dunkerton (1978) (hereafter HM and HD, respectively) showed that a spectrally truncated, β-channel model of the stratosphere would vacillate (i.e., generate repeated sudden warmings) for sufficiently large planetary wave forcing. Further, the vacillation onset could be very finely tuned to the wave forcing. For example, a wave boundary amplitude of 307 geopotential meters (gpm) produced a steady (or nearly steady) solution after 120 days of integration, but a 308 gpm forcing produced a vacillating solution (HD Fig. 1). HD further showed that wave transience was the principal mechanism driving the mean flow changes. The vacillating and steady flow regimes in the HM model are apparently separated by a critical value of the forcing amplitude. The period of the vacillation cycle was also shown by HM to be a function of the forcing amplitude although the cycle is hardly a simple harmonic. In this last regard, the development of the sudden warming at the onset of the cycle occurs rapidly while the rest of the cycle is characterized by a rather slow recovery.

Attempts to associate the development of the sudden warming with baroclinic instability have been unsuccessful principally because baroclinic theories predict a much slower evolution of wave amplitude and response of the mean flow than is characteristic of the sudden warming [see references in Schoeberl (1978) and O'Neill (1980)]. It is presently believed that the stratospheric sudden warming is ultimately derived from the interaction of the mean flow with tropospherically forced planetary waves. Hemispherical, three-dimensional numerical model integrations by Matsuno (1971) and others showed that a mechanistic forcing of the stratosphere with large-amplitude planetary waves will produce a sudden warming. Holton (1976) showed that the mean flow evolution in the stratosphere during the sudden warming was not necessarily confined to a narrow critical line, but takes place over a deep layer. HD showed that wave transience was principally responsible for the forcing of the mean flow during a sudden warming. Recently Davies (1981) has further examined the wave-mean flow interaction with a model similar to HM. He demonstrated quite clearly the close correlations between the changes in the mean flow and wave transience due to the upward flux of potential enstrophy from the troposphere. Potential enstrophy is the square of the potential vorticity.

In this paper we shall examine the response of the mean flow to wave forcing more closely. As with Davies (1981), the analysis is facilitated by examining the potential enstrophy equations. The usefulness of potential enstrophy was also demonstrated by HD who showed that the eddy potential vorticity flux is proportional to the time rate of change of wave potential enstrophy (transience) plus the dissipation rate times the wave potential enstrophy. The channel average potential enstrophy equation is derived in the next section.

In Section 3 we solve the enstrophy equation for the simple wave pulse and for the steady state case. The former provides a general, rigorous, upper limit condition on the amplitude of Rossby waves in the stratosphere. For most waves this limit is distinct from and smaller than that suggested by Lindzen and Schoeberl (1982) using inertial instability of the basic state as the upper limit criterion.

A similar limiting condition is obtained for the steady-state case where the dissipation of wave and mean flow enstrophy are assumed to take the Rayleigh form. Interestingly enough, the steady-state upper limit is half the maximum wave amplitude allowed for the transient pulse. This result has important consequences for the interpretation of steady-state models of the stratosphere.

In Section 4, we turn our attention to the behavior of the zonal mean flow during a sudden warming. In the presence of large-amplitude planetary waves, the mean flow decelerates slowly at first then explosively as the potential vorticity gradient is wiped out across the β-channel by the wave. This process is called wave saturation. Since an explosive deceleration of the mean flow occurs with the catastrophic breakdown of the polar night jet during a sudden warming, it seems likely that wave saturation plays an important role in the sudden warming process.

The transient wave pulse problem, a more realistic simulation of the mean flow behavior during a sudden warming, is treated in Section 5. It is shown that even for relatively small changes in the mean flow, partial wave saturation can amplify the response of the mean flow to the wave forcing. Finally in Section 6 a simple model of wave vacillation is presented.

2. Potential enstrophy

We consider a Rossby wave in a β-channel. If q is the potential vorticity then

$$\frac{dq}{dt} = -D(q - q_0), \tag{1}$$

where $q = f + \tilde{\nabla}^2\psi$, and $d/dt = \partial/\partial t + \mathbf{v}_g \cdot \nabla$ where $\mathbf{v}_g$ is the geostrophic wind, f is the Coriolis frequency,

$$\tilde{\nabla}^2 = \nabla^2 + \frac{f_0^2}{N^2} e^{z/H} \frac{\partial}{\partial z} e^{-z/H} \frac{\partial}{\partial z},$$

and the other terms are defined as follows:

- z — $-H \ln(p/p_0)$; p = pressure, p_0 = reference pressure
- H — atmospheric scale height ($\sim$7 km)
- N — buoyancy frequency ($\sim$0.02 s^{-1})
- f — $f_0 + \beta y$; $f_0 = 2\Omega \sin\theta_0$, $\beta = 2\Omega \cos(\theta_0)/a$
- a — earth's radius
- Ω — earth's rotational frequency
- θ_0 — channel center.

D represents dissipative processes modeled as constant and equal rates of Newtonian cooling and Rayleigh friction. In this study we will take D to be constant; q_0 is the basic state potential vorticity and ψ is the quasi-geostrophic streamfunction ($\psi = \phi/f_0$), where ϕ is the geopotential. Expanding the substantive derivative, taking $\psi = \bar{\psi} + \psi'$ where the overbar denotes zonal average, and linearizing, we have the standard mean flow and planetary wave propagation equations:

$$\frac{\partial \bar{q}}{\partial t} + D(\bar{q} - \bar{q}_0) = -\frac{\partial}{\partial y}[\tfrac{1}{2} \operatorname{Re}(v'^* q')], \tag{2}$$

$$\frac{\partial q'}{\partial t} + im\bar{u}q' + Dq' = -v'\frac{\partial \bar{q}}{\partial y}, \tag{3}$$

where $(\ )' \sim e^{imx}$ and v is the meridional velocity. The asterisk indicates the complex conjugate. Multiplying (3) by q'^* and adding to the complex conjugate of the same equation gives the wave potential enstrophy equation

$$\frac{\partial}{\partial t}|q'|^2 + 2D|q'|^2 = -2 \operatorname{Re}(v'^* q')\frac{\partial \bar{q}}{\partial y}. \tag{4}$$

Note that $(\overline{q'^2}) = \tfrac{1}{2}|q'|^2$. Multiplying (2) by $\bar{q}$, using (4), and integrating across the channel ($q' = 0$ at the β-channel boundaries) yields

$$\frac{1}{2}\left\langle \frac{\partial}{\partial t}\bar{q}^2 \right\rangle + D\langle \bar{q}^2 \rangle - D\langle \bar{q}\bar{q}_0 \rangle = -\frac{1}{4}\left\langle \frac{\partial}{\partial t}|q'|^2 \right\rangle - \tfrac{1}{2}D\langle |q'|^2 \rangle, \tag{5}$$

where $\langle\ \rangle = \int_0^L dy$ and L is the channel width.

