Studies of the Atmosphere Using Aerospace Probings

Annual Report-1969

Space Science and Engineering Center

The University of Wisconsin

November, 1970

COVER PHOTOGRAPHS

The overall picture of cyclone behavior supplied by the ATS spin-scan cameras inspired a cyclone model. One of the characteristics of this model is that a necessary precondition of achieving a logarithmic cloud tracer in the horizontal is the existence of a decreasing logarithmic vertical velocity profile, $w(r) = c_o - c_1 \ln \frac{r}{r_o}$ where $w(r)$ is the azimuthal average of the vertical velocity as a function of radius and c_o and c_1 are constants with limits of LT^{-1} and r_o is a scaling factor.

The upper picture is a dishpan model with holes punched into it to simulate this condition. The marker assumes the form of a log spiral as it comes to an equilibrium state. The lower picture is a rectified version of a cyclone in the North Central Pacific which is also a log spiral. The similarities between the two are striking.

Space Science and Engineering Center
The University of Wisconsin
Madison, Wisconsin

STUDIES OF THE ATMOSPHERE USING AEROSPACE PROBINGS

Annual Report on

ESSA Grant E-230-68-G

1969

The research reported in this document has been supported by the National Environmental Satellite Center of the Environmental Science Services Administration.

November, 1970

304-89

PRINCIPAL INVESTIGATOR

Verner E. Suomi

CONTRIBUTIONS BY

C. Anderson
B. Auvine
J. Bennett
J. Holton
D. Houghton
S. Kaniel
R. Krauss
F. Nicholson
U. Shafrir
B. Shkoller
R. T. Song
J. Wallace
J. Young

University of Wisconsin

PREFACE

The major portion of this annual report is concerned with theoretical studies. Some of these studies were suggested by the satellites' view of the earth which show the tropics as an entity. These results now suggest what one might look for in further studies utilizing the data from meteorological satellites. Thus we have travelled the full circle.

Special mention should be made of the contribution made by those not directly supported by the grant, Drs. J. R. Holton and J. M. Wallace and for use of the powerful computer at the Goddard Institute for Space Studies in New York.

It is again a pleasure to acknowledge the contributions of my associates.

Verner E. Suomi
Principal Investigator

CONTENTS

A METHOD FOR ESTIMATING CYCLONE VERTICAL MOTIONS FROM SATELLITE CLOUD PHOTOGRAPHS

Francis Hilaire Nicholson

Abstract

The cyclone is treated from the point of view of a symmetrical three-dimensional model, occupying that portion of the atmosphere below the level of nondivergence.

A tangential and radial wind field is postulated, which allows both the trajectories and cloud patterns to be logarithmic spirals.

From the interaction of these two spirals, using the hypothetical wind field, a vertical velocity field is established by means of considerations of conservation of mass.

A physical analogue is generated in the laboratory which gives remarkably similar results to those observed in the model.

The properties of the cyclone model are examined, particularly with regard to the pressure gradient, the vorticity patterns, and the relative packing of isotherms along the fronts within the cyclone.

Contents

1. INTRODUCTION

From the satellite photographs, cloud spirals can be observed to be a definite property of both midlatitude and tropical cyclones. These spirals very frequently take an approximately logarithmic form. This study is designed to test a simple kinematic cyclone model for relevance to the observed behavior of one particular cyclone observed in the mid-north Pacific in January of 1967.

The idea is to fit mathematically a curve to the spiral observations from the satellite for this particular cyclone over a period of five days and to observe the behavior of this fitted curve. We intend to demonstrate that a relatively simple model of a combination vortex-sink will enable us to relate the movement of this curve to the vertical-velocity profile in a realistic manner, even with other data being absent.

1.1 A History of the Study of Spiral Bounds

To the author's knowledge, the derivation of vertical velocity by investigation of spiral systems from the satellite has not been previously attempted, although the general field of storm spirals itself has been well studied. Shaw (1923) stated that the spiral bands in cyclones were logarithmic spirals. Fujiwhara (1921) has studied the tendencies of vortices to achieve symmetry.

Fujita (1958), Kessler (1957), and Senn and Hiser (1959) have all studied rain bands in hurricanes.

For years radar meteorologists, for example Kessler, have been fitting logarithmic spirals to the hurricane rain bands observed on PPI scopes (Plan Position Indicator). Twitchell et al. (personal communication) did the same with aerial photographs of Typhoon Ida.

Krishnamurti (1966) has studied the hurricane from the viewpoint of conservation of mass and momentum and has related these two to the spiral bands observed in the system.

Tang et al. (1964) point out that theirs is the first three-dimensional model of a cyclone developed, previous theoretical formulations having been limited to two dimensions. They compute both radial and azimuthal wave numbers allowing a complete description of the horizontal pattern formed by the cloud bands. The patterns predicted by their theories are predicted "in excellent agreement with the cloud patterns formed in the real atmosphere" (Tang, Brooks and Watson, 1964).

Tang et al. make an attempt "to relate the orientation, width, and spacing of the vortex cloud bands to the horizontal and vertical variations of wind and temperature in the vicinity of these clouds."

Glaser (1961) suggested that a dying cyclone is evidenced by the evolution of an asymmetric spiral band pattern into a symmetric arrangement of nearly concentric circular bands.

Tang et al. conjecture that the evolution of a cyclone's cloud pattern can be explained primarily by the fields of wind and/or vertical motion.

"The cyclone bands and clear sector in the Tiros photographs have the same configuration as the moist and dry tongue of an isentropic chart of an occluding cyclone. Soundings show that the clear sector contains a cold, dry air mass between warmer, moist air masses, which Nagle and Serenbreny (1962) attributed to advection.

"Tropical cyclones have cyclonic cloud bands and clear sectors on a smaller scale and with greater symmetry than those of extratropical cyclones. A typical cloud band of mature hurricane forms an approximately logarithmic spiral (that is, inflow angle independent of the distance from the center)" (Tang et al., 1964).

2. THE KINEMATIC MODEL

The model which is postulated is essentially very simple, as shown in Fig. 1a. It is three-dimensional, axially symmetric, pertaining only to the atmosphere below the level of nondivergence. We use polar coordinates. The atmosphere is assumed to be incompressible in the horizontal. It is not steady state, meaning that the local derivative of any of the quantities describing the model is not necessarily constant with time.

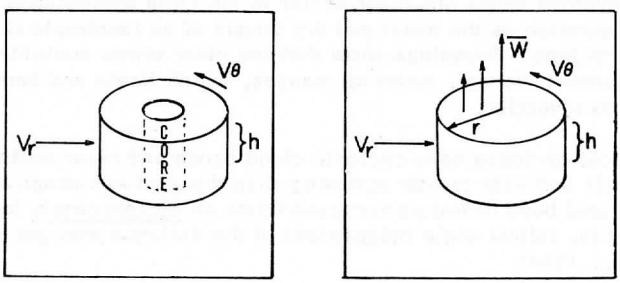
The model is divided into two sections: the small inner section, called the core, and the section surrounding this, the main body of the storm.

The model, not being steady state, is subject to unbalanced torques per unit mass affecting the tangential velocity within the model.

2.1 The Main Storm

The objective of this model is the specification of the volume flux for all three coordinates in a cylindrical coordinate system. Not the least of these is the specification of a vertical velocity field, w. In order to do this we must specify a tangential and radial velocity field and seek to justify them from the geometry they produce. If it is established that these are not arbitrary but indeed have a sound basis in actual fact, then we may with some confidence derive a vertical-velocity field corresponding to these other two from considerations of mass conservation.

We will first treat the main body of the storm only and ignore the core. For purposes of simplification we will assume that the core extends from the very center of the storm outward a small radial distance. The main body of the cyclone extends from r_0 to infinity.



a) With core b) With volume flux

Fig. 1. Cyclone Model

A successful model must fulfill the boundary condition that for $r > r_0$, the tangential velocity, V_θ, must be some multiple of the radial velocity V_r in order to account for logarithmic spirals observed in the cloud patterns. A further boundary condition requires that the vertical velocity, w, be linearly proportional to height, z. This implies that the divergence of the horizontal wind field is assumed to be constant in z.

Thus $w = cz$; $\frac{\partial}{\partial z}(\nabla \cdot V_H) = 0$ where w is the vertical velocity and c is a constant in units of T^{-1}. In order to relate the radial velocity of the cyclone to the vertical velocity, it is necessary to obtain the area average of the vertical velocity, $\bar{w}$.

$$\bar{w} = [\int_{r_0}^{r} w(r)dr + K]\frac{1}{r}$$

where K is the total volume efflux of the core.

The cylinder (Fig. 1b) has the volume influx in through the sides equal to the volume efflux out of the top such that $-2\pi rhV_r = \pi r^2 \bar{w}$, which simplifies to $V_r = \frac{-r\bar{w}}{2h}$.

Since at any level the radial velocity $-V_r$ can be related to the area average of $\bar{w}$, we may then ultimately relate it to the vertical velocity itself.

Because of horizontal incompressibility, the radial velocity $-V_r$ is related to the area average of the vertical velocity $\bar{w}$ by $-V_r 2\pi rh = \pi r^2 \bar{w}$. From the definition of $\bar{w}$ we have $r\bar{w} = \int_0^r w(r)dr$ where $w(r)$ is the vertical velocity at some radius r.

We can obtain $w(r)$ from the definition of $\bar{w}$. Differentiating with respect to r, we have

$$\bar{w} + r\frac{d\bar{w}}{dr} = w(r) \quad \text{or} \quad \bar{w} + r\frac{d\bar{w}}{dr} = w(r)$$

$$\bar{w} = \frac{-2hV_r}{r} .$$

2.2 Cyclone Parameters

The function we will choose for V_r is $V_r = -r(B - c_2 \ln \frac{r}{r_0})$ where B and c_2 are constants with units of T^{-1}. $\bar{w} = 2h(B - c_2 \ln \frac{r}{r_0})$ so $w(r) = 2h(B - c2 \ln \frac{r}{r_0}) - 2c_2h$ or $w(r) = \bar{w} + 2c_2h$ for the complete cyclone from $z = 0$ to $z = h$. The function we choose for V_θ is $V_\theta = r(A - c_1 \ln \frac{r}{r_0})$ where A and c_1 are constants with units of T^{-1}. By letting $V_\theta = DV_r$, or

more precisely, $A - c_1 \ln \frac{r}{r_0} = D[B - c_2 \ln \frac{r}{r_0}]$, we will achieve the conditions in a model by which not only the streamlines are logarithmic spirals, but the patterns in the cloud fields are log spirals as well (Malkus and Riehl, 1960).

3. THE LOGARITHMIC SPIRAL

Since the understanding of logarithmic spirals is of fundamental importance in this work, we will now review in detail the properties of these spirals as they pertain to the body of this work.

Log spirals can be treated most conveniently in two different coordinate systems. These are polar coordinates and the coordinates of the logarithm of the radius vs. the angle theta. A spiral of this type can be expressed in polar coordinates in a straightforward and simple manner. In the semilog coordinates the spiral assumes its most simple geometry—a straight line.

We will first discuss the applications in the polar coordinate system. The slope of any curve in polar coordinates is given by the tangent of the angle of intersection ψ of the radius vector $\vec{r}$, with the tangent $\overline{CD}$ to the curve $\widehat{AB}$ at the point of intersection of the vector $\vec{r}$ and the curve, $\widehat{AB}$, as is shown in Fig. 2. This tangent is given as $r \frac{d\theta}{dr}$ (Thomas, p. 292). When the slope is a constant relative to the radius as a function of θ, then the curve is defined by the slope and a known point on the curve.

The equation of a logarithmic spiral is $\alpha \ln \frac{r}{r_0} = (\theta - \theta_0)$ where the coordinates of the known point are R_0, θ_0, and α is a dimensionless constant. Differentiating this equation we obtain $\alpha d \ln r = d\theta$ or $r \frac{d\theta}{dr} = \alpha = \tan \psi \equiv$ slope.

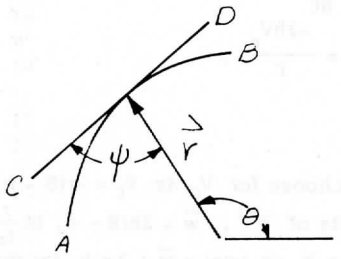


Fig. 2. Slope of a curve in polar coordinates

That a log spiral has a constant slope is easily visualized if we plot the curve on semilog paper so that the ordinate is in terms of the ratio of r at θ to r_0 at θ_0 and the abscissa is $\theta - \theta_0$ in degrees or radians. For such a plot, a logarithmic spiral becomes a straight line, which constitutes a simple test of fit. Clearly, another logarithmic spiral of different properties would also be a straight line on the semilog graph, but with a different slope.

It becomes evident then that a ray in polar coordinates is a degenerate form of a logarithmic spiral because it has a constant slope of zero. The equation for a ray is $\theta = c$ where c is a dimensionless constant. The slope is $r\frac{d\theta}{dr} = 0$. A tracer of ink or cloud may initially be a ray, and thereafter be twisted into a spiral. On the semilog graph the straight line that was the ray would simply rotate, changing its slope.

It follows then that time sequential positions of the tracer can be plotted on a log r vs. theta coordinate system as we have been discussing. If the tracer produces a logarithmic shape we have a series of straight lines with different slopes. It is the change in slope which will provide the information about the motion field. One can visualize that the motion normal to the spiral is the motion which alters the spiral.

If the angular velocity $\frac{d\theta}{dt} = A - c_1 \ln\frac{r}{r_0}$ as in Fig. 3 where A and c_1 are constants with units of T^{-1}, and the velocity along the semilog coordinate is $\frac{d \ln r}{dt} = B - c_2 \ln\frac{r}{r_0}$ and the two are related by $\frac{d\theta}{dt} = D\frac{d \ln r}{dt}$ where D is a simple proportionality factor, then the vector sum of these two is the trajectory, also a logarithmic spiral with a magnitude linearly proportional to $\ln\frac{r}{r_0}$. $|T| = P(B - c_2 \ln\frac{r}{r_0})$ where $|T|$ is the magnitude of the trajectory and P is a proportionality factor.

In this case $P = \text{sectan}^{-1} D$ where D is $d\theta/d \ln r$, $d\theta/d \ln r = V_\theta/V_r = \Gamma/Q$ so $P = \sec(\tan^{-1}\frac{\Gamma}{Q})$, where $\Gamma = 2\pi r V_\theta$ and $Q = 2\pi r V_r$.

In Fig. 4 we have the case of a spiral at Time I being transformed into another spiral at Time II, each particle of the spiral moving under the influence of the trajectories, which are also log spirals and which are parallel lines in this θ, $\ln\frac{r}{r_0}$ coordinate system.

We are concerned with the component of the trajectory normal to the spiral, as this is the portion which rotates the spiral. This normal component itself has a theta component which can be measured. This component is $\Delta\theta$. $d\theta/dt$ can also be measured, as can the angle ϕ as shown in Fig. 6 which the intermediate spiral makes with the θ coordinate.

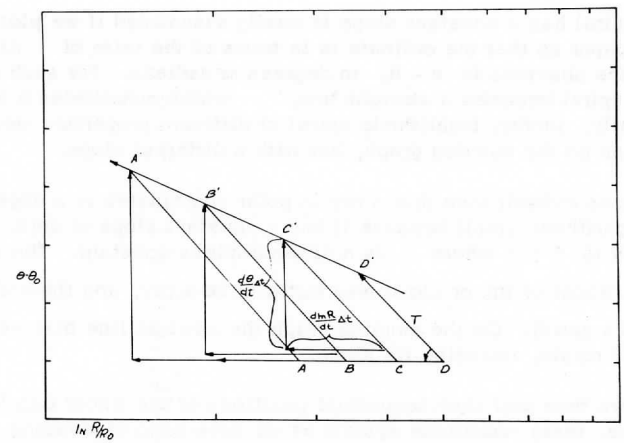


Fig. 3. Fluid motion field of theoretical model

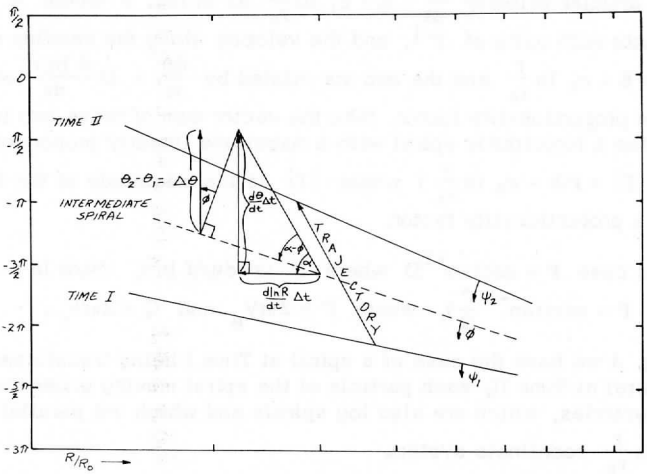


Fig. 4. Movement of spirals under the influence of trajectories

The length of the trajectory itself over the time Δt is given by $\theta \Delta t \csc \alpha$ where $\alpha = \tan^{-1} d\theta/d \ln R$. The product of this trajectory length and the sine of the angle between the trajectory and the intermediate spiral is the normal displacement to the intermediate spiral. The angle referred to is $\alpha - \phi$ where α is $\tan^{-1} \frac{\Gamma}{Q}$ where $\frac{\Gamma}{Q} = \frac{d\theta}{dt} \Delta t / \frac{d \ln R}{dt} \Delta t$.

We can now utilize this to derive the vertical velocity w anywhere in the model. For a particular radial value, r, the displacement of the spiral at Time I to that at Time II along the theta coordinate is

(3.1) $$\theta_2 - \theta_1 = \dot{\theta} \Delta t \csc \alpha [\sin (\alpha - \phi)] \cos \phi$$

where $\phi = \tan^{-1} \frac{\theta_2' - \theta_1'}{\ln r_1/r_2}$ on the intermediate spiral, and $\dot{\theta} \Delta t$ is the angular displacement at r, $\frac{d\theta}{dt}$.

$$\sin(\alpha - \phi) = \sin \alpha \cos \phi - \cos \alpha \sin \phi$$

$$\theta_2 - \theta_1 = \dot{\theta} \Delta t [\cos^2 \phi - \operatorname{ctn} \alpha \sin \phi \cos \phi]$$

$$\frac{\theta_2 - \theta_1 + \dot{\theta} \Delta t \cos^2 \phi}{\dot{\theta} \Delta t \sin \phi \cos \phi} = \operatorname{ctn} \alpha = \frac{Q}{\Gamma}$$

$$Q = \Gamma \left[\frac{\theta_2 - \theta_1}{\dot{\theta} \Delta t \sin \phi \cos \phi} + \operatorname{ctn} \phi \right]$$

$$Q = \Gamma \operatorname{ctn} \phi \left[\frac{\theta_2 - \theta_1}{\dot{\theta} \Delta t \cos^2 \phi} + 1 \right]$$

but

(3.2) $$\bar{w} = \frac{Qh}{\pi r^2} \quad \text{and} \quad \Gamma = (\bar{\zeta} + f) \pi r^2$$

$$\bar{w} = h(\bar{\zeta} + f) \operatorname{ctn} \phi \left[\frac{\theta_2 - \theta_1}{\dot{\theta} \Delta t \cos^2 \phi} + 1 \right]; \quad w = \bar{w} - 2c_2 h$$

(3.3) $$\therefore \; w = h(\bar{\zeta} + f) \operatorname{ctn} \phi \left[\frac{\theta_2 - \theta_1}{\dot{\theta} \Delta t \cos^2 \phi} + 1 \right] - 2c_2 h$$

where $\bar{\zeta}_r$ is the area average of the relative vorticity.

4. THE LABORATORY MODEL

We have developed a theoretical model by considerations of what the conditions would be that would cause the generation of a logarithmic spiral. We now wish to test these conditions to see if they do indeed generate a log spiral, using a laboratory model. This physical analogue is not identical to the theoretical model in all its boundary conditions, but we will demonstrate that the differences are unimportant.

4.1 Apparatus

The apparatus consists of two disphans nested together, a lazy susan, a small bobber, and some ink. The inner dishpan, labelled #1 in Fig. 5, has holes punched into its bottom at hyperbolic intervals so as to approximate the w field of the model, i.e., $w = 2h(B_0 - c_2 \ln \frac{r}{r_0})$ where B_0 and c_2 are constants with units of T^{-1}. (When $r \simeq r_0$ this is a good approximation.) The #1 pan with holes in it is set into another, #2, and this set is filled with water. The combination is set onto a lazy susan and rotated until the water is in solid rotation. This is indicated by the little float in the water.

Once the water is in solid rotation, a marker of ink is put into the water in the form of a ray, itself a degenerate form of spiral. Then the inner pan is lifted from the combination and allowed to drain. The tangential velocity of the fluid is seen to change quickly to the velocity field prescribed by the theoretical model. There are discrepancies between the simple lab model and the theoretical model.

4.2 Problems in Modeling

4.2.1 Changing Pressure Head

Since the dishpan has only a finite amount of water, the pressure head in the water will be a logarithmic function of time. One might ask if this has a direct bearing on the geometry and produces in some manner a logarithmic spiral in the pan.

Whereas it is true that the fluid moves under the influence of the pressure head in the horizontal, and the horizontal pressure head is related to the radial flow by a logarithmic arrangement of the vertical velocity distribution in the horizontal, the rate at which the log spiral winds up is controlled by the changing pressure head. By changing the time scale, however, we can simulate a constant pressure head to compensate for this change. The theoretical model is not itself steady state, so that it is not necessary to have a constant

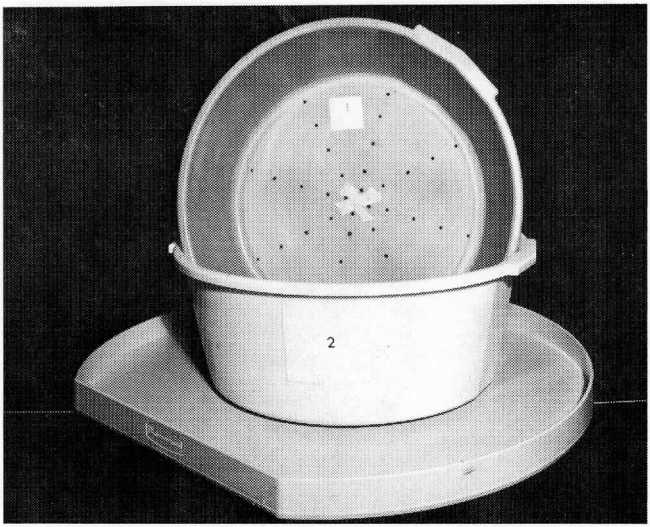


Fig. 5. Laboratory apparatus

pressure head, but by making the pressure head a constant over a short time interval we do not alter the horizontal distribution of pressure for any point in time. The time rate of change of the pressure head, therefore, is not the cause of the spiral, but rather the horizontal distribution as caused by the arrangement of the w field.

4.2.2 Friction

Lateral surface friction also exists in the dishpan, whereas it does not exist in the theoretical model. From rest it takes about 5 minutes to get the water in the dishpan into solid rotation. It takes only 30 seconds to drain. This suggests that the effect of friction is small.

4.2.3 The Tracer Difficulties

During the course of the experiment, we learned that it was very difficult to obtain a true ray for the initial tracer work. Therefore we approximated the ray into two sections. Then we treated the resulting spiral in two sections also. These spirals which were obtained in a series of dishpan experiments are plotted on semilog vs. angle theta paper. An example is shown in Fig. 6.

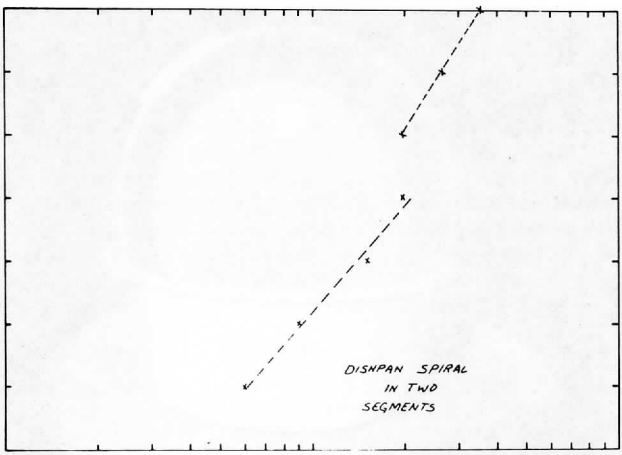


Fig. 6. Two segments of a dishpan spiral as seen in Plate XI

Little difficulty was experienced in obtaining the coordinates of the tracer in the dishpan experiment. Figure 6 shows that two sections of the tracer form fairly good logarithmic spirals since they oscillate about straight lines in this form of presentation.

4.3 The Relative and Absolute Tangential Velocities

In the dishpan experiment, before the pan with holds is lifted from the pan in which it is nested, the motion of the fluid is very nearly solid rotation where the tangential velocity of the fluid particles is proportional to the radius. However, after the inner pan is lifted from the other and allowed to drain, we note that the tangential velocity everywhere is now nearly constant. We need in addition to express this tangential velocity as a velocity relative to the spinning lazy susan much the same as we express winds relative to the earth's surface.

This is expressed as $V_{\theta_r} = V_{\theta_A} - \Omega r$ where Ω is the angular rotation rate of the dishpan, r the radial distance from the center, and V_{θ_A} is the "absolute" tangential velocity of the fluid with respect to the floor of the laboratory.

V_{θ_r}, therefore, is the relative velocity of the fluid to the dishpan. On the surface of the earth the rotation is relative to the earth's vertical component of rotation.

The tangential velocity of the cyclone therefore for $r > r_0$ is to be described as $V_{\theta_r} = V_{\theta_A} - [f/2]r$ where f is the coriolis parameter. Since the absolute tangential velocity is given by $V_{\theta_A} = r(A - c_1 \ln \frac{r}{r_0})$ the relative velocity would then be $V_{\theta_r} = r(A - \frac{f}{2} - c_1 \ln \frac{r}{r_0})$ where $\frac{f}{2} = \Omega \sin \phi$, where ϕ is latitude.

Thus the vertical component of the earth's rotation in the cyclone corresponds to the lazy susan rotation in the model.

5. TEST OF MODEL USING SATELLITE CLOUD PHOTOS

5.1 Technique of Analysis

The objective of the analysis procedure is to obtain from the patterns produced by clouds in cyclones certain parameters which can be utilized in the model. In particular, the patterns we see appear to have the form of logarithmic spirals. Our objective is to approximate the cloud band to a logarithmic spiral derived from these patterns, and fit logarithmic spirals to them.

The ideal spiral would appear to have edges which are zero order discontinuities in the brightness field. This means that there would be no shading of the cloud at the edge, but a very sharp and distinct change from black to white. The spirals, therefore, would have edges which are unambiguous.

Another requirement of the model is that all of the spirals would be very close approximations of one another. This means that if one spiral were rotated around the cyclone until it reached another spiral, the two would coincide.

If a single cloud-band spiral is a true representative of all of the other tracer spirals found in a cyclone, then the implication of this is that the cyclone is kinematically symmetrical in an axial sense.

The edges of clouds as seen in photos do not go from sudden brilliance to sudden darkness but rather through a transition zone of gray approaching black. The cloud brightness is expressed in digital form; assuming uniform illumination over the entire storm it is possible to establish a "cloud edge" on the basis of a specific numerical threshold value attached to the isolines of brightness which describe the cloud.

Since the illumination is never perfectly uniform, the preferable method, assuming a low noise-to-signal ratio, is to use the brightness gradient for the edge of the cloud. This obviates the necessity of having uniform illumination but makes it necessary to have clean data, i.e., data which is relatively free

of noise; otherwise the maximum will not only not be unambiguous, but will not even be unique, i.e., there may be several of them.

