Excitation of Transient Rossby Waves on the Stratospheric Polar Vortex and the Barotropic Sudden Warming

J. G. ESLER

Department of Mathematics, University College London, London, United Kingdom

R. K. SCOTT

Northwest Research Associates, Inc., Bellevue, Washington

(Manuscript received 15 September 2004, in final form 9 March 2005)

ABSTRACT

The excitation of Rossby waves on the edge of the stratospheric polar vortex, due to time-dependent topographic forcing, is studied analytically and numerically in a simple quasigeostrophic *f*-plane model. When the atmosphere is compressible, the linear response of the vortex is found to have two distinct components. The first is a spectrum of upward-propagating waves that are excited by forcing with temporal frequencies within a fixed "Charney–Drazin" range that depends on the angular velocity at the vortex edge and the vortex Burger number. The second component of the response is a barotropic mode, which is excited by forcing with a fixed temporal frequency outside the Charney–Drazin range. The relative magnitude of the two responses, in terms of total angular pseudomomentum, depends on the ratio of the horizontal scale of the forcing to the Rossby radius. Under typical stratospheric conditions the barotropic response is found to dominate. Nonlinear simulations confirm that the linear results remain relevant to understanding the response in cases when strongly nonlinear Rossby wave breaking ensues. It is shown that a sudden warming, or rapid increase in vortex angular pseudomomentum, can be generated at much lower forcing amplitudes when the barotropic mode is resonantly excited compared to when the upward-propagating waves are excited. A numerical simulation of a "barotropic sudden warming" due to excitation of the barotropic mode by a relatively weak topographic forcing is described.

1. Introduction

There has been greatly renewed interest in sudden stratospheric warmings since the spectacular event in the Southern Hemisphere of September 2002 (*J. Atmos. Sci.*, special issue, March 2005). Synoptically, the sudden warming event was observed in isentropic maps of Ertel's potential vorticity (Newman and Nash 2005), as well as in satellite ozone measurements (Baldwin et al. 2003), to occur following an elongation of the polar vortex in the lower and middle stratosphere and to culminate in the vortex splitting into two parts. Although the dynamical cause of the warming has been widely discussed (Charlton et al. 2005; Harnik et al. 2005; Manney et al. 2005), its exact nature, whether it be anomalous conditions in the troposphere, preconditioning of the vortex, or otherwise, remains obscure.

Dynamically there is a clear link between upward-directed Eliassen–Palm (EP) flux at the tropopause level and stratospheric warming events (Edmon et al. 1980; Dunkerton and Baldwin 1991; Polvani and Waugh 2004). It is usually assumed that it is dynamical forcing, due to planetary-scale Rossby waves originating in the troposphere, that generates the upward EP flux, thereby causing the warming. However, even in models, the relationship between measures of the dynamical forcing (e.g., geopotential height fields at the tropopause level) and the upward EP flux is not well understood. For example, the 2002 Southern Hemisphere warming coincided with a strong burst of upward EP flux, but Rossby wave amplitudes at the tropopause level were not exceptional (see, e.g., Charlton et al. 2005, their Fig. 10). Additionally, Scott and Polvani (2004) provide direct modeling evidence that the condition of the stratosphere may itself strongly influence

Corresponding author address: J. G. Esler, Department of Mathematics, University College London, 25 Gower Street, London, WC1E 6BT, United Kingdom.
E-mail: gavin@math.ucl.ac.uk

the magnitude of the upward EP flux. The relative importance of stationary and transient Rossby waves in determining the flux is also unclear (Scinocca and Haynes 1998). Kushner and Polvani (2005) have recently highlighted a spontaneous sudden warming in a general circulation model containing only transient waves, whereas following Matsuno (1971), most mechanistic model simulations of warmings have been forced with stationary waves only.

Previous theories of sudden stratospheric warmings include the "linear resonance" theory of Tung and Lindzen (1979a,b). In their theory, which was developed initially for a β-plane channel model, sudden warmings are anticipated to occur when a free-traveling mode in the stratosphere comes into resonance with a forced stationary wave. Tung and Lindzen noted the possibility for the barotropic mode in the atmosphere to become resonantly excited in this manner, and also the possibility that other modes may be excited if a vertical turning surface is present, above which the mode in question is evanescent. As the theory is linear, however, little insight is gained into whether, for a given vertical mode, resonant wave growth can be sustained into the nonlinear regime sufficiently long enough to allow for a realistic warming to develop. To begin to address this issue, Plumb (1981) extended and developed Tung and Lindzen's work by allowing weakly nonlinear modification of the mean flow in the β channel. The consequence of allowing this modification was found to be that unbounded wave growth occurred only if the stationary forcing was initially off resonant, with the mean flow modification acting to bring the free waves closer to resonance as they grow. Plumb named this mechanism "self-tuning resonance," and its occurrence in more realistic stratospheric models has been investigated by, for example, Smith (1989).

In this paper we address the above issues by analyzing a simple quasigeostrophic *f*-plane vortex (Dritschel and Saravanan 1994), described in detail in section 2. The system facilitates both analytic and numerical investigation of the relationship between the details of a prescribed time-dependent lower boundary forcing, the upward EP flux, and the response of the vortex itself. Despite the idealizations made, the system retains many important dynamical aspects of the winter polar stratosphere, and we submit that the relationships obtained here are highly relevant to the situation in the winter polar stratosphere in both hemispheres. Arguably the key strength of the model is that it allows the relevance of linear or weakly nonlinear theories to be assessed directly in comparison with fully nonlinear numerical simulations. Unlike the situation with the β-channel model, fully nonlinear simulations of the vortex model capture what we believe to be the fundamental aspect of the nonlinear dynamics of the winter stratosphere, namely, the dynamics associated with the finite-amplitude distortion of a three-dimensional columnar vortex.

In what follows a linear analysis broadly analagous to that of Tung and Lindzen (1979a) for the β-channel model is carried out for the vortex model, although the treatment is generalized to allow for transient forcing with arbitrary time dependence. Emphasis is placed on the development of linear predictions for the evolution of quantities that satisfy nonlinear conservation properties. Hence one of the main results, described in section 3, is the derivation of an exact analytic expression for the time-integrated upward EP flux at the lower boundary that is valid in the linear limit for a barotropic vortex. The expression has two components, the first of which may be interpreted in terms of the Charney–Drazin theorem applied to the vortex (Charney and Drazin 1961) and corresponds to the excitation of a spectrum of upward-propagating waves. The second component of the expression corresponds to the excitation of a barotropic mode. Importantly, the lower boundary EP flux is found to depend on the entire time history of the forcing, or equivalently it depends on both the instantaneous forcing amplitude and the current state of the vortex itself. The analytic results are verified using a linear numerical model in section 3, and their relevance to the atmospheric situation is demonstrated by considering the fully nonlinear, three-dimensional evolution in a high-resolution numerical model in section 4. Three-dimensional representations of the vortex allow the physical consequences of different forcings to be visualized. In particular, it is shown that forcing applied near the resonant frequency of the barotropic mode can cause the vortex to split, even at relatively low forcing amplitudes. A numerical simulation of such a vortex split, described as a "barotropic sudden warming," is discussed in detail. Section 6 contains our conclusions.

2. Quasigeostrophic *f*-plane vortex model: Formulation and background

a. Analytic formulation

The *f*-plane vortex model to be investigated here was first described by Dritschel and Saravanan (1994) (see also Fyfe and Wang 1997; Waugh and Dritschel 1999; Wang and Fyfe 2000). In its most general configuration, the vortex at a given vertical level is represented by an area $\mathcal{S}(z)$ containing anomalously high quasigeostrophic potential vorticity (PV hereafter) $q = f + \Delta_b + \Delta(z)$ compared with a uniform background value out-

side $\mathcal{S}(z)$ where $q = f + \Delta_b$. Here f is the Coriolis parameter and Δ_b is a constant vorticity that may be used to add a solid body rotation to the flow. The PV distribution may be used to obtain the flow streamfunction ψ, and hence the horizontal velocity $\mathbf{u} = -\boldsymbol{\nabla} \times \psi\mathbf{k}$, by inverting the elliptical operator

$$q(\mathbf{x}, z) = f + \nabla_H^2\psi + \frac{1}{\rho}\left(\rho \frac{f^2}{N^2}\psi_z\right)_z = \begin{cases} f + \Delta(z) + \Delta_b & \mathbf{x} \in \mathcal{S}(z) \\ f + \Delta_b & \text{o/w} \end{cases}, \quad (1)$$

subject to the lower boundary condition

$$f\psi_z + N^2 h_T = 0, \quad \text{on} \quad z = 0. \quad (2)$$

Here ∇_H^2 is the horizontal Laplacian operator, and N is the buoyancy frequency, which we take to be independent of z, although this condition is easily relaxed. The density $\rho(z) = \exp(-z/H)$, where H is a constant scale height. The boundary condition (2) represents a topography-like forcing on the lower boundary and is equivalent to setting the vertical velocity in pressure coordinates to zero on a material surface. See, for example, White (1978) for a discussion of the exact topographic boundary condition in which the physical vertical velocity is set to zero on the material surface. The pressure height field h_T is allowed, in general, to be time dependent so as to model the effects of forcing by transient tropospheric planetary waves.

The wind field obtained by the inversion procedure may then be used to advect the boundary of the vortex edge. Defining the vortex edge in cylindrical polar coordinates (r, ϕ, z) to be located at $r = R(z) + \eta(\phi, z, t)$, the appropriate kinematic condition may be written

$$\left(\frac{\partial}{\partial t} + \mathbf{u} \cdot \boldsymbol{\nabla}\right)(\eta - r) = 0 \quad \text{on} \quad r = R + \eta. \quad (3)$$

Taken together with the inversion equation (1) and lower boundary condition (2), Eq. (3) describes an idealized dynamical model of a polar stratospheric vortex.

One important dynamical simplification owing to our choice of model is the suppression of Rossby wave propagation in the radial direction since Rossby waves are confined to propagate on the vortex edge, which is defined by a single contour. Propagation is therefore in the vertical direction only. Based on linear analysis of the stratospheric mean flow (e.g., Karoly and Hoskins 1982; Harnik and Lindzen 2002) it is often argued that equatorward propagation of Rossby wave activity is important in the stratosphere. However, the surf zone surrounding the vortex in midwinter is well observed to be strongly mixed, implying that at best only very weak PV gradients exist to support such equatorward propagation. It might therefore be argued, and indeed Swanson et al. (1997) made a similar argument for the situation at the extratropical tropopause, that in suppressing radial propagation we are creating a paradigmatic model that is more (rather than less) relevant to the actual nonlinear dynamics of the winter stratosphere.

A further important question concerns the choice of lower boundary condition (2). As stated above, a topography-like forcing at the lower boundary is used to model transient dynamical forcing due to tropospheric planetary wave activity, and physically the model lower boundary is situated at the height of the tropopause. In many mechanistic models, however, it is geopotential height that is specified at the lower boundary, which is equivalent to specifying ψ rather than ψ_z, as here. Some insight into the meaning of each boundary condition can be obtained by exploiting the ideas of piecewise PV inversion (Nielsen-Gammon and Lefevre 1996). According to this view, it appears sensible to associate the flow induced by inverting the boundary anomaly with whatever flow might be induced by tropospheric PV anomalies lying beneath the model boundary, were they resolved by the model. Bretherton (1966) showed that a ψ_z anomaly on a horizontal boundary corresponds to a monopole PV anomaly concentrated at the boundary; hence, the flow induced by a ψ_z boundary anomaly is similar to that obtained from a strong PV anomaly with a relatively small vertical extent, for example, that induced by a Rossby wave at the tropopause level. By contrast, a ψ boundary anomaly corresponds to a PV anomaly with a dipole structure in the vertical (Dritschel and Ambaum 1997). To our knowledge, large-scale PV anomalies with a dipole structure in the vertical do not regularly occur at the tropopause level, so for this reason we consider the ψ_z boundary condition (2) to be more physical. We note, however, that lower boundaries of both types necessarily distort the Green's function used in the inversion of the PV equation (1) in their immediate vicinity (Scott and Dritschel 2005).

b. *Dispersion relation and Charney–Drazin theorem*

Waugh and Dritschel (1999, WD99 hereafter) and Wang and Fyfe (2000) have respectively presented discretized and Boussinesq versions of the linear Rossby wave dispersion relation for the case of a barotropic vortex (Δ and R both z independent). Here the near-analogous derivation of the non-Boussinesq continuous dispersion relation is briefly described, both as a general background and for later use in section 3.

