Baroclinic instability at the winter stratopause

By A. J. SIMMONS

U.K. Universities' Atmospheric Modelling Group, Department of Geophysics, University of Reading

(Manuscript received 23 May 1974)

SUMMARY

Baroclinic instability is examined in an idealized model of the winter stratopause. The upper stratosphere is more stably stratified than the lower mesosphere, and instability can occur when the zonal flow has westerly shear. Using representative parameter values, a maximum doubling time of 2 days is found for a disturbance with zonal wavelength 4000 km and vertical scale 10 km, moving with a velocity within $2 \text{ m s}^{-1}$ of that of the zonal wind at the stratopause. Newtonian cooling reduces growth-rates by some 80% of the damping rate, and becomes important for relaxation times of less than 7 days.

1. INTRODUCTION

Theoretical studies of the winter circulation of the upper atmosphere have concentrated mostly on the behaviour of disturbances of planetary scale. Charney and Drazin (1961) demonstrated that only such disturbances can penetrate significantly from the troposphere into the stratosphere and mesosphere, and subsequent investigations (Matsuno 1970, 1971; Simmons 1974a) have resulted in a description of planetary waves in good agreement with present observations. Although in essence forced from below, these waves are found to extract significant amounts of energy from the stratosphere itself. In some idealized models this is found to be the principal source of disturbance energy (Simmons, loc cit.).

In addition to studies such as the above, it is of interest to consider whether energy may also be released due to instabilities of the circulation of the upper atmosphere. For the lower stratosphere, McIntyre (1972) has re-examined Murray's (1960) model of the polar-night jet. He indeed found baroclinic instability, but growth-rates were small, with doubling times of a week or more. The faster-growing disturbances had zonal wavenumbers between 5 and 9, and vertical scales of only a few kilometres.

Other regions favourable to baroclinic instability are the winter stratopause and mesopause (Green 1972). Climatological sections of zonal wind for winter indicate westerly shear at the mid-latitude stratopause, and easterly shear at the mesopause (Batten 1961, Murgatroyd 1969). Regarding the more (statically) stable layers above and below the mesosphere as acting in a manner similar to rigid horizontal boundaries, Green points out that instability is to be expected at both boundaries, as in examples discussed earlier by him (Green 1960).

In this paper, we present specific calculations for the model suggested by Green. Attention is concentrated on the winter stratopause, where we indeed find significant instability. The mechanism considered here gives rise to a much weaker instability at the winter mesopause.

2. THE MATHEMATICAL MODEL

As is common in the study of baroclinic instability, we assume quasi-geostrophic motion on a mid-latitude $\beta$-plane, and examine the initial growth of small-amplitude perturbations to a zonal current that varies only with height. Our basic model is illustrated in
Fig. 1. The model distribution of zonal wind and static stability.

Fig. 1. The winter stratopause is taken to be at $z = 0$, a level which separates two unbounded fluids of uniform, but different, static stability:

$$N_1^2 \text{ in } z < 0, \quad \text{and} \quad N_2^2 \text{ in } z > 0,$$

with $N_1^2 > N_2^2$. Axes are taken moving with the zonal flow at the stratopause, and this flow is given elsewhere by $\bar{u} = \Lambda z$, $\Lambda$ a constant vertical shear. The density scale height is taken as a constant, $H_s$.

In view of the likely importance of radiative and photochemical damping in the upper stratosphere and mesosphere (see Dickinson 1973, Blake and Lindzen 1973), we do not assume adiabatic motion, but include Newtonian cooling (Dickinson 1969) such that a term $-T''/t_e$ is added to the right-hand side of the equation for rate of change of perturbation temperature, $T''$. The cooling time, $t_e$, is specified independent of height. This and other assumptions of uniformity and unboundedness are justified by our subsequent finding that the faster-growing disturbances possess relatively short vertical scales.