5.2 Photographic Techniques

A photodensitometer can be used to measure the thickness of the silver deposit on the ATS negative. Again isolines of density require uniform illumination. A gradient of density is not as easily found in this technique as the brightness gradient is contaminated by the grain in the silver deposit.

Another technique involves the use of the human eye. The human eye has both advantages and disadvantages. The eye is capable of integrating patterns of a cloud into a complete curve by virtue of the human intellect connected to it. This is a distinct advantage when it comes to the smoothing of data. The difficulty lies in the fact that the human eye is subject to operator bias, and it is necessary to utilize several curves drawn by the same person or curves drawn by several people in order to get a statistical mean for them. In this manner, operator bias is reduced.

6. CORRECTIONS FOR DISTORTIONS DUE TO FORESHORTENING

The photographs and digital data used in this study were taken by ATS-I over the Pacific. The photos and the TV raster show the earth as a disk and a considerable amount of distortion is evident in the photograph. This distortion can be corrected by straightforward but laborious calculations in geometry. The author has developed a complete program to do this and it is available at the Department of Meteorology at the University of Wisconsin (Nicholson, 1967).

It is also possible to change the subsatellite point photographically, but it is absolutely essential that we simulate on a smaller scale the exact geometric situation that existed at the time the photograph was taken in the real world.

When an image of the storm was projected upon the globe through the proper lens at the correct distance, the globe became a screen with a miniature replica of what ATS saw that day (Fig. 7a). The distortions due to foreshortening had been removed. One can then rephotograph the phenomenon of interest and record it on film from a new subsatellite point directly above it. This is what was done in Fig. 6, which shows the unrestored photograph of the cyclone of 5 January 1967 on the left with the subsatellite point at 152W and 0°N. The picture on the right is a view with the geometry corrected.

Fig. 7a

Fig. 7b

Fig. 7. Photographically corrected (a) cyclone compared to uncorrected (b) one, and apparatus (c)

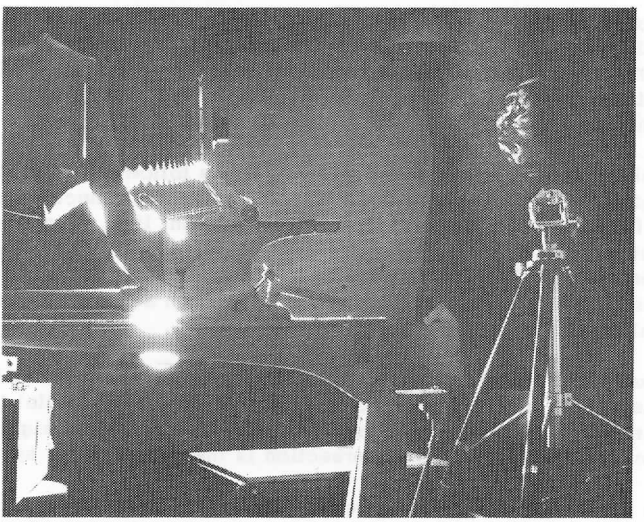


Fig. 7c

7. METHODS FOR ESTABLISHING SPIRAL PARAMETERS

7.1 Method of Maxima and Minima

Once the points of the curve have been established and corrected for foreshortening and other distortions, it is not at all difficult to find the center of the system. An initial guess is made as to the center and then relative to this guess the theta and radius coordinates are assigned and then plotted on semilog vs. angle theta paper. If the plot is a straight line, then the initial guess was correct. If the initial guess was not correct, then the plot will be a curve much like a Bessel function oscillating about the true spiral as is shown in Fig. 8a. In Fig. 8a, Curve I is the plot of Curve I in Fig. 8b, only in a differerent coordinate system—semilog vs. angle theta. Curve II in Fig. 8a corresponds to the plot of Curve I with an initial incorrect guess as to the center of the spiral. This is shown in Curve II in Fig. 8b which is displaced to the upper portion of the diagram. Note how the intersections of the two spirals form a straight line in the polar coordinate system. Spiral II in Fig. 8b has maxima and minima of radii which appear in cyclic fashion. In Fig. 8a the maxima and minima appear as crests and troughs in a wavey pattern of the radial values. The "wave" intersects the straight line at the points of halfway between the crest and the trough of each wave. The true spiral can therefore be plotted and the center identified as the intersection of two circles having radii and centers specified on the log r vs. theta plot. In this case, these values would be the intersections A, B and C in Fig. 8a.

It may be necessary to perform this operation more than once to get a better estimation of the center of a particular spiral. At least three maxima and/or minima are required in this technique to define the spiral. If there is not a sufficient number of these, it is possible to find the center by using other methods.

Probably the simplest of these methods is the following. Construct two sets of parallel lines which are tangent to the spiral. Connect the points of tangency for each set of parallel lines with a straight line. The lines intersect at the center of the spiral. The spiral must be sufficiently tightly wound to do this; otherwise it will be impossible to draw two sets of parallel lines tangent to the spiral. The spiral must therefore wind through at least π radians before this technique can be employed.

An example is given in Fig. 9 where $\overline{AB}$ and $\overline{CD}$ are parallel lines tangent to the spiral, and $\overline{A'B'}$ and $\overline{C'D'}$ are another set of parallel lines tangent to the spiral. The points of tangency can be joined by straight lines for each set of parallel lines. The intersection is the center of the spiral.

Once the points of the curve are established, it is necessary to reorient them. Unless the points of the cyclone lie directly at the subsatellite points,

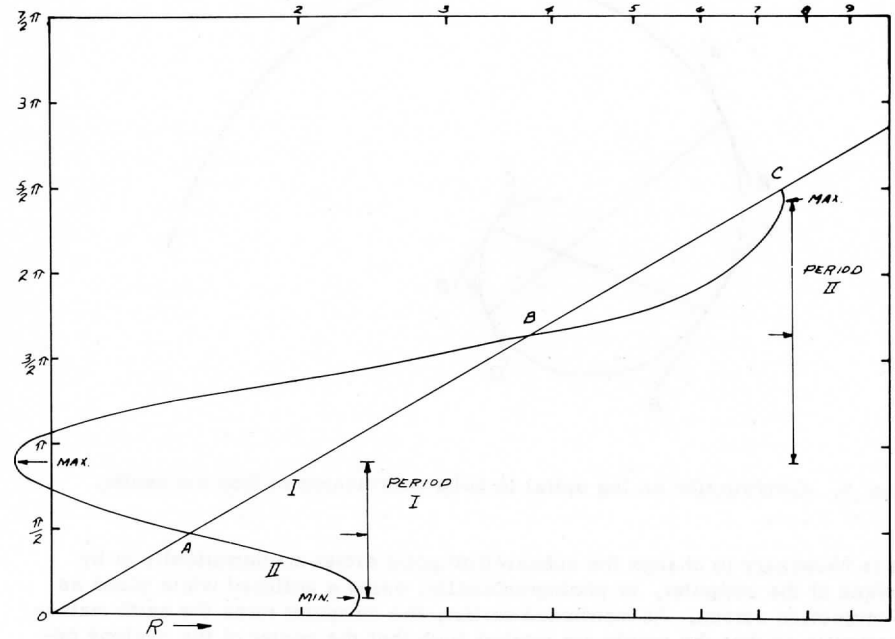


Fig. 8a. Plots of off-center vs. centered log spirals on semilog vs. theta paper.

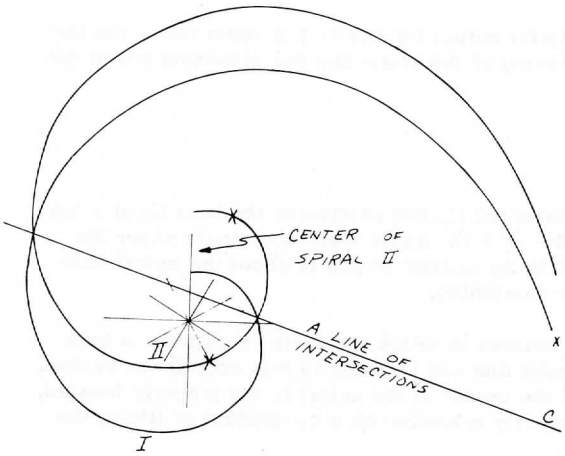


Fig. 8b. Plots of off-center vs. centered log spirals on polar coordinate paper.

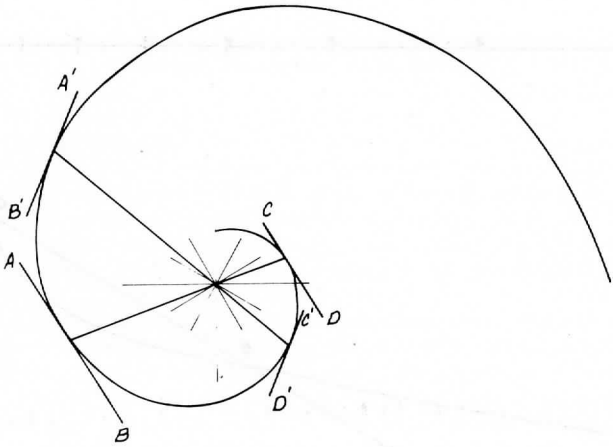


Fig. 9. Construction on log spiral in polar coordinates to find the center.

it is necessary to change the subsatellite point either mathematically or by means of the computer, or photographically, using a polished white globe as a projection screen. As mentioned earlier, the computer turns the earth mathematically so that the points are rotated such that the center of the cyclone becomes the new subsatellite point. For this procedure, landmarks and satellite latitude and longitude are unnecessary, but the radius of the earth and the satellite height are necessary.

It is possible to use the digital output for merely the storm itself and the photographic data for the positioning of the storm and the determination of the proper scale ratios.

7.2 Curve Fitting Technique

Once the data points are corrected for foreshortening the best fit of a log spiral can be calculated. On the ln r vs. angle theta coordinate paper the true spiral is a straight line while the scatter of points about the spiral indicates the data corrected for foreshortening.

Using the method of least squares in which the theta coordinate is kept constant, the best fit of a straight line can be made in this coordinate system. The advantage of this is that if the center of the spiral is not properly located, it will be automatically and correctly relocated as a by-product of fitting the data to a straight line.

All that is necessary, therefore, is to make a guess as to the center of what would be the fitted spiral to obtain polar coordinates for these points and even an off-center guess will cause the data points to oscillate about the true spiral fit on a ln r vs. theta graph. Since they do oscillate about the true spiral fit, then a fit of the oscillating data points will give the same results as that obtained by use of the method of three maxima and/or minima in the oscillation cited previously.

8. TEST ON STORM OF JANUARY 4–7, 1967

The real cyclone which was tested occurs in the north Pacific Ocean in the vicinity of Hawaii during the first week of January 1967. This storm began as a vorticity center aloft into which a cold front moved from the northwest. The result appeared to be very much like dumping ink into the rotating dishpan. A spiral immediately formed on the 4th and lasted until the 7th of January.

Photographs of the cyclone of the 4th through the 7th of January were traced by the hand method. These tracings were then digitized and rotated mathematically by computer and a least squares fit used on the points to achieve the correct spiral. Once the spiral was established, it was compared to spirals at a later time and vertical velocities for the intervening periods were derived. Figure 10 shows the hand tracings of two different spirals of the storm of the 5th of January on semilog paper. These spirals were traced on each side of the cloud band in Fig. 7b. A is on the left, B is on the right of this bank of clouds in the spiral. Each spiral is composed of two separate tracings. The points of the first tracing are represented by circles, the second by dots. The tracings very nearly coincide for the same spiral indicating that operator bias is small.

The spirals, however, are not parallel lines, indicating that the cyclone is not perfectly symmetrical if the lines began as rays at the same time. This is a weakness of the model. This is to be expected if there is to be some divergence in the cyclone leading the vertical motion needed to produce cloud banks. Indeed, this is the cause of fronts. If there were no axial asymmetry, then there could not be frontogenesis because the necessary confluence in the tangential wind field would be absent. Therefore the model is oversimplified in its demand for axial symmetry. Whereas the model may be symmetrical over 2π radians, over shorter angular displacements there may well be a nonzero change of the tangential velocity, for instance, with respect to arc length, or $\frac{\partial}{r\partial\theta}(V_\theta) \neq 0$ but $\int_0^{2\pi} \frac{\partial}{r\partial\theta}(V_\theta)\, d\theta = 0$.

Even so, there is a certain amount of asymmetry which the model can tolerate without giving significantly different results from the perfectly symmetrical model. The effects due to asymmetry are treated in Section 9. Similar tracings for a digital display of a vorticity center near Hawaii are shown in Figs. 10b and c. In Fig. 10c, it can be seen that the trace of spiral IV has

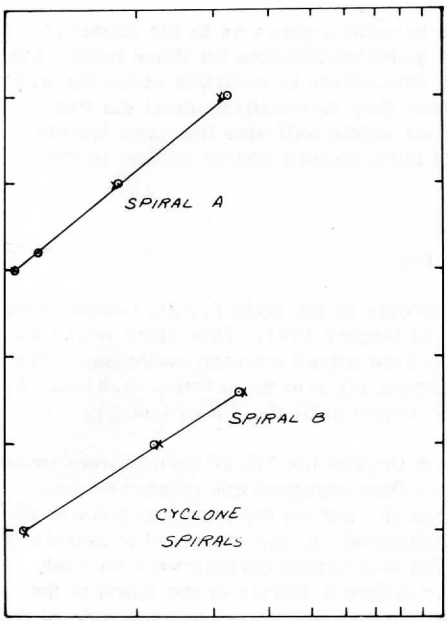


Fig. 10a. Separate cyclone spirals for the same time for 5 January 1967.

a discontinuity which indicates the core of a cyclone first, and then a log spiral. The edge of the core is delineated by a constant value of ln R, or a vertical line on the semilog graph. Curves IV and VI are the boundaries of a clear area. If there were divergence in this area it would show up as a lack of parallelism in the semilog plot of these two curves. This is not immediately apparent in the graph of Fig. 10c.

This data can be expected to provide general trends and perhaps order of magnitude measurements for vertical velocity.

9. FURTHER THEORETICAL CONSIDERATIONS

Among the considerations of the model's properties which we will investigate further are the winding rate of the spiral, the vorticity field implicit in the cyclone, and a relaxation of the initial rigid condition of absolute axial symmetry and the consequences of this.

Fig. 10b. Tropical disturbance NW of Hawaii with overlay of logarithmic spirals.

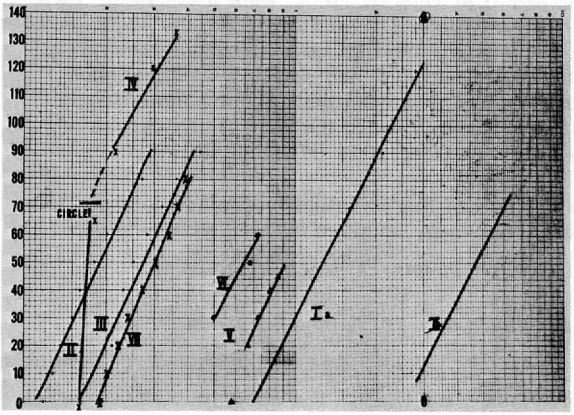


Fig. 10c. Spirals of Hawaiian tropical storm plotted on semilog paper.

9.1 The Vorticity Field

As we saw earlier in the section concerning the kinematic model, the area average of the vertical velocity $\bar{w}$ is related to the vertical velocity w at any point outside of the core by the equation $w = \bar{w} + 2c_2 h$ where c_2 is a constant with units of T^{-1}. Utilizing the same procedure by which we established $w = \bar{w} + 2c_2 h$ we may show that $\zeta_a = \bar{\zeta}_a + K$ where K is a constant with units of T^{-1}, and ζ_a is the absolute vorticity, whereas $\bar{\zeta}_a$ is the area average of the absolute vorticity. The point is, then, that as $\bar{\zeta}_a$ varies, so does ζ_a.

For example, if $\frac{\partial \bar{\zeta}_a}{\partial r} = \phi(r)$ where $\phi(r)$ is some function, then $\frac{\partial}{\partial r}(\zeta_a - K) = \frac{\partial \zeta_a}{\partial r} = \phi(r)$ which is the same function. By letting $\bar{\zeta}_a = \bar{\zeta}_r + f$ where $\bar{\zeta}_r$ is the area average of the relative vorticity we are able to deal with the relative vorticity in the same manner as the absolute vorticity, i.e., by substituting the area average to see how both functions vary. Let us now investigate the mathematics of the vorticity as seen in the model. If we investigate Fig. 6, we may see that by subtracting the displacement of theta at one point from that at another $\Delta\theta_2 - \Delta\theta_1$ and adding $\ln\frac{r_2}{r}\tan\phi$, we are able to obtain the slope of spiral II at Time II in terms of the rate of displacement along the θ coordinate.

$$\tan\psi_2 = \frac{\Delta\theta_2 - \Delta\theta_1}{\ln r_2/r_1} + \tan\psi_1 \quad \text{or} \tag{9.1.1}$$

$$\tan\psi_2 = \left[\frac{\bar{w}}{h(\bar{\zeta} + f)}\tan\phi - 1\right]\frac{\Delta t \cos^2\phi}{\ln r_2/r_1}(\dot{\theta}_2 - \dot{\theta}_1) + \tan\psi_1 \tag{9.1.2}$$

where ψ_2 is the slope of the spiral at Time II. If $\bar{\zeta}$ were a constant, then for hurricanes (tropical cyclones) with small f and large $\bar{w}$, the angle between the tangent to the curve and the radius should be larger for the same time period than for an arctic cyclone with large f and small $\bar{w}$. Experience shows that the opposite is true, which indicates that $\bar{\zeta}$ for a hurricane is much higher than for an arctic storm. In view of the very strong curvature through which intense winds are propagated, this is not surprising.

The relative vorticity is the sum of the curvature term plus the shear term. The relative tangential velocity, however, is given by $V_{\theta_r} = r(A - c_1\frac{r_0}{r}\ln\frac{r}{r_0} - \frac{f}{2})$. The shear term is $\partial V_{\theta_r}/\partial r$ which is $A - \frac{f}{2} - c_1\ln\frac{r}{r_0} - c_1$. The curvature term is $V_{\theta_r}/r = A - c_1\ln\frac{r}{r_0} - \frac{f}{2}$. The sum of the two is $2A - f - c_1(\ln\frac{r}{r_0} + 1)$. This means that as r approaches 0, the relative vorticity approaches a maximum. The absolute vorticity is given

by $2A - c_1 (\ln \frac{r}{r_0} + 1)$ which reaches a maximum when $r = 0$. Thus the vorticity increases toward the center of the storm to a maximum. This is observed in real life (Whitney and Herman, 1968).

9.2 Twisting Rate

The tangent of ψ_2, which is the slope of the spiral, is given by Eq. (9.1.2).

$$(9.2.1) \quad \tan \psi_2 = \left[\frac{(w - 2c_2 h)\tan \phi}{h2(A - c_1 \ln r/r_0) - c_1} - 1 \right] \frac{\Delta t \cos^2 \phi}{\ln r_2/r_1} \times (\dot{\theta}_2 - \dot{\theta}_1) + \tan \psi_1$$

Using the appropriate substitutions, we obtain $\tan \psi_2$ in terms of w, V_{θ_A} and Δt where V_{θ_A} is the absolute tangential velocity.

$$\dot{\theta}_2 = A - c_1 \quad \ln r_2/r_0$$

$$\dot{\theta}_1 = A - c_1 \quad \ln r_1/r_0$$

$$\dot{\theta}_2 - \dot{\theta}_1 = - c_1 \quad \ln r_2/r_0 + c_1 \quad \ln r_1/r_0$$

$$= - c_1 \quad \ln r_2/r_1$$

$$\tan \psi_2 = \left[\frac{(w - 2c_2 h)\tan \phi}{h2(A - c_1 \ln r/r_0) - c_1} - 1 \right] c_1 \Delta t \cos^2 \phi + \tan \psi_1$$

$$(9.2.2) \quad \tan \psi_2 = \left[\frac{(w - 2c_2 h)\tan \phi}{h 2V_{\theta_A}/r - c_1} - 1 \right] c_1 \Delta t \cos^2 \phi + \tan \psi_1 .$$

Assume $\psi_1 = 0$.

When V_{θ_A} is large, $\tan \psi_2$ is small and vice versa. ψ_2 increases as a function of time Δt and is inversely proportional to the absolute tangential velocity divided by the radius.

The twisting rate, therefore, $\frac{d\psi}{dt}$, is a function of the absolute tangential velocity, and hence the intensity of the cyclone itself. The more intense the cyclone, the slower the twisting rate.

9.3 Axial Asymmetry and Occlusions

One of the initial assumptions made in the model was that the model was axially symmetric. For a large cyclone of a diameter of 20 degrees, centered at 40 degrees north, the difference in the coriolis parameter between the northern and the southern tip is in the ratio of 1.555 (the ratio of the sines of 50 and 30 degrees). If $\bar{\zeta}_r$ is asymmetrical with less variance than this ratio between the northern and southern portions of the cyclone (which one would expect starting with air initially at rest on a rotating sphere), then the tangential velocity of more northerly and hence cooler air V_{θ_N} will be greater than that of more southerly and warmer air V_{θ_S}.

$$\frac{V_{\theta_N}}{V_{\theta_S}} = \frac{\pi r^2(\bar{\zeta} + f)N}{\pi r^2(\bar{\zeta} + f)S} > 1.$$

This implies that the cooler air must travel more swiftly in this model and overtake the warmer air. As this happens, the spiral winds up and the interface of the cooler and warmer air becomes more evenly distributed as a function of azimuth so that the model becomes more axially symmetric.

In point of fact, the relative vorticity of a storm does tend to be symmetrically distributed, whereas the absolute vorticity is slightly asymmetric, increasing in the direction of the colder air, i.e., initially toward the north.

This being the case, a plan view of a cyclone with isotherms running initially east-west, Fig. 11, the isotherms would be in a region of confluence in the western portion and diffluence in the eastern. This means that the isotherms would become more closely packed in the area of confluence than that of diffluence. This is illustrated in Fig. 11, in a simplified view of a cyclone which begins as a wave along a set of closely packed isotherms.

10. CONCLUSIONS

We have postulated a cyclone model in which a time line of air parcels form logarithmic spirals similar to the patterns made by clouds in a cyclone. The model can relate the vertical velocity of the cyclone to the change in the geometry of the cloud patterns.

The model specifies the relative tangential velocity as $V_{\theta_r} = r(A - c_1 \ln\frac{r}{r_0} - \frac{f}{2})$ where $\frac{f}{2}r$ is the solid rotation rate of the earth under the storm and A and c_1 are constants with units of T^{-1}. The model also specifies the radial velocity as $V_r = r(B - c_2 \ln\frac{r}{r_0})$. From these we derived a

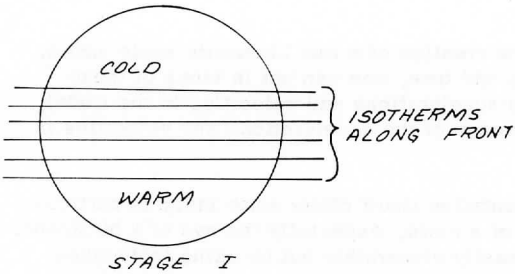
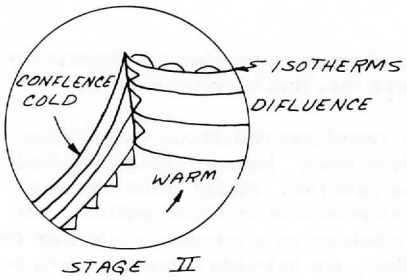


Fig. 11. Plan view of isotherms in cyclone as perturbed by the flow field.



vertical velocity as a function of r as $w = h(\bar{\zeta} + f)\,\mathrm{ctn}\,\phi\left[\frac{\theta_2 - \theta_1}{\dot{\theta}\Delta t\,\cos^2\phi} + 1\right]$ $+ 2hc_2$ where $\bar{\zeta}$, θ_2, θ, $\dot{\theta}$ and ϕ can be measured by the satellite and r_0 and c_2 can be calculated graphically. From these equations we developed further considerations concerning the vorticity field. A and c_1 can be solved for, given the tangential wind velocities. Also r_0 can be obtained graphically from the spirals.

The relative vorticity is expressed by the equation $\zeta_r = 2A - c_1 \frac{r_0}{r}(\ln \frac{r}{r_0} + 1)$.

A dishpan experiment shows that the conditions postulated in a cyclone would, when applied to a dishpan model, produce log spirals in a tracer. The dishpan model further demonstrates that even though the holes in the bottom of the pan confined the drainage locally, the integrated effect was to obtain movement throughout the entire pan so as to produce the observed spirals. A similarity can be observed in mesoscale phenomena such as the thundershowers which occur in the spiral bands of a hurricane. Whether or not the thunderstorms which occur in a hurricane actually produce a synoptic effect has not yet been verified by testing, at least not by this author.

The model points the way to the creation of a true kinematic scale model. Using model ratios of length, mass and time, one can get in terms of these further model ratios. The observed accelerations and velocities in the model, therefore, would lead us to a knowledge of the accelerations and velocities in the atmosphere.

The technique of spiral representation itself offers some insights particularly valuable in locating the core of a storm, especially the eye of a hurricane. In many hurricanes the eye is not easily discernible but by using techniques indicated in this paper one can find the eye quickly and easily. A knowledge of the wind field in the theoretical model also gives us a grasp of the angular momentum budget.

Since the calculated torque minus the observed torque would indicate the torque due to friction, further insights along this line become feasible.

Finally, we have shown that from the model one can obtain quantitative data from the shapes observable in the cloud field. Johnson (1968) obtained horizontal velocities from the synchronous satellite. Sikdar (1969) derived mesoscale vertical velocities from the change in size of cirrus outflow. We now are in a position to derive vertical velocities on a synoptic scale from the change in shape of the cloud patterns. Thus, the satellite representations are useful for much more than mere exercises in aesthetics.

In all candor, however, we must say that the effort to test the model using the many examples which the satellite provides us with has been only cursory. What one needs is data—not from the open oceans where data is sparse, but from a data rich area. Only then will one be able to fully test the model.

ACKNOWLEDGMENTS

I wish to express my deepest gratitude to Professor Verner E. Suomi without whose inspiration and great patience this thesis would have been an impossibility.

I also want to thank Professors John A. Young and Frank Sechrist for the many hours they spent in reading this work and in giving helpful advice and suggestions. I am particularly indebted to Mr. Jack I. Kornfield whose great insight helped make this work possible.

I wish to thank Dr. D. N. Sikdar for his continued confidence and encouragement, and without whose help it would have been most difficult to complete this work. I am indebted to Mr. David E. Cadle for his expert technical assistance, and very generous cooperation.

Plate 1

Finally, I wish to acknowledge the great debt I owe to my wife Kathleen and my son Kenneth, whose confidence and affection were a constant source of inspiration to me.

This work was supported by ESSA on grant E-230-68G.

BIBLIOGRAPHY

Erickson, C. O. and Hubert, L. F., 1961: "Identification of Cloudforms from TIROS 1 Pictures," Meteorological Satellite Laboratory Report, No. 7, U. S. Weather Bureau, p. 68.

Fujita, T. "Mesoanalysis of Hurricane Rainbands," Proceedings of the Technical Conference on Hurricanes, American Meteorological Society, p. C6-1.

Fujiwhara, S. "The Natural Tendency Towards Symmetry of Motion and Its Applications as a Principle of Meteorology," QJRMS, 47, p. 287 (1921).

Glaser, A., 1961: "TIROS I: An Operational Evaluation of a New Meteorolgical Tool," Chapter 1 of "Contributions to Satellite Meteorology," GRD Research Notes, No. 36, Vol. II, p. 1-138, edited by F. R. Valovcin.

Johnson, M. H., 1968: A Photogrammetric Technique for Finding Winds from Satellite Photos, in "Studies in Atmospheric Energetics Based on Aerospace Probings," Annual Report, 1968, Department of Meteorology, University of Wisconsin.

