Charney Problem with a Generic Stratosphere

MANKIN MAK, a SIYU ZHAO, b,c AND YI DENG b

a Department of Atmospheric Sciences, University of Illinois at Urbana-Champaign, Champaign, Illinois
b School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, Georgia
c Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, Los Angeles, California

ABSTRACT: This paper reports a comprehensive instability analysis of a 3D Charney-like model with an observationally compatible generic stratosphere. It is found that the values of a single nondimensional parameter $\Gamma = \frac{\kappa^2}{\beta DN_f^2}$ (detailed definition in text), in conjunction with representative values of four other nondimensional parameters, would dictate the existence of multiple branches of unstable modes as a function of the zonal wavenumber. Prototype Charney mode, Green mode, and two additional structurally distinct modes (Charney+ mode and tropopause mode) are identified. The latter result from the additional strong influence of the tropopause. The dynamical nature of all modes is delineated in terms of their meridional fluxes of heat and potential vorticity. The three-dimensional structure of the ageostrophic velocity field in each mode is presented to identify its potential of inducing frontogenesis. Optimal-mode analyses are also performed to ascertain how a disturbance with a predisposed structure would explosively develop toward each type of normal mode. In view of the pivotal role of $\Gamma$ in baroclinic instability and for historical reason, we name it the Charney number. It is most instructive to think of $\Gamma$ as a ratio of the meridional gradient of PV associated with the basic flow in the troposphere, $\lambda_f^2/(DN_f^2)$, to that associated with Earth’s rotation, $\beta$.

KEYWORDS: Atmospheric circulation; Baroclinic flows; Dynamics; Instability; Stratospheric circulation; Waves, atmospheric

1. Introduction

The study of extratropical cyclogenesis as posed by Charney (1947) is referred to as the “Charney problem.” His analytic finding was a groundbreaking advance in our understanding of cyclogenesis from the perspective of baroclinic instability. Those cyclones and the self-induced smaller-scale disturbances within are critically important for short-term weather forecast as well as for maintaining the global climate. However, it is pertinent to recognize that Charney’s model does not have a bona fide stratospheric region because its basic state consists of a single constant-shear zonal flow and one uniform stratification in the whole vertical domain. Consequently, the amplitude of the prototype Charney mode monotonically decreases with increasing elevation. To further elucidate the nature of cyclogenesis in a simplest possible model, we seek to delineate the impacts of a generic stratosphere.

Apart from the perspective of energetics, it is most succinct to interpret the physical nature of baroclinic instability in terms of potential vorticity (PV) dynamics. The notion of equivalent potential vorticity (Bretherton 1966) enables us to interpret the instability in the Charney problem as a positive feedback process involving an equivalent PV anomaly at the surface and an interior PV anomaly associated with the beta effect (Robinson 1989). Such interpretation has been rephrased as an interaction between a pair of counterpropagating Rossby waves (CRW) constructed as a combination of a pair of amplifying/decaying normal modes (Heifetz et al. 2004). It has been referred to as “wave resonance mechanism” for short (Mak 2011, section 8.B.5.6). Another interpretation of baroclinic instability invokes the notion of overreflection of Rossby waves about a critical layer (Lindzen et al. 1980).

Lindzen (1994) incorporated the beta effect in the Eady (1949) model by modifying the basic wind or and the basic stratification so that the basic potential vorticity gradient remains zero. But some key aspects of Charney’s unstable modes cannot be reproduced in such a framework. The limitation is to be expected a priori because instability in such a modified Eady model can only stem from resonant interaction between the PV anomalies at the two boundary surfaces. Consequently, there would be still short-wave cutoff and there would not be multiple branches of unstable modes.

There have been a number of works that seek to extend Charney’s study. Green (1960) found unstable modes of planetary scale that could propagate into the stratospheric region in the Charney model. Geisler and Garcia (1977) investigated the instability of several vertical profiles of the zonal wind up to 100 km. They reported that the Green mode has the structure of a trapped internal normal mode with large amplitude. Such finding is likely biased by their upper boundary condition, $\psi_{\text{top}} = 0$, which would reflect any disturbances arriving from below. Straus (1981) analyzed the instability of a global winter solstice zonal wind that is a function of both latitude and height up to 70 km. But posing such instability study goes beyond the scope of the Charney problem since the meridional shear in the basic flow gives rise to an additional dynamical process. Wang et al. (1985) investigated the impacts of Newtonian cooling and Ekman layer dissipation on the instability in Charney’s model. Those two additional processes were found to have only minor quantitative impacts as seen in their Fig. 8. Harnik and Lindzen (1998) examined the impact

DOI: 10.1175/JAS-D-20-0027.1

© 2021 American Meteorological Society. For information regarding reuse of this content and general copyright information, consult the AMS Copyright Policy (www.ametsoc.org/PUBSReuseLicenses).
of a basic PV gradient on baroclinic instability in a troposphere–stratosphere model with several basic flows and basically confirmed Green’s finding.

An underestimate of the sharpness of a tropopause can be a significant factor underlying systematic errors in numerical forecast models (Saffin et al. 2017). It warrants further investigating the Charney problem with a model that incorporates an observationally compatible stratosphere. That would require introducing a basic state that has a tropopause with two signature characteristics: sharply increasing stratification and sharply decreasing vertical shear with height. It follows that the meridional gradient of the basic PV would sharply decrease with height across a tropopause like a Heaviside function. In contrast, the mean flow and stratification in the rest of the stratosphere vary by relatively much less. It is therefore reasonable to hypothesize the latter to be of secondary importance. This is the conceptual basis underlying the design of our basic state (a constant $U$ in the model stratosphere, a constant shear in the model troposphere and a much larger stratification in the stratosphere than in the troposphere). We refer to it as a generic stratosphere in a Charney-like model. The use of such an idealized stratosphere would help highlight the key effects. It turns out that our model stratosphere is identical to the one used in Rivest et al. (1992, Fig. 1) who analytically investigated the dynamics of upper-tropospheric synoptic-scale waves as neutral normal modes in a modified Eady model. However, since their solution is time independent, it does not offer insight into the genesis of such waves. Our model is more general in that it incorporates the beta effect, has an exponentially decreasing density with height, nor is the model tropopause treated as an impenetrable surface. The model is amenable to examining the issue of genesis of tropopause disturbances in the context of baroclinic instability.

A number of tasks are performed in this study. The first task is to establish the influences of all external parameters in modifying the characteristics of the multiple branches of unstable modes in the Charney problem (Burger 1962; Miles 1964; Kuo 1973). The second task is to specifically ascertain the impacts of the tropopause. The third task is to document the structure of the modal 3D ageostrophic velocity field, which inherently can give rise to frontogenesis in a counterpart semigeostrophic model. The fourth task is to closely examine the transient development of each normal mode.

The model formulation and method of analysis are described in section 2. The various aspects of the model results are presented in section 3. The applications of optimal-mode analysis to the various modes are presented in section 4. The paper ends with some concluding remarks in section 5.