As shown in appendix A, the resulting instability problem may be reduced to that of finding the complex root of a single equation involving confluent hypergeometric functions. Solutions have been determined numerically, and supplemented by approximate analytical solutions. These are presented in appendix B.

3. ADIABATIC SOLUTIONS

Unstable disturbances occur in the present model much as in that of Charney (1947), the role of the lower boundary in Charney’s model being played here, less efficiently, by the lower layer of more stable fluid. Graphs showing the dependence of phase-speed and growth-rate on wavenumber are displayed in Fig. 2. For these calculations the motion is assumed adiabatic, and we have used the values: $f = 10^{-4} \text{s}^{-1}$, $\beta = 1.6 \times 10^{-11} \text{ m}^{-1} \text{s}^{-1}$, $N_1^2 = 5 \times 10^{-4} \text{s}^{-2}$, $N_2^2 = 2.5 \times 10^{-4} \text{s}^{-2}$, $\Lambda = 4 (\text{m s}^{-1}) \text{ km}^{-1}$ and $H_s = 7 \text{ km}$. Also illustrated are analytical solutions obtained using the short-wavelength approximation described in appendix B. Agreement is seen to be good, particularly with respect to growth-rate, for which the analysis has been carried out to a higher order, and the analytical solutions may be used to gain some idea of the sensitivity of the solutions presented here to the value of various parameters.
From Fig. 2 it is seen that phase-speeds for all wavelengths differ by less than $2 \text{ m s}^{-1}$ from the speed of the zonal flow at the stratopause. The fastest growth-rate is $0.33 \text{ (day)}^{-1}$, corresponding to a doubling time of $2.1 \text{ days}$, and occurs at a wavelength of $4200 \text{ km}$. For comparison with Charney's model, we note that a fastest growth-rate of $0.71 \text{ (day)}^{-1}$

Figure 3. The variation with height of the amplitude, $|\phi|$, phase, $\theta(z)$, and dimensionless poleward eddy heat transport, $\frac{1}{\psi} |\phi|^2 \theta_n$, for the fastest-growing adiabatic disturbance.
occurs at a wavelength of 6000 km when the fluid in \( z < 0 \) is replaced by a rigid horizontal boundary at \( z = 0 \).

Aspects of the structure of the fastest-growing disturbance are presented in Fig. 3. The disturbance has a maximum 2 km into the mesosphere and penetrates considerably further into the mesosphere than into the stratosphere, a result in part due to the decreased static stability of the upper layer, and in part due to the decrease of density with increasing height. The disturbance has an amplitude in excess of half its maximum value over a height of 11 km. Significant phase tilts are confined closer to the stratopause, and the heat transport mostly takes place in a layer some 2 km deep.

In agreement with a general result due to McIntyre (personal communication), the eddy heat transport is equatorward in \( z < 0 \), and in this region there is thus a transfer of energy from the disturbance to the mean state. The vertical shear of the zonal flow in \( z < 0 \), far from enhancing the instability, is seen to hinder disturbance growth. The conversion from zonal to eddy available potential energy in \( z > 0 \), in addition to supplying energy for the growth of the disturbance in \( z > 0 \), has to supply energy both for the growth in \( z < 0 \), and for overcoming the reversed conversion in this lower region. Consistently, growth-rates are less than found when a rigid boundary is placed at \( z = 0 \).

For a maximum eddy velocity of 10 m s\(^{-1}\), the maximum dimensional poleward heat transport is 31 m degC s\(^{-1}\). This is of the order of the average transport at low levels in the mid-latitude troposphere, and in excess of the transport in the stratosphere below 30 km (see Newell et al. 1969, Fig. 12). Eddy transports at the winter stratopause owing to baroclinic instability are thus likely to be locally of significant magnitude. On account of their limited vertical scale, they are, however, unlikely to contribute significantly to the overall heat transfer in the stratosphere and lower mesosphere.