Kessler, E. "Outer Precipitation Bands of Hurricane Edna and Ione," Bulletin of the American Meteorological Society, Vol. 38, p. 335 (1957).

Krishnamurti, T. N. "Dynamical Structure of a Steady State Hurricane I," Bulletin of the University of California in Los Angeles, Contribution No. 127, p. 32 (1966).

Malkus, J. S. and H. Riehl: "On the Dynamics and Energy Transformations in Steady State Hurricanes," Tellus, Vol. 12, No. 1 (1960).

Nagle, R. E. and Serebreny, S. M., 1962: "Radar Precipitation Echo and Satellite Cloud Observations of a Maritime Cyclone," J. Appl. Meteorol., 1, 279-295.

Nicholson, F. H. "The Theoretical and Mathematical Justification for Navigation and Analysis from ATS Satellites," unpublished (1967).

Senn, H. V. and H. W. Hiser. "On the Origin of Hurricane Spiral Rain Bands," Journal of Meteorology, Vol. 16, p. 417 (1959).

Sikdar, D. N., 1969: Convective Heat Transport over the Tropical Mid-Pacific as Estimated from a Geosynchronous Satellite Altitude. Ph. D. dissertation, University of Wisconsin.

Tang, Yen, Edward M. Brooks, and Bruce F. Watson. "Theoretical and Observational Studies of Vortex Cloud Patterns," CGA Technical Report, No. 64-2-G.

Whitney, Linwood F., Jr., Herman, Leroy D., 1968: "The Nature of Intermediate Scale Cloud Spirals," ESSA Technical Report NESC-45.

SEVERE LOCAL STORM RESEARCH

Brian A. Auvine and Dr. Charles E. Anderson

While many researchers have made case studies of the squall line and associated severe local weather, the advent of satellite technology has made possible a more comprehensive examination of such phenomena.

A particularly well defined outbreak of severe weather in the Ohio River Valley occurred on April 23, 1968. As one can see from Fig. 1, this outbreak is evidenced in the satellite photograph as a curved arc of bright cumulonimbus cloud complexes with associated cirrus blowoff. Severe weather and several tornadoes were coincident with the largest of these cloud masses.

The presence of this cumulonimbus arc gives rise to several questions: is this line a reflection of some sort of synoptic or dynamic process, and if so, can this same process explain the location of the most severe weather? How do these complexes correspond to the region of precipitation (i.e. the radar echoes)? To answer the latter question, radar observations from the Cincinnati radar were transposed onto an accurately gridded brightness contour analysis of the satellite photo. In Figs. 2, 3 and 4 one can see that the radar echoes, instead of being located in the brightest region of the cloud, are found in the western part or even on the edge of the entire cloud mass. This location could be explained by the fact that cirrus tops of the active convective areas are blown off far to the east of the precipitation area, thus leaving most of the echoes in the western sector of the cloud mass. Also, according to Fujita's model of an active cumulonimbus (Fujita, 1962), precipitation may be partly displaced into the cloud-free area to the west of the cloud due to the dynamic character of the updraft. Thus it would appear reasonable to expect the radar echo to appear to the west of the brightest cloud region.

Before proceeding to the synoptic aspect of this investigation, a word should be said about the Airways Reporting Network, the network used in the following analysis. The Airways subsynoptic network is particularly well suited to a study of the April 23 cloud masses, since we are here dealing with a subsynoptic phenomenon, that is, a phenomenon between the synoptic and the mesoscale. The compatibility between the network used and the type of analysis attempted is important since patterns of quantities such as divergence and vorticity are very sensitive to the resolution capabilities of the data (Rodriguez, 1964).

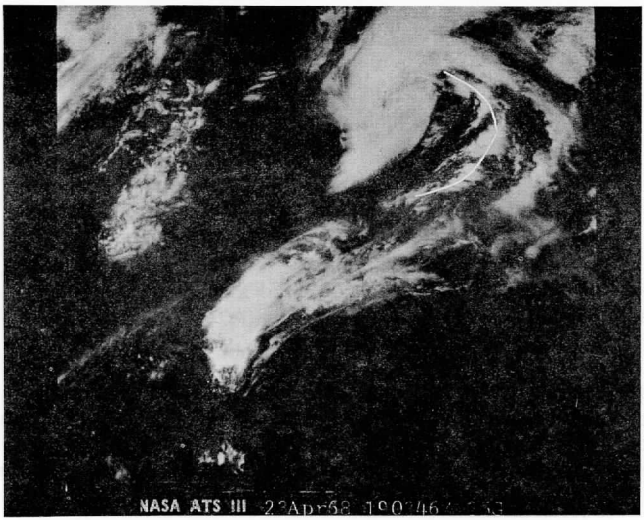


Fig. 1. ATS-III satellite photograph for 1853Z, April 23, 1968. The squall line clouds are connected by the curved line.

Since one might expect that a convergence analysis would have great applicability to the present case, such an analysis was performed by first obtaining a u, v breakdown of the reported winds and then calculating the convergence by means of the equation: $\partial u/\partial x + \partial v/\partial y = \text{Div}$. In Fig. 5, one such analysis for 1900Z is shown. It is clear that an area of significant convergence is present whose trough line extends roughly from north to south across the data area. Transposing these convergence trough lines onto a brightness contour analysis of the satellite photo, one can see (in Figs. 2, 3 and 4) that such convergence lines are closely associated and travel at about the same speed as the arc of storm activity, and that the more intense convergence over Northern Kentucky, Ohio and Michigan corresponds to the greater storm activity found in these same regions.

In order to determine whether this convergence line was in fact the cause of the squall line rather than a result of it, convergence analyses were also performed at 0600Z and 1500Z, 12 and 3 hours before the period of major activity, respectively. At both times, a line of convergence continued to be in evidence, thus proving the convergence line to be independent of the storms.

It is not surprising to find a convergence field associated with convective cloudiness. Kasahara (1967) and Cressman (1946), among others, have found

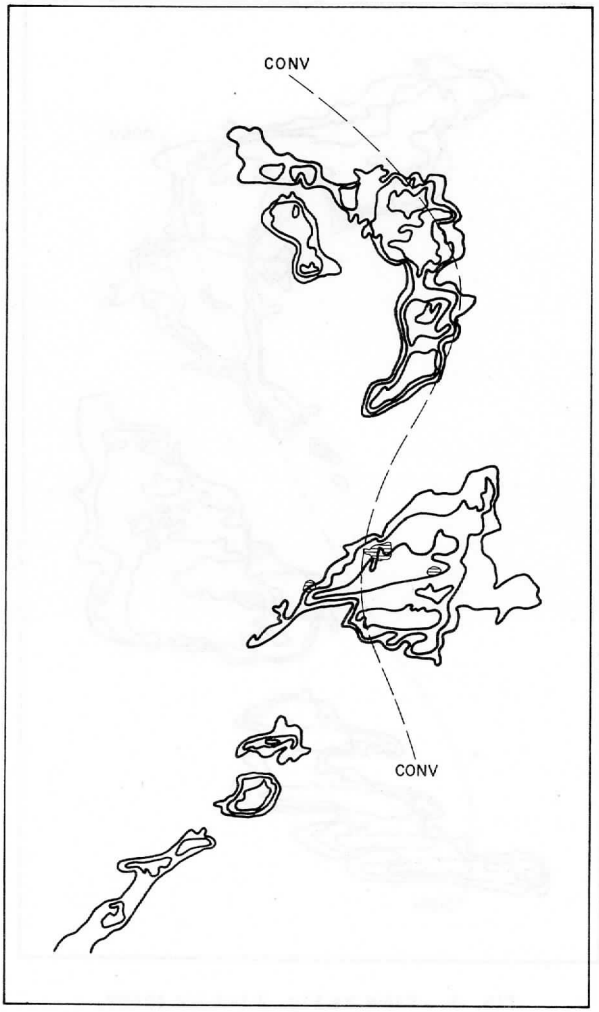


Fig. 2. Brightness contour analysis of an ATS-III satellite photograph. Shaded areas are the transposed radar echoes observed at the same time from Cincinnati. Dotted line represents the line of maximum surface convergence.

Fig. 3. Same as Fig. 2 but for 1900Z.

Fig. 4. Same as Fig. 2 but for 2000Z.

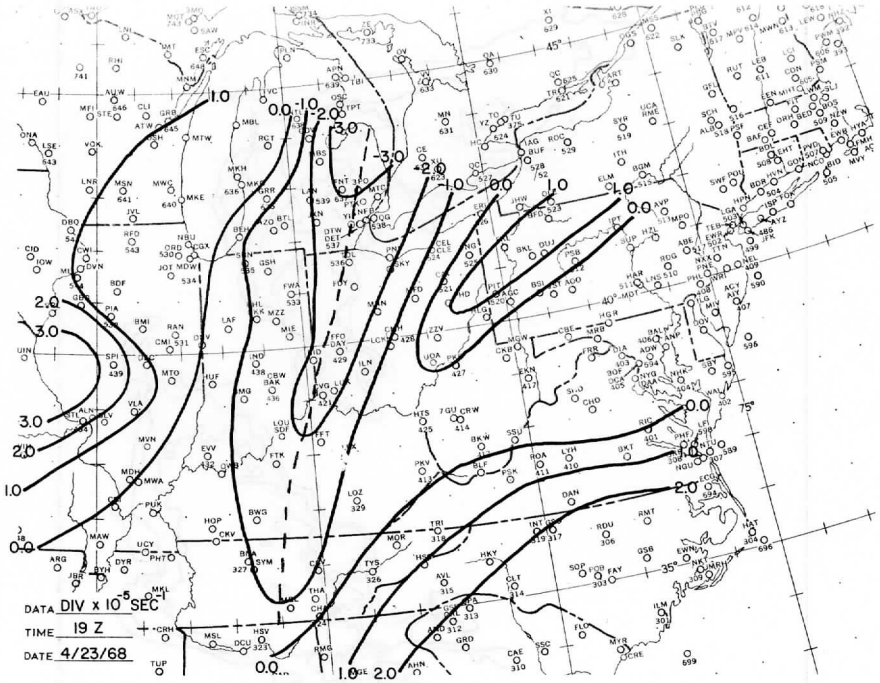


Fig. 5. 1800Z surface convergence analysis. Units used are 1×10^{-5} sec^{-1}.

in their theoretical studies that convergence is a major cause of significant convection. Case studies and objective analyses of storm situations have led to a similar conclusion: that surface convergence or upper air divergence is a good objective indicator of severe weather (Endlich, 1968). In the April 23 case, however, the obviously close relation between the line of convergence and the cloud arc leads one to the question of what factors might be responsible for the line of convergence in the first place.

Since Tepper (1950) first proposed the concept of the pressure jump line, others (for instance, Ogura (1962)), have postulated a gravity wave explanation to account for squall lines. Recently, Drs. Houghton and Sechrist (University of Wisconsin) have proposed that squall lines are essentially a gravity wave phenomenon generated by the high level jet. As this high speed air progresses into an unstable environment, a crescent shaped mound of air is created in advance and along both sides of the isotach maximum, this wave in turn producing a line of convergence. In a situation of conditional

instability, such convergence will produce intense convection and cause the formation of a squall line.

Although this theory has yet to prove itself operative in every severe weather situation, the April 23, 1968, outbreak of storms seems to conform plausibly with a gravity wave hypothesis, not only because a jet does follow behind the cloud arc, but also because we have clear evidence of the postulated line of convergence.

More case studies will be necessary before it can be shown that a cause and effect relationship exists between the jet and the line of convergence. However, the juxtaposition of these two phenomena in this clearly defined storm situation is encouraging and worthy of further investigation.

REFERENCES

1. Cressman, G., 1946: The influence of the field of horizontal divergence on convective cloudiness, Jour. Met., 3, 85-88.

2. Endlich, R. and Mancuso, R., 1968: Objective analysis of environmental conditions associated with severe thunderstorms and tornadoes, Mon. Wea. Rev. 96, 342-350.

3. Fujita, T. and Byers, H., 1962: Model of a hail cloud as revealed by photogrammetric analysis, Tech. Rep. #3, U. Chicago, 9 pp.

4. Kasahara, A. and Asai, T., 1967: Effects of an ensemble of convective elements on the large scale motions of the atmosphere, Jour. Met. Soc. Japan, 45, 280-291.

5. Ogura, Y. and Charney, J., 1962: A numerical model of thermal convection in the atmosphere, Proc. Intern. Sym. on Num. Wea. Pred., Tokyo, Nov. 7-13, 1960, 431-451.

6. Rodriguez-Feo, L., 1964: Vertical structure of the divergence field over severe storms, Meteorologische Abhandlungen, 32/4, 57 pp.

7. Tepper, M., 1950: A proposed mechanism of squall lines: The pressure jump line, Jour. Met., 7, 21-29.

THREE-DIMENSIONAL, TIME DEPENDENT NUMERICAL EXPERIMENTS WITH DRY AND MOIST, SHALLOW AND DEEP CONVECTION MODELS

Uri Shafrir, [1] Shmuel Kaniel, [2] and Boris Shkoller [3]

Goddard Institute for Space Studies
New York, New York

Abstract

The Boussinesq and anelastic systems of equations for shallow and deep atmospheric convection, respectively, are solved numerically, both for dry and moist situations. Numerical techniques are discussed in some detail, and the results of several experiments are exhibited.

Contents

[1] On leave from the Institute of Planetary and Space Science, Tel-Aviv University, Tel-Aviv, Israel.

[2] On leave from the Department of Mathematics, Hebrew University, Jerusalem, Israel.

[3] On leave from the Department of Applied Mathematics, Weizmann Institute of Science, Rehovot, Israel.

I. INTRODUCTION

The problem of atmospheric convection has been treated relatively extensively in the literature in recent years. Theoretical investigations of dry convective motions resulted in analytic solutions, where assumptions of symmetry, steady-state or linearity were made (Gutman (1961), Lorenz (1962), Platzman (1965)). However, it has long been recognized that, because of the moist, unsteady, nonlinear, asymmetric nature of most convective scales in the atmosphere, numerical solutions carry the only hope for a "realistic" simulation of these phenomena. Such solutions were offered in papers by Malkus and Witt (1959), Asai (1960), Charney and Ogura (1960), Lilly (1962), Ogura (1962, 1963a) and Orville (1965, 1968). A comprehensive review of numerical simulation of atmospheric convection was published by Ogura (1963b); a more recent review was published by NCAR in 1967 (Thermal Convection: A Colloquium).

Still, all available numerical models are lacking in two cardinal areas, where no advances were attempted until now: One possible (and obvious) contribution could have been to break away from symmetry conditions (two-dimensional, or axisymmetrical) and to treat the meteorologically more significant asymmetrical cases (resulting, for example, from an asymmetrical initial disturbance, or from the presence of wind shear). This was not done, because computations of asymmetrical convection require a fully three-dimensional model, which was nonexistent until now because of lack of adequate numerical techniques, as well as computing capacity. The other avenue of approach could have been to develop the early, shallow-convection models into deep, moist convection modeling in two dimensions; however, as pointed out by Ogura and Phillips (1962), this requires the phase-change part of the computations to be introduced in a consistent and simultaneous way into the model, evidently a requirement not easy to meet. One may also mention that the introduction of an altitude-dependent density profile excludes the possibility of using the stream function and vorticity as dependent variables, rather than the velocity components plus pressure.

This paper attempts to advance simultaneously along both fronts: it presents a fully three-dimensional numerical model for deep, moist atmospheric convection. Thus, it creates a tool which should enable the meteorologist to experiment with some of the most important of atmospheric phenomena.

II. THE METEOROLOGICAL PROBLEM OF A CONVECTIVE CLOUD

The meteorological motivation for developing a fully three-dimensional, time-dependent model of a convective cloud is manifold. When flying over tropical regions, whether over land (South America) or over water (the Pacific, Caribbean or the Atlantic Ocean), even an untrained observer is easily convinced that the convective cloud is the basic building block of mesoscale atmospheric phenomena.

The logic behind the grouping of three or four types of clouds, into repetitive shapes like streets (straight, or gracefully curved), rings, horse-shoes, and so on, is still a mystery to us. However, the untrained observer will further note that whatever laws of nature dictate the rhythmical creation and destruction of these patterns, they persistently choose three scales of convection (low level "puffs"; medium size, shallow; and deep convective clouds), plus the cirrus blowoff, as their building material in this never-ending, fascinating game.

It then becomes obvious that the first order of business for the curious, trying to unveil the secrets of tropical cloud patterns, is the understanding of the various scales of single cell atmospheric convective activity.

The second load of motivation for modeling convective clouds comes from computer simulation of the general circulation of the atmosphere and from numerical weather prediction. Both types of studies face basically a similar dilemma: as the modeling is usually done in geometric space (at least in the horizontal), by finite difference methods, the size of the grid determines the resolution of the numerical solutions. The present state of development of electronic computers dictates that the horizontal size of the grid be several hundred kilometers, with a minimum of about 100 kilometers achieved to date in some special, high resolution studies. This, obviously, excludes small, as well as mesoscale, phenomena from being resolved directly within the numerical solution to the set of equations describing the behavior of the atmosphere on a global, or even a hemispheric, scale. If we define the predictability of the atmosphere to be inversely proportional to the deviation of an a-priori computed set of parameters characterizing the state of the atmosphere (such as wind, pressure, temperature, moisture, etc.) from their a-posteriori measured values, one can only guess what is the contribution of the exclusion from the model of small and mesoscale phenomena to increasing this deviation as the computations proceed in time. If, using today's numerical models, the atmosphere is "predictable" (to within an agreed-upon value of the allowed deviation) to a period of umpteen days, we do not feel that one can say with any degree of certainty by how much this period can be lengthened if small and mesoscale effects are included in the prediction model; however, energetic and other considerations tell us that the correct inclusion of these effects will indeed result in the lengthening of this period, thereby making the atmosphere "more predictable," or predictable for a longer period of time. It therefore follows that, in the spectrum of thrusts being planned today for improving numerical weather prediction, the inclusion of subgrid phenomena should occupy a prominent place.

We now pick up again the discussion of the adequacy of computation facilities where we left it: namely, with the statement that today's computers are not large and fast enough to handle a numerical model of the whole atmosphere with a resolution of better than 100 kilometers or so. The obvious conclusion from this statement seems to be that the problem of including what

today are "subgrid phenomena" in general circulation and numerical weather prediction models will automatically resolve itself when the size of the grid in these studies is cut by two or three orders of magnitude. This conclusion may arouse optimism, which we consider false on two accounts. First, the day when electronic computers which are fast, big and efficient enough to handle a global model of the atmosphere with a spatial resolution of several hundred meters seems to be quite far; as a matter of fact, it is so far that we know of no future date when the availability of such a machine can be anticipated with any degree of certainty. Second, a scale analysis of the various atmospheric motions (see, for example, Ogura (1963b)) reveals that each of the scales present renders itself to a somewhat different mathematical description. In other words, even if a supercomputer was available today, we presently do not know of a single set of equations which describes all the various scales of meteorologically significant motions, and them alone, in a way which will make obtaining these various solutions simultaneously, and on a global scale, practical. The conclusion from the foregoing discussion seems to be that not only do we have today to solve different systems of equations in order to obtain solutions to various scales of atmospheric motions, but we will probably need to continue this practice right through the times when—hopefully—the availability of supercomputers will make it practical to obtain high resolution solutions to the problem of numerical weather prediction.

The second motivation, then, for numerical studies of the various scales of atmospheric convection is to parameterize subgrid phenomena for inclusion in general circulation and numerical weather prediction studies.

The third motivation for such studies comes, one might say, for their own sake. All through the historical development of meteorology as a science, cloud physicists and atmospheric dynamicists identified themselves with different "groups" or "clans" within the profession, with very little dialogue between them. The reason was, of course, that the subject matter of their concern was in such a preliminary stage of development, that any attempt to mate these various disciplines of meteorology, in order to obtain a better and more comprehensive understanding of atmospheric phenomena, was bound to render meaningless results. This situation has been rapidly changing in recent years, and, we dare say, is no longer true today. A large body of experience has been accumulated in numerical studies of both large- and small-scale dynamics of atmospheric motions, since the breaks between transporting the wounded in 1916-1918, when ambulance driver Richardson worked out by hand the first example of numerical weather prediction; thus the mathematical and numerical tools seem presently ready for a three-dimensional study of deep atmospheric convection. On the other hand, recent years have seen unprecedented advances in our understanding of microphysical processes in clouds. In particular, solutions to problems of condensation and accretion of single cloud droplets, as well as the time dependent development of a spectrum of cloud droplets into raindrops, have reached a stage where their mating with cloud dynamics no longer seems immature.

These three avenues of motivation—the understanding of a single convective element, as the basic building block of mesoscale tropical patterns; parameterization of subgrid phenomena in general circulation and numerical weather prediction studies; and the mating of cloud dynamics and microphysics—have emerged, then, to consist the reason for trying to obtain three-dimensional numerical solutions to the problem of moist, deep atmospheric convection.

III. THE MATHEMATICAL MODEL

The basic set of what later came to be known as the "anelastic equations" was originally derived by Batchelor in 1953. In carrying out an analysis aimed at deriving conditions for dynamic similarity of atmospheric motions, he found that the simple assumption that the distributions of pressure and density are, at all times, very close to the distributions of these parameters in an adiabatically stratified atmosphere leads to a set of equations suitable for studies of the various scales of convection in the atmosphere.

Ogura and Charney (1960) have derived independently an identical set of equations, by filtering acoustic waves out of the "primitive" set of hydrodynamic (e. g., Navier-Stokes) equations for an ideal fluid. Ogura and Phillips (1962) have shown that the apparently simpler set of approximations made by Batchelor are actually identical with those of Ogura and Charney and, in particular, that an assumption about the time scale must be made in order to derive these equations.

Orville (1965) has derived the additional terms and equations necessary to complete the set of equations for the case of moist convection.

The following set of equations is based on the derivations by Batchelor, Ogura and Charney, Ogura and Phillips, and Orville.

$$\frac{\partial u}{\partial t} + u\frac{\partial u}{\partial x} + v\frac{\partial u}{\partial y} + w\frac{\partial u}{\partial z} - \gamma \nabla^2 u = -\alpha \frac{\partial \pi}{\partial x} \tag{3.1}$$

$$\frac{\partial v}{\partial t} + u\frac{\partial v}{\partial x} + v\frac{\partial v}{\partial y} + w\frac{\partial v}{\partial z} - \gamma \nabla^2 v = -\alpha \frac{\partial \pi}{\partial y} \tag{3.2}$$

$$\frac{\partial w}{\partial t} + u\frac{\partial w}{\partial x} + v\frac{\partial w}{\partial y} + w\frac{\partial w}{\partial z} - \gamma \nabla^2 w = -\alpha \frac{\partial \pi}{\partial z} + \beta(\theta + 0.608 q_v' - q_\ell) \tag{3.3}$$

$$\frac{\partial \theta}{\partial t} + u\frac{\partial \theta}{\partial x} + v\frac{\partial \theta}{\partial y} + w\frac{\partial \theta}{\partial z} = -\frac{\delta}{\bar{T}} \frac{dq_v}{dt} + k \nabla^2 \theta \tag{3.4}$$

$$\rho\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y}\right) + \frac{\partial(\rho w)}{\partial z} = 0 \tag{3.5}$$

$$\frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} + w\frac{\partial q}{\partial z} = \nu \nabla^2 q_v \qquad (3.6)$$

In these equations, u, v, w are the x, y, z components of the flow; π is the deviation of the pressure, and θ is the corresponding deviation in potential temperature from an adiabatically stratified atmosphere; q is the mixing ratio of water to air, and is the sum of the two mixing ratios for water vapor and liquid water, denoted by q_v and q_ℓ, respectively; q_v' is the deviation of q_v from its time independent value $\overline{q_v}(z)$, specified initially; γ, k and ν are the momentum, heat and moisture diffusion coefficients, respectively, both nonnegative quantities. These equations were nondimensionalized by the transformations

$$t^* = \tau t, \quad x^* = Dx, \quad y^* = Dy, \quad z^* = Dz, \quad u^* = u\frac{D}{\tau},$$

$$v^* = v\frac{D}{\tau}, \quad w^* = w\frac{D}{\tau}, \quad \pi^* = \pi, \quad \theta^* = \theta\Theta, \quad k^* = k\frac{D^2}{\tau}, \qquad (3.7)$$

$$\alpha = \frac{\tau^2 \Theta C_p}{D^2}, \quad \beta = \frac{\tau^2 g}{D}, \quad \gamma^* = \gamma\frac{D^2}{\tau}, \quad \delta = \frac{L}{C_p}, \quad \nu^* = \nu\frac{D^2}{\tau}$$

where D and τ are the characteristic length and time, respectively, and starred quantities are dimensional. (The set derived by Batchelor was for the case $q = q_v = q_\ell = 0$, and $\gamma = k = \nu = 0$.)

These equations, with the appropriate boundary conditions, are believed to be a mathematical model describing deep, moist atmospheric convection.

When solving these equations numerically by finite difference methods within a finite cube, one can minimize wall effects by keeping them as far as possible (determined, actually, by practical considerations of computer capacity and speed) from the disturbance, and by imposing no-stress boundary conditions. These are, then:

At the walls x = const.:

$$u = 0, \quad \frac{\partial v}{\partial x} = 0, \quad \frac{\partial w}{\partial x} = 0, \quad \frac{\partial \theta}{\partial x} = 0, \quad \frac{\partial q}{\partial x} = \frac{\partial q_\ell}{\partial x} = \frac{\partial q_v}{\partial x} = 0.$$

At the walls y = const.:

$$\frac{\partial u}{\partial y} = 0, \quad v = 0, \quad \frac{\partial w}{\partial y} = 0, \quad \frac{\partial \theta}{\partial y} = 0, \quad \frac{\partial q}{\partial y} = \frac{\partial q_\ell}{\partial y} = \frac{\partial q_v}{\partial y} = 0 \qquad (3.8)$$

At the walls z = const.:

$$\frac{\partial u}{\partial z} = 0, \quad \frac{\partial v}{\partial z} = 0, \quad w = 0, \quad \theta = 0, \quad q = q_{\ell} = q_{v} = 0 .$$

So far, no condition is imposed upon π.

This problem is mathematically well posed when initial conditions are specified, i.e., when u, v, w, θ, q, q_v are specified for time $t = 0$.

Let us examine some properties of the differential system, which will serve as guidelines for choosing an adequate difference scheme.

The equations remain invariant under the following transformations:

$$\begin{aligned} &x = -x^*, \quad y = y^*, \quad z = z^*, \\ &u = -u^*, \quad v = v^*, \quad w = w^* \\ &x = x^*, \quad y = -y^*, \quad z = z^*, \\ &u = u^*, \quad v = -v^*, \quad w = w^* \\ &x = x^*, \quad y = y^*, \quad z = -z^*, \\ &u = u^*, \quad v = v^*, \quad w = -w^* . \end{aligned} \tag{3.9}$$

It is easy to check that the boundary conditions, too, remain invariant under the transformations shown in (3.9). These symmetries enable us to continue the solution across any wall.

The symmetry relations imply the boundary conditions for π:

$$\begin{aligned} &\text{At walls } x = \text{const.}: \quad \frac{\partial \pi}{\partial x} = 0 \\ &\text{At walls } y = \text{const.}: \quad \frac{\partial \pi}{\partial y} = 0 \\ &\text{At walls } z = \text{const.}: \quad \frac{\partial \pi}{\partial z} = 0 . \end{aligned} \tag{3.10}$$

The last of these boundary conditions merits special attention, as it is strictly fulfilled only when the reference atmosphere has an adiabatic lapse rate. In this case, the potential temperature is composed of two parts: the constant (reference) Θ, plus the (relatively small) disturbance θ, appearing in Eqs. (3.3) and (3.4). However, when the reference atmosphere is defined by a lapse rate which is different than the adiabatic one, there is an additional term in the definition of the potential temperature; this term is dependent upon z, but not on x, y or t, and is computed from the given (diabatic) lapse rate at the initial time. Thus:

$$\theta = \Theta + \bar{\theta}(z) + \theta'(x, y, z, t)$$
$$\pi = \Pi(z) + \bar{\pi}(z) + \pi'(x, y, z, t) \quad . \tag{3.11}$$

$\Pi(z)$ is computed from Θ, and $\bar{\pi}(z)$ from $\bar{\theta}(z)$. Therefore, Eq. (3.3) still holds, when θ' and π' are substituted for θ and π; similarly, the last of the boundary conditions (3.10) can be used for π' .