Eq. (5) is the channel integral, potential enstrophy equation. We may simplify (5) further by defining $\bar{q} = \bar{q}_0 + \bar{q}_p$, where $\bar{q}_p$ is the change in $\bar{q}$ due to wave forcing of the mean flow. Eq. (5) then becomes

$$\frac{1}{2}\left\langle \frac{\partial}{\partial t}\bar{q}_p^2 \right\rangle + \left\langle \frac{\partial}{\partial t}\bar{q}_0\bar{q}_p \right\rangle + D\langle \bar{q}_p^2 \rangle + D\langle \bar{q}_0\bar{q}_p \rangle = -\frac{1}{4}\left\langle \frac{\partial}{\partial t}|q'|^2 \right\rangle - \tfrac{1}{2}D\langle |q'|^2 \rangle = F. \tag{6}$$

With a little manipulation (6) can also be written as

$$\left\langle \bar{q}\left(\frac{\partial}{\partial t}\bar{q}_p + D\bar{q}_p\right)\right\rangle = F. \tag{7}$$

We will refer to F as the wave forcing; it consists of both "transience" and "dissipation" in the nomenclature of HD.

Even though a linearized wave equation has been used to obtain (5) this assumption is not actually required and (5) can be obtained directly from the potential vorticity equation (1). Multiplying (1) by q, using the geostrophic continuity equation, and channel averaging gives

$$\left\langle \frac{1}{2}\frac{\partial q^2}{\partial t} \right\rangle = \langle D(qq_0 - q^2) \rangle,$$

where wave and mean flow perturbations are assumed to vanish at the y boundaries. Letting $q = \bar{q} + \Sigma_s q'_s$ where s is the zonal wave index, the general form of (5) is

$$\left\langle \frac{1}{2}\frac{\partial \bar{q}^2}{\partial t} + \frac{1}{4}\frac{\partial}{\partial t}\Sigma_s|q'_s|^2 \right\rangle = \langle D(\bar{q}_0\bar{q} - \bar{q}^2 - \tfrac{1}{2}\Sigma_s|q'_s|^2\rangle.$$

3. Limiting cases

We shall now consider some very simple problems to understand the implications of (7). We assume as in HM that the wave and mean flow solutions are separable in their y and z dependence. We take $\psi' = Z_1(z, t)\sin ky$ where $k = \pi/L$. It is apparent from (2) that $\bar{q}_p = Z(z, t)\sin ky \cos ky$. HM further expanded $\bar{q}_p$ in a cosine series retaining only the first term, but that expansion is not required for our purposes. Now $\bar{q}_0 = f_0 + \beta y + \tilde{\nabla}^2\bar{\psi}_0$, so for a constant value of $\bar{u}_0$ we write (7) as

$$(A + BZ)\left(\frac{\partial Z}{\partial t} + DZ\right) = -I\left(\frac{1}{4}\frac{\partial|Z'|^2}{\partial t} + \frac{1}{2}D|Z|'^2\right) = F, \quad (8)$$

where

$$A = \beta\int_0^L y \sin ky \cos ky dy = -\frac{L^2\beta}{4\pi},$$

$$B = \int_0^L \sin^2 ky \cos^2 ky dy = \frac{L}{8},$$

$$I = \int_0^L \sin^2 ky dy = \frac{L}{2}.$$

a. Transient wave limit

We first consider the case of a transient wave pulse switched on at $t = 0$ in the absence of dissipation. Setting $D = 0$ in (8) we obtain

$$\frac{\partial}{\partial t}(AZ + \tfrac{1}{2}BZ^2 + \tfrac{1}{4}I|Z'|^2) = 0. \quad (9)$$

Integrating (9) from $t = 0$ to t yields

$$AZ(t) + \tfrac{1}{2}BZ^2(t) + \tfrac{1}{4}I|Z'(t)|^2 = 0,$$

where $Z(0)$, $Z'(0) = 0$. Solving for $Z(t)$ we obtain

$$Z(t) = -A/B - (A^2 - BI|Z'(t)|^2/2)^{1/2}B^{-1}, \quad (10)$$

where the negative root is chosen since A is negative and $Z(t)$ should be zero at $t = 0$. Since $Z(t)$ cannot be complex we obtain two limits

$$Z(t) \leqslant -A/B = 2\beta/k, \quad (11a)$$

$$|Z'(t)| \leqslant -\sqrt{2}A(BI)^{-1/2} = \sqrt{2}\beta/k. \quad (11b)$$

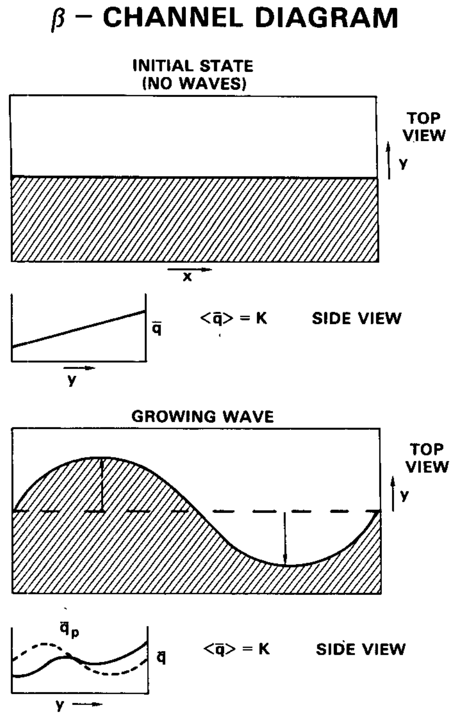

FIG. 1. A schematic diagram of the potential vorticity in a β-channel. Top of figure shows the initial state; bottom shows the result of wave mixing the gradient in $\bar{q}$.

Eq. (11b) gives an upper limit criterion for wave amplitudes from enstrophy conservation. The physical mechanism which is expressed by (11b) can be understood as follows: As a wave grows in amplitude it gains potential vorticity (p.v.) by redistributing the p.v. of the basic state. In order to do this, a background gradient of p.v. must be available. As the wave grows, however, it reduces the basic state p.v. gradient so that p.v. transport becomes more difficult. This process is illustrated in Fig. 1. Note that the total p.v. (wave plus mean flow) of the system remains constant.

At some point, for a large enough wave, the basic state p.v. gradient becomes such that wave p.v. transport must stop. This is the limit given by Eq. (11). In fact, it is easy to show that for (11) $\langle(\overline{v'q'})\partial\bar{q}/\partial y\rangle = 0$.

The upper limit given by (11) is very large; in fact, the Rossby number for $Z = 2\beta/k$ is $\sim 2/\pi$ which is outside the range of validity for the quasi-geostrophic scaling. In the rest of the text we shall refer to the limits given by (11) as wave saturation.