### 4. Solutions with Newtonian Cooling

In view of uncertainties regarding damping rates near the stratopause (see Blake and Lindzen 1973), we have determined solutions with a variety of cooling times. Growth-rates as functions of wave number are shown in Fig. 4 for the same parameter values as used in the preceding section. Newtonian cooling is seen to have little influence on the wavelength of the most unstable disturbance. Its doubling time is increased from the adiabatic value of 2-1 days to 3-1 days for a cooling time of 7 days, and 6 days for a cooling time of 3-5 days.

![Figure 4. The dependence of growth-rate on wavenumber for solutions with Newtonian cooling.](image)
Newtonian cooling acts directly to dissipate perturbation potential energy, but influences the perturbation kinetic energy only indirectly through the modified structure of the disturbance. The reduction in growth-rate due to Newtonian cooling with time-scale \( t_c \) is found generally to be less than the decay rate, \( 1/t_a \), that would occur for a temperature perturbation in the absence of dynamical effects. For large \( t_c \), and \( N_1/N_2 = \sqrt{2} \), it is shown in appendix B that growth-rates are reduced by 0-81/\( t_c \) in the short-wavelength limit. Reductions of a similar order are found for all wavelengths in the numerical solutions shown in Fig. 4. Evaluation of the energetics of the fastest-growing disturbances shows the reduction in growth of disturbance energy due to Newtonian cooling to be accounted for in part by the direct dissipation of potential energy, and in part by a reduced conversion from zonal to eddy available potential energy. This contrasts with the case of steady forced disturbances, which are found to be maintained in the presence of an increased Newtonian cooling largely by an increased conversion of energy from the mean state (Simmons 1974a).

5. The influence of an upper region of easterly shear

Observations made by rocket at Fort Churchill (59° N), and discussed by Leovy and Ackerman (1973), are indicative of the presence of synoptic-scale disturbances near the stratopause. These observations are discussed further in section 7. For the present, it is relevant to note that they do not indicate a general westerly shear of the zonal wind extending to distances much above the stratopause.

The occurrence of westerly shear at the stratopause is essential for instability of the order found here. However, although the disturbance amplitude shown in Fig. 3 is of significant magnitude some 10 km into the mesosphere, most of the energy conversion occurs within 2 km of the stratopause. It is thus unclear how far the westerly shear must extend into the mesosphere in order that significant growth-rates can occur. We have therefore determined solutions for a zonal flow

\[
\bar{u} = \Lambda z, \quad -\infty < z < H, \\
= -\Lambda z + (\Lambda + \Lambda_x)H, \quad z > H,
\]

with the distribution of static stability as before. The dependence of growth-rate on wavenumber is shown in Fig. 5 for three values of \( H \). The cooling time is 7 days, and the easterly shear above \( z = H, \Lambda_x \), is taken as 2-5 (m s\(^{-1}\))km\(^{-1}\), a value which precludes any instability.

![Figure 5](image_url)

Figure 5. The dependence of growth-rate on wavenumber for disturbances to a zonal flow with reversed vertical shear a distance \( H \) above the stratopause.
associated primarily with the reversal of the vertical shear. Other parameter values are as before.

When the westerly shear extends 10 km into the mesosphere, growth-rates as shown in Fig. 5 are almost identical with those for a cooling time of 7 days shown in Fig. 4. For smaller $H$, there is a stabilization of longer wavelengths, and the wavelength of maximum instability is shifted towards a value such that the vertical scale of the disturbance is of order $H$. Maximum growth-rates are not, however, reduced, but in fact show an increase, this amounting to a factor of a third when $H = 2.5$ km.