It should be emphasized that this is an approximation, introducing an inconsistency into the set of assumptions leading to the derivation of Eqs. (3.1) - (3.6). There, the atmosphere was assumed to be adiabatically stratified, and the deviations of the potential temperature and pressure, to be small. If we allow a lapse rate other than the adiabatic one, the potential temperature may vary with height by several degrees per kilometer; thus, deviations from a constant value in the reference atmosphere itself (even before introducing further disturbances) can easily become of the order of 50% near the top of a deep convective element (at, say, 15 km).

This apparent contradiction can be partially justified when $\bar{\pi}(z)$ and $\bar{\theta}(z)$ are assumed to balance each other at all times, and thus be removed from Eq. (3.3). We are left, then, only with the disturbances θ' and π', which are indeed small deviations from the reference (diabatic) atmosphere. This line of reasoning only amplifies another drawback in the derivation of these equations, namely, that as the disturbance develops, the mixing and overturning which follow result in the atmosphere reaching a "terminal" state which, necessarily, is different from its initial one. This can probably be ignored for small-scale, shallow convection; we do not know by how much this inconsistency affects the results when deep and moist convection is involved.

Finally, note that the sum of $\bar{\theta}(z)$ and $\theta'(x, y, z, t)$ enters into Eq. (3.4). This results in an additional term for the diabatic case, which carries most of the responsibility for the generation of internal gravity waves when the computational domain extends into the stratosphere.

The set of Eqs. (3.1)-(3.6) has, when certain conditions are fulfilled, a number of integrals of the motion. Some of these conservation properties possess immediate physical meaning, and will thus be involved in the ensuing discussion of numerical techniques to be chosen.

We define kinetic and potential energy, respectively:

$$E_k = \iiint \frac{1}{2} \rho(u^2 + v^2 + w^2)\,dxdydz \tag{3.12}$$

$$E_p = -\beta \iiint \rho z \theta \; dxdydz \tag{3.13}$$

And the momenta of θ:

$$\bar{\theta} = \iiint \rho\theta \, dxdydz \tag{3.14}$$

$$\overline{\theta^2} = \iiint \rho\theta^2 \, dxdydz \tag{3.15}$$

We also define the integrals:

$$\begin{aligned} A &= \gamma \iiint (u\nabla^2 u + v\nabla^2 v + w\nabla^2 w)\rho \, dxdydz \\ B &= -k\beta \iiint \rho z \nabla^2 \theta \, dxdydz \\ C &= \delta\beta \iiint \frac{\rho z}{\bar{T}} \frac{dq_v}{dt} \, dxdydz \\ D &= k \iiint \rho \nabla^2 \theta \, dxdydz \\ E &= 2k \iiint \rho\theta \nabla^2 \theta \, dxdydz \\ F &= \beta \iiint \rho w(0.608 q_v' - q_\ell) \, dxdydz \end{aligned} \tag{3.16}$$

It then follows that:

$$\frac{\partial}{\partial t}(E_p + E_k) = A + B + C + F \ . \tag{3.17}$$

Equation (3.17) can be established by multiplying Eqs. (3.1), (3.2) and (3.3) by ρu, ρv and ρw, respectively, multiplying Eq. (3.4) by $\rho\beta z$ and integrating by parts. Then the nonlinear terms, as well as the pressure in Eqs. (3.1)-(3.3), are integrated out.

In particular, if $\gamma = k = 0$, and for the dry case ($q = q_\ell = q_v \equiv 0$), energy is conserved:

$$E_p + E_k = \text{const.} \tag{3.18}$$

Upon integration of Eq. (3.4), for the dry case:

$$\frac{d}{dt}\bar{\theta} = D \tag{3.19}$$

and, since $k = 0$, potential temperature is conserved:

$$\bar{\theta} = \text{const.} \tag{3.20}$$

Similarly, multiplying Eq. (3.4) by θ and integrating:

$$\frac{d}{dt}\overline{\theta^2} = E \tag{3.21}$$

but since $k = 0$, the square of the potential temperature is conserved:

$$\overline{\theta^2} = \text{const.} \tag{3.22}$$

Consider now the scalar $e = \frac{1}{2}(u^2 + v^2 + w^2)$. This scalar satisfies the same boundary conditions as π does. Therefore, if we subtract $\frac{\partial e}{\partial x}$, $\frac{\partial e}{\partial y}$, $\frac{\partial e}{\partial z}$, from Eqs. (3.1), (3.2) and (3.3), respectively:

$$\frac{\partial u}{\partial t} + v(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}) + w(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}) - \gamma\nabla^2 u = -\frac{\partial}{\partial x}(\alpha\pi + e) \tag{3.23}$$

$$\frac{\partial v}{\partial t} + u(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}) + w(\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}) - \gamma\nabla^2 v = -\frac{\partial}{\partial y}(\alpha\pi + e) \tag{3.24}$$

$$\frac{\partial w}{\partial t} + u(\frac{\partial w}{\partial x} - \frac{\partial u}{\partial z}) + v(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}) - \gamma\nabla^2 w = -\frac{\partial}{\partial z}(\alpha\pi + e)$$

$$+ \beta(\theta + 0.608q'_v - q_\ell) \tag{3.25}$$

Equations (3.23)-(3.25) are identical to (3.1)-(3.3). The motivation for solving the equations in their latter form was, originally, that the conservation of energy (Eq. (3.18)) seemed better assured due to the fact that the nonlinear terms do not contribute locally to the generation of kinetic energy, which is not the case when using Eqs. (3.1)-(3.3). A-posteriori, we are not quite sure how important this factor is.

We will replace $(\alpha\pi + e)$ by a modified pressure to be denoted by f. By adding Eqs. (3.23)-(3.25), after multiplying them by ρ and differentiating, respectively, with respect to x, y and z, one gets the following diagnostic equation for f:

$$\frac{\partial}{\partial x}(\frac{\partial f}{\partial x}) + \frac{\partial}{\partial y}(\frac{\partial f}{\partial y}) + \frac{1}{\rho}\frac{\partial}{\partial z}(\rho\frac{\partial f}{\partial z}) =$$

$$= \frac{\partial}{\partial x}[-v(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}) - w(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}) + \gamma\nabla^2 u]$$

$$+ \frac{\partial}{\partial y}[-u(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}) - w(\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}) + \gamma\nabla^2 v] \tag{3.26}$$

$$+ \frac{1}{\rho}\frac{\partial}{\partial z}\{\rho[-u(\frac{\partial w}{\partial x} - \frac{\partial u}{\partial z}) - v(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}) + \gamma\nabla^2 w + \beta(\theta + 0.608q'_v - q_\ell)]\}.$$

IV. THE NUMERICAL MODEL

We are looking for a finite difference analogue to the set of differential Eqs. (3.23)-(3.25) and (3.5)-(3.6), so that the solution to the finite difference set would be a good approximation of the solution to the differential set of equations. Therefore, it is insufficient to require that the solution to the finite difference set be stable and accurate in itself, but we have also to insure that it approximates the solution to the differential equations. There are two generally acceptable ways for testing how close the two solutions are: conservation of integral properties of the flow, and telescoping of the grid. The philosophy behind the first point is quite simple: Since the set of differential equations possesses some conservative properties for some special cases (i.e., dry convection without diffusion), then the presence of these conservation properties in the solution to the finite difference analogue may be interpreted to mean that the two solutions are in fairly close agreement. This viewpoint led to the development of conservative finite difference schemes (see, for example, Arakawa (1966); Galloway (1968)]. However, we note that a-priori, "a conserving finite difference scheme" does not promise to be either stable or accurate and, furthermore, generally does not conserve all the integrals of the flow. For example, a conserving finite difference scheme, possessing the properties described by Eqs. (3.18) and (3.20), will not, in general, possess also the property described by Eq. (3.22). Therefore, in such a case, even though the solution to the finite difference analogue will conserve total energy as well as potential temperature, it will not conserve the square of the potential temperature, thus rendering erroneous "solutions" which may have nothing in common with the sought after solution to the set of differential equations.

Theoretically, it may be possible to devise such a finite difference scheme that will conserve all the important integrals of the motion; such a scheme, if stable, will be also a good approximation to the solution of the differential set. We have tried this approach and found it to be impractical because of its complexity. We have, then, decided to take the opposite approach, namely: to obtain a stable and accurate solution to a three-dimensional finite difference analogue to the set of differential equations, using conservation properties as an a-posteriori test on how well this solution approximates the solution to the differential set.

In addition to this test, we have used telescoping of the grid as well as the degree of accuracy within which the equation of continuity (3.5) is fulfilled, as further tests of the degree of proximity of our numerical results to the desired solution of the differential set. This last test becomes possible in view of the fact that Eq. (3.5) is never used directly in our numerical procedure.

The above comments relate, for the most part, only to the case of dry,

diffusionless convection, as it is only then that the total energy, as well as $\overline{\theta}$ and $\overline{\theta^2}$ are independent of time.

The inclusion of moisture poses another type of problem: The set of Eqs. (3.23)-(3.25) and (3.4)-(3.6), plus the additional equations relating the saturation vapor pressure to the instantaneous balance between q_v and q_ℓ, should be solved simultaneously in order to yield consistent results. This was never done before, to the best of our knowledge. Traditionally, a "separation of variables" was performed, whereby the set (u, v, w, π, θ) was advanced one step in time, then the set (q, q_ℓ, q_v) was computed for the new time; a temperature correction was then applied, and the procedure repeated for the next time step. This procedure, which does not take account of the interdependence of the saturation vapor pressure, the temperature and the pressure, can be shown to be consistent only for the case of shallow convection, i.e., when the Boussinesq approximation is made (see, for example, Ogura and Phillips (1962)]. In particular, it does not include a "feedback loop" to take account of effects of the dynamic pressure, which may be appreciable in the case of deep convection.

It therefore seemed profitable to use an implicit time advance method, thereby correcting and readjusting variables affected by the moisture cycle, within the iterations performed while advancing the variables from time t to time $t + \Delta t$. On the surface of it, this looks like a very costly procedure in terms of computer time; this would indeed be the case in general. However, the numerical procedure used in this study has the property that the length of the time step is proportional to the number of iterations performed within it, for a given accuracy of the solution. This means that, while not losing computer time by utilizing an implicit time advance method, we gain the advantage of being able to solve Eqs. (3.23)-(3.25) plus Eqs. (3.4)-(3.6) simultaneously, which at the moment seems the only consistent and practical way of obtaining numerical solutions to the deep, moist convection problem.

In view of the fact that integrals of the motion were one of the very few indicators available for testing the "quality" of the numerical solutions, the basic (and, as it turned out, the most difficult and time-consuming) part of this study consisted of constructing a model for the anelastic set of equations, for the case of dry, diffusionless convection. We will therefore present the numerical details of our model in two parts: first the dry, then the moist convection case.

1. Finite difference analogue to the anelastic equations for the case of dry convection

The following discussion relates to the case where $q = q_\ell = q_v = 0$.

We start by inscribing a rectangular net in the cube. The interval lengths may be different for x, y and z. In fact, throughout this study $2\Delta z = \Delta x = \Delta y$. All the functions computed are defined on the mesh points (i, j, k, s) where $i = 0, 1, \ldots, N_x$; $j = 0, 1, \ldots, N_y$; $k = 0, 1, \ldots, N_z$; $s = 0, 1, 2, \ldots,$. Each such point corresponds to $(i\Delta x, j\Delta y, k\Delta x, s\Delta t)$. Sometimes we might omit the dependence on t or on x, y or z.

As in the continuous case, we extend the definition of the computed functions u, v, w, θ, π to any triplet of integers (i, j, k), using the corresponding symmetry relations. For example, (3.9) will result in

$$u(nN_x + i, j, k) = - u(nN_x - i, j, k)$$

$$v(nN_x + i, j, k) = v(nN_x - i, j, k)$$

and so on. The computation itself is done only for points inside the fundamental cube. The symmetry relations are used, however, when differentiating at points close to the boundary.

Let us now define the difference operators, which will approximate the derivatives $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$. Define

$$\nabla_x^{(m)} f(i, j, k) = \frac{f(i+m, j, k) - f(i-m, j, k)}{2m\Delta x} \tag{4.1}$$

with similar definitions for $\nabla_y^{(m)}$ and $\nabla_z^{(m)}$. For all functions involved, there are either symmetry or antisymmetry relations. Thus, for example,

$$\nabla_z^{(3)}\theta(1, 1, 1) = \frac{\theta(1, 1, 4) - \theta(1, 1, -2)}{6\Delta z} = \frac{\theta(1, 1, 4) + \theta(1, 1, 2)}{6\Delta z}$$

$$\nabla_x^{(2)}\pi(1, 2, 3) = \frac{\pi(3, 2, 3) - \pi(-1, 2, 3)}{4\Delta x} = \frac{\pi(3, 2, 3) - \pi(1, 2, 3)}{4\Delta x} \tag{4.2}$$

We now fix the operators ∇_x, ∇_y, ∇_z, the approximants of $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}$, to be linear combinations of $\nabla_x^{(m)}$, $\nabla_y^{(m)}$, $\nabla_z^{(m)}$, etc. We might use the ordinary 2-point difference $\nabla_x = \nabla_x^{(1)}$ or (Collatz, 1959),

$$\text{4 points: } \nabla_y = \frac{1}{3}[4\nabla_y^{(1)} - \nabla_y^{(2)}], \text{ or} \tag{4.3}$$

$$\text{6 points: } \nabla_z = 1.5\nabla_z^{(1)} - 0.6\nabla_z^{(2)} + 0.1\nabla_z^{(3)}. \tag{4.4}$$

It should be emphasized that there are no interrelations between the difference operators ∇_x, ∇_y, ∇_z. If, for example, the nature of the problem is such that

one expects the solution to be more sensitive in the z direction than in the x or y directions, then one might try a 6-point formula for ∇_z, while saving time on the other directions and using, say, 2-point formulae in x and y.

Now define, for any quadruple u, v, w, θ, the vector A_1, A_2, A_3, A_4 as follows:

$$
\begin{aligned}
A_1(u,v,w,\theta) &= v(\nabla_y u - \nabla_x v) + w(\nabla_z u - \nabla_x w) - \gamma(\nabla_x\nabla_x + \nabla_y\nabla_y + \nabla_z\nabla_z)u \\
A_2(u,v,w,\theta) &= u(\nabla_x v - \nabla_y u) + w(\nabla_z v - \nabla_y w) - \gamma(\nabla_x\nabla_x + \nabla_y\nabla_y + \nabla_z\nabla_z)v \\
A_3(u,v,w,\theta) &= u(\nabla_x w - \nabla_z u) + v(\nabla_y w - \nabla_z v) - \gamma(\nabla_x\nabla_x + \nabla_y\nabla_y + \nabla_z\nabla_z)w - \beta\theta \\
A_4(u,v,w,\theta) &= u\nabla_x\theta + v\nabla_y\theta + w\nabla_z\theta - k(\nabla_x\nabla_x + \nabla_y\nabla_y + \nabla_z\nabla_z)\theta
\end{aligned}
\tag{4.5}
$$

Observe that the A_i, $i = 1, \ldots, 4$, satisfy the same symmetry relations as u, v, w, θ (this can be seen after establishing the consistency of the system (3.23)-(3.25), and (3.4)-(3.5), or by direct verification.

In order to achieve second order accuracy in time, we devise the following implicit time-advance method: Compute, between the time points t and $t + \Delta t$, the auxiliary functions u^*, v^*, w^*, θ^*, f^* at the time point $t + \frac{1}{2}\Delta t$

$$
\begin{aligned}
u^* - u(t) &= \frac{-\Delta t}{2} A_1(u^*, v^*, w^*, \theta^*) + \nabla_x f^* \\
v^* - v(t) &= \frac{-\Delta t}{2} A_2(u^*, v^*, w^*, \theta^*) + \nabla_y f^* \\
w^* - w(t) &= \frac{-\Delta t}{2} A_3(u^*, v^*, w^*, \theta^*) + \nabla_z f^* \\
\theta^* - \theta(t) &= \frac{-\Delta t}{2} \cdot A_4(u^*, v^*, w^*, \theta^*)
\end{aligned}
\tag{4.6}
$$

$$
\nabla_x\nabla_x f^* + \nabla_y\nabla_y f^* + \frac{1}{\rho}\nabla_z(\rho\nabla_z f^*) = -\nabla_x A_1^* - \nabla_y A_2^* - \frac{1}{\rho}\nabla_z(\rho A_3^*) . \tag{4.7}
$$

This is an implicit system which is solved by iterations, provided the solution to (4.7) is computed for each iteration. This equation will be dealt with separately later. The iterative process is:

$$
\begin{aligned}
u^*_{(n+1)} - u(t) &= \frac{-\Delta t}{2}\{A_1(u^*_{(n)}, v^*_{(n)}, w^*_{(n)}, \theta^*_{(n)}) + \nabla_x f^*_{(n)}\} \\
v^*_{(n+1)} - v(t) &= \frac{-\Delta t}{2}\{A_2(u^*_{(n)}, v^*_{(n)}, w^*_{(n)}, \theta^*_{(n)}) + \nabla_y f^*_{(n)}\} \\
w^*_{(n+1)} - w(t) &= \frac{-\Delta t}{2}\{A_3(u^*_{(n)}, v^*_{(n)}, w^*_{(n)}, \theta^*_{(n)}) + \nabla_z f^*_{(n)}\} \\
\theta^*_{(n+1)} - \theta(t) &= \frac{-\Delta t}{2} \cdot A_4(u^*_{(n)}, v^*_{(n)}, w^*_{(n)}, \theta^*_{(n)})
\end{aligned}
\tag{4.8}
$$

$$
\begin{aligned}
&(\nabla_x\nabla_x + \nabla_y\nabla_y + \frac{1}{\rho}\nabla_z\rho\nabla_z) f^*_{(n+1)} \\
&\quad = (-\nabla_x A_1 - \nabla_y A_2 - \frac{1}{\rho}\nabla_z \rho A_3)(u^*_{(n+1)}, v^*_{(n+1)}, w^*_{(n+1)}, \theta^*_{(n+1)}) \cdot
\end{aligned}
\tag{4.9}
$$

The iterative process is terminated when

$$
\frac{\max\{|\nabla_x f^*_{(n+1)} - \nabla_x f^*_{(n)}| + |\nabla_y f^*_{(n+1)} - \nabla_y f^*_{(n)}| + |\nabla_z f^*_{(n+1)} - \nabla_z f^*_{(n)}|\}}{|\overline{\nabla_z f^*_{(n+1)}}| + |\overline{\nabla_y f^*_{(n+1)}}| + |\overline{\nabla_z f^*_{(n+1)}}|} < \varepsilon \tag{4.10}
$$

where subscript n denotes the serial number of the iteration. If condition (4.10) is not fulfilled within the first ten iterations, Δt is halved, and the computation starts again at time t. ε was taken to be 10^{-4} in all production runs, after several trial runs yielded identical results with $\varepsilon = 10^{-5}$. This smaller ε did not result, generally, in an appreciable increase in the number of iterations per time step. The max is taken on all grid points (i, j, k). The last iterants are set, then, to be the starred quantities defined in (4.6)-(4.7).

After the computation is completed, we advance explicitly to time $t + \Delta t$, by equating the right-hand side of (4.6) to $u(t+\Delta t) - u^*$, $v(t+\Delta t) - v^*$, and so on, resulting in:

$$
\begin{aligned}
u(t+\Delta t) &= 2u^* - u(t) \\
v(t+\Delta t) &= 2v^* - v(t) \\
w(t+\Delta t) &= 2w^* - w(t) \\
\theta(t+\Delta t) &= 2\theta^* - \theta(t) \\
f(t+\Delta t) &= 2f^* - f(t)
\end{aligned}
\tag{4.11}
$$

Obtaining the solution to the diagnostic pressure equation (4.9) takes most of the computation time. It is therefore necessary to choose an efficient procedure for the solution of this type of equation. We used both the Fast Fourier Transform (FFT) in the x, y, z variables, and the transform in the x, y variables and Gauss elimination in the z variable.

When solving for the Boussinesq system (ρ = const.), either method can be used. However, when solving for the anelastic system [$\rho = \rho(z)$], FFT only in x and y and elimination in z is possible, due to the appearance of a first order derivative of π in (4.9).

2. The numerical process in solving for moist convection

When adding moisture as a dependent variable to the set of equations describing dry convection, Eqs. (3.25) and (3.4) include additional terms on the right-hand side, and Eq. (3.6) is added to the system. Auxiliary equations are needed, in order to determine the partition of q into q_ℓ and q_v. The complete system of equations for the moist convection case is as follows:

$$\frac{\partial u}{\partial t} + v\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right) + w\left(\frac{\partial u}{\partial z} - \frac{\partial w}{\partial x}\right) - \gamma\nabla^2 u = -\frac{\partial f}{\partial x} \tag{4.12}$$

$$\frac{\partial v}{\partial t} + u\left(\frac{\partial v}{\partial x} - \frac{\partial u}{\partial y}\right) + w\left(\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}\right) - \gamma\nabla^2 v = -\frac{\partial f}{\partial y} \tag{4.13}$$

$$\frac{\partial w}{\partial t} + u\left(\frac{\partial w}{\partial x} - \frac{\partial u}{\partial z}\right) + v\left(\frac{\partial w}{\partial y} - \frac{\partial v}{\partial z}\right) - \gamma\nabla^2 w = -\frac{\partial f}{\partial z} + \beta(\theta + 0.608 q_v' - q_\ell) \tag{4.14}$$

$$\frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} + w\frac{\partial q}{\partial z} = \nu\nabla^2 q_v \tag{4.15}$$

$$\frac{\partial \theta}{\partial t} + u\frac{\partial \theta}{\partial x} + v\frac{\partial \theta}{\partial y} + w\frac{\partial \theta}{\partial z} - k\nabla^2\theta = -\frac{\delta}{\bar{T}(z)}\left(\frac{\partial q_v}{\partial t} + u\frac{\partial q_v}{\partial x} + v\frac{\partial q_v}{\partial y} + w\frac{\partial q_v}{\partial z}\right) \tag{4.16}$$

$$\begin{aligned}
&\rho\left[\frac{\partial}{\partial x}\left(\frac{\partial f}{\partial x}\right) + \frac{\partial}{\partial y}\left(\frac{\partial f}{\partial y}\right)\right] + \frac{\partial}{\partial z}\left(\rho\frac{\partial f}{\partial z}\right) \\
&= \beta\frac{\partial}{\partial z}(\rho\theta) + \rho\left[\frac{\partial}{\partial x}(\nabla^2 u) + \frac{\partial}{\partial y}(\nabla^2 v)\right] + \gamma\frac{\partial}{\partial z}(\rho\nabla^2 w) - \\
&- \rho\left\{\frac{\partial}{\partial x}\left[v\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right) + w\left(\frac{\partial w}{\partial x} - \frac{\partial u}{\partial z}\right)\right]\right. \\
&\left. + \frac{\partial}{\partial y}\left[w\left(\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y}\right) + u\left(\frac{\partial u}{\partial y} - \frac{\partial v}{\partial x}\right)\right]\right\} -
\end{aligned} \tag{4.17}$$

$$- \frac{\partial}{\partial z} \{\rho[u(\frac{\partial w}{\partial x} - \frac{\partial u}{\partial z}) + v(\frac{\partial v}{\partial z} - \frac{\partial w}{\partial y})]\} + \beta \frac{\partial}{\partial z}[\rho(0.608 q'_v - q_\ell)]$$

$$q = q_v + q_\ell \tag{4.18}$$

$$e_s = 6107.8 \exp[\frac{17.2694(T - 273.16)}{T - 35.86}] \tag{4.19}$$

$$q_{vs} = \frac{0.622 e_s}{p - e_s} \tag{4.20}$$

Equation (4.19) was derived by Murray (1967). In addition, the pressure and temperature are (see, for example, Ogura and Phillips (1962)]:

$$p = p_0 \pi^{1/\kappa}, \text{ where } \pi = \Pi + \bar{\pi}(z) + \pi' \text{ and } \pi' = \frac{1}{\alpha}(f-e) \tag{4.21}$$

$$T = \pi\theta, \text{ where } \theta = \Theta + \bar{\theta}(z) + \theta' \tag{4.22}$$

The boundary conditions are as in Eqs. (3.8) and (3.10). The initial conditions are: $u = v = w = 0$ at $t = 0$; θ and q $(= q_v)$ are specified over the whole region of integration.

The numerical procedure is as follows: In the iterative cycle, as described by Eqs. (4.8) and (4.9) within a given time step, iterate Eqs. (4.12)-(4.15) once. Compute q_{vs}, with the aid of Eqs. (4.19)-(4.22), and set

$$q_v = r q_{vs} \text{ (but always } \leq q) \tag{4.23}$$

r is the relative humidity where condensation is allowed to begin. Compute q_ℓ from (4.18), and iterate once the remaining time dependent Eq. (4.16). Solve the diagnostic pressure Eq. (4.17) (for $f = \alpha\pi + e$). Repeat the iterative cycle within the current time step, until condition (4.10) is fulfilled, and then advance to the next time step.

3. Some numerical comments

The auxiliary starred functions in Eqs. (4.6) and (4.7) approximate the solution, at time $t + \frac{1}{2}\Delta t$, only to the first order. Nevertheless, the two-step cycle is accurate to second order. The solution of the implicit scheme is relatively fast, about 3-4 iterations per time step. The scheme is flexible with respect to differences in space. The computed results show that there is a considerable gain in accuracy, provided that a multipoint differentiation is used. This is important, in particular, for three-dimensional problems, where the bisection of a mesh results in an eight-fold increase in the computation

time. As far as stability is concerned, we feel that since the equations are nonlinear, the analysis is not pertinent.

A most important feature of the scheme is its consistency. We will demonstrate this by considering a diffusionless Boussinesq system. Consider the pressure equation in its differential form:

$$-\nabla^2 \pi = \left(\frac{\partial u}{\partial x}\right)^2 + \left(\frac{\partial v}{\partial y}\right)^2 + \left(\frac{\partial w}{\partial z}\right)^2 + 2\frac{\partial u}{\partial y}\cdot\frac{\partial v}{\partial x} + 2\frac{\partial u}{\partial z}\cdot\frac{\partial w}{\partial x} + 2\frac{\partial v}{\partial z}\cdot\frac{\partial w}{\partial y} \qquad (4.24)$$

This equation is obtained by differentiating Eqs. (3.1)-(3.3) with respect to x, y, z, respectively, then adding them and using the continuity Eq. (3.5) (for ρ = const.). When a difference equation analogue to (4.24) is constructed, known difference analogues of the Laplacian are generally used on its left side, and differentiation is performed on the right-hand side in any acceptable manner. If we solve the resulting equation, and substitute the derivatives of the pressure back into Eqs. (3.1)-(3.3), Eq. (3.5) will not, in general, be satisfied. The only way to satisfy Eq. (3.5) is to express the difference analogue of Eq. (4.24) in the form (4.9). There we have evaluated the Laplacian by repeated differentiation, and computed the right-hand side by a definite order of differentiation (without the use of identities, that are true for the differential equations, but may not be true for the difference equations!).