2. Formulation of an instability analysis

We consider quasigeostrophic disturbances in a rotating Boussinesq fluid on a beta plane, $f = f_0 + \beta y$. The domain is $0 \leq x < \infty$, $0 \leq y \leq Y$, $0 \leq z \leq z_{\text{top}}$. Although the zonal length of a model should correspond to a certain reference latitude circle, we extend it to infinity for computational convenience, so that we may do computation for a continuous range of zonal wavenumbers, rather than a discrete set of multiple values of a certain minimum wavenumber. The model boundaries $z = 0$, $y = 0$, and $y = Y$ are rigid surfaces across which there is no flow. We treat $z_{\text{top}}$ as a free surface. The background density is $\rho_0 = \rho_0 \exp (-z/H)$, where $H$ is a scale height. The basic state has a simple wind profile: $U = \lambda z (0 \leq z \leq D)$ and $U = \lambda D (z \geq D)$ with $z_{\text{top}}$ much higher than the tropopause $D$. The basic state also has a simple profile of Brunt–Väisälä frequency: $N = N_1$ for $z < D$ and $N = N_2$ for $z > D$. This modified Charney model is referred to as Charney++ model for short.

The flow in this model is governed by a prognostic equation and a supplemental diagnostic equation. The prognostic equation expresses conservation of OG potential vorticity $P = \beta y + \phi_{x x} + \phi_{y y} + (f_0^2/\rho_0) (\partial/\partial z)(\rho_0 N^2 (\partial \phi/\partial z))$ following the movement of a quasigeostrophic flow component $(\alpha = -\phi_y, \nu = \phi_x)$, where $\phi(x, y, z, \lambda)$ is the geostrophic streamfunction. Considering a perturbation, we decompose $P$ and $\phi$ into a basic component and a departure component: $P = Q + q$ and $\phi = \Phi + \phi$, where

$$Q = \beta y - \frac{f_0^2}{H N^2} \frac{\partial \Phi}{\partial z} + \frac{\partial}{\partial x} \left( \frac{f_0^2}{\rho_0} \frac{\partial \Phi}{\partial z} \right),$$

$$q = \phi_{x x} + \phi_{y y} - \frac{f_0^2}{H N^2} \frac{\partial \phi}{\partial z} + \frac{\partial}{\partial x} \left( \frac{f_0^2}{\rho_0} \frac{\partial \phi}{\partial z} \right).$$

The linearized potential vorticity equation governing such a perturbation is then

$$\left( \frac{\partial}{\partial t} + U \frac{\partial}{\partial x} \right) \left[ \phi_{x x} + \phi_{y y} + \frac{1}{\rho_0} \frac{\partial}{\partial z} \left( \frac{f_0^2}{\rho_0} \frac{\partial \phi}{\partial z} \right) \right] + \phi_x \left[ \beta + \frac{1}{\rho_0} \frac{\partial}{\partial z} \left( \frac{f_0^2}{\rho_0} N^2 \phi_{y y} \right) \right] = 0.$$  

(3)

For the basic state described above, we can rewrite (3) for the tropospheric region $0 \leq z \leq D$ as

$$\left( \frac{\partial}{\partial t} + \lambda z \frac{\partial}{\partial x} \right) \left[ \phi_{x x} + \phi_{y y} - \frac{f_0^2}{H N_1^2} \frac{\partial \phi}{\partial z} + \frac{\partial}{\partial z} \left( \frac{f_0^2}{\rho_0} \frac{\partial \phi}{\partial z} \right) \right]$$

$$+ \phi_x \left( \beta + \frac{f_0^2}{H N_1^2} \right) = 0.$$  

(4)

In the stratospheric region $D < z \leq z_{\text{top}}$, (3) takes on the form

$$\left( \frac{\partial}{\partial t} + \beta D \frac{\partial}{\partial x} \right) \left[ \phi_{x x} + \phi_{y y} - \frac{f_0^2}{H N_2^2} \frac{\partial \phi}{\partial z} + \frac{\partial}{\partial z} \left( \frac{f_0^2}{\rho_0} \frac{\partial \phi}{\partial z} \right) \right]$$

$$+ \phi_x (\beta) = 0.$$  

(5)

On the basis of the thermodynamic equation, the constraint $w = 0$ at the bottom surface yields a lower boundary condition

$$\phi_{x x} = \lambda \phi_x.$$  

(6a)

The two lateral boundary conditions are

$$\text{at } y = 0, Y: \quad \nu = \phi_x = 0,$$

leading to a Dirichlet boundary condition.
\( \phi = 0. \quad (6b) \)

From a causality point of view, it would be desirable to impose a radiation condition at \( z = z_{\text{top}} \) so that a solution would make up of only upward-propagating waves. In an instability modal analysis, radiation condition is a conditional functional constraint that would involve the unknown eigenvalue itself [e.g., expressed by Eq. (2.8) in Wang et al. 1985]. It follows that without prior knowledge of the eigenvalue itself, one cannot separate the upward-propagating from the downward-propagating wave components. Implementing such constraint would require applying an iterative algorithm. We opt for treating the upper boundary as a free surface, viz.,

\[ z = z_{\text{top}}, \quad \phi_z = 0. \quad (7) \]

This Neumann condition is less restrictive than a Dirichlet condition because it permits disturbances to pass through \( z = z_{\text{top}} \) only requiring the vertical tilt to vanish.

The model tropopause, \( z = D \), is an internal boundary where there is zero-order discontinuity in both the basic stratification and the basic shear. They give rise to a zero-order discontinuity in the basic meridional PV gradient, \( f_0^2 \lambda / (HN_1^2) \). Apart from requiring \( \phi \) to be continuous at \( z = D \), we also need to impose a self-consistent matching condition. It can be obtained by integrating (3) term by term over a very short distance \( 2 \delta \) across the tropopause and taking the limit \( \delta \to 0 \). The required matching condition for \( \phi \) at \( z = D \) is

\[
\left( \frac{\partial}{\partial t} + \lambda D \frac{\partial}{\partial x} \right) \left( \frac{f_0^2}{N_2^2} \phi_z \bigg|_{D^+} - \frac{f_0^2}{N_1^2} \phi_z \bigg|_{D^-} \right) + \frac{f_0^2}{N_1^2} \lambda \phi_z = 0. \quad (8)
\]

In other words, the total derivative of the difference in the stratification-weighed potential temperature at the tropopause is required to be proportional to the difference of heat flux in the meridional direction at that level. Equations (4) to (8) constitute a complete set of governing equations for the evolution of \( \phi \).

The supplemental diagnostic equation of a OG model is known as “omega equation.” Its linearized form instantaneously relates the vertical component of the ageostrophic velocity field \( w \) to the geostrophic streamfunction:

\[
\nabla^2 w + \frac{f_0^2}{N_2^2} \frac{\partial}{\partial z} \left[ \frac{1}{\rho_i} \frac{\partial}{\partial z} (\rho_i w) \right] = 2 \frac{f_0}{N_2^2} (U_x \nabla^2 \phi_x) + \frac{f_0}{N_2^2} \beta \phi_{cz}. \quad (9)
\]

In the tropospheric layer \( 0 \leq z \leq D \), (9) becomes

\[
\nabla^2 w + \frac{f_0^2}{N_1^2} \frac{\partial}{\partial z} \left[ \frac{1}{\rho_i} \frac{\partial}{\partial z} (\rho_i w) \right] = 2 \frac{f_0}{N_1^2} (\lambda \nabla^2 \phi_x) + \frac{f_0}{N_1^2} \beta \phi_{cz}. \quad (10)
\]

The lower boundary condition is

\[ z = 0, \quad w = 0. \quad (11) \]

In the stratospheric layer \( D \leq z \leq z_{\text{top}} \), (9) becomes

\[
\nabla^2 w + \frac{f_0^2}{N_2^2} \frac{\partial}{\partial z} \left[ \frac{1}{\rho_i} \frac{\partial}{\partial z} (\rho_i w) \right] = \frac{f_0}{N_2^2} \beta \phi_{cz}. \quad (12)
\]

The corresponding upper boundary condition is

\[ z = z_{\text{top}}, \quad w_z = 0. \quad (13) \]

Apart from being continuous at \( z = D \), \( w \) must also satisfy a matching condition. It is again derived by integrating (9) term by term over \( 2 \delta \) across the tropopause and taking the limit \( \delta \to 0 \). It can be shown that such matching condition for \( w \) is

\[
\frac{1}{N_2^2} (w_z |_{D+\delta} - w_z |_{D}) - \frac{1}{N_1^2} (w_z |_{D-\delta} - w_z |_{D}) = 0. \quad (14)
\]

In other words, the stratification-weighed difference of the vertical gradient of vertical velocity above the tropopause relative to that at the tropopause must be equal to the corresponding quantity below the tropopause. We will use (10) through (14) to determine \( w \) whenever the \( \phi \) field is known.