6. INSTABILITY AT THE WINTER MESOPAUSE

Our calculations do not support the suggestion that a similar instability occurs as a result of the conjunction of easterly shear and an increase in static stability at the winter mesopause. The difference is due to compressibility. For a Boussinesq fluid, it is easy to verify that growth-rates and phase-speeds are unchanged in the model of section 2 when the static stabilities $N_1^2$ and $N_2^2$ are inter-changed, and the sign of the vertical shear reversed. However, for a finite value of the density scale height, $H_s$, Pedlosky's (1964) necessary conditions for instability show that the mechanism proposed by Green (1972) can work only when the easterly shear, $\Lambda$, satisfies

$$\Lambda < \frac{\beta N_1^2 H_s}{f^2}.$$  

Taking $\beta = 1.6 \times 10^{-11}$ m$^{-1}$ s$^{-1}$, $N_1^2 = 2.5 \times 10^{-4}$ s$^{-2}$, $H_s = 7$ km and $f = 10^{-4}$ s$^{-1}$, we find stability when $\Lambda > 2.8$ (m s$^{-1}$)km$^{-1}$. With $N_2^2 = 5 \times 10^{-4}$ s$^{-2}$, the fastest-growing adiabatic disturbance occurs for an easterly shear of 1.5 (m s$^{-1}$)km$^{-1}$, and a wavelength of 7000 km. Its growth-rate is only 0.079 (day)$^{-1}$. No unstable modes are found for any value of the vertical shear when Newtonian cooling is included with a time-scale of 10 days or less.

7. CONCLUSIONS

The instability that has been considered in this paper is associated with growth-rates rather less than found for the corresponding tropospheric model of Charney (1947). Nevertheless, the growth-rates which have been determined indicate a strong likelihood that baroclinic instability will give rise to disturbances with zonal wavelengths of about 4000 km and vertical scales close to 10 km, concentrated at the winter stratopause, and moving with a speed close to that of the zonal flow there. In the absence of dissipation, such disturbances double in amplitude over a two-day period when the vertical shear is 4 (m s$^{-1}$)km$^{-1}$. With Newtonian cooling, doubling times of under three days are found for relaxation times greater than seven days. The doubling time is extended to six days for a relaxation time of 3$\frac{1}{2}$ days.

Recent calculations of time-scales over which temperatures relax towards equilibrium values have tended to shorten previous estimates. Dickinson (1973) finds a time-scale of 5 days for infra-red cooling near the stratopause. Values as short as 1$\frac{1}{2}$–2$\frac{1}{2}$ days are found with the inclusion of photochemical effects (Blake and Lindzen 1973), the precise influence of photochemistry depending on uncertain reaction and cooling-rate coefficients, and becoming less near the winter pole. Such relaxation times have a significant effect on computed growth rates, and uncertainties in damping-rate thus directly result in uncertainties regarding the growth-rate of possible instabilities. Spatial scales and phase-speeds are, however, insensitive to the amount of dissipation.
The instability considered here depends critically on the presence of westerly shear at the stratopause, but this need not extend more than a few kilometres into the mesosphere. We have entirely neglected any variation with latitude of either the basic state or the disturbance. Our solutions may yield a first approximation to the growth-rate, phase-speed and vertical structure of disturbances to a zonal flow with meridional variation of planetary scale (Simmons 1974b), but the influence of more marked meridional variations is unclear.

Detailed observation of disturbances of the scale indicated by our results is difficult. Vertical and zonal scales are both too small to be resolved by temperature soundings made using the current series of Nimbus satellites (see Chapman et al. 1974). Since the predicted disturbances move with the speed of the zonal flow near the stratopause, periods relative to the surface of between 1/4 and 1 day are implied by a wavelength of 4000 km and stratopause velocity between 50 and 100 m s\(^{-1}\). Data presented by Green (1972) show transient velocities to be relatively large near the stratopause (and also near the mesopause). The periods associated with these transients are, however, unknown. Rocket observations at a single station analysed by Leovy and Ackerman (1973) reveal disturbances with vertical scales similar to those found here. Periods were apparently in the range of 1/2 to 4 days, but the phase-speeds determined herein suggest that these periods may be subject to aliasing owing to the use of daily observations, as recognized by Leovy and Ackerman. In view of such difficulties, the results of further observations of the winter stratopause are awaited with interest.