Following WD99 and Wang and Fyfe, the kinematic condition for a near-circular vortex may be linearized as

$$\left(\frac{\partial}{\partial t} + \Omega_e \frac{\partial}{\partial \phi}\right)\eta - v = 0, \quad \text{on} \quad r = R, \tag{4}$$

where $\Omega_e = (\Delta + \Delta_b)/2$ is a constant angular speed at the vortex edge and $\mathbf{u} = (u, v)$, where u and v are azimuthal and radial velocities, respectively. The vorticity jump across the vortex edge may also be related to the vortex edge displacement, following, for example, Swanson (2000),

$$[\psi_r]_{r=R^-}^{r=R^+} = -\Delta\eta. \tag{5}$$

Looking for normal mode solutions of the form

$$\eta = \mathrm{Re}\epsilon R \exp\left[\frac{z}{2H} + ik\phi + imz - i\omega_s(m)t\right], \tag{6}$$

with nondimensional amplitude ϵ, integer azimuthal wavenumber k, vertical wavenumber m, and frequency $\omega_s(m)$, we first note that the normal mode streamfunction satisfies $\psi = \Delta R\tilde{\psi}(r, m, k)\eta$, where the radial structure function

$$\tilde{\psi}(r, m, k) = \begin{cases} I_k[\mathcal{B}(m)r/R]K_k[\mathcal{B}(m)] & \text{for} \quad r < R \\ K_k[\mathcal{B}(m)r/R]I_k[\mathcal{B}(m)] & \text{for} \quad r > R, \end{cases} \tag{7}$$

where

$$\mathcal{B}(m) = \frac{fR}{N}\left(m^2 + \frac{1}{4H^2}\right)^{1/2},$$

and I_k and K_k are kth-order modified Bessel functions. The dispersion relation may now be evaluated from (4),

$$\omega_s(m) = k\Omega_e - \Delta k K_k[\mathcal{B}(m)]I_k[\mathcal{B}(m)]. \tag{8}$$

Following Charney and Drazin (1961), the dispersion relation may be rearranged to emphasize the range of possible forcing frequencies ω that lead to the excitation of vertically propagating waves with real m. These turn out to be

$$k[\Omega_e - \Delta I_k(\mathrm{Bu}^{-1/2})K_k(\mathrm{Bu}^{-1/2})] = \omega_s^- \le \omega \le \omega_s^+ = \Omega_e k, \tag{9}$$

where Bu is the Burger number for the vortex

$$\mathrm{Bu} = \mathcal{B}(0)^{-2} = \frac{4L_R^2}{R^2} \tag{10}$$

and $L_R = NH/f$ is the Rossby radius. The vortex Burger number clearly has a defining role in determining the response to transient forcing. As Bu $\to 0$, $\omega_s^- \to \omega_s^+$, closing the "window" of available forcing frequencies. In the Boussinesq limit Bu $\to \infty$, the window is at its widest, with $\omega_s^- \to \omega_0$ defined below. For the stratospheric parameter values used by WD99, the Burger number Bu $\approx 4/9$, and the range of phase speeds that allow vertical propagation, measured at the vortex edge, is 16.1–58.9 m s^{-1} for wave 1 and 27.9–58.9 m s^{-1} for wave 2, with 58.9 m s^{-1} being the azimuthal velocity at the vortex edge. Interestingly, this means that stationary waves are evanescent in the experiments of WD99, with vertical propagation due to transient waves with eastward phase speeds generated during the mountain switch on.

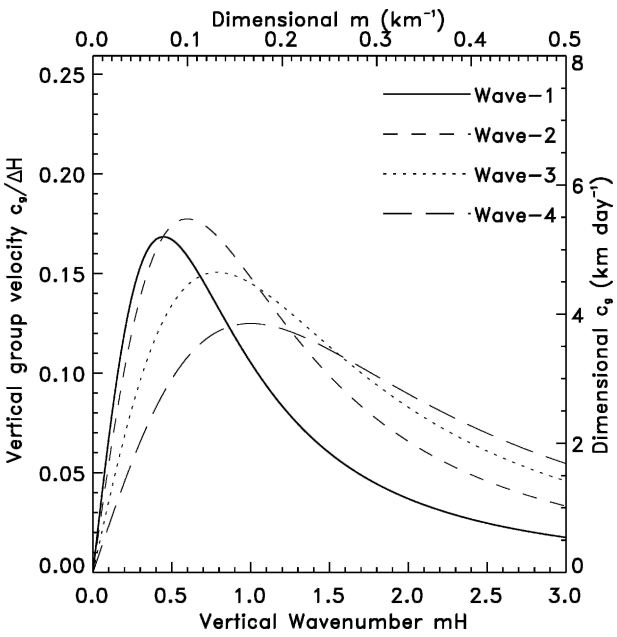


FIG. 1. Vertical group velocity c_g versus vertical wavenumber m (solid curve) for the first four azimuthal wavenumbers $k = 1, 2, 3,$ and 4. Parameter settings are as for WD99 (vortex radius $R = 3L_R = 2700$ km, scale height $H = 6.14$ km, $\Delta = 0.4f$, $\Delta_b = 0.4f$.)

An advantage of using the continuous dispersion relation (8), as opposed to the discretized spectrum given by WD99, is that the vertical group velocity $c_g(m) = \partial\omega_s/\partial m$ can be calculated for the spectrum of upward-propagating waves [see the appendix, Eq. (A1)]. Figure 1 shows the vertical group velocity for the first four azimuthal wavenumbers $k = 1, 2, 3,$ and 4 using the vortex parameter settings of WD99 ($L_R = 900$ km, $R = 3L_R$, $H = 6.14$ km, $\Delta = 0.4f$, $\Delta_b = -0.1f$). Note that the maximum vertical group velocity at around 5.5 km day^{-1} is approximately equal for wave 1 and wave 2, and occurs for very long vertical wavelengths (9–14 $H \approx$ 54–84 km).

An important point that is emphasized further in section 3 is that the vertically propagating normal modes (6), referred to hereafter as the Charney–Drazin spectrum (CDS), do not form a complete set. In addition to

this spectrum there is a barotropic mode with no vertical dependence:

$$\eta_0 = \epsilon \exp\{k\phi - i\omega_0 t\}, \quad \omega_0 = k\Omega_e - \frac{\Delta}{2}.$$

In this case $\psi_0 = \Delta R \eta_0 \tilde{\psi}_0$ with

$$\tilde{\psi}_0 = \frac{1}{2k} \begin{cases} \left(\frac{r}{R}\right)^k & \text{for} \quad r < R \\ \left(\frac{R}{r}\right)^k & \text{for} \quad r > R \end{cases}. \tag{11}$$

Note that the frequency of this mode ω_0 is distinct from the Charney–Drazin frequency range $[\omega_s^-, \omega_s^+]$, provided the atmosphere is compressible (as opposed to Boussinesq, with Bu $\to \infty$). At the WD99 parameter settings, frequency ω_0 corresponds to phase speeds at the vortex edge of -19.6 m s^{-1} for wave 1 and 19.6 m s^{-1} for wave 2. In sections 3 and 4 we will demonstrate that in a realistic compressible atmosphere the barotropic mode plays a major role in the dynamical response of the vortex to transient forcing.

c. *Wave activity conservation and the Eliassen–Palm flux*

Dritschel and Saravanan (1994) derived a nonlinear angular pseudomomentum conservation relation for the general *f*-plane vortex model by considering the conservation of total angular impulse

$$\mathcal{I} = \frac{1}{2} \int \rho(z) r^2 q(\mathbf{X}, t)\, d^3\mathbf{X}. \tag{12}$$

By subtracting the angular impulse of the undisturbed vortex and considering the situation at fixed z (see also Wang and Fyfe 2000) a wave activity conservation relation of the form

$$\frac{\partial A}{\partial t} + \frac{\partial F}{\partial z} = 0 \tag{13}$$

may be derived. In this equation

$$\begin{aligned} A(z,t) &= \frac{1}{8}\Delta(z)\rho(z) \oint_{C(z)} (x^2 + y^2)(x\, dy - y\, dx) \\ &\quad - \frac{\pi}{4}\Delta(z)\rho(z)R(z)^4 \\ &= \frac{1}{8}\Delta(z)\rho(z) \int_0^{2\pi} [(R+\eta)^2 - R^2]^2\, d\phi \end{aligned} \tag{14}$$

can be interpreted as a wave activity, and

$$F(z,t) = \frac{f^2}{N^2} \int_0^{2\pi} \int_0^{\infty} \rho(z) \frac{\partial \psi}{\partial \phi} \frac{\partial \psi}{\partial z}\, r\, dr\, d\phi \tag{15}$$

as its vertical flux. Note that the final equality in (14) holds only when the vortex is not displaced over the pole, that is, when η is well defined. The vertical flux $F(z, t)$ is recognizable as the horizontal integral of the vertical component of the Eliassen–Palm flux for this model (e.g., Edmon et al. 1980). As Wang and Fyfe (2000) have shown for the Boussinesq case, a monochromatic vertically propagating wave F satisfies a group velocity condition:

$$F = c_g A, \tag{16}$$

where the vertical group velocity $c_g = \partial\omega_s/\partial m$. An explicit proof of this identity, expanding on that given by Wang and Fyfe, is given in the appendix.

Defining $\mathcal{A}$ to be the vertical integral of $A(z, t)$, integration of (14) with respect to z reveals that

$$\frac{d\mathcal{A}}{dt} = F(0,t); \tag{17}$$

that is, the rate of change of $\mathcal{A}$ is proportional to the upward EP flux integrated over the lower boundary. A further related dynamical quantity is the azimuthal mean angular momentum

$$\mathcal{M} = \int \rho r \bar{u}\, d^3\mathbf{X}, \tag{18}$$

where the overbar denotes azimuthal mean. Integrating the angular impulse equation by parts reveals that

$$\frac{d}{dt}(\mathcal{M} + \mathcal{A}) = 0. \tag{19}$$

The sum of angular pseudomomentum and momentum therefore remains constant, even under the action of time-dependent topographic forcing at the lower boundary. A positive lower boundary EP flux therefore acts to transfer angular momentum to angular pseudomomentum, and it follows that the rate of change of $\mathcal{M}$ is also related to the lower boundary EP flux through $d\mathcal{M}/dt = -F(0, t)$. Because the lower boundary EP flux determines important aspects of the vortex development, that is, the evolution of $\mathcal{A}$ and $\mathcal{M}$, a theoretical understanding of its dependence on the details of the transient lower boundary forcing is central to the dynamics of the problem. In section 3 we develop an analytic expression for the time-integrated lower boundary EP flux

$$\mathcal{F} = \int_{-\infty}^{\infty} F(0,t)\, dt$$

in the linear wave limit for a barotropic vortex forced by an arbitrary time-dependent pulse, and in section 4

we compare this prediction with the results of nonlinear integrations of a numerical model of the vortex.

3. The lower boundary Eliassen–Palm flux

a. Forcing with general time dependence

To facilitate the calculation of the time-integrated lower boundary EP flux $\mathcal{F}$ in the linear limit, it is assumed that the vortex is forced with a finite pulse at the lower boundary of the form

$$h_T(r, \phi, t) = \frac{h}{2} T(t) J_k(lr) \exp(ik\phi). \tag{20}$$

To further aid the mathematical development below, note also that we may express the time evolution function $T(t)$ in terms of its Fourier transform

$$T(t) = \int_{-\infty}^{\infty} \hat{T}(\omega) \exp(-i\omega t)\, d\omega.$$

A more general lower boundary forcing might be constructed by summing over azimuthal wavenumbers k and integrating over different values of l in the Bessel function argument (i.e., creating a "Bessel function transform" of the topography). The parameter l essentially assumes the role of a radial wavenumber in the cylindrical geometry.

To determine the EP flux it is helpful to separate the flow into a topographic and response component, that is, $\psi = \psi_T + \psi_R$ and $q = q_T + q_R$, which are defined as follows:

- The topographic component, denoted by subscript T, satisfies the lower boundary condition,

$$N^2 h_T + f\psi_{Tz} = 0, \quad \text{on} \quad z = 0, \tag{21}$$

  but has zero PV everywhere in the interior; that is, $q_T \equiv 0$.

  The response component, denoted by subscript R, satisfies the homogeneous lower boundary condition $\psi_{Rz} = 0$ on $z = 0$, but has a PV distribution in the interior identical to that of the vortex

$$q_R(\mathbf{x}) = \begin{cases} f + \Delta(z) + \Delta_b, & \mathbf{x} \in \mathcal{S}(z) \\ f + \Delta_b, & \text{o/w.} \end{cases}$$

Dividing the flow into these two components means that the kinematic equation (4) for the vortex edge can be written as a forced equation,

$$\left(\frac{\partial}{\partial t} + \Omega_e \frac{\partial}{\partial \phi}\right)\eta_R - v_R = v_T \quad \text{on} \quad r = R, \tag{22}$$

where the azimuthal velocity v is obtained from the streamfunction using $v = -(1/r)\psi_\phi$. The topographic component of the flow can be considered to play a similar role to the "forced stationary wave" in the theories of Tung and Lindzen (1979a,b) and Plumb (1981), although the mathematical treatment here is not exactly analogous.