The 3D ageostrophic velocity field satisfies the mass conservation requirement \( \bar{u}_t + \bar{v}_t + (1/\rho)(\partial / \partial z)(\rho \bar{w}) = 0 \), where \((\bar{u}, \bar{v}, \bar{w}) = (u, v, w) \) is the horizontal ageostrophic velocity. We may define a velocity potential \( \eta \) such that \( \bar{u} = \partial \eta / \partial y \) and \( \bar{v} = \eta / \partial x \). The mass conservation equation would then have the form

\[
\eta_{xx} + \eta_{yy} + \frac{1}{\rho_i} \frac{\partial}{\partial z} (\rho_i \bar{w}) = 0. \quad (15)
\]

One can thereby readily determine \((\bar{u}, \bar{v})\) in terms of \( w \).

### 3. Normal mode instability analysis

#### a. Nondimensionalization

We measure horizontal distance, vertical distance, and time in units of the tropospheric radius of deformation \( N_1 D / f \), the height of the tropopause \( D \) and \( N_1 / (k f) \), respectively. The analysis is performed using nondimensional variables defined as \( \bar{x} = x f / (N_1 D) \), \( \bar{y} = y f / (N_1 D) \), \( \bar{z} = z / D \), \( \bar{t} = t k f / N_1 \), \( \bar{\phi} = \phi f / (N_1 D^2) \), \( \bar{w} = w f / (D f_u \lambda) \), and \( \bar{\rho} = \rho / \rho_o \). The nondimensional set of governing equations is given in an appendix, revealing that the properties of a perturbation in this model are controlled by five nondimensional external parameters: \( Y = y f / (N_1 D) \), \( \bar{z}_{\text{top}} = z_{\text{top}} / D \), \( A = N_2^2 / N_1^2 \), \( B = D / H \), and \( C = \beta D N_2^2 / (\lambda f_u^2) \).

The nondimensional normal-mode solution of \( \phi \) and \( w \) in this model, after dropping the tilde for convenience and setting \( Y = \pi \), is

\[
\phi = \chi(z) \sin(ly) \exp[i(kx - \sigma t)], \quad (16a)
\]

\[
w = \xi(z) \sin(ly) \exp[i(kx - \sigma t)], \quad (16b)
\]

where \( l = 1, 2, 3, \ldots \) This solution has a wave form in the \( x \) direction. It suffices to focus our analysis on those modes that have a broadest possible meridional structure, namely, those with a meridional wavenumber \( l = 1 \) in conjunction with a relevant range of the zonal wavenumber \( k \). Upon substituting (16a) into Eqs. (4)–(7), we get a system of ordinary differential equations that govern the amplitude function \( \chi(z) \). A finite-difference form of those equations yields a matrix equation:
The relevant values of four of our nondimensional parameters are listed in section 3a. Relevant values of external parameters are readily solved with MATLAB. Upon substituting (16a) and (16b) into Eqs. (10)–(14), we would get a system of ordinary differential equations that relates $\xi$ to $\chi$. The finite-difference form of those equations is a set of simultaneous algebraic equations written in matrix form as

$$ P \times \xi = S \times \chi. $$

With (17) and (18) we can determine the instability properties of all normal modes and examine their spatial structures in physical space.

b. Instability properties: Growth rate and phase speed

Relevant values of the external parameters are $\beta = 1.6 \times 10^{-11} \text{m}^{-1} \text{s}^{-1}$, $D = H = 10^4 \text{m}$, $z_{\text{top}} = 10^5 \text{m}$, $f_0 = 10^{-4} \text{s}^{-1}$, $N_1^2 = 10^{-4} \text{s}^{-2}$, $N_2^2 = 4 \times 10^{-4} \text{s}^{-2}$, and $\lambda = 3 \times 10^{-3} \text{s}^{-1}$. It has been verified that results are not sensitive to the specific value of $z_{\text{top}}$ provided that it is sufficiently high such as 100 km. The units of horizontal distance, vertical distance, and time are then $10^5 \text{m}$, $10^2 \text{m}$, and $0.3 \times 10^3 \text{s}$, respectively. We use 801 levels to depict the vertical domain so that the grid spacing is merely 125 m. Since we are investigating synoptic-scale disturbances, we focus on waves with $k$ values of order unity. We examine results for zonal wavenumber $0 < k \leq 5$. The relevant values of four of our nondimensional parameters are $z_{\text{top}} = 10$, $Y = \pi$, $A = 0.25$, and $B = 1$. The remaining external nondimensional parameter $C = \beta DN_1^2/(\lambda f_0^2)$ must have a decisive influence on the instability. Its typical value is about 0.5 and may have a considerable range. Our task is to compute the eigenvalues in a parameter space: $1 \times 10^{-2} \leq C^{-1} \leq 10$ and $0 < k \leq 5$ in conjunction with $z_{\text{top}} = 10$, $Y = \pi$, $A = 0.25$, and $B = 1$.

Our computations establish that, at each point in this parameter space, the set of eigenvalues $\sigma$ consist of one complex conjugate pair and 799 real positive values. It means that there is one intensifying mode, one decaying mode and many neutral modes. They are all eastward propagating. An example of the eigenvalues for $k = 2$ and $C^{-1} = 0.6$ is shown in Fig. 1. The complex eigenvalue values are $(\sigma_r, \sigma_i) = (0.37, \pm 0.28)$, meaning that the unstable mode has a phase speed of $\sigma/k = 0.185$ and a growth rate of about 0.28. In dimensional term, this mode propagates in the zonal direction at a speed of $\sim 6 \text{m s}^{-1}$ and intensifies with an $\epsilon$-folding time of $\sim 1$ day. Such values are compatible with the observed characteristics of extratropical cyclones. The neutral modes are thought of as a discrete version of a continuum spectrum.

Next, we investigate the systematic variations of $\sigma = \sigma_r + i \sigma_i$ with the zonal wavenumber $k$ and basic shear parameter $C^{-1}$. Kuo (1973) only showed the variation of the eigenvalues for the Charney model as a function of a composite parameter $r = (\delta + 1)/(\delta^2 + 4 \mu^2)^{1/2}$ [Pedlosky 1987, Eq. (7.8.38)], where $\delta$ and $\mu$ are equivalent to $C$ and $k$, respectively, in our notations. His spectrum of modes indicates successive branches of unstable modes separated by discrete neutral modes at the points of integer values of $r$. It would be more instructive to examine how the eigenvalue would separately vary with $C^{-1}$ and $k$ for our Charney$+$ model. We begin by making two complementary scans. The first scan covers a range of $k$ values for a fixed value of $C^{-1}$. This is done by using a very small increment.