ACKNOWLEDGMENTS

The model considered here was suggested by Dr. J. S. A. Green and reached the author via Dr. R. S. Harwood and Dr. M. E. McIntyre. Comments from Dr. J. R. Holton and financial support from the Natural Environment Research Council are gratefully acknowledged.

REFERENCES


McIntyre, M. E. 1971 'A dynamical model of the stratospheric sudden warming,' Ibid., 28, pp. 1479–1494.


APPENDIX A  THE EIGENVALUE PROBLEM

Taking \( t \) as time, and \( x \) and \( z \) as eastward and upward co-ordinates on a mid-latitude \( \beta \)-plane, we introduce the quasi-geostrophic stream function

\[
Re\{ \phi(z)e^{i(kx-ct)} \}
\]

for a perturbation of zonal wavenumber \( k \). With \( f \) the Coriolis parameter, \( \beta \) its northward gradient, and other parameters as in section 2, we introduce the dimensionless parameters

\[
\delta = \frac{N_1}{N_2}, \quad \mu = \frac{f}{N_2 k H_s}, \quad \beta_1 = \frac{\beta N_2^2 H_s}{f^2 \Lambda} \quad \text{and} \quad \chi = \frac{N_2}{f \Lambda c}.
\]

We scale \( z \) by \( f/N_2 k \) and \( c \) by \( \Lambda f/N_2 k \).

The governing equation for \( \phi \) (Dickinson 1969) then becomes

\[
\begin{aligned}
(z - c')(\phi_{zz} - \mu \phi_z - \phi) + \{\mu(1 + \beta_1) - i \chi\} \phi &= 0, \quad z > 0, \\
(z - c')(\phi_{zz} - \mu \phi_z - \delta^2 \phi) + \{\mu(1 + \delta^2 \beta_1) - i \delta^2 \chi\} \phi &= 0, \quad z < 0,
\end{aligned}
\]

where \( c' = c - i \chi \).

Boundary conditions are

\[
\phi \to 0 \quad \text{as} \quad z \to \pm \infty
\]

and continuity of pressure and vertical velocity across \( z = 0 \) gives

\[
1 + c\left( \frac{\phi_z}{\phi} \right)_{z=0-} = \delta^2 \left\{ 1 + c' \left( \frac{\phi_z}{\phi} \right)_{z=0+} \right\}
\]
Solutions to Eqs. (A.1) are well known in terms of confluent hypergeometric functions (Charney 1947). Imposing the boundary conditions (A.2) we obtain

\[ \phi = \zeta e^{-iz} U(1 - \gamma, 2, \zeta), \quad z > 0, \]
\[ \phi' = \zeta' e^{iz} U(1 - \gamma, 2, \zeta'), \quad z < 0, \]

(A.4)

where

\[ \zeta = 2 \sqrt{1 + \frac{\mu^2}{4}(z - c')}, \quad \zeta' = -2 \sqrt{1 + \frac{\mu^2}{4}(z - c')} \]
\[ \lambda = \sqrt{1 + \frac{\mu^2}{4} - \mu/2}, \quad \lambda' = \sqrt{1 + \frac{\mu^2}{4} + \mu/2} \]
\[ \gamma = \frac{\mu(1 + \beta_1) - i\lambda'}{2 \sqrt{1 + \frac{\mu^2}{4}}}, \quad \text{and} \quad \gamma' = -\frac{\mu(1 + \delta^2 \beta_1) - i\lambda'}{2 \sqrt{1 + \frac{\mu^2}{4}}} \]

The notation of Abramowitz and Stegun (1964) has been used for the confluent hypergeometric function \( U(a, b, z) \).