The topographic streamfunction is obtained by inverting (1) for zero PV subject to (2), giving

$$\psi_T = \frac{fL_R^2 h}{H(2\alpha H - 1)} J_k(lr) \times \int_{-\infty}^{\infty} \hat{T}(\omega) \exp\left(\frac{z}{2H} - \alpha z + ik\phi - i\omega t\right) d\omega,$$

with

$$\alpha^2 = \frac{N^2 l^2}{f^2} + \frac{1}{4H^2}.$$

If this is to be used to force Eq. (22), it is convenient to express it in terms of the normal modes of the unforced problem (4) that satisfy the homogeneous lower boundary condition $\psi_z \equiv 0$ on $z = 0$. Applying an appropriate Fourier transform to $\exp(-\alpha z)$,

$$\psi_T = \frac{fL_R^2 h}{H} J_k(lr) \int_{-\infty}^{\infty} \hat{T}(\omega) \exp\left(\frac{z}{2H} + ik\phi - i\omega t\right) \times \left\{\frac{1}{2L_R^2 l^2} \exp\left(-\frac{z}{2H}\right) + \frac{2}{\pi}\int_0^{\infty} \hat{\psi}_T(m) \cos[mz + \gamma(m)]\, dm\right\} d\omega, \tag{23}$$

where

$$\hat{\psi}_T(m) = \frac{m}{(1 + 4m^2H^2)^{1/2}(\alpha^2 + m^2)} \quad \text{and}$$

$$\tan\gamma(m) = \frac{1}{2mH}.$$

One important point to note is the presence of a barotropic mode component in the Fourier transform (first term inside the curly brackets). The barotropic mode component has no z dependence and excites the barotropic mode response that is discussed below.

Writing the topographic streamfunction in the form (23) allows the response disturbance field η_R to be calculated from (22),

$$\eta_R = \frac{kfL_R^2 hJ_k(lR)}{RH} \int_{-\infty}^{\infty} \hat{T}(\omega) \exp\left(\frac{z}{2H} + ik\phi - i\omega t\right) \left\{ \frac{\exp(-z/2H)}{2L_R^2 l^2(\omega - k\Omega_e + \Delta/2)} + \frac{2}{\pi} \int_0^{\infty} \frac{\hat{\psi}_T(m) \cos[mz + \gamma(m)]}{\omega - k\Omega_e + \Delta k I_k[B(m)] K_k[B(m)]} \, dm \right\} d\omega. \tag{24}$$

The asymptotic behavior of the ω integral in (24) may be determined for the limit $t \to \infty$ using standard complex methods, given some appropriate constraints on the analyticity of the topography Fourier transform $\hat{T}(\omega)$. The constraints necessary for the integral to converge are essentially equivalent to requiring the forcing to decay at large times. Causality is satisfied by making the transformation $\omega \to \omega + i\varepsilon$ and then allowing $\varepsilon \to 0$ (e.g., Lighthill 1967). This has the effect of moving the singularities in the integrals into the lower half of the complex plane and is consistent with $\eta_R \to 0$ as $t \to -\infty$. The following expression for the long-time behavior of the vortex edge is then obtained:

$$\eta_R \sim \frac{kfL_R^2 hJ_k(lR)}{RH} \left\{ \frac{\pi}{L_R^2 l^2} \hat{T}(\omega_0) i \exp(ik\phi - i\omega_0 t) + 4 \int_0^{\infty} \hat{T}[\omega_s(m)] \hat{\psi}_T(m) \cos[mz + \gamma(m)] i \exp\left[\frac{z}{2H} + ik\phi - i\omega_s(m)t\right] dm \right\}, \tag{25}$$

where ω_0 and $\omega_s(m)$ are the barotropic mode frequency and the Charney–Drazin dispersion relation, as defined in section 2.

The orthogonality of the spectrum of vertical modes, including with respect to the barotropic mode, may be exploited in order to evaluate an expression for the total wave activity $\mathcal{A}$ in the limit $t \to \infty$, by integrating $A \approx \int_0^{2\pi} \Delta\rho R^2 \eta_R^2/2 \, d\phi$ with respect to z. Recalling that the time-integrated lower boundary EP flux $\mathcal{F}$ and the change in the azimuthal mean angular momentum $\mathcal{M}$ are related,

$$\mathcal{F} = \mathcal{A}|_{t=\infty} = -(\mathcal{M}|_{t=\infty} - \mathcal{M}|_{t=-\infty}),$$

an expression for $\mathcal{F}$ is obtained:

$$\mathcal{F} = \mathcal{F}_0 + \mathcal{F}_{\mathrm{CDS}} = \pi^2 k^2 f^2 h^2 J_k(lR)^2 \Delta \left\{ \frac{\pi |\hat{T}(\omega_0)|^2}{2Hl^4} + \frac{4L_R^4}{H^2} \int_0^{\infty} |\hat{T}[\omega_s(m)]|^2 \hat{\psi}_T(m)^2 \, dm \right\}. \tag{26}$$

The flux in (26) is divided into two parts, the first term in the brackets being the part of the EP flux $\mathcal{F}_0$ that contributes to the excitation of the barotropic mode, and the second part $\mathcal{F}_{\mathrm{CDS}}$ contributes to the excitation of waves within the Charney–Drazin upward-propagating spectrum. The expression (26) is valid for general time-dependent topographic forcing of the form (20).

The fact that $F(0, t) = 0$ whenever $h_T(t) = \psi_{Tz}(0, t) = 0$ means that (26) can be exploited to obtain explicit expressions for the time evolution of the lower boundary EP flux. The integrated flux up to a given fixed time $\int_0^t F(0, \tilde{t}) \, d\tilde{t}$ may be evaluated by replacing $|\hat{T}(\omega)|^2$ in (26) by $|\hat{T}_*(\omega, t)|^2$, where

$$\hat{T}_*(\omega, t) = \frac{1}{2\pi} \int_{-\infty}^{t} T(\tilde{t}) \exp(i\omega\tilde{t}) \, d\tilde{t}. \tag{27}$$

From this, an expression for the lower boundary flux $F(0, t)$ is obtained by straightforward differentiation with respect to t. Explicitly, $F(0, t)$ can be evaluated by replacing $|\hat{T}(\omega)|^2$ in (26) by $T(t)[\hat{T}_*(\omega, t) \exp(-i\omega t) + \text{c. c.}]$.

One important point about these expressions for the lower boundary EP flux is that they are nonlinear in the lower boundary forcing. In particular, $F(0, t)$ is not proportional to $T(t)$—the function that describes the time evolution of the forcing—but also involves an expression including $\hat{T}_*(\omega, t)$, which is a functional depending on the history of $T(t)$ up to that particular time. At least in this linear limit, therefore, the flux $F(0, t)$ cannot be said to be determined instantaneously by the forcing. It, rather, depends on an interaction between the forcing and the state of the vortex itself, with the state of the vortex being determined by the time history of the forcing.

b. *Impulsive forcing*

Insight into the linear response of the vortex may be obtained by considering an impulsive topographic forcing with the form

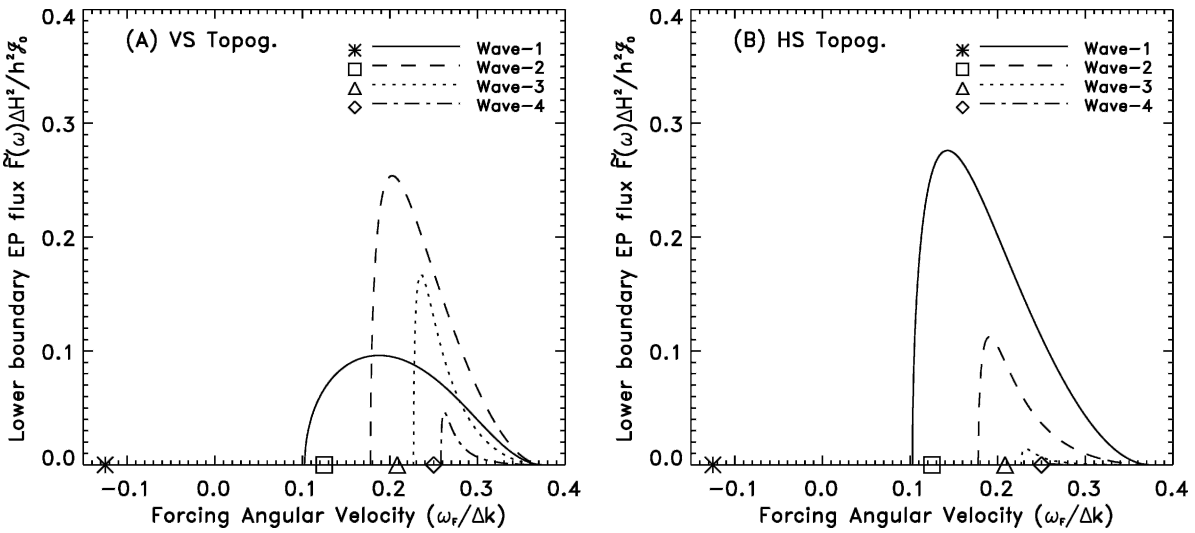


FIG. 2. Nondimensional contribution to lower boundary EP flux $\tilde{F}(\omega)H/h^2f^2R^2$ due to lower boundary topographic forcing at different angular frequencies $\omega_r/\Delta k$ for azimuthal wavenumbers $k = 1, 2, 3$, and 4. Points on the frequency axis show the resonant frequencies of the barotropic mode for $k = 1, 2, 3$, and 4; these frequencies give a singular response: (a) Vortex-scale (VS) ($lR = 2.427$) and (b) hemispheric-scale (HS) ($lR = 1.162$) forcing.

$$h_T = \frac{h}{2}\,\delta(t)J_k(lr)\exp\{ik\phi\}.$$

In this case, $\hat{T}(\omega) \equiv 1$; that is, the forcing is not biased toward any particular frequency. The vortex response to impulse forcing therefore provides a useful measure of the typical relative magnitudes, in terms of pseudomomentum, of the barotropic and Charney–Drazin spectrum responses, as well as a way of objectively comparing the response to forcing at different frequencies within the Charney–Drazin limits given by (9).

Considering the Charney–Drazin spectrum in detail first, we note that the expression for $\mathcal{F}_{\mathrm{CDS}}$ in Eq. (26) may alternatively be written as an integral in ω,

$$\mathcal{F}_{\mathrm{CDS}} = \int_{\omega_s^-}^{\omega_s^+} \frac{|\hat{T}(\omega)|^2\tilde{F}(\omega)}{c_g[m_s(\omega)]}\,d\omega, \quad \text{where}$$

$$\tilde{F}(\omega) = \frac{4\pi^2 J_k(lR)^2k^2f^2L_R^4\Delta h^2}{H^2}\,\hat{\psi}_T[m_s(\omega)]^2.$$

Here $m_s(\omega)$ is the inverse of the dispersion relation (8) and $c_g(m)$ is the vertical group velocity (see the appendix). The function $\tilde{F}(\omega)$ is a measure of the contribution from each frequency ω to both the final wave activity and the change in azimuthal mean angular momentum of the vortex. Expressing $\mathcal{F}_{\mathrm{CDS}}$ in this form also makes it explicit that transient forcing at frequencies outside the Charney–Drazin range $[\omega_s^-, \omega_s^+]$ and away from the barotropic frequency ω_0 does not lead to excitation of Rossby wave activity on the vortex.

Figure 2 shows $\tilde{F}(\omega)$ for zonal wavenumbers $k = 1, 2, 3$, and 4 for two fixed values of the forcing scale lR, which correspond to the "vortex scale" and "hemispheric scale" forcing (see also section 4). The vortex forcing (with $lR = 2.427$) is so called because the peak amplitude of the topographic forcing for $k = 2$ occurs at a distance (3398 km) comparable to the vortex radius (2700 km). By contrast, the hemispheric forcing ($lR = 1.162$) has peak amplitude at 7214 km (again for $k = 2$). Vortex parameters are again as for WD99. Also plotted on the frequency axis are the barotropic frequencies for $k = 1, 2, 3$, and 4 (symbols). In the case of the vortex scale forcing, and for an impulsive forcing with fixed magnitude, it is notable that the largest amplitude vortex excitation is for azimuthal wavenumber $k = 2$, and for forcing frequency much closer to the lower limit of the Charney–Drazin range than the upper limit. The vortex response, in terms of angular pseudomomentum, falls away rapidly with increasing wavenumber. For the hemispheric scale forcing, azimuthal wavenumber 1 provides the dominant response, and is several times larger than the response due to the vortex-scale forcing. This is despite the fact that the height of the hemispheric-scale forcing is substantially lower at the vortex edge (although the peak topographic amplitude is, of course, the same in each case). Rossby waves with low azimuthal wavenumbers are thus preferentially excited by forcing with radial scale comparable to or greater than the vortex radius.

An important question concerns the relative magnitudes of the barotropic and Charney–Drazin responses.

Defining

$$\mathcal{R} = \frac{\mathcal{I}_0}{\mathcal{I}_{\mathrm{CDS}}}$$

to be the ratio of angular pseudomomentum imparted to the barotropic mode to that imparted to the Charney–Drazin spectrum, it can be shown by direct integration of $\mathcal{I}_{\mathrm{CDS}}$ in (26) that in the case of impulsive forcing

$$\mathcal{R} = \frac{(4L_R^2 l^2 + 1)^{1/2}[(4L_R^2 l^2 + 1)^{1/2} + 1]^2}{4L_R^4 l^4}. \qquad (28)$$

Some important aspects of this expression include that

- the ratio $\mathcal{R}$ depends only on the relative scale of the lower boundary forcing and the Rossby radius, as quantified by the nondimensional number $L_R l$. In particular, $\mathcal{R}$ is independent of the azimuthal wavenumber k, as well as any details of the vortex itself.
- the values of $\mathcal{R}$ for the vortex- and hemispheric-scale forcing are approximately 9.35 (for $L_R l = 0.809$) and 76.00 (for $L_R l = 0.387$), respectively, that is, significantly greater than unity in both cases. This means that for any large-scale transient forcing, with a wide distribution of frequencies, one would expect the dominant contribution to the lower boundary EP flux to be associated with the barotropic mode.

c. *Forcing at fixed frequency: The switch-on problem*

A contrasting limit to the limit of impulsive forcing described above is that of forcing at a fixed frequency switched on at $t = 0$. Tung and Lindzen (1979a) restricted attention to forcing at the stationary wave frequency ($\omega_F = 0$) in their study of the β channel. Results analogous to Tung and Lindzen's may be derived for the three-dimensional vortex using Eq. (26) in conjunction with (27), with $\hat{T}_*$ calculated for a lower boundary forcing that is switched on at $t = 0$ and evolves periodically in time with a fixed frequency ω_F thereafter.