Lindzen and Schoeberl (1982) derived a less rigorous upper limit condition for the wave amplitude assuming that $\bar{q}$ does not change sign, the condition for inertial instability. Eq. (11) gives a smaller upper limit value (as will be shown) and $\bar{q}$ may change sign as long as the integral constraints given by (11) hold.

b. Steadiness limit

The steady-state problems assumes that after some long time the wave and mean flow come into equilibrium so time derivatives in (8) may be neglected. Thus, provided $D \neq 0$,

$$(A + BZ)Z = -\tfrac{1}{2}l|Z'|^2. \qquad (12)$$

Note that D disappears from (12) since the damping rates on the wave and the mean flow are the same. More realistic damping formulations give qualitatively similar results as discussed in Appendix I. Solving for Z yields

$$Z = -A/(2B) - (A^2 - 2Bl|Z'|^2)^{1/2}/(2B). \qquad (13)$$

As before, not all values of Z' are allowed since Z must be real. Therefore,

$$|Z'| \leq -A(2Bl)^{-1/2} = \beta/\sqrt{2}k. \qquad (14)$$

Eq. (14) thus gives a "steadiness" criterion for the mean flow in the presence of steady, dissipating waves. In other words, if (14) is violated a steady-state solution is not possible.

Note that the limit given by (14) is one-half that given by (11b). In other words, the "steadiness" criteria for this example occurs at half the wave p.v. at saturation. To determine if the value of $|Z'|$ given above is physically realizable in models, we assume Z' has the form

$$|Z'| = \left(m^2 + k^2 + \frac{(lf_0)^2}{N^2} + \frac{f_0^2}{(2HN)^2}\right)\frac{h_0 g}{f_0}, \qquad (15)$$

where m is (s/a) $\cos\theta_0$, l the vertical wavenumber, h_0 the wave amplitude, and the β-channel is taken to extend from 30° to 90° centered at 60°. Assuming as in Lindzen and Schoeberl (1982) that k is the smallest scale, then

$$|Z'| \approx \frac{h_0 g k^2}{f_0}.$$

The upper limit condition becomes

$$|h_0'| \leq \sqrt{2}\,\frac{f_0\beta}{gk^3} \approx 2000 \text{ gpm}. \qquad (16a)$$

This limit, of course, is a very rough estimate of the maximum amplitude of Rossby waves. Note that (16a) depends on k^3 so the value will vary widely depending on the horizontal scale of the wave and the validity of the separability assumption.

For a steady state,

$$|h_0'| \leq f_0\beta/\sqrt{2}gk^3 \approx 1000 \text{ gpm}. \qquad (16b)$$

The inertial instability limit given by Lindzen and Schoeberl (1982) gives

$$|h_0'| < \frac{f_0^2}{k^2 g}\sin^2\theta \approx 5500 \text{ gpm}, \qquad (16c)$$

TABLE 1. Minimum stable altitude for mean flow in the presence of vertically propagating, dissipating, steady planetary waves.*

λ_z (km)	$h_s(m)$ = 5	$h_s(m)$ = 10	$h_s(m)$ = 50
20	32.1 / 31.2	22.4 / 21.5	— / —
30	41.5 / 39.8	31.8 / 30.1	9.3 / 7
50	51.0 / 47.9	41.3 / 38.1	18.8 / 15.6
∞	62.1 / 55.9	52.4 / 46	29.8 / 23.6

* m is the zonal wavenumber $[m = s/(a\cos\theta_0)]$, λ_z the vertical wavelength $[l = 2\pi/\lambda_z]$, h_s the surface height and $\theta_0 = 60°$. Values above the diagonal refer to $s = 1$, those below to $s = 2$.

the conditions (16a) and (16c) being equivalent when $k = \beta/f_0$.

In numerical models such as HM, waves typically exceed 1000 gpm in amplitude. An estimate of the altitudes where $|Z'| \approx \beta/\sqrt{2}k$ for different vertical wavelengths are shown in Table 1. We assume $h_0 = h_s e^{z/2H}$ which is a crude approximation for the wave amplitude at a given height neglecting the possibility of significant wave reflection or dissipation below the level z. The altitudes in Table 1 should thus be regarded as minimum altitudes below which a steady solution is possible.

From van Loon *et al.* (1973), h_s is typically less than 50 gpm during winter and $\lambda_z \approx$ 20–30 km (although λ_z increases rapidly with altitude). The calculations shown in Table 1 indicate that a steady wave-mean flow solution can be achieved below 10–30 km for average winter conditions. For the stratosphere, this result implies that the lower stratosphere should be "quieter" than the upper stratosphere during winter which is consistent with observations (Schoeberl, 1978).

When the condition given by (14) is violated then the mean flow cannot be steady. Since Z is always positive the perturbed mean zonal wind will be negative and thus planetary wave forcing of the mean flow through dissipation or transience (when positive) will decelerate the mean flow. This deceleration will continue until the wave amplitude is reduced. Thus the condition $|Z'| \leq \beta/\sqrt{2}k$ provides an automatic limiting mechanism for the p.v. of waves in the atmosphere. While the wave vorticity may temporarily exceed this limit (up to the saturation limit) it is unlikely that the time-mean p.v. will do so. (This can be shown rigorously as will be presented in a subsequent publication.) Once the mean zonal flow begins to decelerate it is doubtful that a new steady solution could be established in the lower strato-

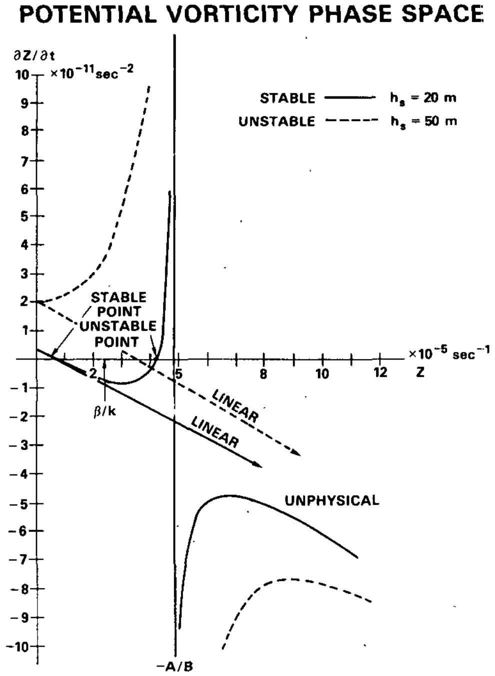

FIG. 2. Solution paths in potential vorticity phase space at 30 km for two surface-forced planetary waves with heights h. Solutions start at $Z = 0$ and move to right in the physical region. For $h_s = 20$ m, the solution halts at the "stable point" and remains there. For $h_s = 50$ m, the solution is explosively unstable. The curves labeled "linear" are linearized solutions to Eq. (17) which always predict a stable solution.

sphere over the 3–4 month winter period, but it may be possible for numerical models with perpetual winters to restabilize.

4. Evolution of the mean flow during wave saturation

The analysis in the previous section treated the limits on wave-mean flow interaction suggested by the enstrophy equations. A more detailed, although still highly simplified, analysis of wave-mean flow interaction is presented in this section.