Substituting the solution (A.4) into (A.3) we obtain the eigenvalue equation

\[ 1 + c\lambda' - \gamma' \frac{U(1 - \gamma', 2, 2c\sqrt{\delta^2 + \frac{\mu^2}{4}})}{U(1 - \gamma, 2, 2c\sqrt{\delta^2 + \frac{\mu^2}{4}})} = \delta^2 \left\{ \frac{U((1 - \gamma, 1, -2c\sqrt{1 + \frac{\mu^2}{4}})}{U((1 - \gamma, 2, -2c\sqrt{1 + \frac{\mu^2}{4}})} \right\}. \]

(A.5)

This has been solved numerically for \( c \), the confluent hypergeometric functions being evaluated from series definitions. As an initial estimate of \( c \), we utilized the approximate analytical solutions presented below.

**APPENDIX B **APPROXIMATE ANALYTICAL SOLUTIONS

Approximate analytical solutions may be obtained for short-wavelength disturbances using either the formulation due to McIntyre (1970), or a direct expansion of (A.5) for small values of \( \mu \). A detailed derivation using the latter method has been given elsewhere (Simmons 1972). For adiabatic motion, we find, for \( \delta > 1 \) and \( (1 + \beta_1) > 0 \)

\[ \text{Re}(c) = c_0 \left\{ 1 + \mu \left( \frac{c_0}{2} - \frac{E + E'}{2 - c_0} \right) \right\} + 0(\mu^2), \]

and

\[ \text{Im}(c) = \frac{\pi \mu c_0}{(2 - c_0)} (1 + \beta_1)e^{-2c_0} \left\{ 1 - \mu \left( \frac{c_0^2 + F + 2E'(1 + \beta_0 - c_0)}{(2 - c_0)(1 - c_0)} - \frac{3c_0E}{2 - c_0} \right) \right\} + 0(\mu^2), \]
where

\[ c_0 = \frac{\delta - 1}{\delta} \]

\[ E = (1 + \beta_1) e^{-2c_0} Ei(2c_0) \]

\[ E' = \frac{1}{\delta^2} (1 + \delta^2 \beta_1) e^{2d_{c_0}} E_1(2\delta c_0) \]

and

\[ F = (1 + \beta_1)(\gamma_E + \log 2c_0) \]

Here \( \gamma_E \) is Euler’s constant, and \( Ei \) and \( E_1 \) are exponential integrals (Abramowitz and Stegun 1964, Chapter 5).

The solution given above has been seen in Fig. 1 to compare well with a numerical solution over a significant range of wavenumbers. The short-wavelength approximation is complemented at longer wavelengths by expansions which may be performed near integral values of the parameter \( \gamma \). For Fig. 1, \( \gamma = 1 \) when \( k = 0.62 \times 10^{-6} \text{ m}^{-1} \).

The neutral mode \( (c = 0) \) which occurs at this parameter value has zero amplitude in \( z < 0 \), and is identical in \( z > 0 \) with that for the problem of Charney (1947). Approximate analytical solutions in the neighbourhood of \( \gamma = 1 \) have been determined in the manner of Miles (1964), and agree well with numerical results.

The short-wavelength analysis may be extended for values of the dimensionless cooling parameter, \( \chi \), small compared with unity. This is a relevant expansion since with \( f = 10^{-4} \text{ s}^{-1}, N_2^2 = 2.5 \times 10^{-4} \text{ s}^{-2}, \Lambda = 4 \text{ (m s}^{-1})\text{km}^{-1} \) and \( t_c = 7 \text{ days}, \chi = 0.065 \). To lowest order, it is found (Simmons 1972) that the growth-rate is reduced by the dimensional amount

\[ \frac{1}{t_c} \left\{ 1 - \frac{c_0}{2 - c_0} (e^{-2c_0} Ei(2c_0) + e^{2\delta c_0} E_1(2\delta c_0)) \right\}. \]

For the value \( \delta = \sqrt{2} \) used throughout, this is \( 0.81/t_c \).