For the case of a fixed frequency, constant amplitude forcing switched on at $t = 0$,

$$|\hat{T}_*(\omega, t)|^2 = \left| \frac{1}{2\pi} \int_0^t \exp[i(\omega - \omega_F)t]\, d\tilde{t} \right|^2 = \frac{1 - \cos(\omega - \omega_F)t}{2\pi^2(\omega - \omega_F)^2}.$$

This result may be used in conjunction with Eq. (27) to yield an expression for the evolution of the integrated EP flux up to time t. Of particular interest is the case where the barotropic mode is forced resonantly; that is, $\omega_F = \omega_0$. In this case, the time-integrated EP flux associated with the barotropic mode is

$$\int^t F_0(0, \tilde{t})\, d\tilde{t} = \pi k^2 f^2 h^2 J_k(lR)^2 \Delta \frac{t^2}{8Hl^4}. \qquad (29)$$

This result shows that the total angular pseudomomentum associated with the barotropic mode grows quadratically in time; hence, the wave amplitude of the barotropic mode grows linearly in time when it is forced resonantly. This is consistent with Eq. (60) of Tung and Lindzen (1979a), which states that, in the case of the β channel, the northward heat flux associated with the barotropic mode grows linearly in time when it is forced resonantly. The angular pseudomomentum of the vertically propagating waves, with frequencies in the Charney–Drazin spectrum, may easily be bounded in this case using the result that

$$|\hat{T}_*(\omega, t)|^2 \le \frac{1}{\pi^2(\omega - \omega_F)^2},$$

which upon using it is straightforward to show that

$$\frac{\int_0^t F_{\mathrm{CDS}}(0, \tilde{t})\, d\tilde{t}}{\int_0^t F_0(0, \tilde{t})\, d\tilde{t}} \le \frac{1}{t^2(\omega_s^- - \omega_0)^2 \mathcal{R}}, \qquad (30)$$

where $\mathcal{R}$ is as given in (28). This result shows that, for typical stratospheric vortex parameters, at times $t \gg 1/(\omega_s^- - \omega_0)$ ($\sim$1.5 days for the parameters used above) the angular pseudomomentum is dominated by the contribution of the barotropic mode rather than the propagating waves.

Naturally the linear phase where wave amplitudes grow linearly in time cannot be sustained, and a full analysis of the switch-on problem at near-resonant frequencies requires nonlinearity to be considered, following Plumb (1981). Below, we restrict attention to finite transient pulses of topographic forcing in order to examine linear theory and test its relevance to the full nonlinear vortex model.

d. *Comparison with linear numerical model results*

To both verify the analytic results above, test the extent to which the results apply to a barotropic vortex of finite height (as opposed to a semi-infinite vortex), and ensure that the vertical resolution used in the nonlinear model below adequately resolves a continuous vortex, we employ a linear numerical model similar to that described in WD99 (see section 4a of WD99 for a detailed description of the numerical algorithm). The formulation of the linear model necessarily differs from

the analytic formulation described above as the domain is made finite with a vertical lid set at $D = 12H$ (73.68 km), and the vortex is discretized in the vertical with N levels. The boundary condition on the upper boundary is $\psi_z = 0$ on $z = D$. The linear model may be used to obtain a discrete set of N vertical structure eigenmodes, together with N eigenfrequencies. In the case of the barotropic vortex the lowest eigenfrequency corresponds to the frequency of the barotropic mode ω_0 above, and the remaining $N - 1$ eigenfrequencies correspond to the Charney–Drazin frequencies $\omega_s(m_n)$ for the vertical wavenumbers $m_j = j\pi/D$, $1 \le j \le N - 1$. The expression for $\mathcal{F}_{\mathrm{CDS}}$ must, therefore, be replaced by a summation

$$\mathcal{F}_{\mathrm{CDS}} = \sum_{j=1}^{\infty} |\hat{T}(\omega_s(m_j))|^2 \hat{\psi}_T(m_j)^2 \delta m, \quad \delta m = \frac{\pi}{D}.$$

Further, this expression will only be accurate for a vortex of finite height, provided D is large enough that $\exp\{-(\alpha - 1/2H)D\} \ll 1$, which holds to a good approximation for topographic forcing on the planetary scale or below, when $D = 72$ km, as here.

To test the analytic prediction (26) the linear numerical model is forced with a pulselike forcing with amplitude describing a Gaussian evolution in time and with a constant angular speed ω_F/k:

$$h_T(r, \phi, t) = \frac{h}{2} J_k(lr) \exp\left\{ -\left(\frac{t}{\tau}\right)^2 + ik\phi - i\omega_F t \right\}. \quad (31)$$

The parameter τ is the time scale over which the lower boundary forcing grows and decays. This topographic forcing may be expressed in the form (20) with

$$\hat{T}(\omega) = \frac{\tau}{2\sqrt{\pi}} \exp\left\{ -\frac{\tau^2(\omega - \omega_F)^2}{4} \right\}. \quad (32)$$

The linear model was integrated from $t = -10$ to $+10$ days, using 200 time steps per day and with parameters $k = 2$, $lR = 2.427$ (vortex-scale forcing) and $\tau = 2$ days. Four separate angular speeds for the topographic forcing are considered: $\omega_F/k = 0.025$, 0.125, 0.225, and 0.325Δ (corresponding to azimuthal velocities of 3.9, 19.6, 35.3, and 51.1 m s^{-1} at the vortex edge, respectively). At the resolution $N = 120$, corresponding to that of the nonlinear numerical model described below, we were able to obtain the correct frequency for the first five vertical modes to within an accuracy of 0.5%. This accuracy was significantly improved when a higher vertical resolution was used.

Figure 3 shows a comparison between the actual and predicted nondimensional pseudomomentum spectra of the upward-propagating vertical modes, with the discrete points showing the amplitude of the results from the linear model, and the curves the predictions from (26). Based on the dispersion relation (8), it might be expected that forcing concentrated at frequencies near the upper limit of the Charney–Drazin range $\omega_s^+ = 0.375\Delta k$ will primarily lead to excitation of short waves, and forcing around $\omega_s^- \approx 0.178\Delta k$ will primarily lead to excitation of long waves. In addition, the maximum response in terms of pseudomomentum should arise from forcing at frequencies corresponding to the peak of the $k = 2$ curve in Fig. 2a, at $\omega_F \approx 0.225\Delta k$. This is, indeed, what is found, as the $\omega_F/k = 0.325\Delta$ experiment (diamonds) excites a spectrum dominated by relatively short waves ($mH \sim 1$–1.5), the $\omega_F/k = 0.225\Delta$ experiment (squares) elicits the largest amplitude response, with $mH \sim 0.5$–1, and the $\omega_F/k = 0.125\Delta$ case (triangles) excites longer waves with lower amplitudes. In the final experiment (stars) $\omega_F = 0.025\Delta k \ll \omega_s^-$, and there is very little excitation of the upward propagating spectrum.

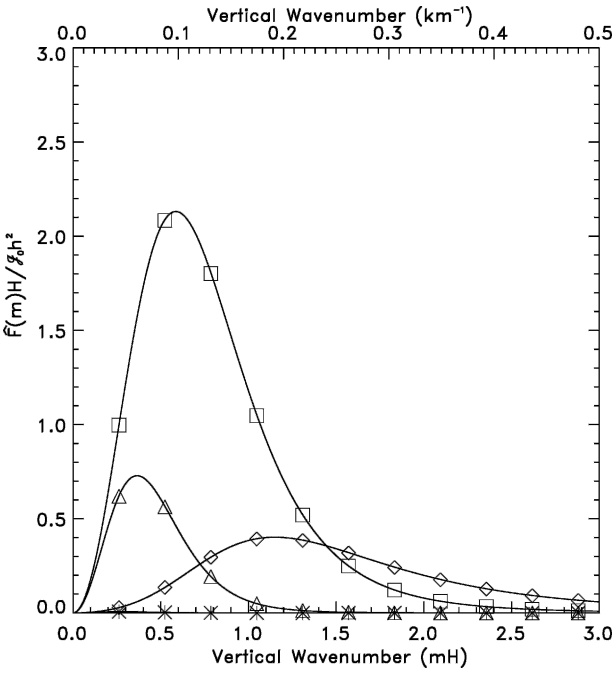


FIG. 3. Nondimensional angular pseudomomentum spectra $\tilde{F}[\omega_s(m)]H/J_0h^2$ as a function of vertical wavenumber m as predicted by linear theory for a transient Gaussian pulse of the form (31) for different peak forcing frequencies $\omega_F/k = 0.025\Delta$ (stars), 0.125Δ (triangles), 0.225Δ (squares), and 0.325Δ (diamonds). The symbols correspond to the result of the equivalent linear model calculations with a lid at $D = 12H = 73.68$ km; hence, only a discretized set of wavenumbers is available.

The angular pseudomomentum of the barotropic mode at the end of each calculation agreed with the prediction $\mathcal{F}_0$ to within 0.5% accuracy in all cases examined. The ratio of the final barotropic pseudomo-

mentum to the total pseudomomentum ($\mathcal{I}_0/\mathcal{I}_{\rm CDS}$) ranged from 0.0102 ($\omega_F/k = 0.325\Delta$), 1.571 ($\omega_F/k = 0.225\Delta$), 57.14 ($\omega_F/k = 0.125\Delta$), to 1110.1 ($\omega_F/k = 0.025\Delta$) for the four linear experiments.

4. Nonlinear model results

a. Nonlinear numerical model formulation

Below, the results from a series of integrations of a nonlinear numerical model of the *f*-plane vortex are described. The model employs the contour advective semi-Lagrangian (CASL) algorithm (Dritschel and Ambaum 1997) to solve Eqs. (3) and (1) in a circular cylindrical domain together with the lower boundary condition (2), an upper boundary condition $\psi_z = 0$ on $z = D$, and the no normal-flow condition $\psi_\phi = 0$ on the outer wall. Full details are given in Macaskill et al. (2003). The model is used in a similar formulation to that described by Scott et al. (2004) except that we set the vertical domain size $D = 12H$ (73.68 km), use 120 model levels, and choose a radial domain size of $30L_R$ ($=10R$). The vortex is discretized at each level by representing the vortex edge by a single circular contour made up of discrete points or nodes. The initial node density is 51 nodes per contour, except where stated otherwise below. The streamfunction is obtained from the PV by inverting Eq. (1) on a grid of 96 radial and 192 azimuthal points, although a grid with 4 times this density is used to store the PV during the integration.

The presence of an outer boundary at the cylinder edge also imposes a constraint on the horizontal scale of the lower boundary pulse described in section 3 below. In order for the analysis of section 3 to apply in the case of the cylinder we require $J_k(10lR) = 0$; that is, the height of the topographic lower boundary pulse must vanish at the outer boundary. This imposes constraints on the possible values of l. For an azimuthal wavenumber $k = 2$, we choose $lR = 1.162$ (referred to as hemispheric-scale forcing below) and $lR = 2.427$ (vortex-scale forcing). These choices ensure that the third and seventh zeroes of J_2 respectively coincide with the outer boundary.

b. Background and experimental setup

Linear theory predicts that the amplitude of an upward-propagating wave grows exponentially with height $\sim \exp(z/2H)$. Considering the case of an initially localized pulse of wave activity on a semi-infinite vortex, no matter how small the initial pulse, at high altitudes the waves will grow to have an amplitude comparable to the vortex radius R, and nonlinearity will become important. Linear theory therefore clearly predicts its own breakdown, and there is an obvious question concerning its relevance to the nonlinear situation. However, it might be expected that in response to a pulse of transient forcing, most of the wave activity $A(z, t)$ will be initially excited near the lower boundary. At the lower boundary the relatively high density means that the vortex is less easily deformed, and linear theory may therefore often be accurate. Subsequently, the wave activity may be conservatively redistributed in the vertical according to (13), but in the absence of dissipation the total wave activity $\mathcal{A}$ will not be changed from its initial value—even in the event of strong nonlinearity and wave breaking. Following this argument, the expression (26) for $\mathcal{F} = \mathcal{A}|_{t=\infty}$ might be expected to be accurate even at moderate forcing amplitudes—a hypothesis that is tested in detail below.