Apply, for example, a two-point differencing formula. The continuity Eq. (3.5) becomes:

$$\frac{u(i+1,j,k) - u(i-1,j,k)}{2\Delta x} + \frac{v(i,j+1,k) - v(i,j-1,k)}{2\Delta y} + \frac{w(i,j,k+1) - w(i,j,k-1)}{2\Delta z} = 0 \qquad (4.25)$$

The usual expression for the Laplacian is:

$$\nabla^2\pi \cong \frac{\pi(i+1,j,k) + \pi(i-1,j,k) - 2\pi(i,j,k)}{\Delta x^2} + \frac{\pi(i,j+1,k) + \pi(i,j-1,k) - 2\pi(i,j,k)}{\Delta y^2} + \frac{\pi(i,j,k+1) + \pi(i,j,k-1) - 2\pi(i,j,k)}{\Delta z^2} \qquad (4.26)$$

While the consistent expression for the Laplacian should be:

$$\nabla^2\pi \cong \frac{\pi(i+2,j,k) + \pi(i-2,j,k) - 2\pi(i,j,k)}{4\Delta x^2}$$

$$+ \frac{\pi(i,j+2,k) + \pi(i,j-2,k) - 2\pi(i,j,k)}{4\Delta y^2} \qquad (4.27)$$

$$+ \frac{\pi(i,j,k+2) + \pi(i,j,k-2) - 2\pi(i,j,k)}{4\Delta z^2}$$

If we are interested only in the solution of (4.24), as a separate problem, we will prefer (4.26) to (4.27). Our aim, though, is to solve (4.24), while fulfilling (3.5), and one way to insure this is to use the apparently less accurate, but more consistent, expression (4.27). We feel that this point is the source of large errors in the known schemes for solving this differential system.

As was mentioned earlier, conservation laws were used to check on the degree of correspondence between the computed solution to the difference equations and the desired solution to the differential equations. However, this can be done only for cases of diffusionless, dry convection.

In order to have some checks on cases where diffusion is present, the following definitions are made:

$$E_r(E_p/E_k) = \frac{E_p(t+\Delta t) - E_p(t) - \Delta t \cdot B(t+\tfrac{1}{2}\Delta t)}{E_k(t+\Delta t) - E_k(t) - \Delta t \cdot A(t+\tfrac{1}{2}\Delta t)}$$

$$E_r(\bar{\theta}) = \frac{\bar{\theta}(t) + \Delta t \cdot D(t+\tfrac{1}{2}\Delta t)}{\bar{\theta}(t+\tfrac{1}{2}\Delta t)} \qquad (4.28)$$

$$E_r(\overline{\theta^2}) = \frac{\overline{\theta^2}(t) + \Delta t \cdot E(t+\tfrac{1}{2}\Delta t)}{\overline{\theta^2}(t+\tfrac{1}{2}\Delta t)}$$

where A, B, D, E are defined in Eqs. (3.16). Equations (4.28) are computed at every time step, and their proximity to unity is an indicator, similar to that of the conservation tests in the dry, diffusionless case.

Finally, a similar test for cases where both diffusion and moisture are present is made by computing

$$E_r(E_p + E_k) \qquad (4.29)$$

$$= \frac{E_p(t+\Delta t) + E_k(t+\Delta t) - \Delta t[A(t+\tfrac{1}{2}\Delta t) + B(t+\tfrac{1}{2}\Delta t) + C(t+\tfrac{1}{2}\Delta t) + F(t+\tfrac{1}{2}\Delta t)]}{E_p(t) + E_k(t)}$$

As it turned out, $E_r(E_p/E_k)$, $E_r(\bar{\theta})$, $E_r(\overline{\theta^2})$ and $E_r(E_p + E_k)$ stayed very close to 1.0 throughout the computations.

V. RESULTS OF NUMERICAL EXPERIMENTS: DRY CONVECTION

Throughout these experiments, the following constants (in c. g. s.) were used:

$$
\begin{aligned}
T_0 &= 288°k & \kappa &= R/C_p = \frac{1}{3.5} \\
P_0 &= 10^6 & L &= 2.49234 \times 10^{10} \\
R &= 2.8704 \times 10^6 & \tau &= 600 \\
\Theta &= 288°k & D &= 3.2 \times 10^5 \\
C_p &= 3.5R & g &= 981
\end{aligned}
\qquad (5.1)
$$

Table 1 summarizes results of eleven dry convection runs. Following are details of the various parameters appearing in this table:

Dimensions. Three different types of experiments were run:

a. Two-dimensional, symmetrical experiments in half-plane, denoted 2S. In this case, $v \equiv 0$ at all times, and all other variables are independent of y.

b. Three-dimensional, symmetrical experiments in one-quarter space, denoted 3S. In this case:

$$
\begin{aligned}
u(l_i, l_j, l_k) &= v(l_j, l_i, l_k) \\
w(l_i, l_j, l_k) &= w(l_j, l_i, l_k) \\
\theta(l_i, l_j, l_k) &= \theta(l_j, l_i, l_k)
\end{aligned}
\qquad (5.2)
$$

These identities are used throughout the computations, in order to save time. A trial run, where these identities were not used, yielded identical results.

c. Three-dimensional, asymmetrical experiment, denoted 3. In this experiment no symmetry relations are used, as the initial disturbance is asymmetrical.

Grid. Four types of grids were used:

#1 Coarse grid for two-dimensional, symmetrical experiments (2s):

Nx = 17; Nz = 33. This corresponds to Δx = 200 meters; Δz = 100 meters.

#2 Fine grid for two-dimensional, symmetrical experiments (2S): Nx = 33; Nx = 65; Δx = 100 meters; Δz = 50 meters.

#3 Three-dimensional, symmetrical experiments (3S): Nx = Ny = 17; Nz = 33; $\Delta x = \Delta y$ = 200 meters; Δz = 100 meters.

#4 Three-dimensional, asymmetrical experiments (3): Nx = Ny = 17; Nz = 33; $\Delta x = \Delta y$ = 200 meters; Δz = 100 meters.

Space differencing. The number of neighboring points participating in space differencing in any given direction: 2, 4 or 6, corresponding to Eqs. (4.1) (for m = 1), (4.3) or (4.4), respectively.

Diffusion. Three cases were run: $\gamma = k = 0$; $\gamma = k = 0.001$ (or $\cong 1.5 \times 10^5$ in c.g.s. units; these values are similar to the ones used by Ogura (1963a)); and $\gamma = k = 0.002$.

Initial disturbance. Three different functions were used for $\theta(t = 0)$:

#1 $$\theta = \bar{\theta}(z) + \theta_0 \exp\left[-\frac{(i-Nx)^2 + (j-Ny)^2}{30.2^a}\right] \sin^2\left(\frac{\pi k}{12.2^b}\right) \qquad (5.3)$$

For grids #1 and #3, a = b = 0; the disturbance extends vertically to 13 layers.

For grid #2, a = 2 and b = 1; the disturbance extends vertically to 25 layers.

#2 The equation for θ is similar to #1 above, with the only difference that the argument of the function $\sin^2$ is $(\pi k)/(6.2^b)$.

For grids #1 and #3, a = b = 0; the disturbance extends vertically to 7 layers.
For grid #2, a = 2 and b = 1; the disturbance extends vertically to 13 layers.

#3 $$\theta = \bar{\theta}(z) + \theta_0 \exp\left[-\frac{(i-11)^2 + (j-6)^2}{5}\right] \sin^2\left(\frac{\pi k}{6}\right) \qquad (5.4)$$

Note that disturbances #1 and #2 are symmetrical (when used in a half-plane in two dimensions, or a quarter-space in three dimensions), while disturbance #3 is asymmetrical. $\theta_0 = 1\,^\circ C$ in all three initial disturbances. $\bar{\theta}(z) = 0$ for runs 1-11.

Density. For the Boussinesq system (denoted "B"), ρ = const. For the anelastic system (denoted "A"), we used the following dependence of the density on height, which corresponds to a hydrostatically stratified atmosphere:

$$\rho = \rho(z) = \left[-\kappa g \left(\frac{R\Theta}{p_0^{\kappa}} \right)^{1/(\kappa-1)} z + \rho_0^{\kappa/(1-\kappa)} \right]^{(1-\kappa/\kappa)}, \quad \text{where } \rho_0 = \frac{p_0}{RT_0} \tag{5.5}$$

In discussing the results of the eleven experiments appearing in Table 1, we will first comment on the integral properties of the flow in each experiment (Eqs. (3.12)-(3.22)), and then discuss in some detail results from one of these experiments.

TABLE 1

Summary of Parameters for Eleven Dry Convection Experiments

Experiment no.	Dimensions	Grid	Space Differencing	Diffusion	Initial Disturbance	Density
1	2S	#1	2	0.	#1	B
2	2S	#1	4	0.	#1	B
3	2S	#1	6	0.	#1	B
4	2S	#2	6	0.	#1	B
5	2S	#1	2	0.002	#2	B
6	2S	#1	6	0.002	#2	B
7	2S	#2	6	0.002	#2	B
8	2S	#2	6	0.	#2	A
9	3S	#3	6	0.	#2	B
10	3S	#3	6	0.002	#2	B
11	3	#4	6	0.001	#3	B

Figure 1 shows the absolute magnitude of the ratios between the increments in potential and kinetic energies between the present and initial times, for Experiments 1-4. As can be seen, as we progress from a two-point formula for space differentiation to four- and six-point formulae, this ratio improves (Experiments 1-3). One may question this statement, in view of the shape of curves 2 and 3, but we interpret this to mean that, had the experiment been continued beyond 15 minutes, a better average (closer to 1.00) would have been maintained for Experiment 3 for a longer period of time than for Experiment 2. Experiment 4 is identical with Experiment 3, except that the grid intervals were halved; the results of this are exhibited in a striking manner in Figure 1; Curve 4 stays very close to unity to the end of the experiment.

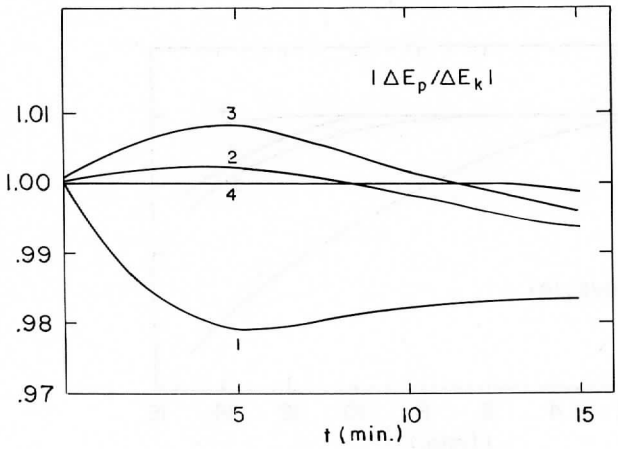


Fig. 1: $|\Delta E_p/\Delta E_k|$, Experiments 1-4.

Figure 2 shows the ratio between the total energy at time t and initially. Because of the way potential energy is defined, the total energy is a negative quantity throughout the experiment, and this is why an increase in total energy with time shows up as a decrease (from unity) of this ratio. The progression from the least sophisticated scheme (Experiment 1) to the most sophisticated one (Experiment 4) manifests itself as a monotonic improvement in this ratio; in particular, large improvements are noted between Experiments 1 and 2, and between Experiments 3 and 4.

Figure 3 is a diagram of the ratio between θ^2 at times t and 0, both averaged over the whole region, as defined by Eq. (3.15); this ratio turned out to be the most sensitive indicator of the stability of the numerical solutions in the various experiments. Again, we find a monotonous improvement when moving from Experiment 1 to Experiment 4, which stays very close to unity throughout the experiment. A similar diagram, showing the ratio between the average potential temperatures (as defined in Eq. (3.14)), between times t and 0 for Experiments 1-4, is not shown, as these ratios were always practically 1.00 for all four experiments.

Figures 4-7 are similar to Figures 1-3, but for Experiments 5, 6 and 7; these show the effects of multipoint differentiation and telescoping of the grid, for schemes including diffusion of both momentum and temperature.

In Figure 4, the ratio between the increments of the potential and kinetic energies is always > 1.0, as could be expected; one notes the oscillatory nature of this function, particularly for Experiments 5 and 7, with Experiment 7 showing a first maximum already at $t = 5$ minutes.

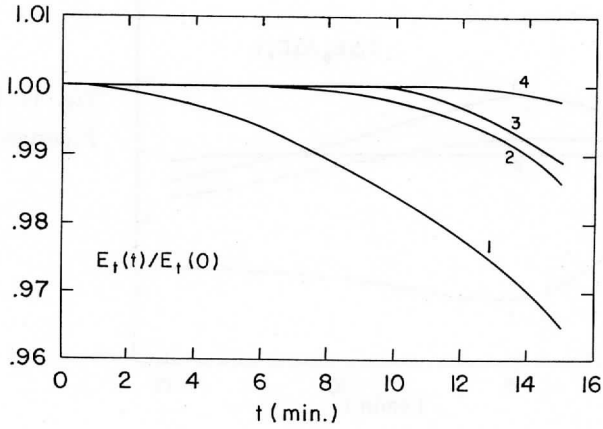


Fig. 2: $E_t(t)/E_t(0)$, Experiments 1-4.

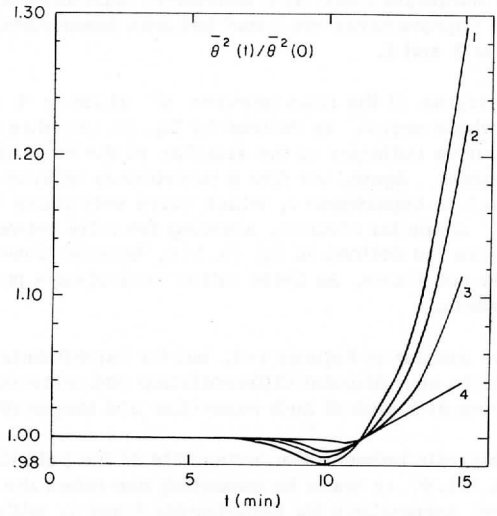


Fig. 3: $\overline{\theta^2}(t)/\overline{\theta^2}(0)$, Experiments 1-4.

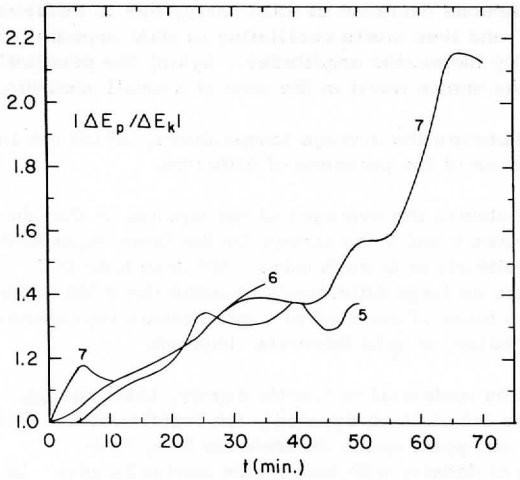


Fig. 4: $|\Delta E_p/\Delta E_k|$, Experiments 5-7.

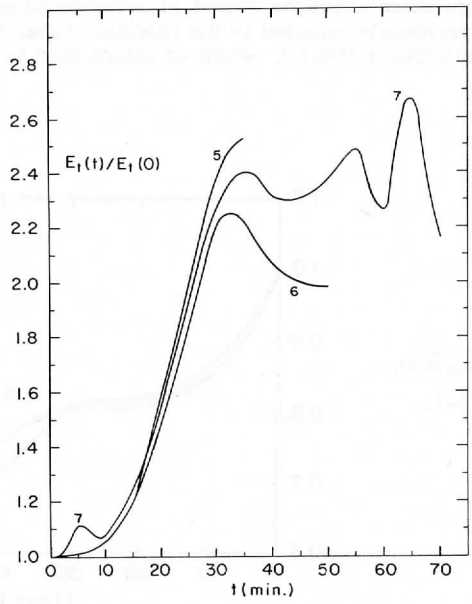


Fig. 5: $E_t(t)/E_t(0)$, Experiments 5-7.

In Figure 5, showing the ratios between total energies, a rather smooth increase (resulting from a monotonous decrease of total energy due to diffusion) continues up to 30-35 minutes, and then starts oscillating in what appears (for Experiment 7) to be exponentially increasing amplitudes. Again, the peculiarity of Experiment 7 at $t = 5$ minutes shows itself in the form of a small maximum.

Figure 6 shows the ratios between the average temperatures, which are indeed expected to decrease because of the presence of diffusion.

Figure 7 shows the ratios between the averages of the squares of the temperatures; note that both in Figures 6 and 7 the curves for the three experiments are very similar in shape and quite close to each other. We also note that throughout Figures 4-7, there are no large differences between the three experiments, and there certainly is no trace of any kind of a monotonous improvement as the space differentiation formulae, or grid intervals, improve.

Figure 8 shows all four ratios (potential to kinetic energy, total energy, average temperature and average temperature squared), for Experiment 8. This experiment is characterized by six-point space differences; fine, two-dimensional grid; and variation of density with height (the anelastic set). As can be seen, all curves stay very close to unity until about 12 minutes' time, where the $\overline{\theta^2}$ curve starts to increase exponentially, until it blows up around 15 minutes. Note that the curve for $\bar{\theta}$ is completely unaffected! This raises a disturbing question about the degree of accuracy of numerical experiments in convection, previously reported in the literature (see, for example, Ogura and Charney (1960); Ogura (1962)), where no information is given on the behavior of $\overline{\theta^2}$.

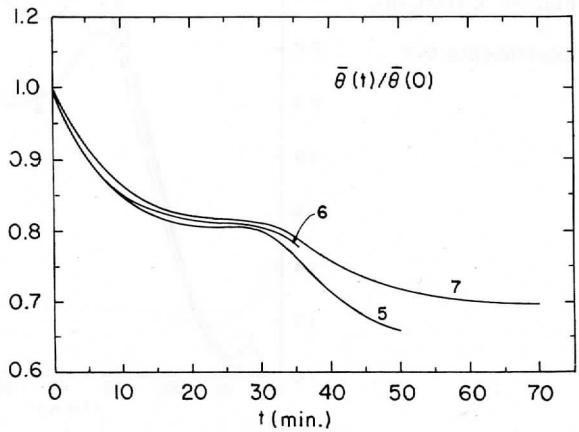


Fig. 6: $\bar{\theta}(t)/\bar{\theta}(0)$, Experiments 5-7.

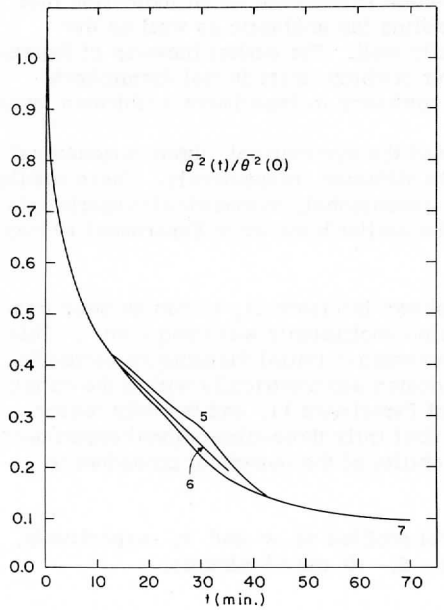


Fig. 7: $\overline{\theta^2}(t)/\overline{\theta^2}(0)$, Experiments 5-7.

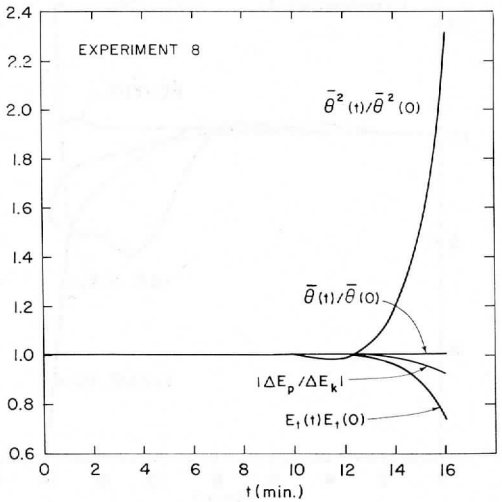


Fig. 8: $|\Delta E_p/\Delta E_k|$, $E_t(t)/E_t(0)$, $\overline{\theta^2}(t)/\overline{\theta^2}(0)$, $\overline{\theta}(t)/\overline{\theta}(0)$, Experiment 8.

Experiment 8, which is very similar to Experiment 4 (for the Boussinesq set), shows a very similar behavior, which is regarded as an indication that the proposed scheme is capable of handling the anelastic as well as the Boussinesq systems of equations equally well. The earlier blow-up of Experiment 8 may be the result of the stronger gradient in its initial disturbance (seven layers) as versus the initial disturbance in Experiment 4 (thirteen layers).

Figures 9 and 10 show the results of the symmetrical, three-dimensional Experiments 9 and 10, without and with diffusion, respectively. These results are very similar to the results of (two-dimensional, symmetrical) Experiments 4 and 7, respectively. The reason for the earlier blow-up in Experiment 10 may be its coarser grid.

The results of Experiment 11 are shown in Figure 11; as can be seen from Table 1, a smaller value for the diffusion coefficients was used (.001). This is the only experiment started by an asymmetric initial disturbance (actually, the disturbance is symmetric, but is located asymmetrically within the cube); the results are quite similar to those of Experiment 11, and the only reason they are shown is because this is the first truly three-dimensional experiment, which was run in order to test the capability of the numerical procedure to handle fully three-dimensional flows.

Figures 12 and 13 show the vertical profiles of w and θ, respectively, at the center of the disturbance, for t = 5, 10 and 15 minutes.

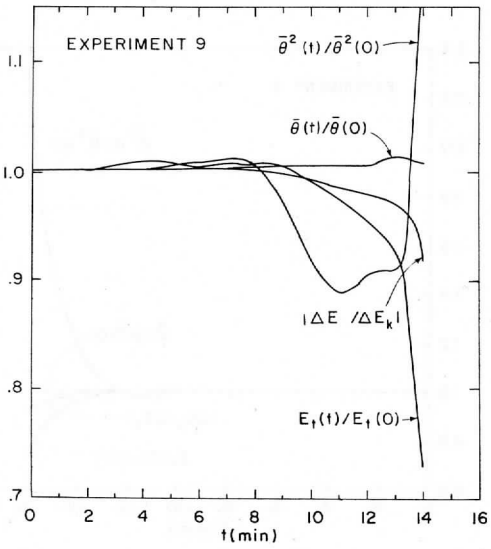


Fig. 9: $|\Delta E_p/\Delta E_k|$, $E_t(t)/E_t(0)$, $\overline{\theta^2}(t)/\overline{\theta^2}(0)$, $\bar{\theta}(t)/\bar{\theta}(0)$, Experiment 9.

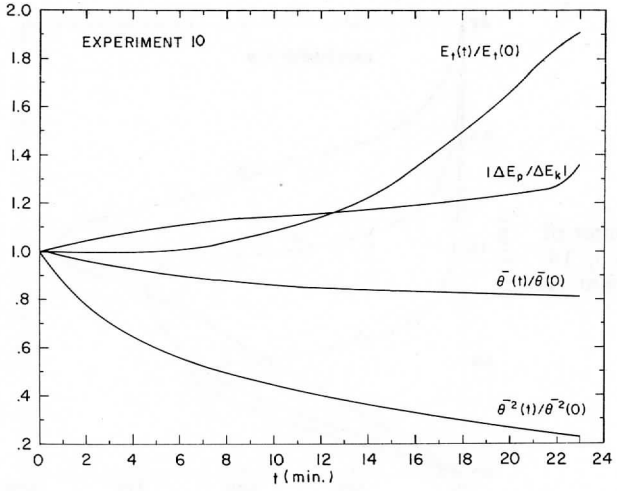


Fig. 10. Similar to Figure 9, but for Experiment 10.

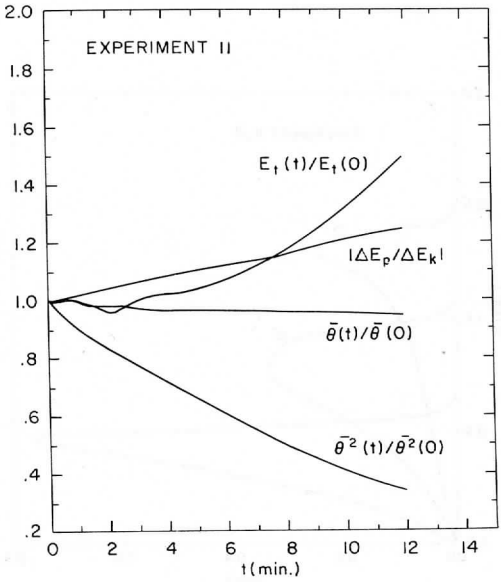


Fig. 11: Similar to Figure 9, but for Experiment 11.

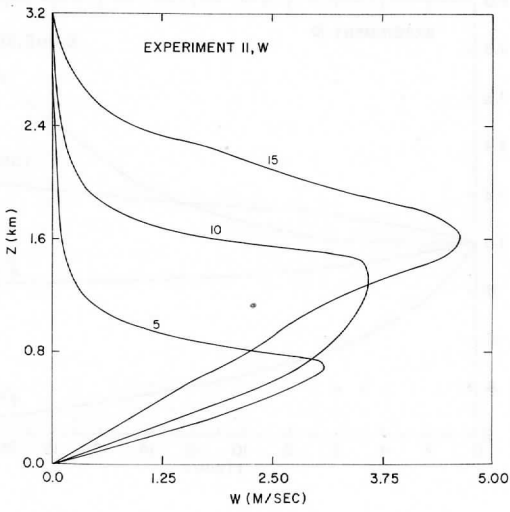


Fig. 12: w(z) at the center of the disturbance, for t = 5, 10 and 15 minutes, Experiment 11.

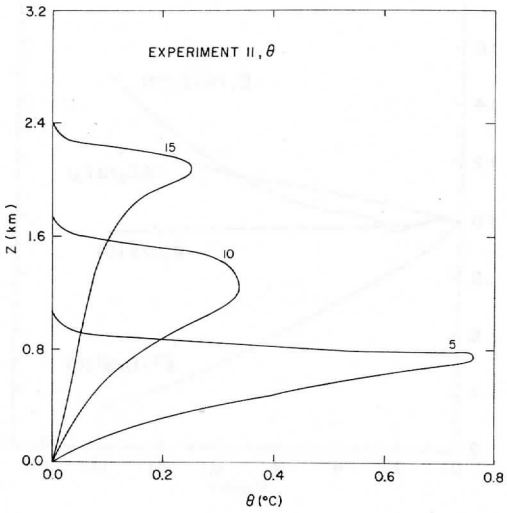
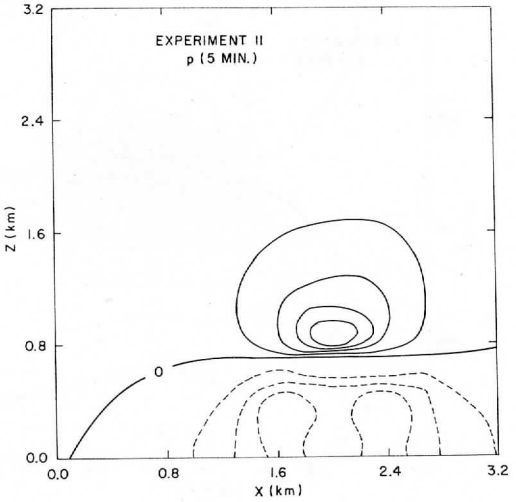
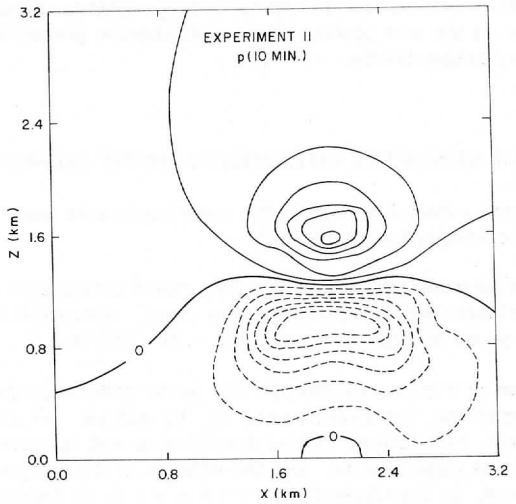


Fig. 13: θ(z) at the center of the disturbance, for t = 5, 10 and 15 minutes, Experiment 11.