As shown by HD, the forcing of the mean flow occurs through both a wave transience and dissipation during vacillation (see HD, Fig. 1). Transience in the numerical model probably results from the vertical propagation of a wave front initiated in the troposphere or in the lower stratosphere (Matsuno, 1971). As the wave propagates vertically, the density decrease with height amplifies the wave perturbation. Thus by the time the wave front reaches the upper stratosphere, the wave amplitude is quite large, and the vorticity perturbation associated with the wave reaches the same order of magnitude as the vorticity of the background flow.

In this section we will analyze the time-dependent behavior of (8) holding F fixed which implies that the wave p.v. grows as $\sqrt{t}$ (in the absence of dissipation) on an isobaric surface. We will also assume that there is no feedback of the change in the mean flow on $\bar{u}$ in the wave equation (3). In other words F is not changed by the variation in Z. This is, of course, very unrealistic as the limits given by (11) are approached. The results of this model near saturation should therefore be regarded as illustrative. Sections 5 and 6 deal more directly with the problem of $\bar{u}$ decreasing in the wave equation as $\bar{q}_p$ increases.

From Eq. (8) we may write

$$\frac{\partial Z}{\partial t} + DZ = \frac{F}{(A + BZ)}. \tag{17}$$

Within the HM model, the wave amplitude is switched on to a constant value over ~10 days which we will use to estimate F, i.e., $F = -\frac{1}{2}I(\sigma + D)|Z'|^2$, where $\sigma = (10 \text{ days})^{-1}$. We take $D = (10 \text{ days})^{-1}$. Z' is held fixed in time so C represents the maximum forcing of the mean flow due to the switch-on process.

Analysis of the nonlinear equation (17) is aided by Fig. 2 which shows a plot of the solution path in p.v. phase space, i.e., $\partial Z/\partial t$ vs Z. We have used $\lambda_z = 50$ km, and taken h_s values at the surface to be 20 and 50 gpm, respectively. We neglect the dissipation in the computation of q' from ψ' as in the previous section. Z values are plotted for $z = 30$ km, $s = 1$. Initially all solutions begin at the $Z = 0$ line on the figure. Since the solution point at $Z = 0$ is above the $\partial Z/\partial t = 0$ axis the solution point moves to the right (Z becomes positive) along the curve. If the point crosses $\partial Z/\partial t = 0$ then the solution point must stop moving and the solution is stable (i.e., a steady-state mean flow is possible). In the 50 gpm case the curve does not cross the $\partial Z/\partial t = 0$ line so no steady solution is possible; in fact, $\partial Z/\partial t$ rapidly goes to ∞ (saturation) as the RHS of (17) becomes singular (i.e., $Z \rightarrow |A/B|$).

To illustrate the difference between the nonlinear equation (17) and predictions by a linear model we have also plotted the linearized solutions to (17) obtained by setting $B = 0$. "Linear" in this context implies that the basic state can absorb any amount of wave forcing. The lines labeled "linear" in the figure always cross the Z axis so the linear model predicts a steady solution for any wave amplitude.

Note that the 20 gpm curve for the nonlinear solution crosses the $\partial Z/\partial t = 0$ line twice, but only the first point to the left of the axis is stable since if the solution is perturbed from the second point it continues to move away. The second point corresponds to the other root of (12).

The critical wave amplitude for the steady mean flow solution from (17) is given by the condition

$$|Z'| \leq -A[2B(\sigma D^{-1} + 1)I]^{-1/2}.$$

Using (15) and $l = 0$, we compute $h_0 \approx 245$ gpm. According to HM, the wave amplitudes achieved prior to the vacillation cycle are in excess of 400 gpm for $s = 1$, so this highly simplified computation predicts that a steady solution should not be possible in the HM model for those wave amplitudes. Note that it may not be appropriate to assume that the wave has a sinusoidal vertical structure during the switch-on process. More likely the wave front has a vertical dependence such that $\psi' \approx e^{-\alpha z}$. Taking $\alpha^{-1} = 20$ km we obtain $h' \approx 300$ gpm, not too different from the results above.

The prediction of a rapid change in the mean flow p.v. even though F is nearly constant is also nicely illustrated in Davies' (1981) computation. His Fig. 2 compares wave enstrophy and mean flow enstrophy during a sudden warming cycle. Prior to the development of the sudden warming, he shows a relatively slow, linear increase in wave enstrophy followed by a slow decrease at 45 km over about 12 days. This implies that F would change slowly over the warming period. In response to the wave forcing, the mean flow enstrophy changes slowly at first then builds up suddenly within ~1 day. This is exactly the type of behavior predicted by Eq. (17) as the mean flow moves toward saturation.

Eq. (17) predicts that a singular value for Z will develop in a finite time provided F is artificially maintained (although this could never happen in practice). An equation with this kind of behavior is called "explosive" (Cap, 1976, p. 35). We may determine the finite time τ to reach the singular acceleration point by integrating (17) analytically from some initial mean flow state Z_I to $-A/B$, the singular point. This gives us an idea of how long the wave saturation process may take if we assume the wave forcing is constant. We obtain

$$\tau = -\frac{A}{\sqrt{R}}\tan^{-1}\left(\frac{2DBZ + AD}{\sqrt{R}}\right) + \frac{1}{2D}\ln(BDZ^2 + ADZ - F)\Bigg|_{Z_I}^{-A/B}, \quad (18)$$

where $R = -4DBF - A^2D^2$, $R > 0$. With the same constants used to evaluate (17) we may plot the time it takes to reach the singular solution vs Z. Although we cannot physically expect Z to become singular since we have not included variation of $\bar{u}$ in the computation of q', for the values used, the time to reach singularity is in qualitative agreement with the observed rapid flow changes during the onset of a sudden warming at different levels (i.e., ~10–50 days). Because the final stage of saturation takes place most rapidly, τ should quite accurately characterize the time scale for partial saturation.

It is important to note that evolution of the mean flow vorticity in response to wave forcing near saturation must be highly dependent on the feedback between F and Z, more specifically, the functional form of $F(Z, Z')$. For example we might choose some parameterization for F close to saturation such as

$$\frac{\partial |Z'|^2}{\partial t} \approx (A + BZ)^n, \quad n \geqslant 1$$

so that Z will experience a linear or slower growth rate. In reality, the functional behavior of Z near saturation is unimportant since at some point before saturation Z will evolve so to block increasing wave fields. For example, Davies (1981) suggested that $|Z'|$ might start to decrease when the relative p.v. gradient $(\beta - \partial\bar{q}/\partial y)$ changes sign since

$$\beta - \partial\bar{q}/\partial y = \tilde{\nabla}^2\bar{u}$$

and $\tilde{\nabla}^2$ is nearly Laplacian. However, Davies' argument neglects the boundary conditions on $\bar{u}$ which introduce constants into the calculation of $\bar{u}$ from $\tilde{\nabla}^2\bar{u}$. The mean zonal wind may therefore become easterly before or after $\beta - \partial\bar{q}/\partial y$ changes sign. Because $\beta - \partial\bar{q}_0/\partial y = 0$ in this model, Davies' condition is not applicable. Nevertheless, we expect problems in the linearization of the wave equation when $\partial\bar{q}_p/\partial y \approx \partial\bar{q}_0/\partial y$ or when $Z > \beta/k$. For that value of Z the wave displacement vector (η') becomes singular at the channel center. Using the linearized formula $\eta' = q'(\partial\bar{q}/\partial y)^{-1}$, this condition implies $\bar{u} \approx 0$ if we assume $\psi'/\bar{u} \approx \eta'$ (which is approximately valid for a slowly changing flow). Not coincidently, the critical line condition $(\bar{u} = 0)$ occurs almost exactly when $|q'|^2$ begins to decrease in Davies' model (although the rapid growth of Z occurs later because the wave forcing is sustained since dissipation is present).