In the main set of numerical experiments using the nonlinear model, we use a barotropic vortex with parameter settings as in WD99. The vortex is forced in each case with a transient topographic lower boundary forcing of the form (31). The model is run from $t = -5\tau$ to $t = 5\tau$ in order that the forcing amplitude is insignificant both at the beginning and also by the end of the experiment, so the late model state is a good approximation for the state at $t = \infty$. The parameters h, ω_F, l, and τ are varied independently between model runs with the topographic forcing amplitude $h = 0.05H$, $0.1H$, $0.2H$, or $0.4H$, and the forcing frequency being varied between $\omega_F = -0.125\Delta k$ and $0.375\Delta k$, with intervals of $0.025\Delta k$. The radial wavenumber is set to $l = 1.162R^{-1}$ hemispheric-scale forcing or $2.427R^{-1}$ vortex-scale forcing, and the pulse time scale is set to $\tau = 2$ or 4 days. For every experiment reported here the azimuthal wavenumber $k = 2$. As will be discussed below, almost all of the experiments result in nonlinear wave breaking and strongly nonlinear development of the vortex at some level (the exceptions being those experiments with $h = 0.05H$ and $\omega_F < -0.075\Delta k$).

c. Vortex response in terms of wave activity

Results from the $\tau = 2$ days vortex-scale (VS) ($lR = 2.427$) forcing set of experiments are summarized in Fig. 4. The final nondimensional angular pseudomomentum $\mathcal{A}|_{t=\infty}H^2/\mathcal{I}_0h^2$ ($\mathcal{A}|_{t=\infty} = \mathcal{F}$) is plotted against forcing frequency $\omega_F/\Delta k$ for the four separate forcing amplitudes. Because we have scaled with h^2, the results can be compared directly with the linear theory predictions of (26) (solid curve). Before considering the nonlinear results it is interesting to compare the full linear response $\mathcal{F}$ (solid curve) with the linear barotropic response $\mathcal{F}_0$ (dotted curve). It is clear that the barotropic response is, by far, the dominant component of $\mathcal{F}$ at

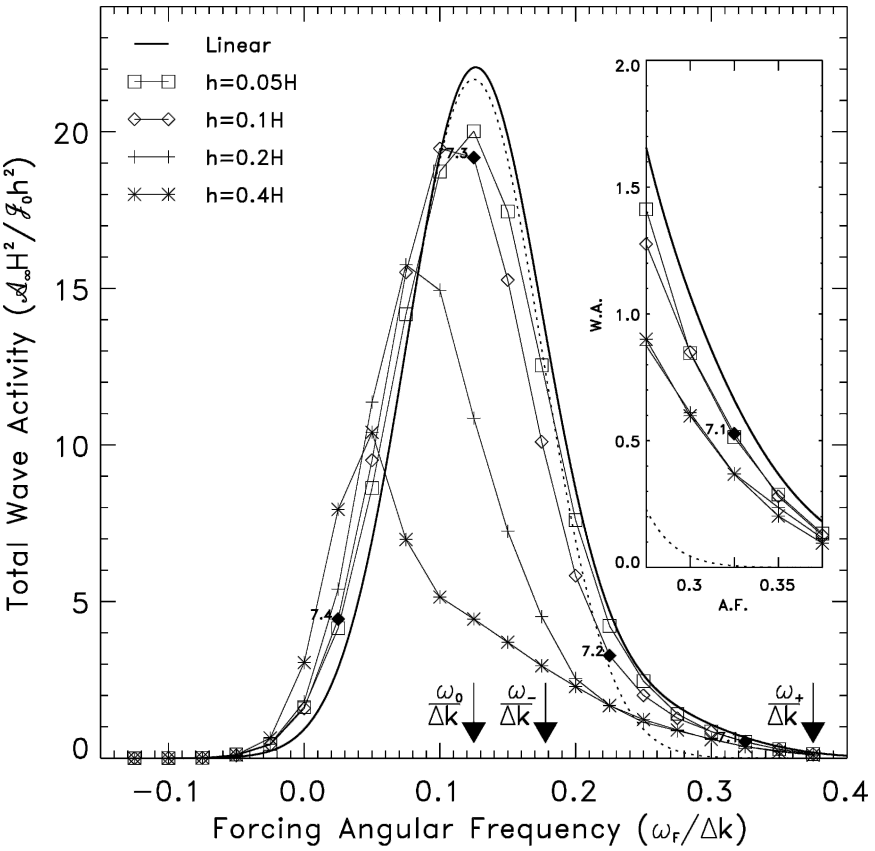


FIG. 4. The linear prediction (solid line) and the nonlinear numerical model results (points) for nondimensional angular pseudomomentum, or equivalently time-integrated lower boundary EP flux $\mathcal{I}H^2/h^2\, J_0$, versus peak forcing angular frequency $\omega_F/\Delta k$ for a Gaussian pulse of the form (31) with $\tau = 2$ days. Different symbols correspond to different topographic forcing amplitudes $h = 0.05$, 0.1, 0.2, and 0.4. The azimuthal wavenumber $k = 2$ and the forcing has the VS (corresponding to $lR = 2.427$). The inset shows a magnification of the right-hand side of the graph, showing the behavior when the peak forcing frequency is in the Charney–Drazin range. The dotted line shows the linear prediction for the barotropic mode only. Note that the experiments discussed in connection with Fig. 7 are marked as solid diamonds labeled 7.1, 7.2, 7.3, and 7.4.

most forcing frequencies, and the peak linear response occurs at the barotropic mode frequency ω_0. The predicted Charney–Drazin spectrum response $\mathcal{F}_{\mathrm{CDS}}$, by contrast, dominates only for frequencies close to ω_s^+ (see inset) and much less angular pseudomomentum is generated by forcing in this frequency range, as might be anticipated by the large value of $\mathcal{R} = 9.35$ associated with the structure of the lower boundary forcing [Eq. (28)].

Comparing the linear prediction with the nonlinear predictions for increasing forcing amplitudes for the $h = 0.05H$ set, the linear prediction is fairly accurate except for around a 10% shortfall in forcing amplitude and a suggestion of a slight frequency shift in the peak response toward lower frequencies. As the forcing amplitude is increased, the frequency shift in the peak response also increases, as does the shortfall in the predicted amplitude.

Figure 5 shows the $\tau = 2$ days hemispheric-scale (HS) ($lR = 1.162$) forcing set of experiments. Comparing the linear predictions, the main difference with the VS set is that the barotropic response dominates to an even greater extent, with the two curves almost indistinguishable except for frequencies near ω_s^+. This is consistent with the higher value of $\mathcal{R} = 76$ from Eq. (28). The amplitude of the maximum response, for the same forcing amplitude, is also greater than that for the VS case. In the nonlinear experiments, as the amplitude is increased, an increasing frequency shift of the response toward lower-frequency forcing is apparent, as with the $\tau = 2$ days VS set. However, unlike the VS set, the maximum predicted amplitude of the response remains

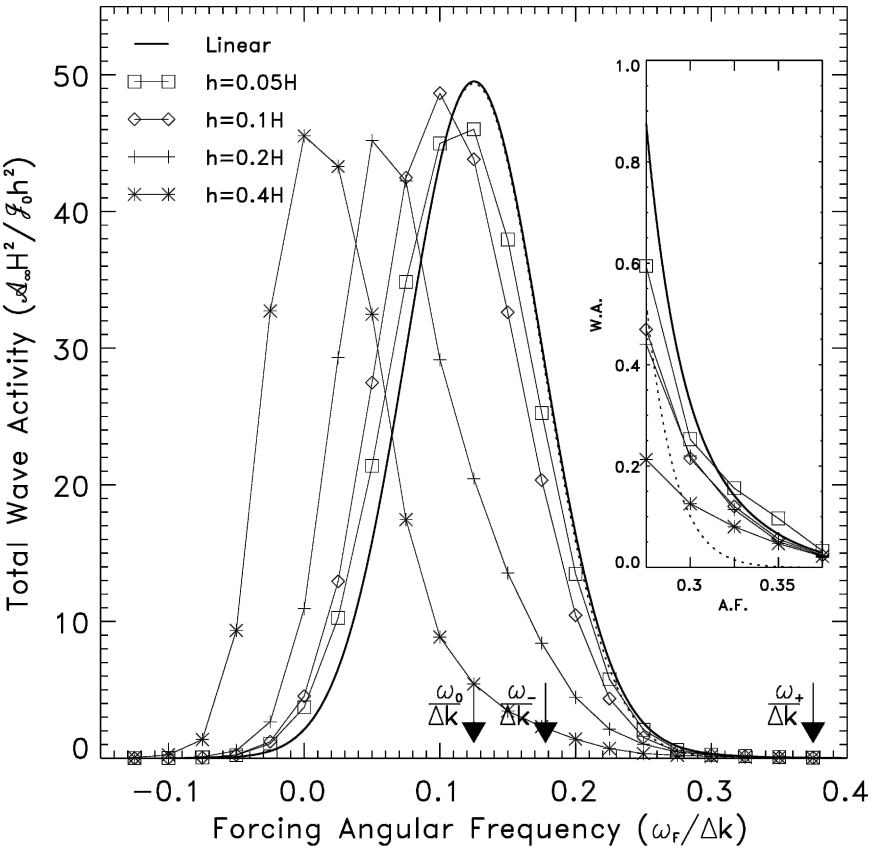
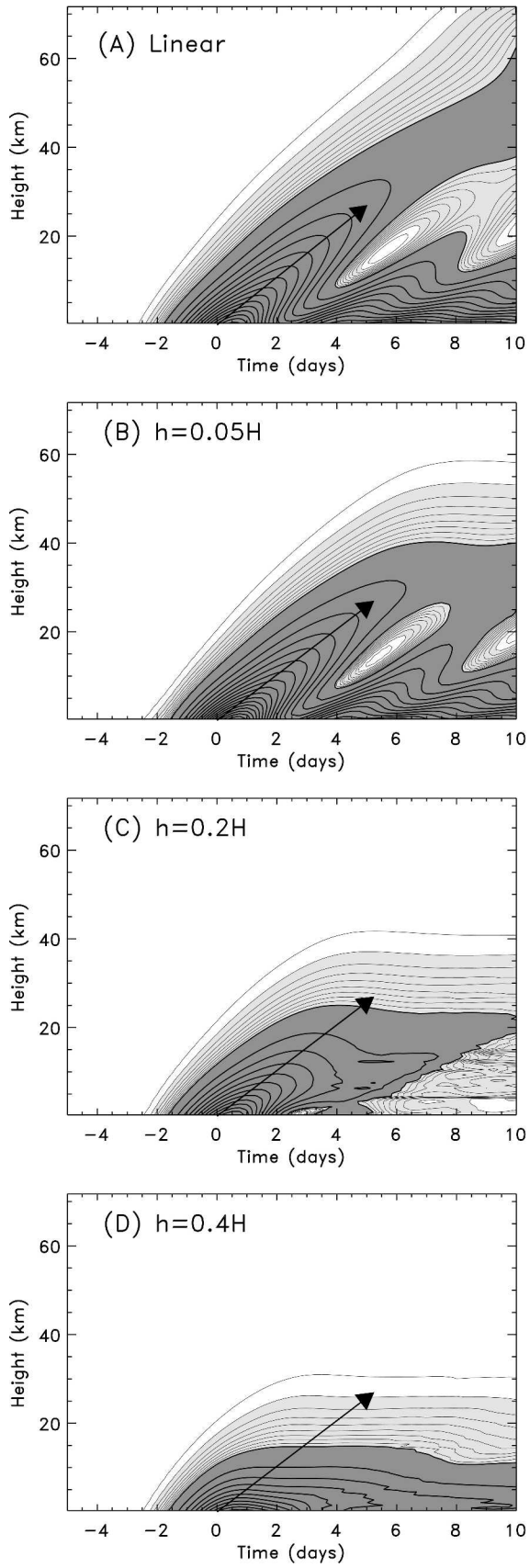


FIG. 5. As in Fig. 4, but for the HS ($lR = 1.162$) forcing.

approximately constant with forcing amplitude. The frequency shift that is apparent in both sets of experiments might be explained by a weakly nonlinear theory (e.g., Plumb 1981), as the mean angular velocity at the vortex edge Ω_e is likely to be reduced by the topographic forcing, thereby reducing the resonant frequency of the barotropic mode. The reduction in amplitude of the response with increasing forcing amplitude in the VS set of experiments might be explained by the structure of the Bessel function topographic forcing. With the VS forcing, there is a valley located around $2.12R$, so once the vortex becomes significantly distorted its outer edges experience weaker effective forcing, whereas with the HS forcing there is no valley near the vortex.

The influence of forcing amplitude on the response can be seen clearly if the evolution of the wave activity $A(z, t)$ is examined in height and time. The group velocity condition (16) suggests that wave activity associated with upward-propagating waves should radiate upward with vertical velocity c_g. The wave activity associated with the barotropic mode, by contrast, decays exponentially with height according to $\exp(-z/H)$, which is easily seen by inserting the barotropic mode structure in Eq. (14). In Fig. 6 we compare results from the $\omega_F = 0.225\Delta k$, $\tau = 2$ days VS set of experiments with a calculation from the linear model with the same forcing. At the forcing frequency in these experiments ($\omega_F = 0.225\Delta k$), both vertically propagating waves and the barotropic mode are excited. Examining the linear response in Fig. 6a, a strong upward-propagating pulse with the expected group velocity (see Fig. 1) of around 5.5 km day^{-1} is observed, and later in the integration it is clear there is also a vertically decaying component to the response. It is also important to note that the topographic velocity field v_T excites both upward- and downward-propagating waves [consider how the vertical structure of v_T decomposes into terms $\sim \exp(z/2H + imz)$ and terms $\sim \exp(z/2H - imz)$] as well as a barotropic response. This explains the interference patterns in $A(z, t)$ that are apparent once the topographic forcing has died away.