FIG. 1. A plot of all the eigenvalues, $\sigma = \sigma_r + i \sigma_i$, in units of $\lambda f_0/N_1$, with their real parts indicated by the abscissa and their imaginary parts by ordinate for the parametric condition ($k = 2$, $l = 1$, $z_{\text{top}} = 10$, $A = 0.25$, $B = 1$, $C^{-1} = 0.6$). Relevant values of these external parameters are listed in section 3a.

For this example, we use $C = 0.6$ and $k = 2$. The finite-difference form of those equations is a set of simultaneous algebraic equations written in matrix form as

$$ \mathbf{P} \times \mathbf{\xi} = \mathbf{S} \times \mathbf{\chi}. $$

With (17) and (18) we can determine the instability properties of all normal modes and examine their spatial structures in physical space.

b. Instability properties: Growth rate and phase speed

Relevant values of the external parameters are $\beta = 1.6 \times 10^{-11} \text{m}^{-1} \text{s}^{-1}$, $D = H = 10^4 \text{m}$, $z_{\text{top}} = 10^5 \text{m}$, $f_0 = 10^{-4} \text{s}^{-1}$, $N_1^2 = 10^{-4} \text{s}^{-2}$, $N_2^2 = 4 \times 10^{-4} \text{s}^{-2}$, and $\lambda = 3 \times 10^{-3} \text{s}^{-1}$. It has been verified that results are not sensitive to the specific value of $z_{\text{top}}$ provided that it is sufficiently high such as 100 km. The units of horizontal distance, vertical distance, and time are then $10^5 \text{m}$, $10^2 \text{m}$, and $0.3 \times 10^3 \text{s}$, respectively. We use 801 levels to depict the vertical domain so that the grid spacing is merely 125 m. Since we are investigating synoptic-scale disturbances, we focus on waves with $k$ values of order unity. We examine results for zonal wavenumber $0 < k \leq 5$. The relevant values of four of our nondimensional parameters are $z_{\text{top}} = 10$, $Y = \pi$, $A = 0.25$, and $B = 1$. The remaining external nondimensional parameter $C = \beta DN_1^2/(\lambda f_0^2)$ must have a decisive influence on the instability. Its typical value is about 0.5 and may have a considerable range. Our task is to compute the eigenvalues in a parameter space: $1 \times 10^{-2} \leq C^{-1} \leq 10$ and $0 < k \leq 5$ in conjunction with $z_{\text{top}} = 10$, $Y = \pi$, $A = 0.25$, and $B = 1$.

Our computations establish that, at each point in this parameter space, the set of eigenvalues $\sigma$ consist of one complex conjugate pair and 799 real positive values. It means that there is one intensifying mode, one decaying mode and many neutral modes. They are all eastward propagating. An example of the eigenvalues for $k = 2$ and $C^{-1} = 0.6$ is shown in Fig. 1. The complex eigenvalue values are $(\sigma_r, \sigma_i) = (0.37, \pm 0.28)$, meaning that the unstable mode has a phase speed of $\sigma/k = 0.185$ and a growth rate of about 0.28. In dimensional term, this mode propagates in the zonal direction at a speed of $\sim 6 \text{m s}^{-1}$ and intensifies with an $\epsilon$-folding time of $\sim 1$ day. Such values are compatible with the observed characteristics of extratropical cyclones. The neutral modes are thought of as a discrete version of a continuum spectrum.

Next, we investigate the systematic variations of $\sigma = \sigma_r + i \sigma_i$ with the zonal wavenumber $k$ and basic shear parameter $C^{-1}$. Kuo (1973) only showed the variation of the eigenvalues for the Charney model as a function of a composite parameter $r = (\delta + 1)/(\delta^2 + 4 \mu^2)^{1/2}$ [Pedlosky 1987, Eq. (7.8.38)], where $\delta$ and $\mu$ are equivalent to $C$ and $k$, respectively, in our notations. His spectrum of modes indicates successive branches of unstable modes separated by discrete neutral modes at the points of integer values of $r$. It would be more instructive to examine how the eigenvalue would separately vary with $C^{-1}$ and $k$ for our Charney$+$ model. We begin by making two complementary scans. The first scan covers a range of $k$ values for a fixed value of $C^{-1}$. This is done by using a very small increment.

FIG. 2. Variations of growth rate $\sigma_r$ and phase velocity $\sigma_i/k$ of the unstable normal modes (a) with their wavenumber $k$ for a particular basic shear $1/C = 0.6$, (b) with the basic shear $1/C$ for a particular wavenumber $k = 4$. The other parameters are $l = 1$, $A = 0.25$, and $B = 1$. $\sigma_r$ and $\sigma_i/k$ are in units of $\lambda f_0/N_1$ and $\lambda D$, respectively. Relevant values of external parameters are listed in section 3a.
$\Delta k = 0.01$. Figure 2a shows that for $C^{-1} = 0.6$ the spectra of growth rate and phase speed is divided into two sections at $k \approx 1.2$. It is compatible with the finding reported in Green (1960).

The most unstable long wave is at $k = 0.8$ and the most unstable synoptic-scale wave is at $k = 2.1$. The growth rate of the former is only about 1/5 of that of the latter; while their phase velocities are comparable. There is no hint of short-wave cutoff at $k \approx 1$. The second scan covers the range $1 \times 10^{-2} \leq C^{-1} \leq 10$ for $k = 4$. This is done by using a very small increment $\Delta(C^{-1}) = 0.001$. The computation resolves the existence of four branches of unstable modes separated by neutral modes at $C^{-1} = 0.14$, 0.07, and 0.015 (Fig. 2b). This result confirms the existence of multiple branches of unstable modes in this Charney+ model.

Figure 3 is constructed by making a very large number of scans for the portion of the $k$ versus $1/C$ plane. It offers a comprehensive view of how the growth rate $\sigma_i$ and phase velocity $\sigma_i/k$ of the unstable modes vary with these two parameters in conjunction with $A = 0.25, B = 1, l = 1$, and $z_{top} = 10$. A logarithmic scale in the ordinate is used for $C^{-1}$ in order to highlight the presence of different branches of unstable modes. Four branches of unstable modes are well resolved for all wavenumbers.

The first branch of unstable modes exists at values of $C^{-1}$ above a well-resolved curve of neutral modes in Fig. 3a. It is of interest to note that the maximum growth rate in this branch reaches about 0.3 at $k \approx 3$ and $C^{-1} \approx 0.7$. For each fixed value of $C^{-1}$, all modes are divided into two subsets of $k$ by a neutral mode as discussed earlier. For larger $C^{-1}$, the most unstable mode shifts toward longer wavelength (smaller $k$), approaching an asymptotic value of $k \approx 1.4$ at $C^{-1} = 10$. There is definitely no short-wave cutoff for any value of $C^{-1}$. Figure 3b shows that the phase speed of the unstable mode is greater for larger $C^{-1}$ and is larger for longer waves.

The result that the growth rate for each $k$ is not necessarily greater for a stronger basic shear (larger $C^{-1}$) is a consequence of the dual influences of “beta.” Beta, like a two-edge sword, has a stabilizing influence as well as a destabilizing influence.