Up to the point when easterly winds appear, the numerical models of Davies and HD indicate that F increases in time so the assumption that F is constant up to some point of partial saturation seems quite reasonable. For example, Fig. 3 (from HD) shows the amplitude of the wave transience (the major component of F) prior to a sudden warming which occurs on day 20. The wave transience computed by HD increases by a factor of 10 from days 18–22 but then rapidly decreases and reverses sign when the wave amplitude field collapses as the easterly winds appear.

From the discussion above we shall use the value $Z = \beta/k$ as the upper limit on the perturbed mean flow in the remainder of this paper. This point is marked on Figs. 2, 3 and 4. Comparing the linear and nonlinear solutions for values of $Z < \beta/k$ in Fig. 2, it is apparent that the effect of partial wave saturation can be important even when Z is well below β/k. In Sections 5 and 6 we will examine the evolution of the flow in the partially saturated regime in more detail.

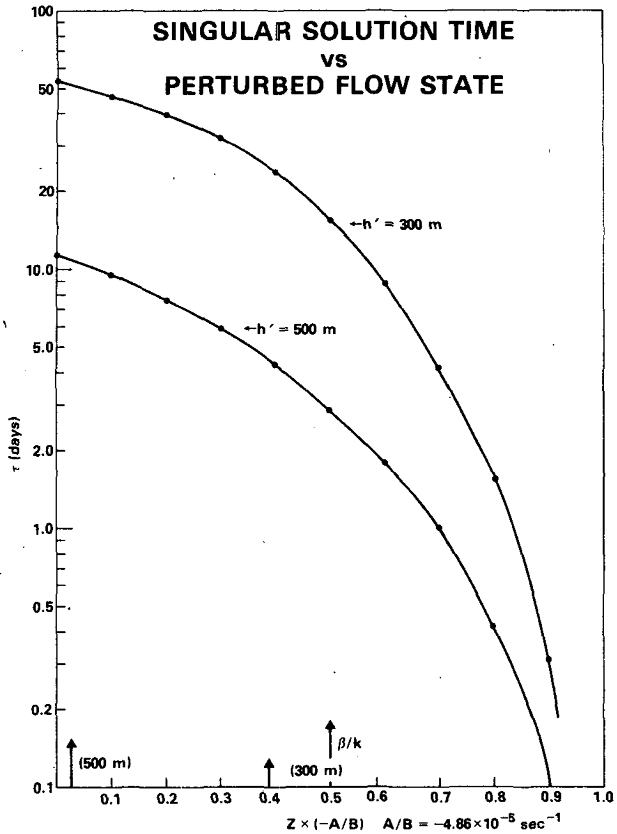

FIG. 3. The time required to reach the singular solution is plotted as a function of the potential vorticity of the initial flow. h' is the wave geopotential height in geopotential meters. Unperturbed mean flow is $Z = 0$. Vertical arrows indicate minima of $\partial Z/\partial t$, and the point β/k.

5. Mean flow response to a wave pulse

As Z (or $\bar{q}_p$) increases $|Z'|$ must eventually decrease as $\bar{u}$ becomes easterly. Neglecting dissipation, a decrease in $|Z'|$ implies that the function F will change sign during the vacillation cycle (for example, Fig. 3, 20–25 day region). In the post-warming period this is functionally equivalent to the passage of the tail end of an upward propagating wave pulse. The pulse tail is produced as the wave amplitude field collapses when easterlies (which block vertical propagation of the stationary wave) appear. By examining the enstrophy changes during the passage of a wave pulse we can implicitly include the effects of changing $\bar{u}$ on q'.

The consequences of the passage of a wave pulse on the mean flow can be understood by examining (5). Integrating (5) from t_2 to t_1 yields

$$\tfrac{1}{2}\langle \bar{q}^2(t_1) - \bar{q}^2(t_2)\rangle = \int_{t_2}^{t_1} D(\langle \bar{q}^2\rangle - \langle \bar{q}\bar{q}_0\rangle)dt$$

$$= -\tfrac{1}{4}\langle |q'(t_1)|^2 - |q'(t_2)|^2\rangle - \int_{t_2}^{t_1} \tfrac{1}{2}D\langle |q'|^2\rangle dt. \quad (19)$$

If t_1 and t_2 are times far removed from the pulse passage then it is apparent that the net changes in the mean flow which occur over the vacillation cycle are due to the integrated effect of damping on the wave and mean field. Thus, while wave transience is the principal source of local changes in the mean flow (as shown by HD), dissipation is the source of changes in the time-averaged flow.

To examine the behavior of (17) during a warming we simulate an amplitude pulse by approximating the wave p.v. as

$$|Z'|^2 = \exp[-(t - t_0)/T]|(15)|^2, \quad (20)$$

where $T = 25$ days^2, $t_0 = 5$ days, and $t_0 < t_{\max}$, $t_{\max}$ being the finite time to singularity shown in Fig. 3. We may solve (17) numerically and plot the phase space path of the solution as in Fig. 2.

Two cases ($D = 0$ and $D \neq 0$) and the linear case for $\lambda_z = 50$ km are plotted in Fig. 4. With the linear case we can quantitatively compare the effects of wave saturation to the "linear" approximation that the mean flow can absorb any amount of wave forcing. For the $D = 0$ case our example also nicely illustrates the non-interaction theorem. Note that Z reaches a maximum exactly at pulse maximum (day 5) and the curves are mirror images about the Z axis.

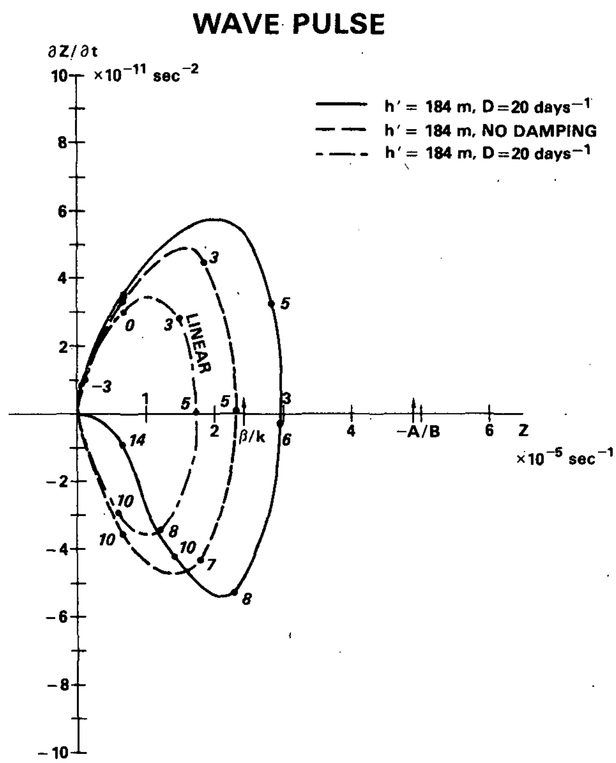

FIG. 4. Solution paths as in Fig. 1 for wave pulse problem. Numbers next to curves indicate the time in days, pulse peak occurs at 5. $l = (50)$ km^{-1}, $s = 1$. The curve labeled "linear" represents a solution to linearized Eq. (17).