The low-amplitude nonlinear experiments (Fig. 6b and also Fig. 7b2) behave in a very similar fashion to the linear experiments except that wave activity cannot propagate far beyond heights where $A(z, t)$ is comparable to $\pi\rho\Delta R^4/4$. At this height, wave amplitudes must be comparable to the vortex radius, indicating that non-



linear wave breaking and the eventual breakdown of the vortex must occur. In those experiments with stronger forcing Figs. 6c and 6d the evidence for upward propagation is weaker, with absorption of wave activity occurring at much lower heights, and with the response in Fig. 6d in particular appearing to be almost purely barotropic.

d. Temporal evolution of the response

It is also of interest to examine in detail the response of the vortex to forcing at different frequencies ω_F while the forcing amplitude and other parameters are kept constant. Figure 7, therefore, shows results from the $h = 0.1H$, $\tau = 2$ days VS forcing set of experiments with each set of panels corresponding to a different forcing frequency ω_F. The left-hand set of panels, (A1)–(C1), shows results from the $\omega_F = 0.325\Delta k$ experiment, (A2)–(C2) the $\omega_F = 0.225\Delta k$ experiment, (A3)–(C3) the $\omega_F = 0.125\Delta k$ experiment, and (A4)–(C4) the $\omega_F = 0.025\Delta k$ experiment. The upper panels in Fig. 7 show a snapshot of the vortex 8 days after the peak amplitude of the lower boundary forcing. The middle panels show the evolution of $A(z, t)$, as in Fig. 6, and the lower panels show the change in the azimuthal mean winds from the beginning to the end of the integration. In the first experiment to be considered, with $\omega_F = 0.325\Delta k$, the forcing frequency is near the upper end of the Charney–Drazin range ($\omega_s^+ = 0.375\Delta k$). As a result, linear theory suggests that a spectrum of relatively short wavelength upward-propagating waves are excited (see curve with diamonds in Fig. 3). These short waves can clearly be observed on the vortex, forming a left-handed corkscrew pattern on the vortex edge, and eventually reaching amplitudes large enough to cause wave breaking at the top of the vortex. From Fig. 1, short waves ($mH \approx 1.25$, where the "diamonds" curve in Fig. 3 is maximum) would be expected to have upward group velocity of around 4 km day^{-1}, and this appears to be a slight overestimate of the velocity associated with the upward-propagating wave activity in (B1) where the arrow denotes 3 km day^{-1}. At the end of the integration, the azimuthal mean wind has been

←

FIG. 6. Evolution of nondimensional wave activity $A(z, t)H/h^2\mathcal{J}_0$ in time t and height z for the set of experiments with $\omega = 0.225\Delta k$, $\tau = 2$ days, and VS ($lR = 2.427$) forcing: (a) linear model; and nonlinear model with (b) $h = 0.05H$, (c) $h = 0.2H$, and (d) $h = 0.4H$. Contour intervals are 0.004 and 0.04 with two different contour intervals being used to allow different regimes to be compared (see also Fig. 7). Note that the intermediate $h = 0.1H$ case can be seen in (B2) in Fig. 7.

significantly decelerated in the vicinity of the vortex edge in the upper half of the vortex only.

Comparing the $\omega_F = 0.225\Delta k$ and $\omega_F = 0.325\Delta k$ experiments, linear theory suggests that in the former case the lower value of forcing frequency ω_F should lead to the excitation of a larger-amplitude upward-propagating spectrum dominated by longer waves (curve with squares in Fig. 3). The snapshot in panel (A2) of Fig. 7 shows that longer waves have indeed been excited and have led to a deeper wave breaking event enveloping the upper part of the vortex. The longer waves can be seen in panel (B2) to have a higher group velocity (arrow denotes 5 km day^{-1}) than the short waves excited in the W325 experiment. Another difference with the $\omega_F = 0.325\Delta k$ experiment is that the base of the vortex is slightly elliptical (the vortex returns to being almost exactly circular by the end of the $\omega_F = 0.325\Delta k$ experiment), indicating that there is some excitation of the barotropic mode. Changes in the azimuthal mean wind, however, are once again largely confined to the upper half of the vortex.

In the $\omega_F = 0.125\Delta k$ experiment, for which the peak forcing frequency equals the frequency of the barotropic mode ($\omega_F = \omega_0$), excitation of the barotropic mode is more evident compared with $\omega_F = 0.225\Delta k$. In Fig. 7 (A3) shows that the vortex is once again enveloped by a deep wave breaking event, consistent with the long linear wave spectrum predicted in Fig. 3 (curve with triangles). The main difference with (A2) is that the base of the vortex is strongly elliptical, owing to stronger excitation of the barotropic mode (the eccentricity at the vortex base is, in fact, 0.829 compared with 0.526). The exponentially decaying signature of the barotropic mode dominates the evolution of $A(z, t)$, although some upward propagation is still apparent at a group velocity of around 4.5 km day^{-1}, approximately the group velocity of the longest waves on the vortex. The barotropic mode is also clearly seen in the azimuthal winds, with a reduction of over 10 m s^{-1} that is apparent at all heights.

Finally, in the $\omega_F = 0.025\Delta k$ experiment, because $\omega_F < \omega_0$, linear theory predicts that there is barely any excitation of the upward-propagating spectrum (curve with stars in Fig. 3) and that the barotropic mode should be excited, but more weakly than in the $\omega_F = 0.125\Delta k$ case. This is consistent with panels (A4)–(C4) of Fig. 7, which show that the vortex is left moderately elliptical (eccentricity 0.669) except at very high (mesospheric) altitudes (not shown) where the upward-propagating waves finally attain sufficient amplitude to cause wave breaking. Barotropic signatures also dominate the $A(z, t)$ and azimuthal wind plots.

e. The barotropic sudden warming

The discussion above of the response of the vortex to forcing applied at different temporal frequencies leaves an important question unanswered: What happens if the vortex is forced resonantly at the frequency of the barotropic mode? One way of investigating this question is by increasing the time scale τ over which the Gaussian topographic forcing grows and decays, thereby forcing the vortex with a narrower band of frequencies. Figure 8 shows results from the $\tau = 4$ days HS forcing set of experiments, and these may be compared with the $\tau = 2$ days HS experiments shown in Fig. 5. As the forcing occurs over a narrower band of frequencies, the range of frequencies that significantly excite the barotropic mode is also much narrower. As in Fig. 5, there is a frequency shift in the maximum response of the vortex with increasing forcing amplitude and it is significantly greater than that observed in the $\tau = 2$ days HS experiments. The maximum amplitude of the response also deviates from that predicted by linear theory (unlike in the $\tau = 2$ days HS case) with maximum angular pseudomomentum recorded in the $h = 0.4H$ experiments around one-third of what might be expected from linear theory. It is perhaps unsurprising, however, that nonlinearity should become more important as the time scale over which the forcing is applied is increased.

The experiments that result in the peak responses for $h = 0.2H$ and $h = 0.4H$ cause the total wave activity $\mathcal{A}$ to exceed the initial vortex angular impulse $\mathcal{J}_0$, and, hence, to be extremely nonlinear. Figure 9 shows four snapshots from the $h = 0.2H$, $\omega = 0.05\Delta k$, $\tau = 4$ days HS experiment at $t = 0$, 4, 10, and 16 days past the time of peak forcing amplitude. The vortex may be observed to become strongly elongated at its base (day 0), to become "pinched" (day 4), and by day 10 to have divided into two distinct vortices at its base, connected only by a thin strip of vorticity. These two vortices remain coherent and, despite some stripping of filamentary material away from each of them, they continue to corotate in a fairly stable manner. We describe this vortex splitting event as a "barotropic sudden warming" since it is caused principally by the resonant excitation of the barotropic mode. Note that it is not necessary for the vortex itself to be barotropic for this type of warming to occur: a barotropic mode will exist whatever the vertical profiles of vorticity jump $\Delta(z)$, vortex radius $R(z)$, or edge velocity $\Omega_e(z)$. In the atmosphere the analogous mode is sometimes described as a Lamb mode (e.g., Salby 1981). In our experiments, barotropic sudden warmings also occur for experiments that are forced by the vortex forcing in the $h = 0.4H$, $\tau = 4$ days

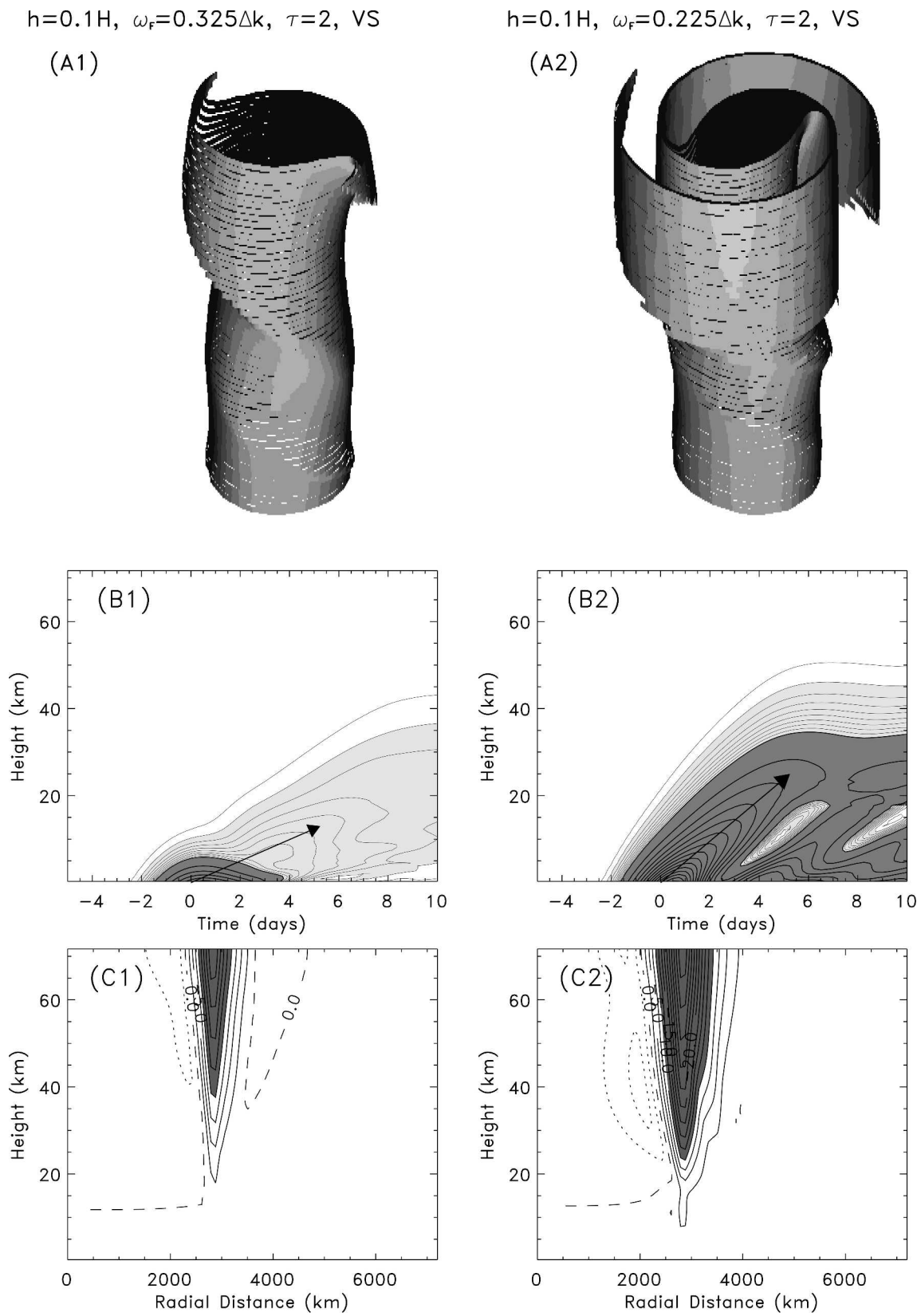


FIG. 7. Illustrating the dynamics of the $h = 0.1H$ $k = 2$ nonlinear experiments with $\omega_F/k = 0.325, 0.225, 0.125$, and 0.025Δ, respectively. The surface plots (A1)–(B1) of the vortex are at $t = +8$ days with the lowest six scale heights ($0 \le z \le 6H = 36.84$ km) shown. Note the similarity between (A1)–(B1) and Fig. 22 of Polvani and Saravanan (2000).