(a)



(b)

Figs. 14(a)-(c): Vertical cross-section, through the center of the disturbance, of the perturbation pressure, for t = 5, 10 and 15 minutes, respectively, Experiment 11. Isobars for Δp = 5 dyn/cm^2; solid = positive; dashed = negative.

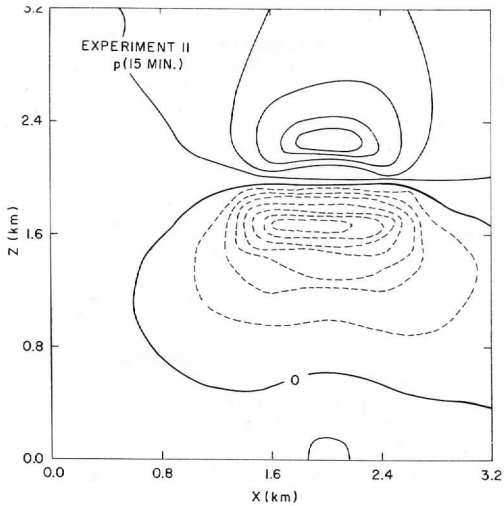


Fig. 14(c)

Figures 14(a)-14(c) show a vertical cross-section, through the center of the disturbance at the x-z plane, of the perturbation pressure, for t = 5, 10 and 15 minutes, respectively.

VI. RESULTS OF NUMERICAL EXPERIMENTS: MOIST CONVECTION

Except where otherwise noted, the same constants were used as in the dry convection experiments (see page 56).

Table 2 summarizes the results of five moist convection runs. In all five cases, the anelastic set of equations was used, where the variation of density with height is given by Eq. (5.5) and $T(z = 0) = 288°k$.

In each one of Figures 15 through 20, some general property of the flow is shown, in comparison, for Experiments 12, 13 and 14. As can be noted from Table 2, the grid, resolution, initial disturbance and relative humidity are identical for these experiments, and the effects of the inclusion of diffusion are tested; hence, in Experiment 12, $\gamma = k = \nu = 0$; in Experiment 13, $\gamma = k = 0.002$, $\nu = 0$; and in Experiment 14, $\gamma = k = \nu = 0.002$. In these experiments, the potential temperature is constant (= 288°k) with height.

Figure 15 shows $|\Delta E_p/\Delta E_k|$ for Experiments 12 through 14. There is a marked difference between the curves for Experiments 12 and 13, as versus

TABLE 2

Summary of Parameters for Five Moist Convection Experiments

Exp. No.	Dimensions	Grid	D (km)	Diffusion			$\frac{\partial \bar{T}}{\partial z}$ °C	Initial Disturbance		Relative Humidity (%)	
				γ	k	ν		Type	θ_0° C	r(t=0)	r (condensation)
12	2S	#1	3.2	0.	0.	0.	-10.0	#2	1	100	100
13	2S	#1	3.2	0.002	0.002	0.	-10.0	#2	1	100	100
14	2S	#1	3.2	0.002	0.002	0.002	-10.0	#2	1	100	100
15	2S	#2	16	0.	0.	0.	-7.2*	#2	5	80	80
16	3	#4	16	0.	0.	0.	-7.2*	#3	5	80	80

*Up to 10 km; $\frac{\partial T}{\partial z} = 0.0$ between 10 and 16 km.

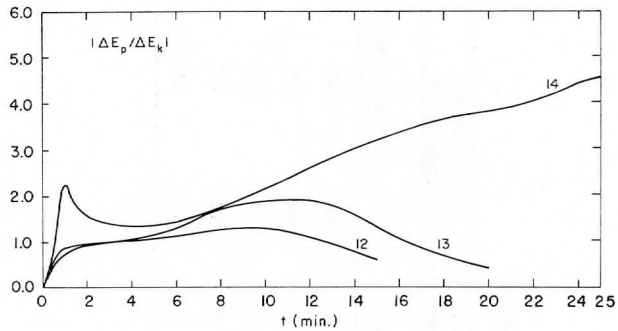


Fig. 15: $|\Delta E_p/\Delta E_k|$, Experiments 12-14.

Experiment 14, whereas Experiments 12 and 13, lacking in at least one of the diffusion coefficients, blow up after 12-15 minutes of moist convection, Experiment 14 stabilizes after initial "adjustments" ($\sim$ 1-2 minutes), with the curve continuing to rise, as expected. A similar phenomenon is detected in Figures 16, 17 and 18, where $E_t(t)/E_t(0)$, and $\overline{\theta}(t)/\overline{\theta}(0)$, and $\overline{\theta^2}(t)/\overline{\theta^2}(0)$ are shown, respectively. The indication for a blow-up in $\overline{\theta^2}(t)/\overline{\theta^2}(0)$ for Experiment 12 should be noted at about 12 minutes (Figure 18), while $\overline{\theta}(t)/\overline{\theta}(0)$ for this experiment is still stable (Figure 17). At least part of the reason for the difference in behavior between experiments where all three diffusion coefficients are present and those where at least one diffusion coefficient is absent is revealed in Figures 19 and 20; Figure 19 shows the ratio between the integrated mixing ratio $\overline{q}(t)$ over the whole volume at time t, and its initial value $\overline{q}(0)$; Figure 20 shows the ratio between the integrated mixing ratio for liquid water, $\overline{q}_\ell(t)$, at time t, and the initial mixing ratio $\overline{q}(0)$. While the inclusion of diffusion in the equation for q causes an almost uniform decrease with time in $\overline{q}(t)/\overline{q}(0)$, thereby also limiting the release of latent heat in Experiment 14, the diffusionless Experiments 12 and 13 convert vapor into liquid water at an exponential rate (Figure 20), thus providing excessive amounts of energy to drive the circulation, and causing the blow-up of the numerical experiment.

Experiment 15 is radically different from Experiments 12 through 14. It is characterized by a fine grid; a height of 16 km; temperature lapse rate of 7.2° C/km up to 10 km, and isothermal stratification above 10 km; $\gamma = k = \nu = 0.0$; amplitude of 5°C in the initial disturbance; and relative humidity of 80% both initially and for the onset of condensation.

This experiment is viewed as a first attempt to numerically simulate deep, moist atmospheric convection. As such, the results turn out to be most interesting indeed.

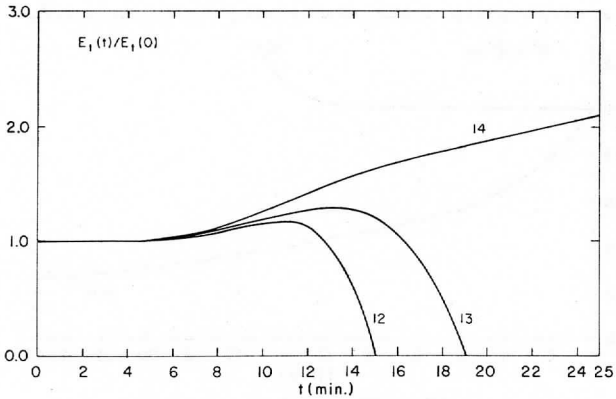


Fig. 16: $E_t(t)/E_t(0)$, Experiments 12-14.

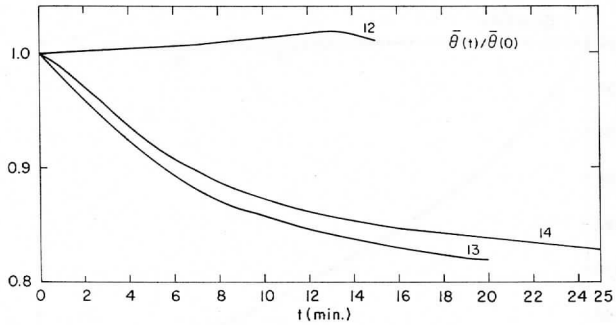


Fig. 17: $\bar{\theta}(t)/\bar{\theta}(0)$, Experiments 12-14.

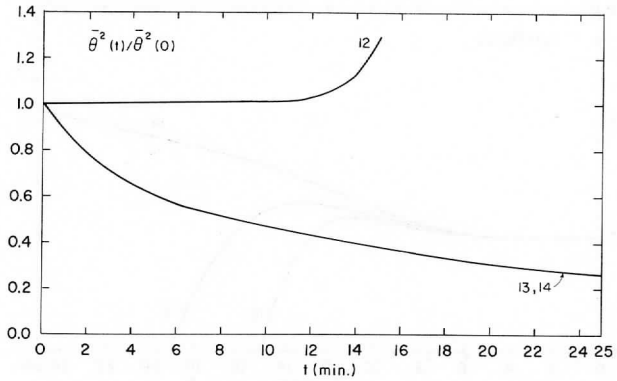


Fig. 18: $\overline{\theta^2}(t)/\overline{\theta^2}(0)$, Experiments 12-14.

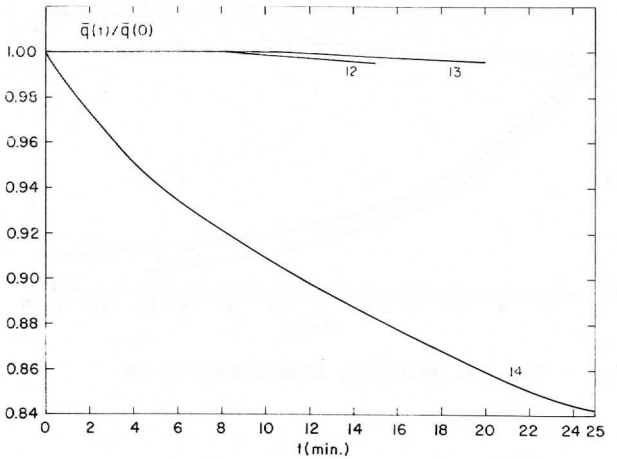


Fig. 19: $\bar{q}(t)/\bar{q}(0)$, Experiments 12-14.

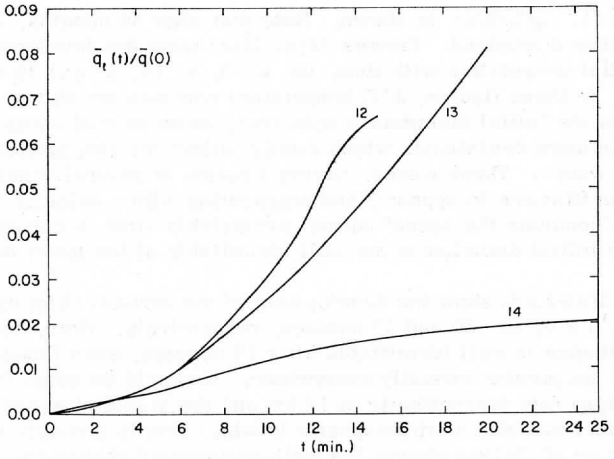


Fig. 20: $\bar{q}_\ell(t)/\bar{q}(0)$, Experiments 12-14.

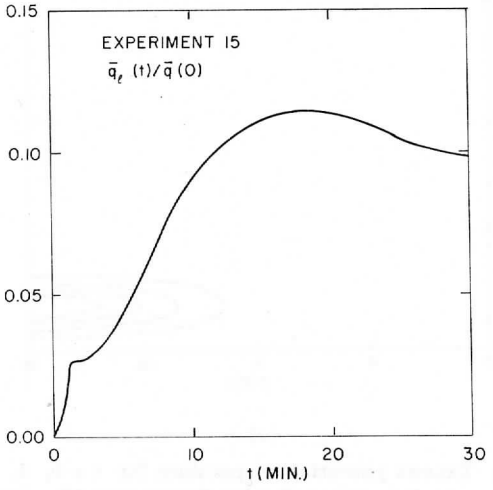
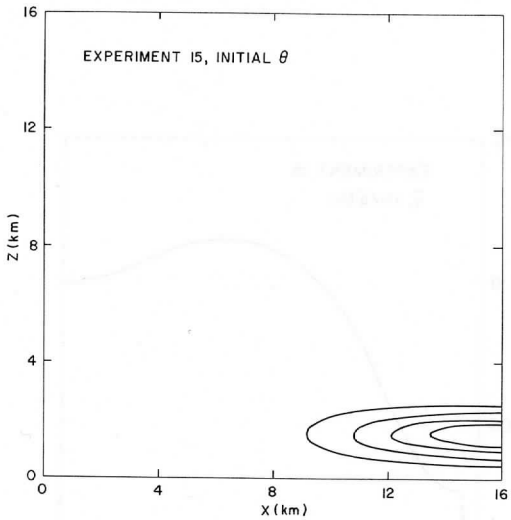


Fig. 21: $\bar{q}_\ell(t)/\bar{q}(0)$, Experiment 15.

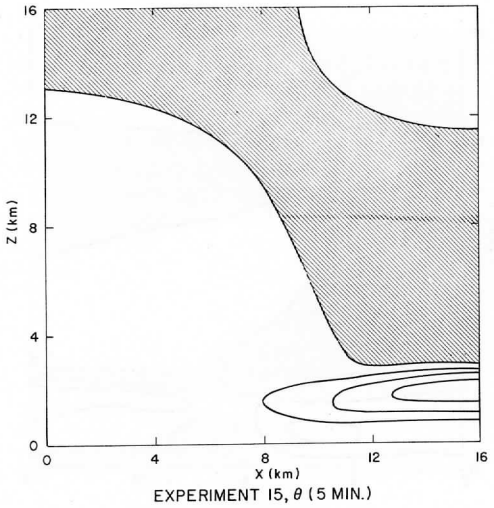
In Figure 21, $\bar{q}_\ell(t)/\bar{q}(0)$ is shown. Note that after 30 minutes, about 10% of the total water condensed. Figures 22(a)-22(e) show the development of the excess potential temperature with time, for t = 0, 5, 10, 20 and 30 minutes, respectively. In these figures, 1°C temperature contours are shown for the development of the initial disturbance with time, while shaded areas denote negative temperature deviations, which mostly reflect the propagation of internal gravity waves. These waves, having a period of several minutes (~ 5 minutes for the first one to appear), and propagating with a velocity of about 1.5 km/min, "dominate the scene" almost exclusively after 30 minutes. The remains of the initial disturbance are still identifiable at the lower left.

Figures 23(a)-23(d) show the development of the vertical velocity field with time, for t = 5, 10, 20 and 30 minutes, respectively. Here, too, the original disturbance is still identifiable after 30 minutes, even though internal gravity waves are present virtually everywhere. It should be noted that the temperature lapse rate discontinuity at 10 km and the stable stratosphere above it result in slight updrafts at stratospheric levels. This is probably the numerical manifestation of "billow clouds," a well-documented phenomenon.



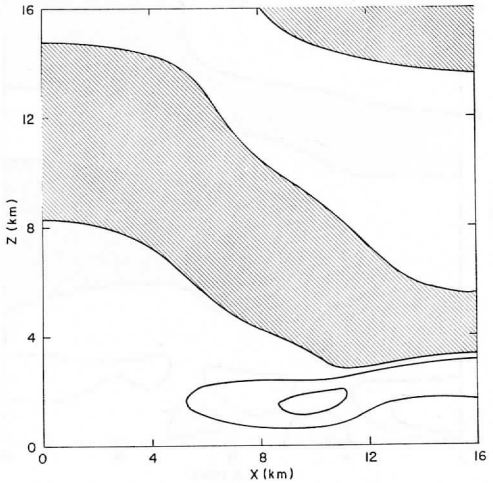
(a)

Figs. 22(a)-22(e): Excess potential temperature for t = 0, 5, 10, 20 and 30 minutes, respectively, Experiment 15. Isotherms for $\Delta\theta$ = 1°C; areas with negative θ are shaded.



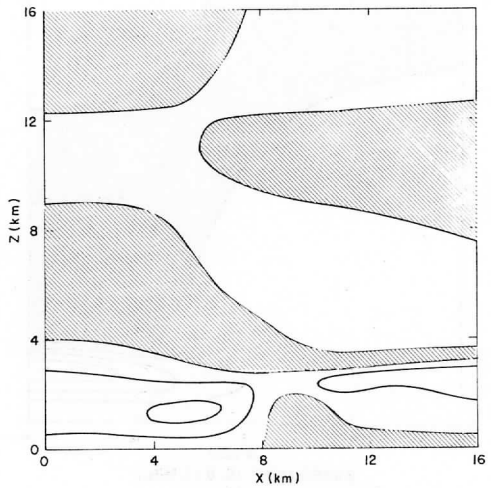
EXPERIMENT 15, θ (5 MIN.)

Fig. 22(b)



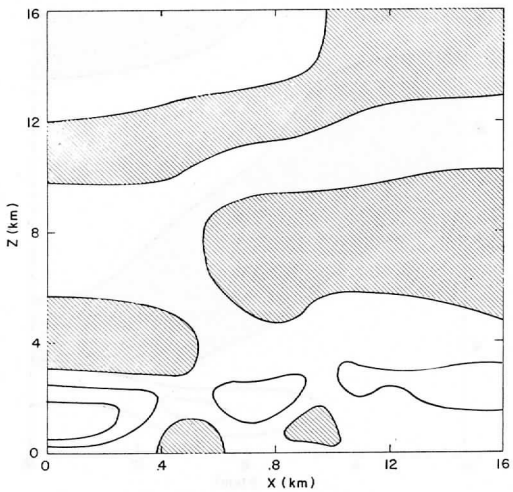
EXPERIMENT 15, θ (10 MIN.)

Fig. 22(c)



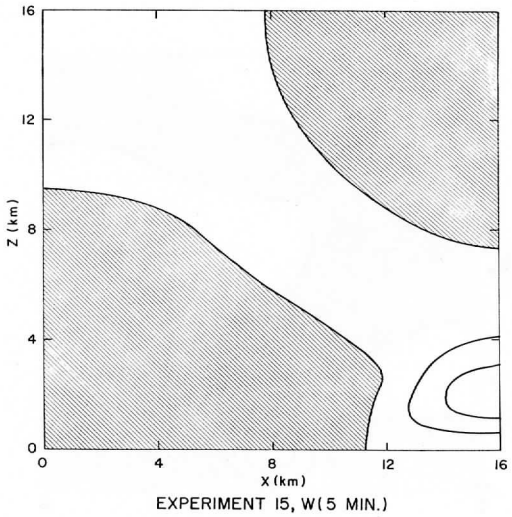
EXPERIMENT 15, θ (20 MIN.)

Fig. 22(d)

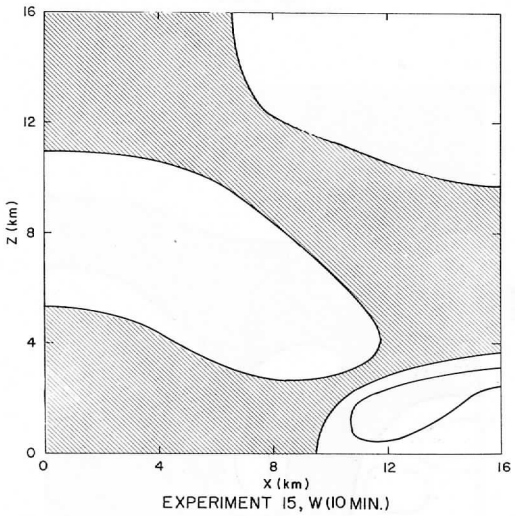


EXPERIMENT 15, θ (30 MIN.)

Fig. 22(e)

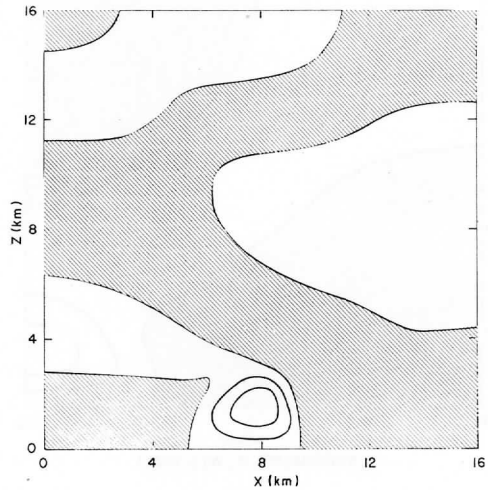


(a)



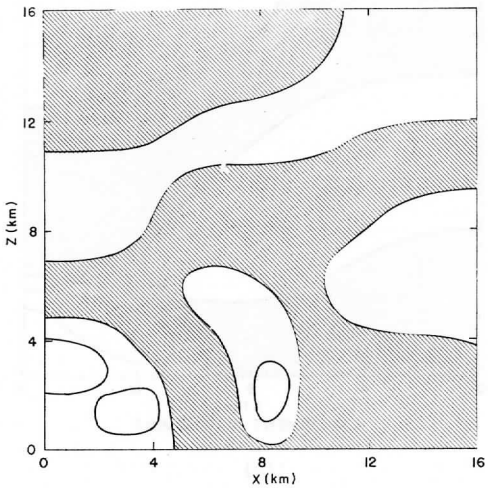
(b)

Figs. 23(a)- 23(d). Vertical velocity for t = 5, 10, 20 and 30 minutes, respectively, Experiment 15. Isopleths for $\Delta w = 1.0$ m/sec; downdraft areas are shaded.



EXPERIMENT 15, W (20 MIN.)

Fig. 23(c)



EXPERIMENT 15, W (30 MIN.)

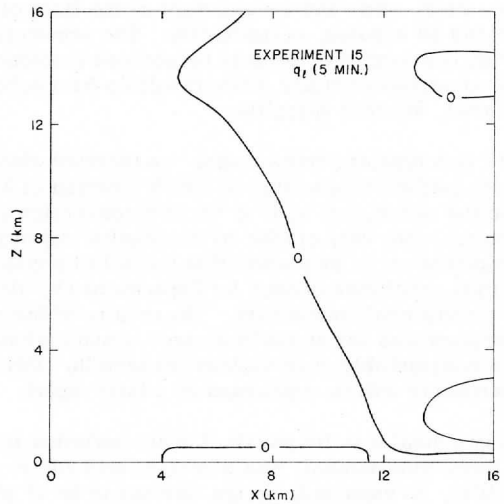
Fig. 23(d)

Figures 24(a)- 24(d) show the development of the field of q_ℓ with time, for t = 5, 10, 20 and 30 minutes, respectively. The accumulation of liquid water accompanying convective activity is recognized (contours every 1.0 g/kg), we well as slight quantities of liquid water resulting from subtle, internal-gravity-wave-induced, vertical velocities.

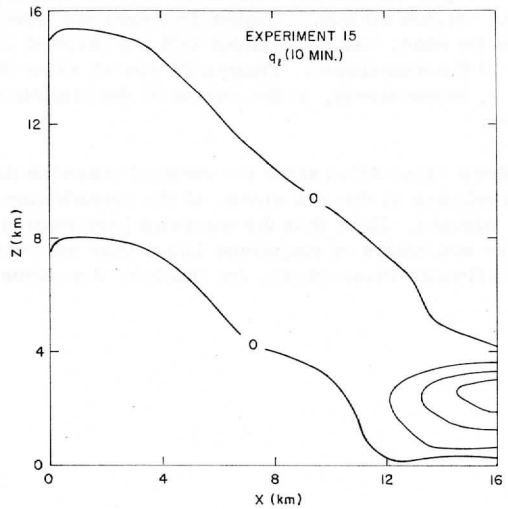
Experiment 15 was repeated three times: as reported above; similar to the above, but with surface temperature of 300°k (instead of 288°k); and, finally, similar to the above, but with $q \equiv 0$ (dry convection). The results of the dry convection run were very similar to the results shown for Experiment 15; this was interpreted as an indication that the set of physical parameters, as well as the initial conditions chosen for Experiment 15, did not contain a very large latent (conditional) instability. The results of the experiment where the surface temperature was set at 300°k showed a much greater release of latent heat, and a considerably more violent convection. Details of this and other similar experiments will be discussed in a later report.

Experiment 16 is similar in its physical characteristics to Experiment 15, except that it is three-dimensional, has a coarser grid and is asymmetrical. Therefore, its results, as expected, do not turn out to be of great interest, when compared with those of Experiment 15. In general, they exhibit the development of an asymmetrical main convection cell, acccompanied by internal gravity waves and "billow clouds." Figure 25 shows the development of $\overline{q_\ell}(t)/\overline{q}(0)$; as can be seen, values of about 15% are reached within the 15 minutes' history of the experiment. Figures 26 and 27 show the vertical profiles of w and θ, respectively, at the center of the disturbance, for t = 5, 10 and 15 minutes.

Finally, Figures 28(a)-28(c) show the vertical cross-section, through the center of the disturbance at the x-z plane, of the perturbation pressure, for t = 5, 10 and 15 minutes. Note that the maximum perturbation pressure in this experiment is over two orders of magnitude larger than the corresponding value at Experiment 11 (Figures 14(a)-14(c)), for shallow, dry convection.



(a)



(b)

Figs. 24(a)-24(b): Liquid water mixing ratio for t = 5, 10, 20 and 30 minutes, respectively, Experiment 15. Isopleths for Δq_ℓ = 1.0 g/kg; the zero isopleth at $q_\ell = 10^{-7}$ g/kg .

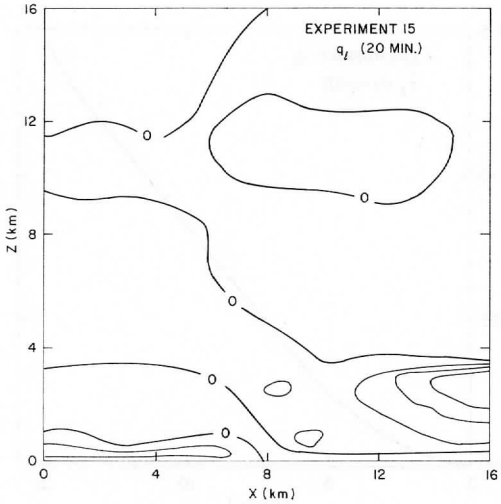


Fig. 24(c)

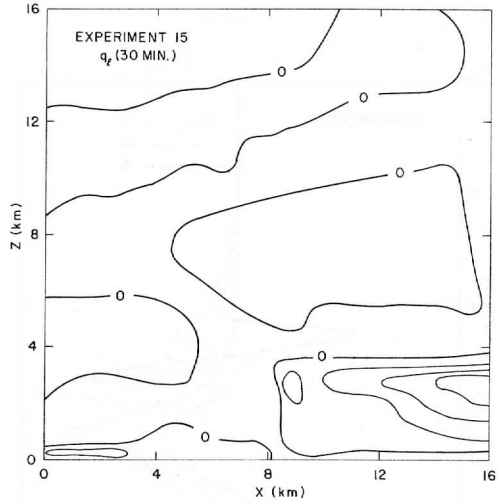


Fig. 24(d)

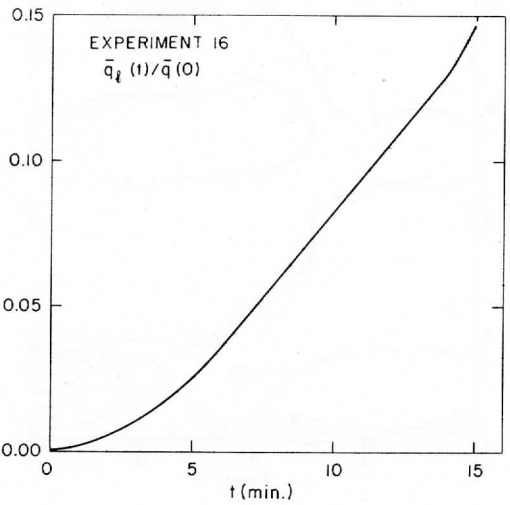


Fig. 25: $\bar{q}_\ell(t)/\bar{q}(0)$, Experiment 16.