As expected the linear and nonlinear [Eq. (17)] solutions are identical for small wave amplitude up to about day −1. After that point the solutions begin to diverge. The nonlinear effects (wave saturation of the mean flow) clearly *enhance* the response of the system to the wave pulse as expected from the discussion in Section 4. Note that both Z and $\partial Z/\partial t$ are about a factor of 1.6 larger for the nonlinear case. Furthermore, divergence between the linear and nonlinear solutions occurs well before the point $Z = \beta/k$ confirming the implication in Section 4 that partial wave saturation of the mean flow is an important effect even for relatively small wave amplitudes.

If dissipation is present in both the wave and mean flow equations, the curve is asymmetric across the Z axis. The net result is $Z > 0$ just after the pulse passage as expected from (19). The presence of dissipation actually increases the mean flow response to pulse passage because of the increase in the value of F. It is interesting to note that linear case is symmetric about the Z axis even though dissipation is present. This is because if $B = 0$ then Z is proportional to $|Z'|^2$ as is obvious from (17). Thus the asymmetry of the nonlinear case with dissipation ($B \neq 0$) about the Z axis is wholly due to the saturation effects. This point will be discussed further in the next section.

6. A parameterized model of stratospheric vacillation

In Section 3 we showed that the stratosphere cannot remain steady in the presence of large-amplitude waves. However, the response of a quasi-geostrophic model to the presence of large-amplitude waves may not necessarily result in a vacillating flow. For example, mean flow changes could simply block large-amplitude waves from propagating into the upper stratosphere. To illustrate some of the processes important in a vacillating flow and to provide a model which contains a more realistic response to changes in $\bar{u}$ and $\partial\bar{q}/\partial y$ in the wave equation (3), a parameterized model of vacillation is presented in this section. In many respects this model contains the essence of the dynamical response of the stratosphere to planetary wave forcing at a single level.

a. Model description

We will let Z', the stationary wave amplitude, be a function of Z in Eq. (17) such that the wave amplitude will suddenly decrease when Z reaches some value. For example, that value of Z could correspond to the appearance of easterlies over a large region. A simple parameterization of this effect is

$$\left.\begin{aligned} |Z'|^2 &= 0, & Z &> Z_2 \\ |Z'| &= \gamma(t')|(15)|^2, & Z &< Z_2 \end{aligned}\right\}. \qquad (21)$$

We take $s = 1$ and $l = (50\ \mathrm{km})^{-1}$. As before, we neglect dissipation in the computation of the wave vorticity (Z') but include its effect in the computation of F when appropriate. To simulate the switch-on of the wave forcing and the restoration of that forcing after the vacillation cycle we let

$$\gamma(t') = H(t' - t_1)(t' - t_1)/t_2\,,$$

where t' is the time measured after Z has exceeded Z_2. H is the step function

$$H(t) = \begin{cases} 1, & t > 0 \\ 0, & t < 0. \end{cases}$$

Imbedded with the function $\gamma(t')$ is the physics of the response of the wave amplitude to the changing mean flow at and below the level for which (18) is solved. We use HM Fig. 17 to determine t_1 and t_2. In their figure, a zero wind region descends and blocks planetary wave propagation for about 30 days, thus $t_1 = 30$ days. The wave forcing reaches peak intensity in about 10–20 days within a cycle. We use a value of 10 days for t_2. We choose Z_2 as β/k which is half the maximum value which Z can obtain and roughly corresponds to the appearance of a zero wind line (Section 4). $D = 5\ \mathrm{days}^{-1}$.

b. Model results

The results of integrating (7) are given in Fig. 5a which shows Z and $|Z'|$ vs t. Three regions are marked on the figure which indicate the important physical processes. The deceleration of the mean flow in the channel by wave transience as the wave amplitude increases is labeled switch-on. Once Z_2 is reached, the wave is turned off which produces a transient reversal and rapid reacceleration of the mean flow (Z decreases). Note that the flow does not return to the initial state because irreversible deceleration during the cycle due to the damping of the waves. Since the waves are shut off for 30 days, the mean flow relaxes almost to its prewarming equilibrium value. This is the "decay" phase indicated in the figure.

The maximum wave amplitude achieved during this cycle is 210 gpm. This is much smaller than the 1200 gpm reached during HM cycle at 50 km. Note however, that vacillation is also present down to 20 km in HM where the maximum wave amplitude is ~400 m. Since we have taken $\nabla^2\bar{\psi}_0$ to be zero and neglected damping in the calculation of $|Z'|$, closer agreement should not be anticipated. Nevertheless, the behavior of Z corresponds quite well to $-\bar{u}$ as shown in HM Fig. 17.

It is clear from the results above that vacillation is not a result of the nonlinearity of (8), but is due to the response function (21). To illustrate again the importance of the nonlinear terms in (8) (the wave saturation effect), we solve the linearized version of