Its stabilizing influence stems simply from the restraining effect of rotation as manifest most clearly in the existence of neutral Rossby wave. On the other hand, $\beta$ is a component of the meridional gradient of the basic PV. Hence, there would be PV anomalies wherever a disturbance has meridional advection.
This makes it possible for an equivalent PV anomaly at the surface under consideration positively interacting with another PV anomaly of the same wavenumber at a certain interior level. When both PV anomalies are phase locked, such interaction would be self-sustainable. It is the net effect of the dual influences of $b$ that selectively determine an optimal value of $C_2$ within each range for each $k$ leading to a local maximum growth rate in each branch. It is noteworthy that this feature in our finding is absent in Green’s results (1960). His Fig. 3 shows that the growth rate of an unstable mode increases monotonically with increasing basic shear.

c. Other instability properties: Structure and fluxes

The unstable modes at different parts of the parameter space under consideration have different structures, suggesting that different modes are attributable to different primary dynamical factors. First of all, the most unstable mode of the first branch has a characteristic westward vertical tilt with an amplitude monotonically decreasing with increasing elevation, virtually unaffected by the presence of the tropopause (not shown for brevity). A case in point is for $k = 5, C_2 = 10.0$ as well as for $k = 4, C_2 = 0.23$. That is structurally very similar to the prototype Charney mode. For somewhat smaller values of $C^{-1}$ the unstable modes still in the first branch have a noticeably different structure. An example is the mode with $k = 2, C^{-1} = 0.6$. It shows the streamfunction field ($\phi$, contours) and the corresponding vertical velocity field ($w$, arrows) on a $x$–$z$ cross section at the middle of the domain, $y = \pi/2$ (Fig. 4). The units of horizontal velocity, vertical velocity, and geopotential are $l/D = 30$ m s$^{-1}$, $Df_0\lambda^3/N_1^2 \approx 10^{-1}$ m s$^{-1}$, and $\lambda Df_0/N_1 \approx 3 \times 10^7$ m$^2$ s$^{-1}$, respectively. This mode is essentially confined in the tropospheric layer, only slightly intruding above the tropopause. The nondimensional amplitude of $w$ and $\phi$ are both about unity. The $\phi$ field has a pronounced westward vertical tilt in the lower troposphere, $0 \leq z < 0.3$, with a progressively diminishing tilt higher above. The amplitude of $\phi$ is strongest at the surface; being $\sim 75\%$ as strong at the tropopause as at the surface (0.7 vs 0.9). The center at the tropopause level is located about 1/4 wavelength to the west of the center at the surface. The ascending region of this mode ($w > 0$) is located slightly to the east of the trough of the streamfunction. This mode differs from the prototype Charney mode mainly in that it has significant amplitude at the tropopause level and is virtually capped below that. We refer to it as a “Charney + mode.”

The energetics of an unstable mode can be inferred from the vertical distribution of its meridional and vertical fluxes of heat averaged over a wavelength. We compute those fluxes per unit mass with the following formulas:

$$[\phi_x\phi_z] = \frac{1}{2} k \sin^2(\theta_y) \left( \chi_x \frac{dx}{dz} - \chi_z \frac{dx}{dz} \right).$$ (19)
where $N^2 = 1$ and $N^2 = 1/A$ in the model tropospheric and stratospheric regions, respectively.

The three panels of Fig. 5 show the vertical distribution of $[\phi, \phi_f]$, $[w\phi_f]$, and $[\nu q]$ of the Charney+ mode shown in Fig. 4. The large northward heat flux in the lower tropospheric layer, $0 \leq z < 0.3$, is consistent with the westward vertical tilt of this mode. It implies a significant transfer of basic APE to EAPE primarily in that layer. Part of that EAPE in turn is converted to EKE over a deeper tropospheric layer as indicated by the substantial positive values of $[w\phi_f]$. Its PV flux has significant negative values in the troposphere with small negative values extending just over the tropopause. These values signify southward flux of PV as required for instability from the perspective of potential vorticity dynamics, $(d/dt)[\eta^2] > 0$, since the meridional gradient of basic PV has positive values.

At somewhat smaller values of $C^{-1}$ in Fig. 3 we enter into a regime of the second branch of unstable modes. We present the structure of a sample of such modes for $(k = 1, C^{-1} = 0.4)$ in two forms. Figure 6a shows that the magnitude of the $\phi$ field itself increases markedly with elevation; so much so that its value in the troposphere is seemingly negligible. That is intrinsically related to the exponential decrease of density with height. We show the structure of $\phi/\bar{\rho}_0$ and the corresponding $w/\bar{\rho}_0$ field in Fig. 6b. The density weighted field highlights its tropospheric origin and still captures its characteristic of upward–eastward propagation. The transfer from basic APE to EAPE in that tropospheric layer is the energy source for the intensification of this baroclinic wave. The vertical length scale of the mode is estimated to be about 40 km. Its $w$ field (arrows) is quite weak, owning to the fact that the stratification in the stratosphere is 4 times greater $(A = 0.25)$. The substantial westward-tilting bands of alternate signs in the stratosphere are a signature of vertical wave propagation associated with negative values of an effective refractive index in the stratospheric region. This mode is structurally very similar to the so-called Green modes (Green 1960). We therefore simply refer to it as a “Green mode.”

The vertical distributions of $\rho_1[\phi, \phi_f]$, $\rho_1[w\phi_f]$, and $\rho_1[\nu q]$ of the Green mode are presented in Fig. 7. These are fluxes per unit volume of air. The $\rho_1[\phi, \phi_f]$ panel confirms that the source of energy for intensification is a shallow layer next to the surface in which some of the basic APE is converted to EAPE. Although there is also some heat flux in the lower half of the stratospheric layer, it does not contribute to the intensification because the model stratospheric layer has no basic meridional thermal gradient. The $\rho_1[w\phi_f]$ panel shows not only significant conversion of EAPE to EKE in the bulk of the tropospheric region, but also in the lower stratosphere. The $\rho_1[\nu q]$ panel confirms that there is a southward flux of wave PV in the troposphere as required for instability. A delta-function-like wave PV flux at the tropopause is noteworthy. It arises from the delta-function-like vertical variation of the meridional gradient of basic PV at the tropopause.

d. Ageostrophic velocity field

As noted in the Introduction, it warrants examining the ageostrophic velocity field inherent in a 3D unstable mode. Measuring the horizontal ageostrophic velocity in unit of
Hence, we may evaluate \( \alpha_{\text{ageostrophic velocity,}} \) and \( \beta_{\text{ageostrophic velocity,}} \) such that \( \alpha = \alpha N_1/(Dl^2) \) and \( \beta = \beta N_1/(Dl^2) \).