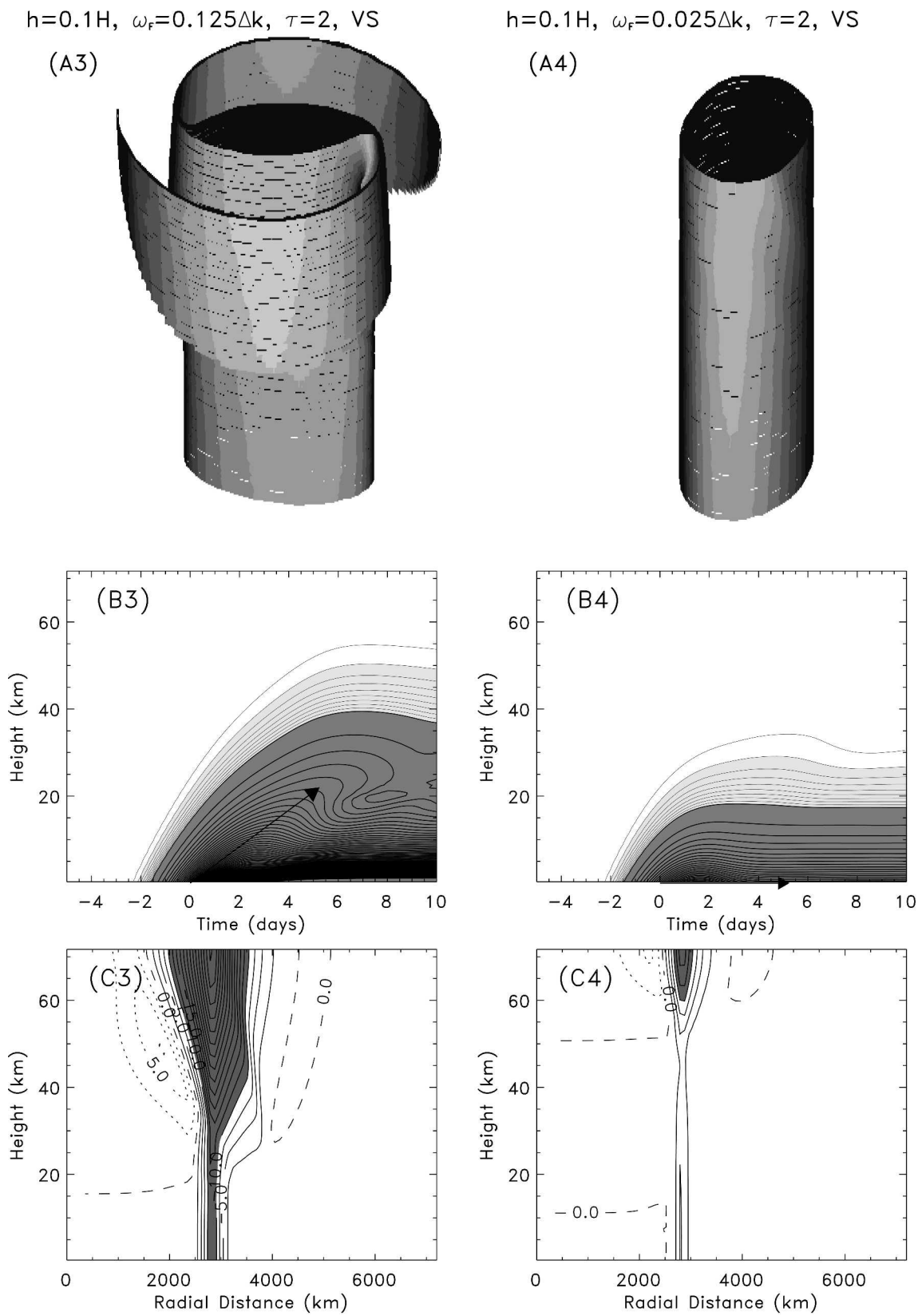


FIG. 7 (*Continued*). The evolution of nondimensional wave activity $A(z, t)H/h^2J_0$, from (B1) to (B4), is as in Fig. 6 and the (c1–c4) Azimuthal mean velocity, (C1) to (C4), is at $t = +10$ days: contour interval 5 m s^{-1}.

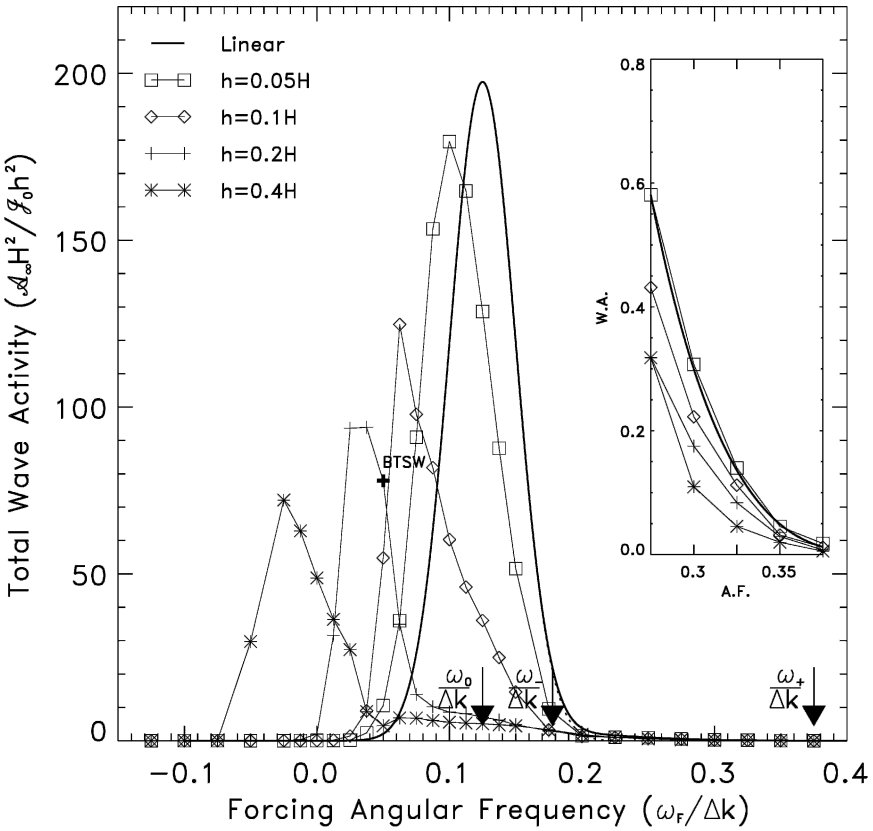


FIG. 8. As in Fig. 5-but for a longer topographic forcing pulse ($\tau = 4$ days). The barotropic sudden warming experiment shown in Fig. 9 is highlighted as the solid cross labeled "BTSW."

VS set, and for some experiments with the shorter forcing time scale of 2 days in the $h = 0.4H$, $\tau = 2$ days HS set. It is worth remarking that a vortex-splitting event analogous to a barotropic sudden warming will occur in an appropriately forced single-layer barotropic model, although for initial vortices with a more general vertical structure than the barotropic vortex discussed here, the analogy is not exact.

5. Discussion and conclusions

In a recent long-time general circulation model experiment, Kushner and Polvani (2005) have shown that a spontaneous stratospheric sudden warming, involving a splitting of the polar vortex, may occur purely as a result of the effects of transient Rossby waves generated by a dynamically active troposphere. In this paper, we have used a relatively simple dynamical model of a polar vortex to investigate theoretically and numerically the conditions that might allow such an event to occur. In doing so, a theoretical connection between the form of the lower boundary forcing, used to model the effect of tropospheric transience, and the response of the vortex has been formulated for the case of a barotropic initial vortex. Despite being valid only in the linear limit the theory gives considerable insight into nonlinear behavior at large forcing amplitudes.

The polar vortex in our simple model supports both upward- and downward-propagating Rossby waves as well as a barotropic mode that has no vertical dependence. When the vortex is forced from below by a transient topographic forcing with a horizontal scale greater than the Rossby radius, the excitation of the barotropic mode dominates the angular pseudomomentum budget. This does not mean, however, that there is no dynamical role for the upward-propagating waves. Because they grow exponentially in amplitude with height, they will break, dissipate, and deposit their angular momentum at high altitude where they will dominate the angular momentum budget. At lower levels the dynamics of the barotropic mode dominate.

The relationship found between the imposed lower boundary forcing and the resulting time-evolving Eliassen–Palm flux at the lower boundary [Eq. (26)] is ex-

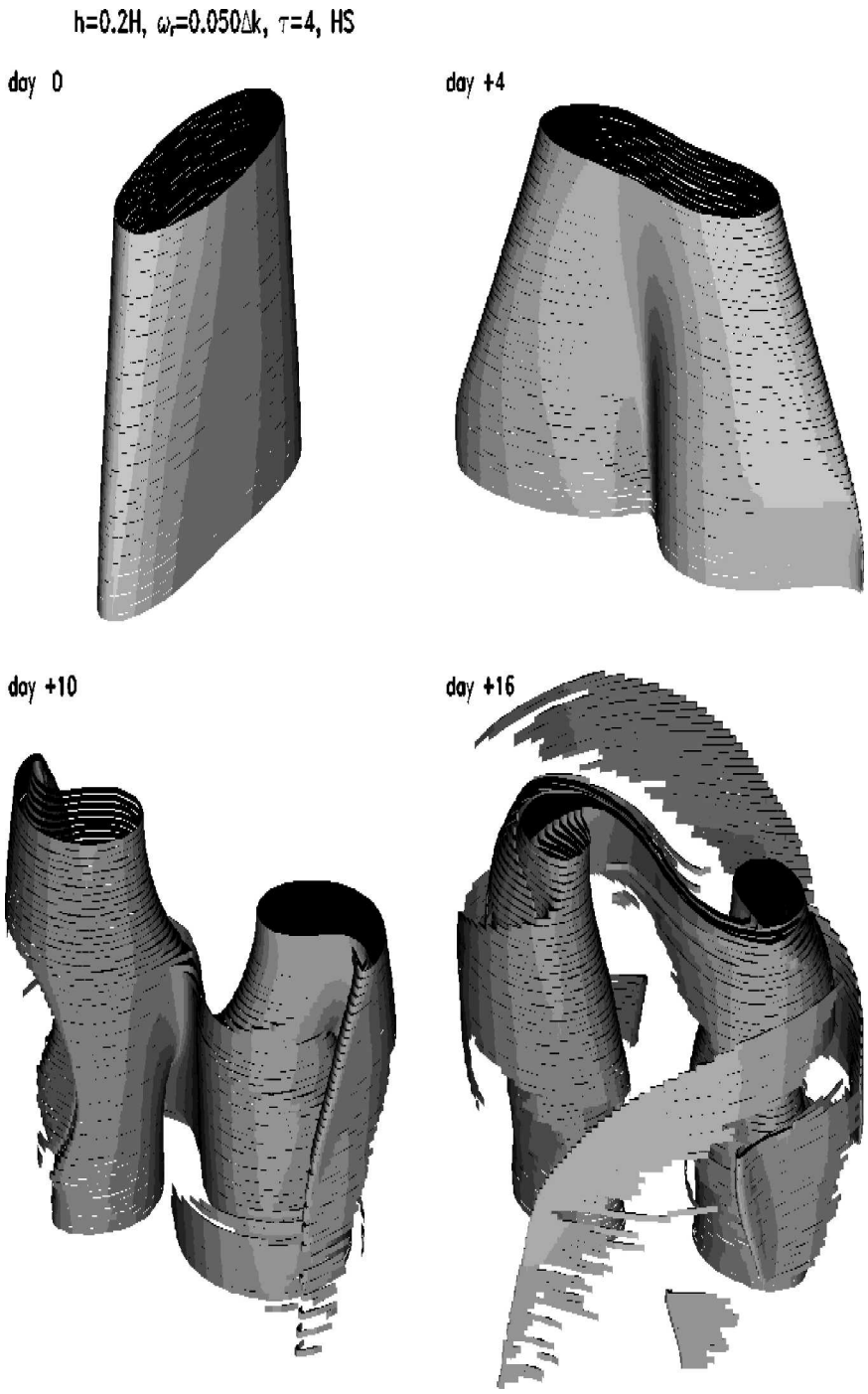


FIG. 9. Surface plots showing the evolution of the vortex during the barotropic sudden warming experiment with $h = 0.2H$, $\tau = 4$ days, HS ($lR = 1.162$) forcing, $\omega_F/\Delta k = 0.05$. The lowest six scale heights ($0 \leq z \leq 6H = 36.84$ km) are shown.

plicitly nonlinear in time. At any instant, therefore, the lower boundary EP flux depends both on the amplitude of the lower boundary forcing and the state of the vortex itself. A consequence for the interpretation of observations is that it is incorrect to describe a peak in Eliassen–Palm flux at the tropopause as being directly "caused" by the dynamics of the troposphere at that instant. Instead, the peak in EP flux occurs because the time history of the tropospheric forcing causes the (elongated or displaced) vortex and the forcing to come

into alignment at that particular time. In the case of wave-2 forcing as here, an "in phase" forcing can act to excite the barotropic mode, for example, causing an upward flux and making the vortex uniformly more elliptical, or the forcing can be out of phase and act to damp the barotropic mode (downward flux) and return the vortex to a more circular profile. O'Neill and Pope (1988) have previously demonstrated temporally nonlinear behavior of the Eliassen–Palm flux in idealized primitive equation experiments, so Eq. (27) may be interpreted as an analytic verification of their result, albeit only in an idealized limit.