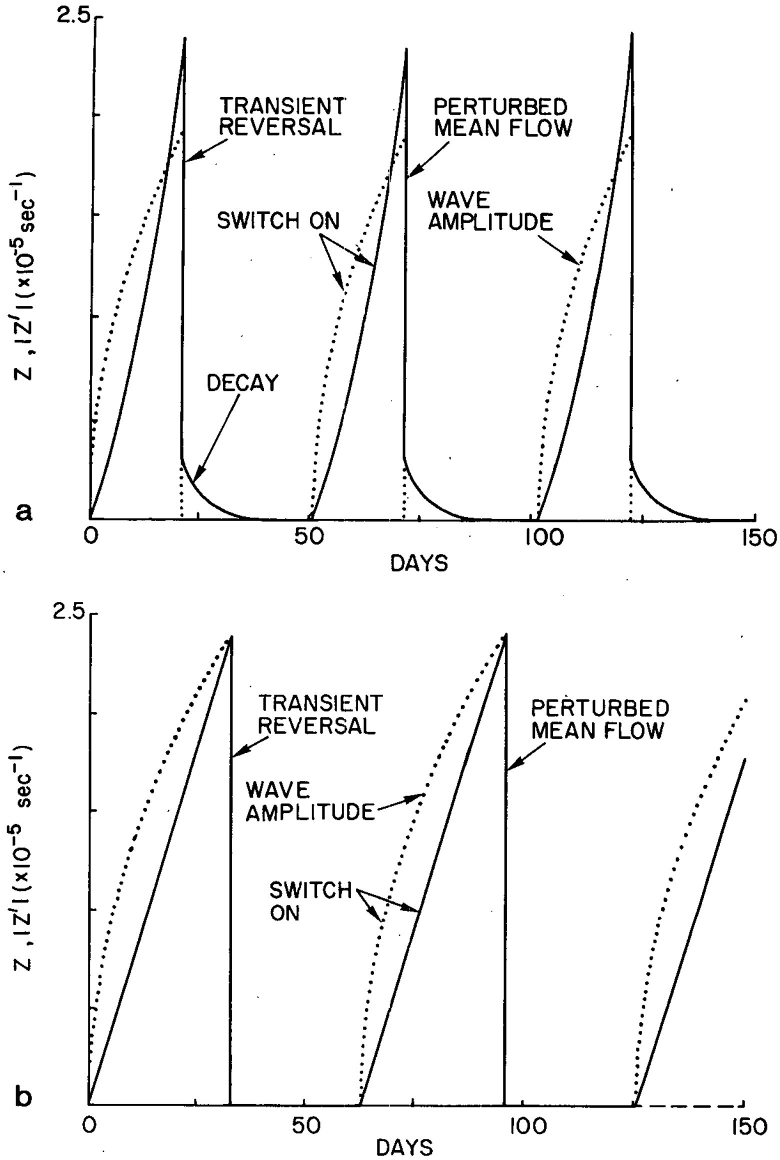

FIG. 5. Perturbed mean flow and wave potential vorticity as a function of t (days) for vacillation model: (a) solution to Eqs. (17) and (22); (b) linear case, $B = 0$.

(8) by setting $B = 0$. As previously mentioned this assumption is equivalent to giving the mean flow an infinite reservoir of p.v. Fig. 5b shows the results.

Clearly from the form of (8) the solution will mirror the function $\gamma(t')$; therefore the decay phase shown in Fig. 5a is absent from Fig. 5b. The maximum wave amplitude achieved in the linear model is 272 gpm, nearly 50 gpm larger than that required by the nonlinear model so the vacillation cycle is longer.

The presence of the decay phase in the HM (their Fig. 6b, days 150–200, 40 km) results verifies the importance of the wave saturation of the mean flow during the warming cycle. If saturation were unimportant the solution would closely resemble the linear response and the mean flow would return to its original value once the wave forcing ceases. We can therefore conclude that the trigger for vacillation [$\gamma(t')$, Z_1 or the point where easterlies appear] lies well within the nonlinear (nearly saturated) regime.

Since the wave amplitude increases linearly with time, this model only reproduces the vacillating solutions of HM. If we chose a function of t which reaches a steady wave amplitude after some time e.g., $\gamma(t') = H(t' - t_1)[1 - \exp(t' - t_1)]$, then the model will also reproduce the sharp switch over between vacillating and steady regimes observed by HD where $Z(Z') = Z_2$ provided $|Z'| < \beta/\sqrt{2}k$.

7. Summary and discussion

The aim of this paper has been to examine the response of the mean flow to wave forcing when the wave amplitude is large. That is, when erosion of the basic state potential vorticity gradient by wave "mixing" has consequences for the continued evolution of the mean flow.

The potential enstrophy equations yield two separate limits on Rossby wave amplitudes. First, it was shown that the wave potential vorticity cannot exceed $\sqrt{2}\beta/k$ under any conditions. This is the saturation limit. Second, a steady mean flow is not possible in the presence of waves if the wave potential vorticity exceeds $\beta/\sqrt{2}k$ (half the value at saturation). These results have important implications for several problems: First, as pointed out by Lindzen and Schoeberl (1982), the propagation properties of Rossby waves are not required to justify the lack of an earth corona as was originally thought by Charney and Drazin (1961); second, it shows that steady-state or time-dependent models which neglect wave-mean flow interaction are unrealistic where the wave amplitude is predicted to be large (e.g., the upper stratosphere during winter). And finally, the "steadiness" limit may provide a partial explanation for the behavior of the vacillation model used by HM. HM found that for low-amplitude forcing, a steady solution could be achieved in their model, but for larger forcing the model remained unsteady. While this kind of behavior is consistent with the "steadiness" theorem derived in Section 3, it may also be possible for the mean flow to change in such a way as to block large amplitude planetary waves from developing and thus permit a steady solution in any model. Possibly the severe horizontal truncation or the imposed boundary conditions used in the HM model forbid the kind of mean flow changes which could permit the evolution of a steady solution. Note that the wavenumber 2 vacillations produced in the numerical model of Schoeberl and Strobel (1980) progressively weaken, suggesting that their model may, in time, achieve a steady solution.

In Section 4 we have considered the response of the mean flow to constant wave forcing which is independent of the basic state. This is an instructive but not very realistic problem. The results, however, verify the steady-state analysis that for large wave forcing a steady solution is not possible. If the wave forcing is maintained, the mean flow changes such that the deceleration rate eventually becomes singular. Interestingly enough the mean flow reaches saturation in a finite time for this case, ~10–50 days, which is called an explosive response, while we cannot, of course, expect the mean flow to actually saturate. It is quite plausible that the rapid changes in the mean flow seen in observations and in models during the sudden warming is enhanced by partial saturation. This effect, for example, appears to be evident in Davies' (1981; Fig. 2) computations.

Because holding the wave forcing fixed is unrealistic we have also considered the problem of the mean flow response to a wave pulse in Section 5. The idea is that as the wave forcing decelerates the mean flow, the wave amplitude will eventually start to decrease (e.g., when easterlies appear), so that the whole process looks to the mean flow like a wave amplitude pulse passing upward through an isobaric surface. Comparing the two cases when wave saturation effects are included and when they are not, it is clearly shown (Fig. 4) that the wave saturation amplifies the response of the mean flow to wave forcing. This amplification is evident even when changes in the mean flow are relatively small.

Finally we have used the potential enstrophy equation to develop a simple parameterized model of vacillation at a single pressure level. We switch-on the wave vorticity proportional to $\sqrt{t}$ and shut it off half way toward mean flow saturation ($Z = \beta/k$). The wave forcing then remains off for a given amount of time and is then turned on again.

The vacillation cycles generated by this model are remarkably realistic given its simplicity. This model also does not vacillate if a steady mean flow solution can be reached below the wave cutoff value of Z. To further demonstrate the importance of wave saturation during the vacillation cycle we repeated the calculations using the linear enstrophy equations. In the linear case the solutions show no "decay phase" and entirely mirror the wave forcing function. Since the decay phase is a prominent feature of observed and modeled stratospheric warmings this result shows that wave saturation plays an important role in all these events.