As noted in section 2, we may introduce a velocity potential \( \eta \) such that \( \tilde{u} = \tilde{u}_x \), \( \tilde{v} = \tilde{v}_y \). Then the nondimensional continuity equation becomes

\[
\eta_{xx} + \eta_{yy} - Bw + w_z = 0. \tag{22}
\]

With \( w = \text{Re}[\xi(z) \sin(ly) \exp(i(kx - \sigma t))] \) associated with the \( \phi \) field of each normal mode, \( \eta \) must be of the form \( \text{Re} \{H(z) \sin(ly) \exp[i(kx - \sigma t)]\} \), where \( H \) is related to \( \xi \) by

\[
H = \frac{1}{k^2 + l^2} \left( -B \xi_x + \xi_z \right). \tag{23}
\]

Hence, we may evaluate \( \tilde{u}(x, y, z) \) with

\[
\tilde{u}(x, y, z) = \frac{1}{k^2 + l^2} k \sin(ly) \left\{ \left( B \xi_x - \frac{d\xi_z}{dz} \right) \sin(kx) + \left( B \xi_t - \frac{d\xi_z}{dz} \right) \cos(kx) \right\}, \tag{24}
\]

\[
\tilde{v}(x, y, z) = \frac{1}{k^2 + l^2} l \cos(ly) \left\{ \left( -B \xi_x + \frac{d\xi_z}{dz} \right) \cos(kx) + \left( B \xi_t + \frac{d\xi_z}{dz} \right) \sin(kx) \right\}. \tag{25}
\]

It is not a simple matter to plot a 3D vector field such as the ageostrophic velocity, \( (\tilde{u}, \tilde{v}, w) \). A practical alternative is to plot the projection of the vector field onto three judiciously chosen mutually orthogonal cross sections. Figure 8 shows such plots of the Charney + mode together with the \( \phi \) field for \( (k = 2, \xi = 0.6) \). Figure 8a reveals a cellular structure on an \( x-z \) plane at the middle of the meridional domain \( y = \pi/2 \) essentially confined in the troposphere. Figure 8b reveals that at the zonal location of maximum ascent \( (kx = 4.8) \), the air in the midupper troposphere begins to diverge toward both the southern and northern boundaries like a fountain as it goes up. Figure 8c reveals the horizontal flow in midtroposphere \( (z = 0.5) \) diverging away both westward and eastward at the location \( (kx \approx 4.8) \) but converging toward a counterpoint \( (kx \approx 2.0) \) to support the compensating descent. On the basis of these 3 plots, one can visualize the overall structure of this 3D vector field without much difficulty. Such overall distribution of ageostrophic velocity in the whole domain would convert EAP to EKE mostly in the lower tropospheric region. We can also infer from the configuration of the vector field with respect to the \( \phi \) field that if the full advective influences of the ageostrophic flow were permitted to feedback upon the quasigeostrophic flow itself as in a counterpart semigeostrophic model, this ageostrophic velocity component would naturally induce 3D surface fronts as well as upper-level fronts during a finite-amplitude stage of development of this unstable wave.

The corresponding plots of the ageostrophic velocity in a Green mode for \( (k = 1, \xi = 0.4) \) show a more complex 3D cellular structure progressively stronger at higher levels (Fig. 9).

---

**Fig. 7.** Vertical distributions at the middle of the domain \( y = \pi/2 \) of the nondimensional meridional heat flux \( \rho_0[\phi_r \phi_b] \), vertical heat flux \( \rho_0[w \phi_b] \), and the meridional PV flux \( \rho_0[v q] \) of the Green mode shown in Fig. 6, where the square brackets stand for averaging over a zonal wavelength. \( \rho_0[\phi_r \phi_b] \), \( \rho_0[w \phi_b] \), and \( \rho_0[v q] \) are in units of \( \lambda^2 N_1^2 \rho_0 f_0 \), \( \rho_0 \lambda^2 D^2 N_1 \), and \( \lambda^2 D \rho_0 f_0 / N_1 \), respectively; relevant values of external parameters are listed in section 3a.
It is characterized by a tilting cellular structure at those levels at $y = \pi/2$. The plot of $(\hat{v}, w)$ at $kx = 3.0$ clearly reveals that this mode originates in the troposphere and propagates upward and exits through the model top. The convergence in the horizontal components of the ageostrophic velocity of this mode at $z = 8.5$ supports the ascending as well as descending flow at $kx = 3.0$. That flow is about half as strong as of a Charney mode at the tropospheric levels (lower panel of Figs. 8 and 9).

e. Tropopause mode

As we closely examine the unstable modes in the second branch for different wavenumbers, we encounter modes with quite distinct characteristic structures. A case in point is the most unstable mode for $(k = 4, C^{-1} = 0.08)$ (Fig. 10). The structure of $\phi$ of this mode is quite unique and fundamentally different from the Charney+ mode. Its amplitude near the tropopause level is ~0.8, whereas at the surface it is 0.35. Therefore, in sharp contrast to the Charney+ mode, the amplitude of the mode under consideration is more than twice as strong at the tropopause as at the surface. It has a very small vertical tilt above $z = 0.2$ unlike the large westward tilt in a shallow layer next to the surface. Furthermore, the position of its upper tropospheric portion is ~180° out of phase with respect to its surface portion. We call this unstable mode a “tropopause mode.”

The tropopause mode is reminiscent of the propagating cyclonic disturbances observed near the tropopause (Petterssen and Smeebye 1971; Hakim 2000; Hakim and Canavan 2005; Tomikawa et al. 2006). Rivest et al. (1992) modified the Eady model by including a stratospheric region identical to the one
used in this study and treating the tropopause as an impervious surface. They found neutral solutions with maximum magnitude at the tropopause. While the structure of our tropopause mode at the upper tropospheric levels is similar to Rivest et al.’s neutral modified Eady mode, its additional shallow surface baroclinic layer signals its ability to intensify in time. As such, it sheds additional light onto the genesis of upper-tropospheric disturbances.

Muraki and Hakim (2001) extended Rivest’s analysis by including a contribution from the next-order OG $^1$ dynamics (Muraki et al. 1999). The resulting disturbance has the more realistic property of cyclonic–anticyclonic asymmetry. Furthermore, if a model happens to have a basic state characterized by strong baroclinicity at an upper tropospheric levels, large static stability in the lower troposphere and large surface roughness, one would naturally expect to find unstable disturbances near the tropopause. That is indeed the case as reported in an analysis by Whitaker and Barcilon (1992) using a primitive equation model with a rigid upper boundary at 16 km.

It warrants emphasizing that this “tropopause mode” for $k = 4$ and the Green mode for $k = 2$ (Fig. 6) belong to the same second branch of unstable modes (Fig. 3). While the tropopause mode is entirely trapped in the troposphere, the Green mode propagates through the stratospheric region. In other words, not all unstable modes within one branch have one and same class of structure. The distributions of $[\phi, \phi_z]$ and $[u_q]$ for the tropopause mode shown in Fig. 11 highlight the role of a very shallow surface layer as the source of this mode from both energetic and PV points of views. There is a much thicker layer where most conversion from EAPE to EKE, $[w\phi_z]$, takes place.

![Fig. 9. Three projections of the 3D ageostrophic velocity field ($u, v, w$) of a Green mode for ($k = 1, l = 1, 1/C = 0.4$) with $A = 0.25$ and $B = 1$; (top left) ($u, w$) vectors on a vertical $x$–$z$ cross section at $y = \pi/2$; (top right) ($v, w$) vectors on a vertical $y$–$z$ cross section at $kx = 3.0$; (bottom) ($u, v$) vectors on a horizontal $x$–$y$ cross section at $z = 8.5$. $\phi$ (contours) are in units of $AN_1D^2f_0/c$; ($u, w$) and ($v, w$) are in units of $\lambda^2D^2f_0/N_1^2$; ($u, v$) is in units of $\lambda^4D^2/N_1^2$; relevant values of external parameters are listed in section 3a. The solid (dashed) lines denote positive (negative) values.](https://example.com/fig9.png)
related to the overturning of the ageostrophic flow within the troposphere.