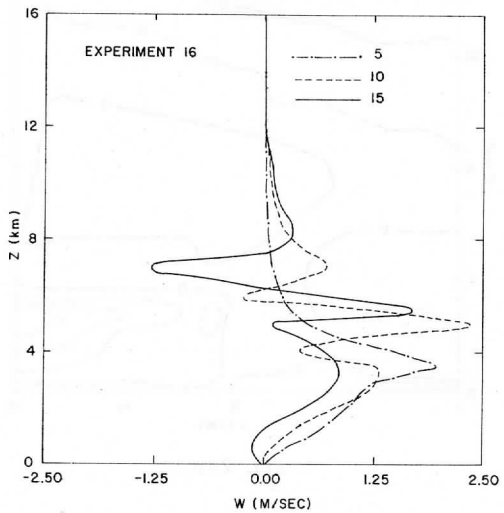


Fig. 26: w(z) at the center of the disturbance, for t = 5, 10 and 15 minutes, Experiment 16.

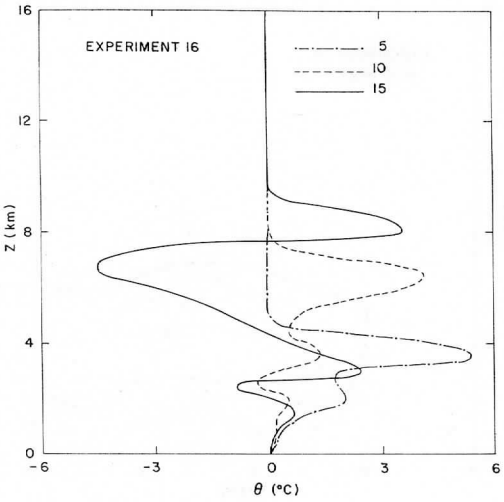
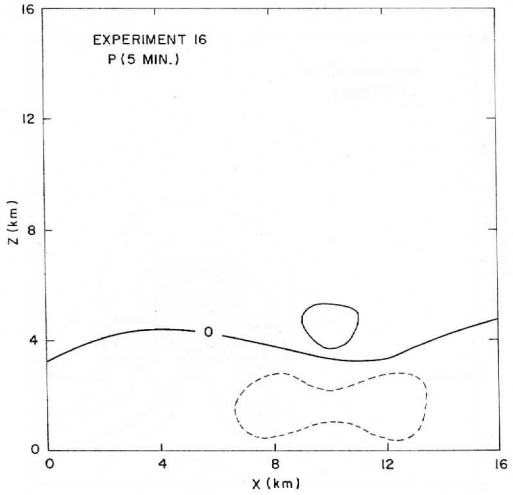


Fig. 27: $\theta(z)$ at the center of the disturbance, for t = 5, 10 and 15 minutes, Experiment 16.



(a)

Figs. 28(a)-28(c): Vertical cross-section, through the center of the disturbance, of the perturbation pressure, for t = 5, 10 and 15 minutes, respectively, Experiment 16. Isobars for Δp = 250 dyn/cm^2; solid = positive; dashed = negative.

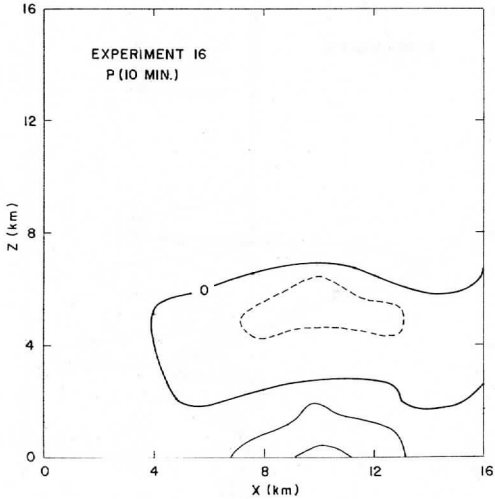


Fig. 28(b)

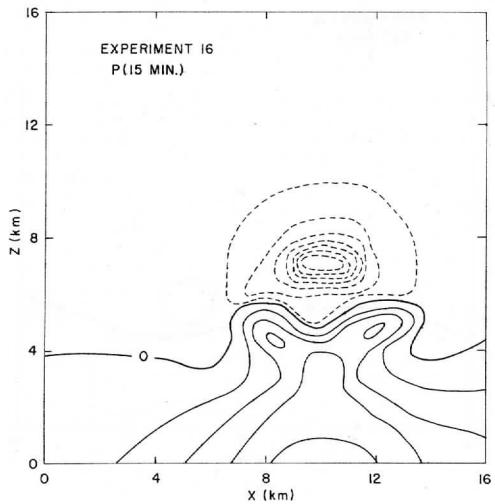


Fig. 28(c)

VII. CONCLUSIONS

In this paper, a method was presented for the numerical solution of the deep, moist atmospheric convection problem. The method was tested for cases where moisture was absent in symmetrical and asymmetrical cases in two and three dimensions, as well as for cases of axial symmetry. It was also tested for the incompressible case (the Boussinesq approximation).

Preliminary results show the development of moist, deep convection in realistic environmental conditions in terms of relative humidity and lapse rate, including the existence of the tropopause and stratosphere. It is shown that internal gravity waves, triggered by the convective activity, cause "billow clouds" to appear a short time after the onset of convection.

It is concluded that it is now possible to experiment numerically with three-dimensional convective flows, characterized by the set of anelastic equations, and taking full account of the interplay between dynamic and thermodynamic variables, caused by the release of latent heat of condensation.

However, the fact that this is possible does not necessarily make it practical as well. It took several hours of an IBM 360/95 computer to produce 15 minutes of results for the moist, fully three-dimensional Experiment 16, for a $17 \times 17 \times 33$ grid. As was concluded from other experiments reported above, this resolution is not nearly sufficient. We feel that a resolution of $33 \times 33 \times 33$, or $33 \times 33 \times 65$, is required in order to obtain meaningful, long-term results. While this may be possible to achieve on the 360/95, with greater effort going into optimalization of the program, the introduction into the model of meteorologically interesting and important (but, eventually, time-consuming) effects, such as wind shear, superimposed large-scale flow field (with net convergence effects), and microphysical effects inside the cloud, may put it out of the realm of practicality for today's computers. However, it seems that even if this happens, we may be short by about one order of magnitude on speed (or throughput), which may be furnished by the next generation of computers, expected to become operational in a couple of years.

Until then, future experiments should concentrate on the introduction of microphysical effects into the model, as well as on investigating the effects of entrainment, and on allowing mass transport through the boundaries. This can be done with the aid of two-, or symmetrical three-dimensional models, which today's computers are capable of handling easily.

ACKNOWLEDGMENTS

The authors are indebted to Professor Verner E. Suomi for his moral and financial support during the initial phase of this research, which was carried out at the Space Science and Engineering Center at the University of Wisconsin on ESSA Grant E-230-68-G. We are grateful to Professor R. W. Hockney, of IBM's Watson Research Center, New York, for letting us use his program for Fast Fourier Transform. Thanks are also due to Mrs. M. Broman, of Computer Applications, Inc., for her help in programming the graphical output of these experiments.

REFERENCES

Arakawa, A., 1966: Computational design for long-term numerical integration of the equations of fluid motion: Two-dimensional incompressible flow. Part I, J. Computational Phys., 1, 119-143.

Asai, T., 1960: Numerical experiment of convection in the model atmosphere. Proc. International Symposium on Numerical Weather Prediction in Tokyo, Tokyo, Meteor. Soc. Japan, 469-476.

Batchelor, G. K., 1953: The conditions for dynamical similarity of motions of a frictionless perfect-gas atmosphere. Quart. J. Roy. Meteor. Soc., 79, 224-235.

Collatz, L., 1959: The Numerical Treatment of Differential Equations, 3rd ed. Berlin, Springer-Verlag, 568 pp.

Galloway, F. M., Jr., 1968: On a finite-difference approach to turbulence problems. Ph.D. dissertation, School of Engineering, Case Western Reserve University, Cleveland, Ohio, 219 pp.

Gutman, L. N., 1961: The theory of cumulus cloudiness. Bull. Acad. of Sciences of USSR (Geophysical Series), Amer. Geophys. Union Translation, 688-698.

Lilly, D. K., 1962: On the numerical simulation of buoyant convection. Tellus, 14, 148-172.

Lorenz, E. N., 1962: Simplified dynamic equations applied to the rotating-basin experiments. J. Atmos. Sci., 19, 39-51.

Malkus, J. S., and G. Witt, 1959: The evolution of a convective element: A numerical calculation. The Atmosphere and the Sea in Motion, New York, Rockefeller Institute Press, 425-539.

Murray, F. W., 1967: On the computation of saturation vapor pressure. J. App. Meteor., 6, 203-204.

NCAR, 1967: Thermal convection: A colloquium. Tech. Notes, TN-24, National Center for Atmospheric Research, Boulder, Colo., 637 pp.

Ogura, Y., 1962: Convection of isolated masses of a buoyant fluid: A numerical calculation. J. Atmos. Sci., 19, 492-502.

_____, 1963a: The evolution of a moist convection element in a shallow, conditionally unstable atmosphere: A numerical calculation. J. Atmos. Sci., 20, 407-424.

_____, 1963b: A review of numerical modeling research on small scale convection in the atmosphere. Severe local storms, Meteor. Monogr. 5, No. 27, Boston, Amer. Meteor. Soc., 65-76.

_____ and J. G. Charney, 1960: A numerical model of thermal convection in the atmosphere. Proc. International Symposium on Numerical Weather Prediction in Tokyo, Tokyo, Meteor. Soc. Japan, 431-452.

_____ and N. A. Phillips, 1962: Scale analysis of deep and shallow convection in the atmosphere. J. Atmos. Sci., 19, 173-179.

Orville, H. D., 1965: A numerical study of the initiation of cumulus clouds over mountainous terrain. J. Atmos. Sci., 22, 684-699.

_____, 1968: Ambient wind effects on the initiation and development of cumulus clouds over mountains. J. Atmos. Sci., 25, 385-403.

Platzman, G. W., 1965: The spectral dynamics of laminar convection. J. Fluid Mech., 23, 481-510.

A NOTE ON INERTIAL INSTABILITY [1]

David D. Houghton and John A. Young
The University of Wisconsin, Department of Meteorology, Madison

Introduction

Blumen and Washington (1969) suggest that "pure" centrifugal instability should be a rare phenomenon in the atmosphere. This conclusion is based upon an analysis of a two-layer hydrostatic model and appears to contradict the concept of inertial instability set forth in textbooks on dynamic meteorology. In addition this result seems contrary to theoretical results concerning stability of a balanced vortex to cylindrically symmetrical disturbances. See, e.g., Eliassen (1951).

Acceptance of this general conclusion of Blumen and Washington (henceforth referred to as BW) would eliminate inertial instability as an explanation for the upper bound on observed synoptic scale wind shears on the equatorward side of jet streams. See Reiter (1961). It is the purpose of this note to discuss certain limitations of the analysis in BW and to offer a simple model which provides a more general understanding of inertial instability as it might occur in the atmosphere.

The special nature of the BW model is easy to understand. We recall that their model consisted of two incompressible layers separated by an interface of variable height. To facilitate analysis BW made the additional assumption that the upper layer was inert. This has the effect of limiting the basic geostrophic current to a special baroclinic form since the basic flow is not zero in the lower layer. The conclusions given in BW were thus inappropriate for general disturbances on a barotropic current for a two-layer model. On the other hand, the governing equations in BW coincide with those derived from a one-layer model using a reduced value of gravity. In this case the basic flow is indeed barotropic, but the disturbances are then constrained to be barotropic also.

It would seem that neither interpretation of the BW model corresponds to a model of general (i.e., baroclinic) disturbances on a barotropic base current.

[1] This research was supported under NSF Grant GA-12112 and E. S. S. A. Grant E-230-68-G.

The consequences of these special features for current stability can be discussed in terms of the interface slope in the mass continuity equation. As mentioned in BW, this basic slope has the effect of stabilizing the disturbances even when the horizontal divergence is zero.

The analysis shown in the next section is an attempt to provide a link between the results of BW (where the response of the pressure field to the disturbance motion is very special) and other results where the response is lessened or even eliminated by the use of particle dynamics.

Mathematical Analysis

The model considered here is hydrostatic, adiabatic and inviscid with no variations in the x-direction. A tangent plane is assumed. In all these respects this model is identical to that of BW. The primary difference is in the vertical structure. The BW model was essentially a one-layer model, and the dynamic interactions did not involve any vertical variation of the horizontal velocity within the lower layer. In the model considered here variations in the vertical are permitted.

The linear model is formulated in pressure coordinates with the mean state assumed to be barotropic, of a constant stability, and with a locally constant horizontal wind shear. The linearized perturbation equations can be written:

$$\frac{\partial u_1}{\partial t} + v_1\left(F\frac{dU_0}{dy} - 1\right) = 0, \tag{1}$$

$$\frac{\partial v_1}{\partial t} + u_1 + \frac{\partial h_1}{\partial y} = 0, \tag{2}$$

$$\frac{\partial \omega_1}{\partial p} + \frac{\partial v_1}{\partial y} = 0, \tag{3}$$

$$\frac{\partial}{\partial t}\left(\frac{\partial h_1}{\partial p}\right) + \omega_1 S = 0. \tag{4}$$

These are the two horizontal equations of motion, the continuity equation and the first law of thermodynamics, respectively. The notation is made as similar as possible to that used by BW:

t : time nondimensionalized by f, the constant Coriolis parameter,

y : the horizontal space coordinate in the y-direction nondimensionalized by the radius of deformation, defined as c/f where c is a reference horizontal phase speed of long gravity waves,

p : the vertical coordinate of pressure nondimensionalized by a reference pressure, p_r,

u_1 : perturbation x-component of velocity nondimensionalized by c,

v_1 : perturbation y-component of velocity nondimensionalized by c,

U_0 : Mean x-component of velocity nondimensionalized by c,

F : Froude number, the scaling magnitude of U_0,

h_1 : perturbation geopotential height of a pressure surface nondimensionalized by the depth c^2/g,

ω_1 : perturbation total pressure derivative nondimensionalized by p_r and f,

S : static stability of the mean state equal to $-\frac{p_r}{\rho_0 c^2} \frac{\partial \ln \theta_0}{\partial p}$ where $\rho_0(y, p)$ and $\theta_0(y, p)$ are the density and potential temperature of the mean state.

The coefficients of the governing equations are all assumed to be constant in space and time in order to facilitate the mathematical analysis. In order to evaluate local wave disturbances the following forms are assumed for the dependent variables:

$$u_1 = \text{Real}(\hat{u}\, e^{i(\ell y + mp - \sigma t)}), \tag{5a}$$

$$v_1 = \text{Real}(\hat{v}\, e^{i(\ell y + mp - \sigma t)}), \tag{5b}$$

$$\omega_1 = \text{Real}(\hat{\omega} e^{i(\ell y + mp - \sigma t)}), \tag{5c}$$

$$h_1 = \text{Real}(\hat{h}\, e^{i(\ell y + mp - \sigma t)}). \tag{5d}$$

Equations (1) to (4) are converted to equivalent equations for the Fourier coefficients $\hat{u}$, $\hat{v}$, $\hat{\omega}$, $\hat{h}$. This set of equations is homogeneous and possesses nontrivial solutions only if the determinant of the coefficients is zero. This condition yields the following frequency equation (if we disregard the geostrophic root, $\sigma = 0$):

$$\sigma = \pm \sqrt{\frac{\ell^2 S}{m^2} - (F\frac{dU_0}{dy} - 1)} \quad . \tag{6}$$

This indicates that instability will occur for

$$F\frac{dU_0}{dy} > \frac{\ell^2 S}{m^2} + 1 . \qquad (7)$$

Discussion

The two terms on the right hand side of the inequality (7) can be related to physical processes. If the first term is ignored, the condition for instability is

$$F\frac{dU_0}{dy} > 1 \qquad (8)$$

which is the result obtained from particle dynamics considerations where the pressure field is assumed to be independent of fluid motions. The first term on the right contains the static stability parameter S and can be related to the baroclinic velocity-pressure interactions in the fluid. If S is negative, then this term always tends to destabilize the flow (see equation (7)); i.e., the effects of static instability are incorporated in this analysis.

For stable stratification the term $\frac{\ell^2 S}{m^2}$ in (7) is proportional to the ratio of the radius of deformation to the horizontal length scale of the wave. This indicates that waves of small horizontal scale will be more stable than those of larger scale. The term also shows that shallow wave modes (large m) will be less stable than deep ones. The deeper wave modes correspond to the more barotropic wave motions, and in this respect the results are consistent with the result of the BW model where the barotropic disturbances are unconditionally stable.

When discussing instability in the real atmosphere, one must consider all possible vertical wave modes as well as the special barotropic one. The instability criterion (7) indicates that for large m the criterion approaches the classical criterion, $F\frac{dU_0}{dy} > 1$, and we can conclude that inertial instability is a relevant concept for baroclinic disturbances in the atmosphere. The analysis does show that the instability would tend to proceed with motions with vertical scales much less than the "total depth" of the atmosphere.

In closing, it should be noted that Eliassen (1951) shows that for a fluid with a baroclinic mean state, the tendency for instability is even greater than indicated by (7). This supports our contention that inertial instability is relevant to atmospheric motions and sets the results of BW into an even more special category. These conclusions also suggest that caution must be exercised in studying primitive equation models with a one-layer representation.

References

Blumen, W. and W. M. Washington, 1969, The effect of horizontal shear flow on geostrophic adjustment in a barotropic fluid. Tellus 21: 167-176.

Eliassen, A., 1951, Slow thermally or frictionally controlled meridional circulation in a circular vortex. Astrophysica Norveqica 5: 19-60.

Reiter, E. R., 1963, Jet Stream Meteorology. University of Chicago Press, Chicago, 515 pp.

THE INFLUENCE OF LATITUDINAL WIND SHEAR UPON LARGE-SCALE WAVE PROPAGATION INTO THE TROPICS

John R. Bennett and John A. Young
University of Wisconsin, Madison

Abstract

The effects of horizontal shear of the mean zonal wind on the lateral propagation of disturbances through the tropics is studied by the use of a one-layer model. The governing equations are reduced to a second-order differential equation for v, the northward component of velocity. The equation is analyzed as an eigenvalue problem and solved numerically for the free modes of the tropics for the case with zero mean flow. These solutions are compared with solutions which are forced at a boundary situated in mid-latitudes for cases with and without a mean zonal flow.

At "critical latitudes," the basic equation has a singularity (where the phase speed of a wave forced at the boundary is equal to the mean flow). The case for forced motions is investigated in more detail by numerically studying the evolution of disturbances as an initial value problem for the case of nondivergent flow.

The horizontal shear is shown to significantly alter the types of mid-latitude motions which can affect tropical motions. In particular, disturbances with large eastward phase propagation are shown to have negligible effect. Disturbances that have phase speeds which are somewhere equal to the mean flow are shown to be absorbed at the critical latitude. Disturbances with phase speeds more westward than the mean flow may be free to propagate into the tropics, providing their wavelengths are not too short.

Introduction

In the last few years there has been renewed interest in the study of large-scale tropical wave motions. To a certain extent the low spatial resolution of the observational network in the tropics has been offset by the use of analysis

of long time series of wind data at a few stations (Rosenthal, 1960; Marayuma, 1967). Theoretical works by Rosenthal (1965), Matsuno (1966), Koss (1967), and Lindzen (1967) have been of great value in interpreting the data since tropical motions may consist of several types of wave disturbances (such as Kelvin waves and mixed Rossby-gravity waves) which do not occur elsewhere. In a recent summary, Wallace (1969) described the results of observations of the Kelvin-type waves, the mixed Rossby-gravity modes (which are vertically propagating and occur mainly in the stratosphere), and two types of tropospheric modes which do not propagate vertically. With the exception of one of the tropospheric modes which has a wavelength of 3000 km, the others have wavelengths in excess of 10,000 km.

The sources of energy for synoptic scale Rossby wave motions at low latitudes is poorly understood at present. Suggested mechanisms have included tropical instabilities associated with latent heating or horizontal wind shear and the lateral coupling of the tropics with mid-latitude waves. For example, Charney and Eliassen (1964) and Rosenthal (1967) first studied the stability of quasi-balanced disturbances driven by organized convection; research on this conditional instability of the second kind is presently being continued by several investigators. Investigations into the possibility of barotropic instability at low latitudes have recently been accomplished by Nitta and Yanai (1969) and Lipps (1970).

The lateral coupling mechanism was proposed by Mak (1969) and demonstrated to be plausible in a simple two-layer stochastic model. A prominent dynamical influence in the model was the horizontal shear of the zonal wind. However, Mak chose to emphasize the statistics of the motions and did not attempt to explain the results in physical detail. This paper is an attempt to isolate the various effects of horizontal shear on this process of energy exchange between the tropics and mid-latitudes.

The propagation of energy through a shear flow arises in other physical problems such as the generation of surface water waves (Miles, 1962) and the upward propagation of gravity waves in the atmosphere. Internal gravity wave theory (Bretherton, 1966) predicts that certain layers (where the mean wind is equal to the trace speed of the wave) tend to absorb wave energy. Dickinson (1969) has recently studied Rossby wave absorption in the immediate vicinity of such a region. Other layers with varying mean flow profiles also redistribute the energy through the processes of reflection and tunnelling.

Charney (1969) discussed the problem of energy propagation into the tropics qualitatively by the use of WKB methods. His main conclusion was that westerly disturbances should not propagate far into an easterly regime. Mak's results are consistent with this idea in that the forced motions in the tropics of his model were found to exhibit predominantly easterly phase propagation, even when the bulk of the mid-latitude wave forcing was westerly.

The above work indicates that the characteristics of the mean flow may highly restrict the character of lateral forcing which is likely to produce effects in the tropics. The purpose of this work is to shed additional light on this interaction. To accomplish this task, the basic model is formulated and a wave equation is derived from it in section 1. Section 2 introduces a method of analysis of the problem involving free and forced solutions and illustrates it for the case of zero mean zonal flow. Section 3 treats forced motion in the presence of a zonal flow for the cases in which there is no wave absorption and comparison is made with the results of section 2. Section 4 contains a treatment of the case in which wave absorption occurs. The process of absorption is shown by solving the establishment of a wave regime as an initial value problem.

1. Formulation of the Model

Since we intend to focus on the effects of horizontal shear on the lateral propagation of wave energy, we will consider a model which allows only quasi-horizontal motion. This limits the applicability of the results to motions occurring in deep atmospheric layers. But, because of the paucity of observations, only these largest modes have been observed in detail. Examples where this approximation is valid are the tropospheric modes which do not propagate vertically which were discussed by Wallace (1969). These wave motions can occur in a simple one-layer model on an equatorial beta-plane such as Matsuno (1966) used. We wish to alter this model systematically by adding the effects of a zonal shear flow.

We consider a single layer of a homogeneous, inviscid fluid with a free surface in hydrostatic equilibrium. The equations for the fluid are linearized about a steady base state consisting of zonal flow, $U(y)$, which varies in the northward direction y. The linearized equations for the two horizontal velocity components (which depend on the two horizontal coordinates and time only) and the equation of continuity are:

$$\frac{\partial u'}{\partial t} + U\frac{\partial u'}{\partial x} + v'\frac{dU}{dy} - fv' + \frac{\partial \phi'}{\partial x} = 0 \tag{1}$$

$$\frac{\partial v'}{\partial t} + U\frac{\partial v'}{\partial x} + fu'\frac{\partial \phi'}{\partial y} = 0 \tag{2}$$

$$\frac{\partial \phi'}{\partial t} + U\frac{\partial \phi'}{\partial x} + v'\frac{\partial}{\partial y}(g'H) + g'H\left(\frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y}\right) = 0 \tag{3}$$

The primed quantities are perturbation variables and the capital letters (U, H) refer to the base state. In (3), g' is a scaled acceleration due to gravity. In our application of the model to the earth's atmosphere, Matsuno (1966) has shown that it may be thought of as a value reduced as follows:

$$g' = g\left(\frac{\Delta\bar{\theta}}{\bar{\theta}}\right)$$

where $\bar{\theta}$ is the mean potential temperature of the column and $\Delta\bar{\theta}$ is its difference over the vertical distance H.

Velocities in the x and y directions are u' and v' and the geopotential of the top surface is ϕ (= g'z). Because we are only considering deep motions, H can be replaced by H_0, an average of H over the whole domain of interest, when it is multiplied by the divergence in Eq. (3). The coriolis parameter, f, is assumed to vary linearly in the northward direction:

$$f = \beta y \tag{4}$$

where $\beta = \frac{df}{dy} = 2.29 \times 10^{-11}\ m^{-1}\ sec^{-1}$ at the equator. From the assumptions that the base state is composed of a steady zonal velocity which varies only with y, it follows that it is geostrophic:

$$\frac{dH}{dy} = -\frac{f}{g'}U \tag{5}$$

Equations (1), (2) and (3) can be made nondimensional by choosing time and length scales of:

$$\hat{T} = (g'H_0)^{-1/4}\beta^{-1/2}, \qquad \hat{L} = (g'H_0)^{1/4}\beta^{-1/2} \tag{6}$$

It follows that the velocity scale is $\hat{L}/\hat{T} = (g'H_0)^{1/2}$, the speed of long gravity waves in the model when β is set equal to zero. Using (4), (5) and (6), the equations for the nondimensional variables take the form:

$$\frac{\partial u'}{\partial t} + U\frac{\partial u'}{\partial x} + v'\frac{dU}{dy} - yv' + \frac{\partial \phi'}{\partial x} = 0 \tag{7}$$

$$\frac{\partial v'}{\partial t} + U\frac{\partial v'}{\partial x} + yu' + \frac{\partial \phi'}{\partial y} = 0 \tag{8}$$

$$\frac{\partial \phi'}{\partial t} + U\frac{\partial \phi'}{\partial x} - yUv' + \frac{\partial u'}{\partial x} + \frac{\partial v'}{\partial y} = 0 \tag{9}$$

Assuming solutions of the form:

$$u' = Re[u(y)e^{i(kx - kct)}]$$

$$v' = Re[v(y)e^{i(kx - kct)}]$$

$$\phi' = Re[\phi(y)e^{i(kx - kct)}].$$

The equations reduce to:

$$ik(U-c)u + v\frac{dU}{dy} - yv + ik\phi = 0 \tag{10}$$

$$ik(U-c)v + yu + \frac{d\phi}{dy} = 0 \tag{11}$$

$$ik(U-c)\phi - yUv + iku + \frac{dv}{dy} = 0 \tag{12}$$

By solving (10) for u and substituting in the last two equations, we obtain:

$$ik(U-c)v + \frac{y}{ik(U-c)}[yv - ik\phi - v\frac{dU}{dy}] + \frac{d\phi}{dy} = 0 \tag{13}$$

$$ik(U-c)\phi - yUv + \frac{1}{U-c}[yv - ik\phi - v\frac{dU}{dy}] + \frac{dv}{dy} = 0\,. \tag{14}$$

Differentiation of Eq. (14) with respect to y and substitution for $\frac{d\phi}{dy}$ from (13) and for ϕ from the undifferentiated form of (14) yields a single equation for v:

$$\frac{d^2v}{dy^2} + \frac{dv}{dy}[-yU - \frac{2(U-c)}{(U-c)^2-1}\frac{dU}{dy}] + v[k^2(U-c)^2 - k^2 + \frac{1}{U-c}$$

$$-y^2 - U - \frac{1}{U-c}\frac{d^2U}{dy^2} + (\frac{dU}{dy})^2[\frac{1}{(U-c)^2}][1 - \frac{1}{1-(U-c)^2}] + y^2U/U-c +$$

$$\frac{1}{(U-c)^2-1}\frac{dU}{dy}\langle yU(U-c) + yU/(U-c) + \frac{dU}{dy} - 2y\rangle] = 0\,. \tag{15}$$

Equation (15) is the general governing equation for waves in the tropics for this model. Because the speeds of different wave types vary so greatly, the terms in (15) are not of uniform significance. For example, the approximation $c \gg U$ may be a good one for gravity waves. On the other hand, our interest in the slower moving waves (quasi-geostrophic away from the equator) leads us to explore the case $c \sim U$. Thus under the conditions

$$c = O(U),\quad c \ll 1,\quad \frac{dU}{dy} = O(\frac{U}{L}),\quad y = O(L)$$

and using the fact that $\frac{1}{(U-c)^2}[1 - \frac{1}{1-(U-c)^2}]$ has magnitude unity, the terms of Eq. (15) can be scaled as follows:

$$\underline{1} \qquad \underline{L^2U} \qquad \underline{U^2} \qquad \underline{k^2L^2U^2} \qquad \underline{k^2L^2} \qquad \underline{L^2/U-c}$$

$$\underline{L^4} \qquad \underline{UL^2} \qquad \underline{U/U-c} \qquad \underline{U^2} \qquad \underline{\frac{UL^4}{U-c}}$$

U^3L^2 $\quad \frac{L^2U^2}{U-c} \quad$ U^2 $\quad UL^2$

U and L are characteristic nondimensional velocities and lengths. For flows in which $\sqrt{g'H_0} = 100\text{m/sec}$ ($\frac{\Delta\bar{\theta}}{\bar{\theta}} = 0.1$, $H_0 = 10$ km) the length scale from (6) is approximately 2100 km. If we consider this length scale as unity and assume $U \approx 0.1$, $k \approx 1.0$, the dominant terms in (15) (with scalings underlined) give:

$$\frac{d^2v}{dy^2} + v(-k^2 - y^2 + \frac{1}{(U-c)}[1 - \frac{d^2U}{dy^2} + y^2U]) = 0 \qquad (16)$$

This equation is the approximate governing equation for the propagation of large-scale, slowly moving wave motions in the tropics.