This work also suggests other caveats for the interpretation of EP flux. Invariably, in the literature (e.g., Andrews et al. 1987) a group velocity condition [e.g., Eq. (16)] is cited as justification for interpreting the flux as being evidence of vertically, and often latitudinally, propagating Rossby waves. However, excitation of the barotropic mode, which is not vertically propagating, leads to an upward flux that exponentially decays with height according to $F \sim \exp[-(\alpha + 1/2H)z]$. It is worth recalling that many researchers scale EP cross sections with $1/\rho$ precisely in order to enhance features that are decaying exponentially. The wave activity associated with upward-propagating waves should remain independent of height, at least in the absence of dissipation, and therefore should not require any rescaling. A temporally averaged EP cross section, in our view, may often then be dominated by the signature of the excitation of the barotropic mode of the vortex. It is also clear from our simple model that to equate horizontal momentum fluxes with local latitudinal Rossby wave propagation may also be misleading. In our model, no radial (latitudinal) propagation of Rossby waves is possible, as the Rossby waves are trapped on the vortex edge. Despite this, significant horizontal momentum fluxes occur during our nonlinear experiments. In circumstances that are dominated by a single vortex, only the horizontally averaged EP flux [as in Eq. (15)] retains a connection with Rossby wave propagation, and only in the vertical direction.

If the vortex is continuously forced with sufficient amplitude at a frequency near the linear frequency of the barotropic mode (allowing for a weakly nonlinear correction), a "barotropic sudden warming" ensues. The vortex divides into two, with both resulting vortices remaining coherent over several scale heights. To our knowledge, this is the first simulation of a vortex-split sudden warming based on the deliberate resonant excitation of a free mode of the vortex. The three-dimensional evolution of the lower part of the vortices in our simulation is reminiscent of realizations of the February 1979 wave-2 sudden warming in the Northern Hemisphere (Manney et al. 1994, see their Fig. 14). By contrast, three-dimensional visualizations of the September 2002 Southern Hemisphere warming reveal that the two vortices eventually form a double spiral (Manney et al. 2005). We believe that, in order to reproduce a qualitatively similar warming in our simple vortex model, it is necessary to introduce significant vertical structure into the vortex basic state.

The idea that sudden warmings correspond to the resonant excitation of free modes of the atmosphere was first discussed in detail, in the context of a β-channel model, by Tung and Lindzen (1979a,b). In the case where flow in the channel is uniform, they highlighted the possibility of the resonant excitation of the barotropic mode, in a manner analogous to that occurring in our vortex model. Here, we have shown that such excitation can lead to a vortex split, a scenario that is obviously not admitted by the geometry of the β-channel model. Tung and Lindzen's work was extended further by Plumb (1981) who considered the weakly nonlinear evolution of free traveling waves under a near-resonant forcing. Plumb showed that the forcing frequency that resulted in the maximum response changed significantly with increasing nonlinearity. Our numerical results show that a similar shift in the peak forcing frequency occurs in the vortex model.

Both Tung and Lindzen (1979a,b) and Plumb (1981) also consider the resonant excitation of propagating wave free modes formed by turning surfaces in the upper stratosphere, as well as the barotropic mode. The idea, discussed in detail in the review of McIntyre (1982), is that the presence of a turning surface (latitudinal and/or vertical) forms a resonant cavity for the propagating waves. However, the relatively low vertical group velocities associated with propagating vortex edge waves (5 km day^{-1}), and the relatively low pseudomomentum imparted to these modes at realistic forcing amplitudes, lead us to conclude that vortex-splitting warmings caused purely by propagating waves are unlikely to occur, at least for our barotropic vortex. Additionally, the character of the vortex wave breaking is invariably different when the propagating modes are excited, with filamentation occurring rather than vortex splitting. For the case of the barotropic mode, however, it is hoped that a theory similar to Plumb's will lead to a deeper understanding of the changes in frequency and amplitude of the peak barotropic mode response that we observed with increasing forcing amplitude.

Clearly it will also be of interest to determine whether our results are significantly modified in the case of vortices that have realistic vertical structure or for primitive equation vortices in spherical geometry. The dynamics associated with other wavenumbers, although

briefly mentioned in section 3, may also merit further investigation (although the nonlinearity that is associated with wave-1 disturbances in the model used here is somewhat unrealistic). Finally, it is hoped that the ideas developed here can be extended to help identify exactly what aspect of the planetary-scale meteorology of September 2002 contributed to the unique sudden warming event in the Southern Hemisphere.

Acknowledgments. The authors are grateful to Darryn Waugh, Peter Haynes, and Lorenzo Polvani and the two anonymous reviewers for their insightful comments on the work and manuscript.

APPENDIX

Group Velocity Property for the Eliassen–Palm Flux

Here we demonstrate that the Eliassen–Palm flux (15) satisfies the group velocity condition (16) as claimed. Taking a monochromatic wave, as given by (6), and neglecting powers of ε higher than ε^2, we have wave activity

$$A = \frac{\pi}{2}\epsilon^2 R^4 \Delta$$

and group velocity

$$c_g = \frac{\partial \omega_s}{\partial m} = \Delta \frac{f^2}{N^2}\frac{mkR^2}{\mathcal{B}(m)}\left(K_{k+1}I_k - K_k I_{k+1} - \frac{2k}{\mathcal{B}(m)}K_k I_k\right), \tag{A1}$$

with Bessel functions all evaluated at $\mathcal{B}(m)$.

It remains to evaluate F from the definition

$$F = \int_0^\infty \int_0^{2\pi} \frac{f^2}{N^2}\rho \psi_z \psi_\phi r \, d\phi \, dr \tag{A2}$$

and $\psi = \Delta R \eta \tilde{\psi}(r, m, k)$. Evaluating the ϕ integral this may be written

$$F = \pi k m \Delta \frac{f^2}{N^2}[K_k(B(m))^2 \int_0^R I_k(B(m)r/R)^2 r \, dr + I_k(B(m))^2 \int_R^\infty K_k(B(m)r/R)^2 r \, dr]. \tag{A3}$$

Using the appropriate Bessel function identities these become

$$F = \frac{\pi}{2}\frac{f^2}{N^2}\epsilon^2 \Delta^2 R^6 k m [I_K^2 K_{k+1}K_{k-1} - I_{k+1}I_{k-1}K_k^2], \tag{A4}$$

where all of the Bessel functions are evaluated at $\mathcal{B}(m)$. This may be manipulated using standard identities to give

$$F = \frac{\pi}{2}\frac{f^2}{N^2}\epsilon^2 \Delta \frac{R^6 m k}{\mathcal{B}(m)}\left(I_k K_{k+1} - I_{k+1}K_k - 2\frac{k}{\mathcal{B}(m)}I_k K_k\right). \tag{A5}$$

From this is clear that $F = c_g A$ for linear monochromatic waves, as stated. (Note that in obtaining these results we have used Bessel function identities that were taken from http://functions.wolfram.com/.)

REFERENCES

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

Baldwin, M., T. Hirooka, A. O'Neill, and S. Yoden, 2003: Major stratospheric warming in the Southern Hemisphere in 2002: Dynamical aspects of the ozone hole split. *SPARC Newsletter*, No. 20, 24–26.

Charlton, A. J., A. O'Neill, W. A. Lahoz, and P. Berriford, 2005: The splitting of the stratospheric polar vortex in the Southern Hemisphere, September 2002: Dynamical evolution. *J. Atmos. Sci.,* **62,** 509–602.

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

Dritschel, D. G., and R. Saravanan, 1994: Three-dimensional quasi-geostrophic contour dynamics, with an application to stratospheric vortex dynamics. *Quart. J. Roy. Meteor. Soc.,* **120,** 1267–1297.

——, and M. H. P. Ambaum, 1997: A contour-advective semi-Lagrangian numerical algorithm for simulating fine-scale conservative dynamical fields. *Quart. J. Roy. Meteor. Soc.,* **123,** 1097–1130.

Dunkerton, T. J., and M. P. Baldwin, 1991: Quasi-biennial modulation of planetary wave fluxes in the Northern Hemisphere winter. *J. Atmos. Sci.,* **48,** 1043–1061.

Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, 1980: Eliassen–Palm cross-sections for the troposphere. *J. Atmos. Sci.,* **37,** 2600–2616.

Fyfe, J., and X. Wang, 1997: Upper boundary effects in a countour dynamics/surgery model of the polar stratospheric vortex. *Atmos.–Ocean,* **35,** 189–207.

Harnik, N., and R. S. Lindzen, 2002: The evolution of a stratospheric wave packet. *J. Atmos. Sci.,* **59,** 202–217.

——, R. K. Scott, and J. Perlwitz, 2005: Wave reflection and focusing prior to the major stratospheric warming of September 2002. *J. Atmos. Sci.,* **62,** 640–650.

Karoly, D. J., and B. J. Hoskins, 1982: Three-dimensional propagation of planetary waves. *J. Meteor. Soc. Japan,* **60,** 109–122.

Kushner, P. J., and L. M. Polvani, 2005: A very large, spontaneous stratospheric sudden warming in a simple AGCM: A prototype for the Southern Hemisphere warming of 2002? *J. Atmos. Sci.,* **62,** 890–897.

Lighthill, M. J., 1967: On waves generated in dispersive systems by travelling forcing effects, with applications to the dynamics of rotating fluids. *J. Fluid Mech.,* **27,** 725–752.

Macaskill, C., W. E. P. Padden, and D. G. Dritschel, 2003: The CASL algorithm for quasi-geostrophic flow in a cylinder. *J. Comput. Phys.,* **188,** 232–251.

Manney, G. L., J. D. Farrara, and C. R. Mechoso, 1994: Simulations of the February 1979 stratospheric sudden warming: Model comparisons and three-dimensional evolution. *Mon. Wea. Rev.,* **122,** 1115–1140.

——, and Coauthors, 2005: Simulations of dynamics and transport during the September 2002 Antarctic major warming. *J. Atmos. Sci.,* **62,** 690–707.

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

McIntyre, M. E., 1982: How well do we understand the dynamics of stratospheric warmings? *J. Meteor. Soc. Japan,* **60,** 37–65.

Newman, P. A., and E. R. Nash, 2005: The unusual Southern Hemisphere stratosphere winter of 2002. *J. Atmos. Sci.,* **62,** 614–628.

Nielsen-Gammon, J. W., and R. J. Lefevre, 1996: Piecewise tendency diagnosis of dynamical processes governing the development of an upper-tropospheric mobile trough. *J. Atmos. Sci.,* **53,** 3120–3142.

O'Neill, A., and V. D. Pope, 1988: Simulations of linear and nonlinear disturbances in the stratosphere. *Quart. J. Roy. Meteor. Soc.,* **114,** 1063–1110.

Plumb, R. A., 1981: Instability of the distorted polar night vortex: A theory of stratospheric warmings. *J. Atmos. Sci.,* **38,** 2514–2531.

Polvani, L. M., and R. Saravanan, 2000: The three-dimensional structure of breaking Rossby waves in the polar wintertime stratosphere. *J. Atmos. Sci.,* **57,** 3663–3685.

——, and D. W. Waugh, 2004: Upward wave activity flux as precursor to extreme stratospheric events and subsequent anomalous surface weather regimes. *J. Climate,* **17,** 3547–3553.

Salby, M. L., 1981: Rossby normal modes in nonuniform background configurations. Part I: Simple fields. *J. Atmos. Sci.,* **38,** 1803–1825.

Scinocca, J. F., and P. H. Haynes, 1998: Dynamical forcing of stratospheric waves by the tropospheric circulation. *J. Atmos. Sci.,* **55,** 2361–2392.

Scott, R. K., and L. M. Polvani, 2004: Stratospheric control of the upward wave flux near the tropopause. *Geophys. Res. Lett.,* **31,** L02115, doi:10.1029/2003GL017965.

——, and D. G. Dritschel, 2005: Quasi-geostrophic vortices in compressible atmospheres. *J. Fluid Mech.,* **530,** 305–325.

——, ——, L. M. Polvani, and D. W. Waugh, 2004: Enhancement of Rossby wave breaking by steep potential vorticity gradients in the winter stratosphere. *J. Atmos. Sci.,* **61,** 904–918.

Smith, A. K., 1989: An investigation of resonant waves in a numerical model of an observed sudden stratospheric warming. *J. Atmos. Sci.,* **46,** 3038–3054.

Swanson, K. L., 2000: Stationary wave accumulation and the generation of low-frequency variability on zonally varying flows. *J. Atmos. Sci.,* **57,** 2262–2280.

——, P. J. Kushner, and I. M. Held, 1997: Dynamics of barotropic storm tracks. *J. Atmos. Sci.,* **54,** 791–810.

Tung, K. K., and R. S. Lindzen, 1979a: A theory of stationary long waves. Part I: A simple theory of blocking. *Mon. Wea. Rev.,* **107,** 714–734.

——, and ——, 1979b: A theory of stationary long waves. Part II: Resonant Rossby waves in the presence of realistic vertical shears. *Mon. Wea. Rev.,* **107,** 735–750.

Wang, X., and J. Fyfe, 2000: Onset of edge wave breaking in an idealized model of the polar stratospheric vortex. *J. Atmos. Sci.,* **57,** 956–966.

Waugh, D. W., and D. G. Dritschel, 1999: The dependence of Rossby wave breaking on the vertical structure of the polar vortex. *J. Atmos. Sci.,* **56,** 2359–2375.

White, A. A., 1978: A note on the horizontal boundary condition in quasi-geostrophic models. *J. Atmos. Sci.,* **35,** 735–740.