4. Optimal mode analysis

Transient growth of baroclinic disturbances is an important issue from both theoretical and practical points of view. De Vries and Opsteegh (2007) examined such problem with a constant-density Charney’s model using a PV building block methodology. De Vries et al. (2009) showed that a reduced system consisting of a pair of CRW together with an additional neutral Rossby wave component would be capable of delineating the transient evolution of initial disturbances. This issue can be addressed more generally with an optimal-mode analysis (Farrell 1989). It enables us to determine a disturbance that would intensify toward a normal mode by a maximum possible extent as measured according to a chosen norm over a specific time interval $\tau$. The qualitative aspect of such development does not hinge upon the chosen norm. The simplest kind of norm is the Euclidian norm known as L-2 norm. It would suffice to use such norm for our purpose. The resulting disturbance is known as an optimal-mode. It amounts to solving an eigenvalue--eigenvector problem under each parametric condition (Farrell 1989; Mak 2011, section 8B.3):

$$B(\tau)a = \gamma B(0)a,$$

where $\gamma$ is an eigenvalue representing the amplification factor for an optimization time interval $\tau$, and $a$ is the projection coefficient of an initial streamfunction in a spectral expansion using a set of normal modes $\chi_j$ ($j = 1, 2, \ldots, J$) as basis vectors. The latter is separately obtained in an instability analysis for the parametric condition under consideration. The other notations are $B(\tau) = \Lambda^{\tau} P^{\tau} D P\Lambda$ and $B(0) = P^0 D P$, $P$ a $J \times J$ matrix with columns made up of $\chi_j$; $P^{\tau}$ the Hermitian of $P$, $\Lambda$ a $J \times J$ matrix with diagonal elements $e^{\rho j}$ and all other elements being zero. Corresponding to the L-2 norm, $D$ is a unit matrix. The mode associated with the largest eigenvalue of (26) is the fastest amplifying mode, referred to as the optimal mode.

**a. Optimal transient evolution toward the Charney+ mode**

Figure 12 is a plot of the amplification factor of all eigenmodes of (26) for $\tau = 4$ associated with a Charney+ mode under parametric conditions ($l = 1, k = 2, A = 0.25, B = 1, C^{-1} = 0.6$). The largest amplification factor in this case has a value of $\sim 118$. The growth rate of the corresponding most unstable normal mode is about 0.28. The amplification factor of the second mode is much smaller, only $\sim 20$.

Figure 13 shows the evolution of the $\phi$ field on an $x$-$z$ plane at $y = \pi/2$ using the optimal mode associated with the Charney+ mode as an initial disturbance. At $t = 0$, this disturbance has an almost uniform westward tilt and spans the whole depth of the troposphere. Its amplitude is normalized to unity, strongest at $z = 0.5$. Its tilt suggests an impending rapid intensification via the so-called Orr’s mechanism. The advective influence of the background zonal shear flow progressively reduces the tilt of this disturbance in time. By $t = 8$ the disturbance has evolved to a structure quite resembling a Charney+ mode. The amplitude of the disturbance has intensified by a factor of 20. For comparison, we also compute the corresponding evolution of the Charney+ mode itself (Fig. 4) initially normalized to the same magnitude as the optimal mode. As expected, the normal mode intensifies in time while its structure remains unchanged except for a zonal translation. At $t = 4$ the magnitude of the normal mode has increased only by about 2.5 folds (not shown for brevity). In other words, the unstable normal mode is only about one-fourth as strong as the corresponding optimal mode on $t = 4$ in this case.

**b. Optimal transient evolution toward the Green mode**

We next examine the transient evolution of the optimal mode toward a Green mode for parametric condition ($l = 1, k = 1, A = 0.25, B = 1, C^{-1} = 0.4$). The amplification factor of the first optimal mode for $\tau = 5$ is much larger, about 393.5. Figure 14 shows the development of the structure of this disturbance at $t = 0, 15$, and 30. This optimal mode is entirely in the troposphere ($0 \leq z \leq 1$). By $t = 15$, its magnitude has intensified 30 times and the strongest part of the disturbance has risen to $z = 25 \text{ km}$. The highest part of this disturbance has reached $z = 6$. By $t = 30$, the disturbance has reached the model top and its strength in the stratosphere has intensified by about 200 folds and is indeed so strong that its presence in the troposphere is virtually invisible. The overall structure confirms that a Green mode has been firmly established by...
This sequence of plots illustrates how a Green mode transiently/rapidly evolves by means of upward propagation while it is continually energized near the surface. The growth rate of the normal mode itself is quite small, $\sim 0.12$. Its intensity increases by no more than 6 times over this time interval, while its structure remains the same except for zonal propagation (not shown for brevity).

c. Optimal transient evolution toward the tropopause mode

With the tropopause mode in mind, we examine the optimal mode under the parametric condition ($l = 1$, $k = 4$, $A = 0.25$, $B = 1$, $C^{-1} = 0.08$) with $\tau = 5$. The amplification factor of this optimal mode is even larger, $\sim 565$. Its evolution depicted at $t = 0, 8, 16, 24$ in Fig. 15 reveals a hitherto unsuspected interesting aspect of development. The optimal mode only spans the troposphere and has a very large westward vertical tilt (indicating an even smaller aspect ratio of the vertical versus horizontal length scales). The disturbance explosively intensifies leading to an intensified amplitude about 18 folds of its original value by $t = 8$ via joint impacts of the Orr’s mechanism and vertical propagation. Its structure at this time has two distinct components. There is an eastward-tilting lower component in the upper troposphere ($0.4 \leq z \leq 1.0$) due to overshooting of the Orr mechanism. There is an additional westward-tilting upper component in the lower stratosphere ($1 \leq z \leq 3$). By $t = 16$, the two components have completely separated. The maximum amplitude is in the lower stratosphere and is $\sim 26$ folds of the original value. The lower component now has nearly no vertical tilt and stays in the upper troposphere. Meanwhile, the upper component continues to propagate deeper into the stratospheric region as a wave packet while retaining its westward vertical tilt. The upper component of this disturbance is clearly a transient Green-like disturbance. This trend of development persists in time. Its intensity further increases to $\sim 55$ folds by $t = 24$. The lower component continues to be a significant element. The upper component eventually exits through the top of the domain at large time.
The lower component evolves toward a tropopause mode continually receiving energy from a very shallow baroclinic layer next to the surface. This figure brings to light that a Green-like disturbance and a tropopause-like disturbance coexist during the transient evolution. While the former is a transient element eventually exiting through the top boundary, the latter ultimately establishes itself as a tropopause mode.

5. Concluding remarks

We have quantitatively investigated all aspects that we can think of about the Charney problem with the inclusion of an observationally compatible generic stratosphere. Figure 3 is a portrait of all its instability properties as a function of the zonal wavenumber \( k \) and the parameter \( C^{-1} \) well resolving the existence of four branches of unstable modes. The structure of the most unstable mode in the first branch is quite similar to a prototype Charney mode and is referred to as “Charney mode.” For somewhat smaller values of \( C^{-1} \), we come across unstable modes still in the first branch that have largest amplitude at the surface and a secondary maximum amplitude at the tropopause. We refer to them as “Charney+ modes” (Fig. 4). At even smaller values of \( C^{-1} \), there are modes structurally similar to a Green mode (Fig. 6). It is noteworthy that Green modes are not necessarily modes of planetary scale as is often stated to be the case. They can appear as modes in both the second and third branches dependent on the value of \( k \) (Fig. 6 versus Fig. 10). In addition, we have encountered modes of synoptic short scale that have decidedly larger amplitude at the tropopause level than at the surface. We refer to them as “tropopause modes,” which may be pertinent to the genesis of the observed upper-tropospheric cyclonic disturbances. The modes in the third and fourth branch are structurally similar to the Green mode (Fig. 6a) in that they have progressively larger amplitude at greater elevations. The dynamical nature of all branches of modes has been diagnosed in terms of their meridional fluxes of heat and PV. The tropopause in this Charney+ model plays an important role because only the tropospheric region has a basic shear flow and the stratification significantly increases across it. It follows that the vertical variation in the gradient of basic PV across the tropopause is a delta function.