We have made a number of simplifying assumptions in this analysis. We have assumed a quasi-geostrophic, separable system of equations (consistent with HM). However, vacillation is also observed in nonseparable, quasi-geostrophic and PE models of the stratosphere (Schoeberl and Strobel, 1980). Sudden warmings take place in these models primarily after planetary waves have been ducted into high-latitude regions by reflecting critical lines which form at lower latitudes (Dunkerton *et al.*, 1981). The ducting of the planetary waves probably concentrates the wave forcing into the polar cap region where β is small and a subsequent explosive deceleration in the mean flow predicted by this theory would occur even for moderate wave forcing.

Acknowledgments. The author would like to thank H. M. Mitchell, R. S. Lindzen, M. E. McIntyre, Darrell F. Strobel, R. Rood and a reviewer for helpful comments and discussions. This work was funded by the National Aeronautics and Space Administration's Upper Atmospheric Research Office and the Office of Naval Research.

APPENDIX

General Linear Damping

In order to simplify the discussions in the paper, the dissipation was taken to be linearly related to the perturbed potential vorticity ($\bar{q}_p$, q'). It is doubtful that such a simple process exists. We now consider a more general, but linear, dissipation form of Newtonian cooling and Rayleigh friction which vary independently with height. The wave equation is

$$\frac{\partial}{\partial t} q' + R\left(\nabla^2 \psi' + \frac{f_0^2}{RN^2\rho}\frac{\partial}{\partial z} a\rho \frac{\partial}{\partial z}\psi'\right) = -v'\frac{\partial \bar{q}}{\partial y}, \quad \text{(A1)}$$

where $a(z)$ is the Newtonian cooling coefficient and R the Rayleigh friction coefficient. The mean flow equation is

$$\frac{\partial}{\partial t}\bar{q} + R\left(\frac{\partial^2 \bar{\psi}_p}{\partial y^2} + \frac{f_0^2}{N^2 R\rho}\frac{\partial}{\partial z} a\rho \frac{\partial}{\partial z}\bar{\psi}_p\right) = -\frac{\partial}{\partial y}(\overline{v'q'}). \quad \text{(A2)}$$

Using the procedure in Section 2 to obtain the channel average enstrophy equations we have

$$\frac{1}{2}\left\langle \frac{\partial}{\partial t}\bar{q}^2\right\rangle + R\left\langle \bar{q}\frac{\partial^2}{\partial y^2}\bar{\psi}_p + \frac{\bar{q}f_0^2}{RN^2\rho}\frac{\partial}{\partial z} a\rho \frac{\partial}{\partial z}\bar{\psi}_p\right\rangle = -\frac{1}{4}\left\langle \frac{\partial}{\partial t}|q'|^2\right\rangle - R\langle\overline{q'\nabla^2\psi'}\rangle - R\left\langle\overline{\frac{q'f_0^2}{\rho RN^2}\frac{\partial}{\partial z} a\rho\frac{\partial}{\partial z}\psi'}\right\rangle. \quad \text{(A3)}$$

To see how this equation differs from the results in Section 3, we examine a simple problem. Let $R = 0$ and assume the waves are steady. Then (A3) becomes

$$\left\langle(\bar{q}_0 + \bar{q}_p)\left(\frac{\partial}{\partial z} a\rho\frac{\partial}{\partial z}\bar{\psi}_p\right)\right\rangle = -\left\langle\overline{q'\frac{\partial}{\partial z} a\rho\frac{\partial \psi'}{\partial z}}\right\rangle. \quad \text{(A4)}$$

Assuming separability as in Section 3 and taking a background constant wind field, (A4) becomes

$$(Z'' + C_1 Z')(Z'' + C_2 Z' + C_3) = F, \quad \text{(A5)}$$

where C_1, C_2 and C_3 incorporate the y integral constants; F represents the wave forcing as before; and Z is the separable part of $\bar{\psi}_p$. If the vertical variation of Z in (A5) is known, we have an algebraic equation similar to (8) in Section 3. In general, though, (A5) must be solved as a nonlinear equation and the condition that Z be real must still hold. Obviously, from the form of (A5) not all values are allowed for real Z as can be shown trivially if $Z \approx e^{kz}$, for example.

From (A4) it is also apparent that if a is constant the steadiness criteria is independent of the Newtonian coefficient. For a coefficient which varies with altitude there is a dependence on the derivative of $\ln a$, a weakly varying function, as can be seen by writing (A4) in the form

$$\left\langle(\bar{q}_0 + \bar{q}_p)\left(\frac{\partial}{\partial z}\rho\frac{\partial}{\partial z}\bar{\psi}_p + \rho\frac{\partial\bar{\psi}_p}{\partial z}\frac{\partial}{\partial z}\ln(a)\right)\right\rangle = -\left\langle\overline{q'\left(\frac{\partial}{\partial z}\rho\frac{\partial}{\partial z}\psi' + \rho\frac{\partial\psi'}{\partial z}\frac{\partial}{\partial z}\ln(a)\right)}\right\rangle. \quad \text{(A6)}$$

Thus the results of Section 3 that the steadiness criteria for the mean flow is independent of the damping are still approximately valid.

REFERENCES

Cap, F. F., 1976: *Handbook on Plasma Instabilities*, Vol. I. Academic Press, 458 pp.

Charney, J. G., and P. G. Drazin, 1961: Propagation of planetary scale waves from the lower atmosphere to the upper atmosphere. *J. Geophys. Res.*, **66**, 83–109.

Davies, H. C., 1981: An interpretation of sudden warmings in terms of potential vorticity. *J. Atmos. Sci.*, **38**, 427–445.

Dunkerton, T., C. P. Hsu and M. E. McIntyre, 1981: Some Eulerian and Lagrangian diagnostics for a model stratospheric warming. *J. Atmos. Sci.*, **38**, 819–843.

Holton, J. R., 1976: A semi-spectral numerical model for wave, mean-flow interactions in the stratosphere: Application to sudden stratospheric warmings. *J. Atmos. Sci.*, **33**, 1639–1649.

——, and C. Mass, 1976: Stratospheric vacillation cycles. *J. Atmos. Sci.*, **33**, 2218–2225.

——, and T. Dunkerton, 1978: On the role of wave transience and dissipation in stratospheric mean flow vacillations. *J. Atmos. Sci.*, **35**, 740–744.

Lindzen, R., and M. R. Schoeberl, 1982: A note on the limits of Rossby wave amplitude. *J. Atmos. Sci.*, **39**, 1171–1174.

Matsuno, T., 1971: A dynamical model of the stratospheric sudden warming. *J. Atmos. Sci.*, **28**, 1479–1494.

O'Neill, A. O., 1980: The dynamics of stratospheric warmings generated by general circulation model of the troposphere and stratosphere. *Quart. J. Roy. Meteor. Soc.*, **106**, 659–690.

Schoeberl, M. R., 1978: Stratospheric warmings: Observations and theory. *Rev. Geophys. Space Phys.*, **16**, 521–538.

——, and D. F. Strobel, 1980: Numerical simulation of sudden stratospheric warmings forced by mountains. *Geophys. Res. Lett.*, **7**, 149–152.

van Loon, H., R. L. Jenne and K. Labitzke, 1973: Zonal harmonic standing waves. *J. Geophys. Res.*, **78**, 4463–4471.