We have presented the structure of the complex 3D ageostrophic velocity field of each mode. Such plots reveal the locations where frontogenesis could take place in a counterpart semigeostrophic model. We have also performed an optimal-mode analysis to delineate the transient evolution
of the optimal mode associated with the various modes. We find that under certain weak shear condition a Green-like structure together with a tropopause-like structure transiently coexist, evolving from an optimal mode. While the former eventually exits through the top boundary, the latter ultimately survives in the troposphere.

It is pertinent to mention that if we use ‘‘\(C\)’’ instead of ‘‘\(C_1\)’’ as the ordinate to replot Fig. 3(i) and (ii) vs \(k\), the resulting figure (not shown for brevity) looks roughly similar to a counterpart figure reported in Mak (2018, Fig. 3.6) for what was dubbed ‘‘equivalent-Charney model.’’ That model is like this Charney+ model except the stratospheric region is conceptually shrunk to an infinitely thin layer. Such drastic simplification vastly simplifies an analysis but does not qualitatively alter the stability properties of this Charney+ model. A number of characteristics reported in this study are unavoidably missing/distorted in that overtly simplified analysis.

It is worth emphasizing that the value of a single nondimensional parameter \(C_1\) dictates the existence of multiple branches of unstable modes in this Charney+ model. It warrants highlighting such pivotal importance of \(C_1\) by giving it an appropriate name. Vallis (2006) named a nondimensional quantity akin to \(C\) itself the ‘‘Charney–Green number.’’ In our opinion, Charney’s contribution in the study of baroclinic instability is groundbreaking in nature, whereas Green’s computational result is supplemental. For this reason, we name \(G_1 = \frac{\lambda f_2}{(\beta DN_1^2)}\) the Charney number. It is dynamically most instructive to think of \(C_1\) as a ratio of the meridional gradient of PV associated with the basic shear flow in the troposphere, \(\frac{\lambda f_2}{(DN_1^2)}\), to that associated with Earth’s rotation, \(\beta\). Since the original Charney model has a single stratification \(N_1\) and basic shear \(\lambda\), a corresponding definition for the Charney number would be \(\Gamma = \frac{\lambda f_2}{(\beta HN_1^2)}\), with \(H\) being the density-scale height, which is typically close to the tropopause height \(D\). Charney number is on par in historical significance with the several other widely accepted nondimensional numbers in atmospheric sciences [Rossby number (Ro); Richardson number (Ri); Burger number (Bu)].

This model can be reduced to the Charney model if we use a basic state defined by \(U = \lambda z\) for \(0 \leq z \leq z_{top}\) and \(N_2 = N_1\). With a slight modification of the original code, we can determine the numerical solutions of the Charney model to be used as a quantitative reference. Such solutions of the unstable modes would serve as a primary reference. It helps identify how the impacts of the generic stratosphere would give rise to
the various unstable modes. It suffices to show the result obtained by scanning over the whole range of Charney number $C^{-1} = \frac{\lambda f_0}{(\beta DN_1^2)}$, $10^{-2} \leq 1/C \leq 10^1$ for a specific wavenumber, $k = 2$. Figure 16a shows the variation of the growth rate $\sigma$, and phase speed $\sigma/k$ with $1/C$. It confirms the existence of multiple branches of unstable modes separated by a set of neutral modes. Figure 16b shows that the most unstable mode in the first branch for $C^{-1} = 0.5$ has a westward vertical tilt with an amplitude largest at the surface monotonically and sharply decreasing upward. This is the structure of a Charney mode. Figures 16c and 16d show the structures of the most unstable mode of the second and third branches for $C^{-1} = 0.15$ and $C^{-1} = 0.093$, respectively. These modes in the higher branches progressively extend to greater heights. The level of primary maximum amplitude in each case is not at the surface, but rather at certain level above. They are different versions of the Green mode. Such unstable modes would be modified by the tropopause of our generic stratosphere. Hence, the tropopause with a large increase in the meridional gradient of the basic potential vorticity is where the Charney+ mode or tropopause mode come from.

We have also made computations designed to separately examine the impacts of a large increase of stratification and a large increase of vertical shear across the tropopause. It is verified that each of these two elements could give rise to a weaker version of the Charney+ mode and tropopause mode. The results are not presented for brevity.

Acknowledgments. MM undertook this project just for fun. The support for SZ and YD comes from the National Science Foundation Climate and Large-Scale Dynamics (CLD) program through Grants AGS-1354402 and AGS-1445956 and the National Oceanic and Atmospheric Administration through award NA16NWS4680013. We wish to thank the three reviewers for their helpful comments. AMS waives the page charge for this manuscript with the condition that all figures are to be in black and white. This support is much appreciated.

APPENDIX

Finite-Difference Form of the Governing Equations

The nondimensional forms of (4) to (7) are

$$0 \leq z \leq 1, \quad \left( \frac{\partial}{\partial t} + z \frac{\partial}{\partial x} \right) (\phi_{xx} + \phi_{yy} - B \phi_x + \phi_{zz}) + \phi_x (C + B) = 0,$$

(A1)
where the definitions of the parameters $A$, $B$, and $C$ have been given in section 3. Substituting the normal mode solution (16a) into (A1) to (A5), we get a system of ordinary differential equations that govern the amplitude function $x(z)$ of an ordinary mode. The resulting equations from (A1) and (A3) are then cast in center-difference scheme. The first derivatives in (A2), (A4), and (A5) are cast in either forward or backward difference schemes. The resulting expressions are collectively put in a matrix form of an eigenvalue–eigenvector equation, (17).

The nondimensional forms of (10) to (14) are

$$0 \leq z < 1, \quad \nabla^2 w - Bw + w_{zz} = 2\nabla^2 \phi_x + C\phi_{xz},$$

$$\text{at } z = 0, \quad w = 0,$$

$$1 < z \leq z_{z_{top}}, \quad \nabla^2 w - ABw_x + Aw_{xz} = AC\phi_{xz},$$

$$0 \leq z \leq z_{z_{top}} \leq z_{top}, \quad \frac{\partial}{\partial t} \left( \phi_x + \phi_y - AB\phi_x + A\phi_{yz} \right) + C\phi_x = 0,$$

$$\text{at } z = 0, \quad \phi_y = 0,$$

$$\text{at } z = z_{z_{top}}, \quad \phi_x = 0,$$

$$\text{at } z = 1, \quad A\phi_{xz} = \phi_x - \phi_z,$$

$$\quad \quad + \phi_x = 0,$$

where substituting (16a) and (16b) into (A6) to (A10), we get a system of ordinary differential equations that relate the amplitude function $w(x, z)$ to $\chi(z)$. The counterpart finite-difference form of those equations is the matrix Eq. (18).

**REFERENCES**