This equation may be thought of as a generalization of the equation derived by Matsuno (for the case $U \equiv 0$) to cases involving mean zonal shear flows. We consider the kinds of solutions to (16) in the following two sections. Section 2 considers the case $U = 0$ whereas section 3 considers the more general case $U(y) \neq 0$.

2. Free and Forced Waves in the Absence of a Mean Zonal Flow

The analysis of the last section developed a second order ordinary differential equation for v, the complex representation of the northward component of velocity of the form:

$$\frac{d^2v}{dy^2} + Q(k, c, U, y)v = 0 .$$

We now wish to use this equation to investigate the response of the tropics to forcing at the northern and southern boundaries. Since the equation is linear we can consider the forcing composed of a superposition of different wave numbers, k and phase speeds, c. The response of the tropics will then be the sum of the responses to the various boundary modes. Moreover, it is possible to separate the effects of the two boundaries by setting $v = 0$ at one boundary and $v \neq 0$ at the other and adding the response to the solutions obtained by reversing the procedure. This procedure of adding solutions obtained from the two boundaries for various frequencies ($= kc$) and wave numbers, k, has been used by Mak (1969) for a two-layer model with internal dissipation. Mak did not explain physically why the eastward-propagating waves at the boundary (which are typical of mid-latitude motions) produced a negligible response in the tropics. In order to isolate the effects of different forcing situations and the effect of the wind shear on the response, we first treat the problem without a zonal wind. Since we expect the forced response to be related to the free modes of the tropics, we begin by examining them.

When the mean velocity is set identically equal to zero, equation (15) reduces to

$$\frac{d^2v}{dy^2} + (k^2c^2 - k^2 - \frac{1}{c} - y^2)v = 0 \quad . \tag{17}$$

As shown in the last section, the term k^2c^2v is smaller than the other terms when we consider the slowly moving large-scale waves. In order to facilitate comparison with Matsuno's results, we will include this term for the time being. The free waves are described by the eigenfunctions of this equation satisfying the homogeneous boundary conditions. In the infinite domain the solutions are $e^{-y^2/2}$ multiplied by Hermite polynomials, $H_n(y)$, and the eigenvalues are:

$$\alpha = k^2c^2 - k^2 - \frac{1}{c} = 2n + 1, \quad n = 1, 2, \ldots$$

where n is the order of the polynomial. For each value of n there are three roots for c, the phase speed. Two roots correspond to gravity waves ($c \sim \pm \sqrt{1 + \frac{2n+1}{k^2}}$) and one to a Rossby wave ($c \sim \frac{-1}{k^2 + 2n + 1}$). Note that the neglect of the k^2c^2 term in (17) has the effect of filtering out the gravity wave solutions, for the characteristic equation then reduces to a linear (not cubic) form whose sole root is that of the Rossby wave: $-k^2 - \frac{1}{c} = 2n + 1$, $n = 1, 2, \ldots$.

As Matsuno pointed out, the low order modes (small n) are trapped near the equator (for $n = 0$, $v = e^{-y^2/2}$) and are relatively insensitive to far distant boundaries. However, the higher order modes found by Matsuno would be affected by the finiteness of the real earth. Theoretical evidence regarding such boundary effects has been presented by Rosenthal (1965) and Koss (1967).

To facilitate comparison with later numerical results, we now study the numerical eigenvalue problem corresponding to the homogeneous boundary conditions $v = 0$ at $y = \pm 10.0$ by finding the eigenvalues and the eigenfunctions of the finite-difference analogue of the homogeneous Eq. (17). Since Eq. (17) is symmetric with respect to the equator, the solutions are either odd or even. Thus, in determining the eigenfunctions we need only to compute the solution for half of the domain. We consider two eigenvalue problems as defined by the boundary conditions:

$$v(0) = 0, \quad v(10) = 0 \qquad \text{(THE ODD CASE)}$$

$$\frac{dv}{dy}(0) = 0, \quad v(10) = 0 \qquad \text{(THE EVEN CASE)}$$

The finite difference scheme gives the following algebraic equation for the value of v at the i^{th} (i.e., $y = i\Delta y$) gridpoint in terms of the values of v on either side of it:

$$\frac{v_{i+1} + v_{i-1} - 2v_i}{(\Delta y)^2} = - Q_i v_i \ .$$

This can be written:

$$\frac{v_{i+1} + v_{i-1} - 2v_i}{(\Delta y)^2} - y_i^2 v_i = - \alpha v_i \quad \text{where} \quad \alpha = -k^2 + k^2 c^2 - \frac{1}{c} \ . \tag{18}$$

The choice $\Delta y = 0.5$ corresponds to 19 internal gridpoints between $y = 0$ and $y = 10$. Once the boundary conditions at $y = 0$ and $y = 10$ are specified, there are 19 algebraic equations for the 19 values of v, written here in matrix form:

$$\underline{A}\,\underline{v} = - \alpha \underline{v} \ . \tag{19}$$

$\underline{A}$ is a 19×19 matrix of the known coefficients and $\underline{v}$ is a column vector of length 19. For the even case $v(0)$ was set equal to v_1, the first value of v north of the equator to approximate the zero derivative at the equator. Equation (19) can be written in the form

$$(\underline{A} + \alpha \underline{I})\,\underline{v} = 0 \tag{20}$$

where I is the unit matrix. Since this is a homogeneous system of linear algebraic equations, for there to be nontrivial solutions for v the determinant of the matrix $(\underline{A} + \alpha \underline{I})$ must vanish. The values of $\alpha \left(= k^2 - k^2 c^2 + \frac{1}{c}\right)$ for which this occurs are the eigenvalues of the matrix $-\underline{A}$.

The eigenvalues and associated eigenvectors of the system (20) were determined numerically. Figure 1 exhibits some eigenfunctions which are symmetric about the equator. We see that each mode is centered near a latitude $y = \alpha^{1/2}$. The highest modes (large α) are found away from the equator and are understandably more nearly quasi-geostrophic in character. In these cases the lowest frequency then reduces to the classical mid-latitude Rossby wave frequency, $\omega = \frac{-\beta k}{k^2 + \ell^2}$ where $\ell^2 = \alpha$.

Additional numerical eigensolutions were found for the free waves for the case with boundaries at $y = \pm 3.0$. It was found that the trapped solutions with smallest α were not changed much by the closer boundaries. They are shown as dashed lines in Fig. 2 and will be discussed shortly. Solutions corresponding to larger values of α were so affected by truncation error as to have little value.

The effect of forcing the tropics in the absence of a zonal flow was investigated by specifying a nonzero value of v at the boundary $y = + 3.0$ and setting $v = 0$ at $y = -3.0$, thus modelling the effects upon the tropical atmosphere of a wave moving at higher latitudes in the Northern Hemisphere.

By arbitrarily setting the nonzero boundary condition and specifying c and k we confine ourselves to looking at the asymptotic ($t \to \infty$) solution to the initial value problem generated by starting a wave traveling at constant speed and amplitude along the northern boundary until the whole tropical atmosphere is pulsing at its frequency and wavelength. By ignoring the transient responses to the forcing at the boundary we are neglecting the way in which the energy initially enters the tropics, preferring to look at the tropics a long time later when a steady energy density has developed.

The forced solutions (with U = 0) are shown as solid lines in Fig. 2 for various values of the phase speed. We see that the ability of the motions to

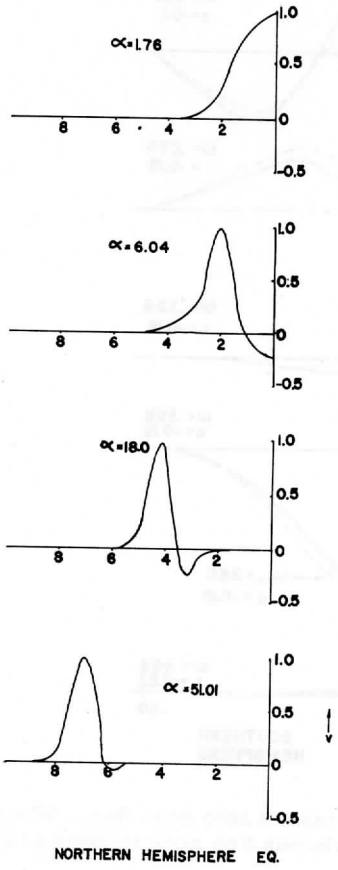


Fig. 1. Free modes of the tropics in the absence of a mean flow which are symmetric about the equator (y = 0). Magnitude of scale of v is arbitrary.

penetrate into the Southern Hemisphere depends strongly on the phase speed (in the cases shown, all are to the west). The maximum response at all latitudes is found for waves whose phase speeds are close to those of a free mode, all of which move westward. In these quasi-resonant cases, we see that the structure of the response is very nearly the same as that of the free mode, as expected. Figure 2 also exhibits two "non-propagating" responses (c -0.16, c -0.22) which do not resemble a wave-like structure in y and which are of small amplitude throughout most of the tropics.

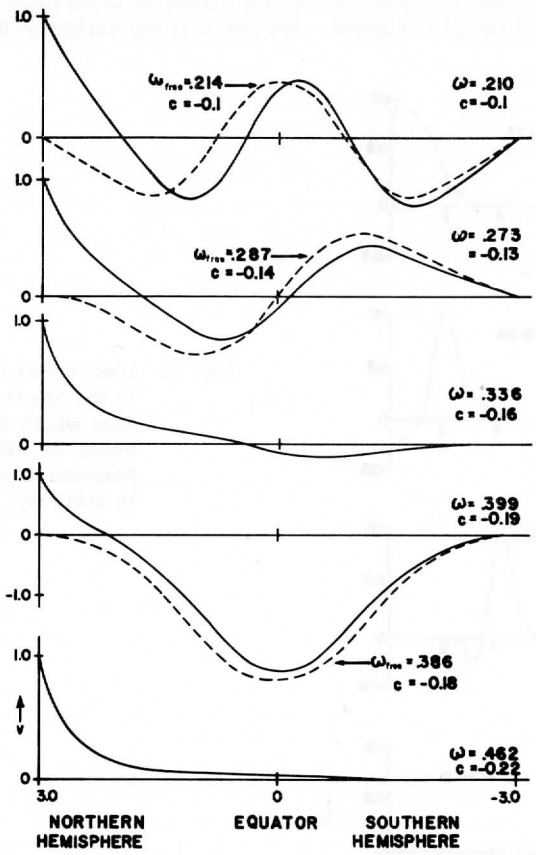


Fig. 2. Forced solutions for v(y) for the case of zero mean flow. Where a resonant response occurs, the analogous free mode is given in dashed line.

3. Forced Waves in the Presence of a Zonal Current

The analysis of section 2 using Eq. (16) cannot completely be extended to the case in which there is a zonal current, because it is not possible to set up a conventional matrix eigenvalue problem for the free modes using (16) alone. It is possible to use the complete set of Eqs. (10-12) as a basis for an eigenvalue approach, but the dimension of the problem then becomes three times larger.

It is of course possible to use Eq. (16) to treat the forced problem, in which c is specified a priori. The problem can be analyzed qualitatively in the following manner: Eq. (16) can be written:

$$\frac{d^2v}{dy^2} + Q(k, U, c, y)v = 0 . \tag{21}$$

The solutions can be broadly classified as being of three types: (1) propagating, (2) evanescent, (3) singular. In regions where Q is positive, there are oscillatory solutions and thus latitudinal propagation. In regions where Q is negative the solutions are exponential and the wave is evanescent since solutions forced at one boundary die off exponentially in y (recall the case $c = -.22$ in Fig. 2). In regions where $U(y)-c$ approaches zero, the coefficient Q gets arbitrarily large since it contains $1/(U-c)$. The equation has a regular singular point there. Since this case is difficult to treat numerically by the method we have employed so far, it will be the subject of the next section. In this section we will try to learn something of the nonsingular solutions.

We will discuss a special case in which a parabolic zonal flow and a wavenumber, k, are specified. A fair approximation of the mean winds throughout the troposphere (Mak, 1969) is:

$$U(y) = -0.04 + 0.0625y^2 \tag{22}$$

where we have again chosen 100 m/sec for $\sqrt{g'H_0}$. This gives a 4 m/sec easterly wind at the equator and a 10 m/sec westerly wind at $y = 1.5$ ($\sim 30°N$).

Figure 3 shows values of Q as a function of c and y when $k = 2.1$ (corresponding to an east-west wavelength of 6000 km) for the zonal flow specified by (22). The zonal flow is also plotted on the diagram with a dashed line. Note that the range of c for which $U(y) - c = 0$ at some latitude is hatched, indicating that the equation becomes singular at the dashed line. Regions where Q is positive for the entire range of y occur only for waves moving westward faster than the maximum easterly zonal current, but more slowly than about 7 m/sec.

Waves moving westward faster than 7 m/sec must first "tunnel" their way through the evanescent region near the northern boundary. Inspection of

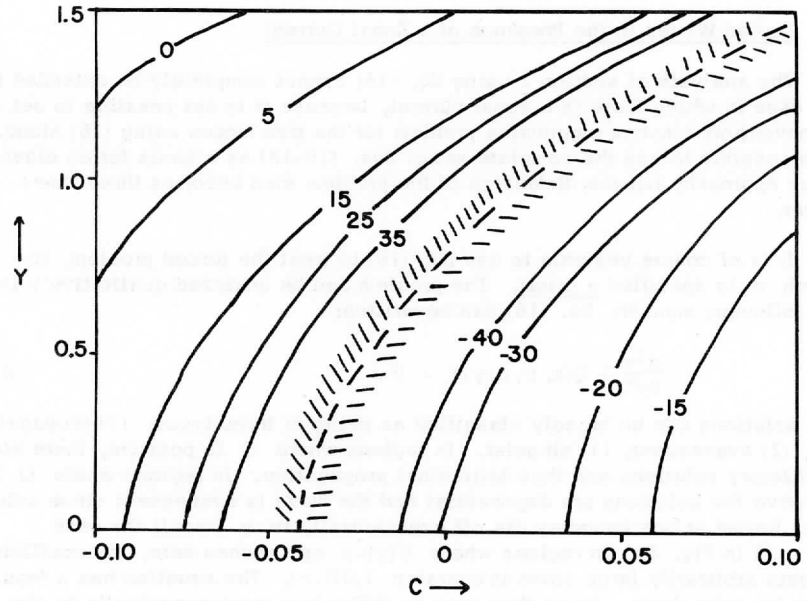


Fig. 3. Values of $Q(y, k, U, c)$ for $k = 2.1$ and mean flow given by dashed line. $y = 1.5$ corresponds to approximately 30° latitude, positive values indicate propagating solutions while negative values indicate evanescent solutions. In the hatched region $Q \sim 1/U-c$ and thus gets arbitrarily large.

Eq. (16) shows that the evanescent region expands downward and in the direction of increasing c as the x-wavelengths decrease (as k increases). Thus, it appears that only the longer waves can propagate freely into the tropics; sufficiently short waves may suffer attenuation for all wave speeds.

Forced solutions were computed for nine different values of c in the propagating region for this current. They are shown in Figs. 4 through 8.

We see that in many propagation cases the wavelength changes significantly with latitude, implying a continuous process of reflection and transmission which also alters the amplitude variations as well. The evanescent region near the left boundary is too weak to detect in most of the cases shown in Figs. 6 - 8. The gross magnitude of the response follows the same general trend as if the free modes were those in the problem treated in section 2. There are four responses for which a value of v is greater than that imposed

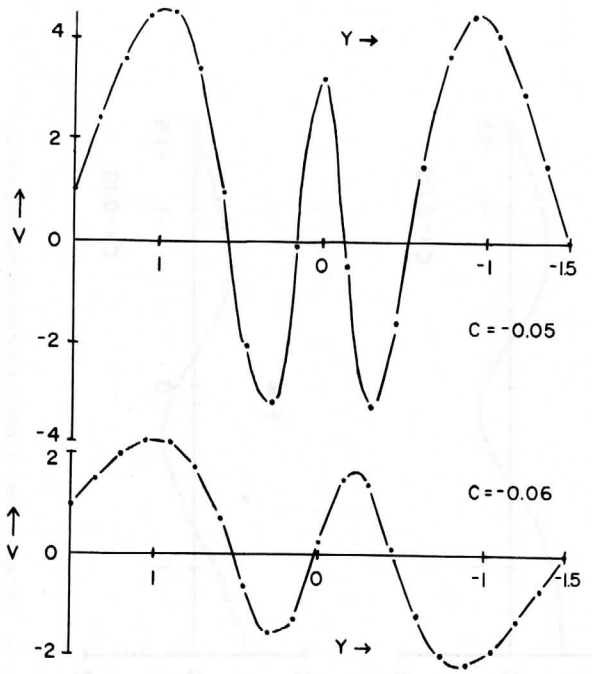


Fig. 4. Solutions of v(y) where v is forced to be 1.0 at northern boundary, for different values of c. The mean flow U(y) and k (= 2.1) correspond to values in Fig. 3.

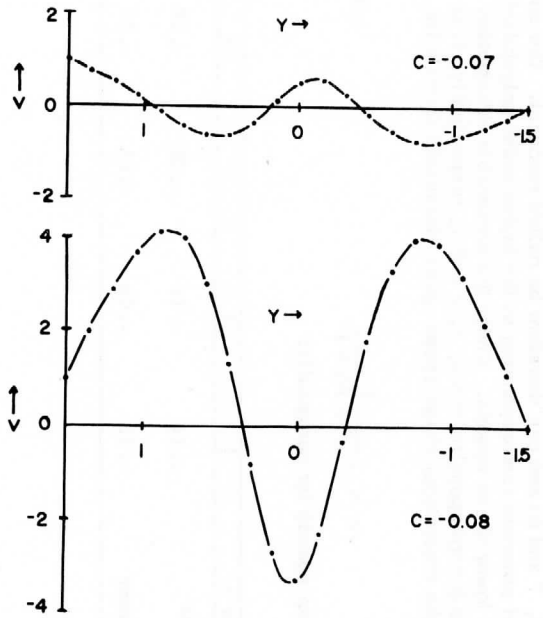


Fig. 5. Same as Fig. 4 but for different values of c.

on the boundary (4, 5, 7 and 8) and can therefore be called resonant. The ones with the most detailed structure (corresponding to the higher order polynomials of Matsuno) have the lower phase speeds. Under the reasonable assumption that Figs. 4, 5, 7 and 8 correspond to n = 4, 3, 2 and 1, respectively, it is possible to compute the theoretical phase speed using Matsuno's formula for free Rossby modes:

$$c = -\frac{1}{k^2 + 2n + 1} \tag{23}$$

and compare it with the speeds for our results:

n →	1	2	3	4
Matsuno wave with no zonal current	-.14	-.10	-.08	-.07
Model with zonal current	-.12	-.08	-.06	-.05

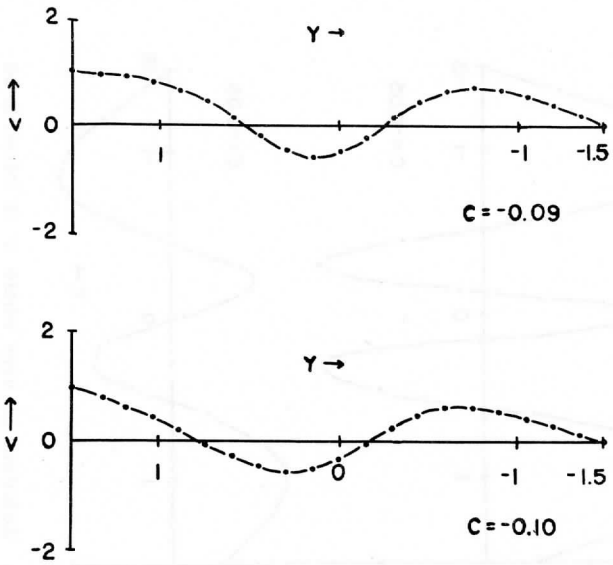


Fig. 6. Same as Fig. 4 but for different values of c.

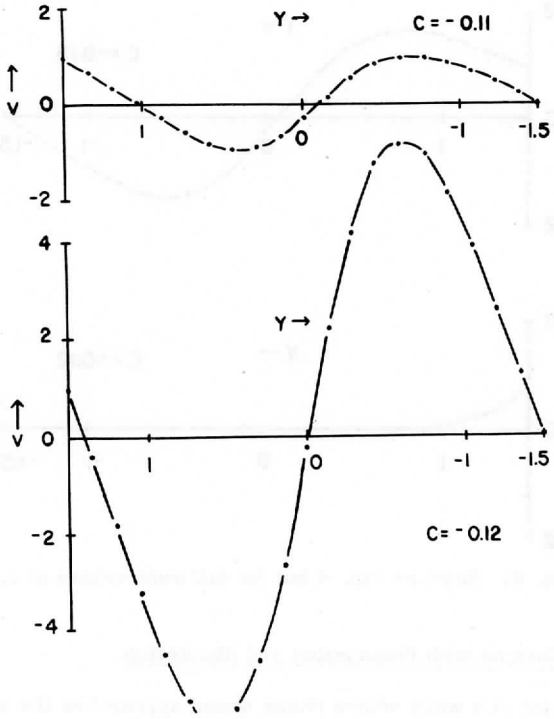


Fig. 7. Same as Fig. 4 but for different values of c.

Note that the free modes are shifted in phase velocity in the direction of the average of the wind over the tropics (U(y) averaged from y = 0.0 to y = 1.5 is +0.034, i.e., westerly). This implies that the net effect of the shear flow is to cause a larger difference between the resonant phase velocity and the velocity of typical mid-latitude forcing than would result if the wind were the same at all latitudes. If the wind were westerly at all latitudes, the free modes would be doppler shifted by approximately the same amount as the forcing motions in mid-latitudes. With shear, however, the mid-latitude motions will be advected by a strong westerly current, in contrast to the free tropical modes which are advected by more easterly winds.

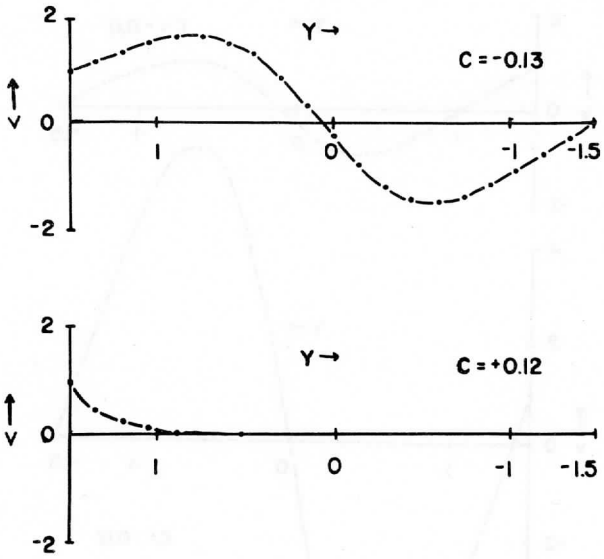


Fig. 8. Same as Fig. 4 but for different values of c.

4. Forced Solutions with Propagation and Absorption

The behavior of a wave whose phase speed approaches the speed of the zonal current has been significant in other treatments of wave propagation through a shear flow. The fact that the linearized equations exhibit a singularity at that point has led investigators to use terms like "critical layer" (Booker and Bretherton, 1967), or "singular line" (Dickinson, 1968). Because of the singularity, attempts to treat the problem using numerical methods have led to some confusion in the literature. Hines and Reddy (1967) discussed the alternative physical interpretations of the critical layer effect in terms of reflection versus absorption processes, while Bretherton (1966) discussed the problem in terms of classical ray tracing theory and found absorption of wave energy. The matter was finally resolved when Booker and Bretherton (1967) showed that an alternative approach, in which a simple model is allowed to achieve a steady state after being started from rest, led to the absorption predicted by Bretherton.

The problem we have is actually both physically and mathematically more like the problem of Dickinson (1968) since it occurs for horizontal shear flow and has the same type of singularity. (The coefficient of v in Eq. (16) goes to infinity as $U \to c$ like $1/(U-c)$, while in the gravity wave theory the analogous coefficient goes like $1/(U-c)^2$.) Dickinson noted that the

singularity appears to invalidate any planetary wave theory in which the waves are artificially confined between reflecting barriers. He states that there are no normal modes of a "singular wave guide"; i.e., a region bounded by critical layers. Dickinson (1969) has also solved the problem considered here for the long wave approximation (i.e., the term k^2 of Eq. 27 is neglected), and obtained results consistent with ours.

The solution of the "asymptotic" problem contains many mathematical difficulties and still fails to answer several fundamental questions. Hines and Reddy considered an atmospheric model composed of discrete layers and matched solutions at the interfaces. This gave controversial results because it could not adequately allow for the fact that the coefficient actually gets arbitrarily large over a very small distance. Jones (1967) solved the "asymptotic" problem by adding a small imaginary component to the phase speed (corresponding to Rayleigh damping) so that $1/(U-c)$ became bounded in magnitude by $1/|c_i|$. Booker and Bretherton attempted to solve an initial value problem in order to generate the asymptotic solution. They used a Laplace transform approach but were only able to evaluate the resulting integral equation for large times, thereby losing vital information about the wave development in the critical layer.

The development in a critical layer or singular line is important for two reasons. First, it gives information about the time scales needed for development and the manner in which the final state is reached. Second, it tells one how long it takes for the linearization to become invalid. The asymptotic solutions imply an infinite energy density in the vicinity of the critical layer and thus cast doubt upon the validity of the solutions. If one can show that this phenomena occurs only after physically unrealistically long times, then these solutions need not be dismissed as a mathematical curiosity.

To facilitate the solution of the mathematical problem, we find it convenient to assume the flow to be nondivergent. This, of course, removes gravity waves from our model but at the length and time scales we have been working with they have not really been important, as we saw in the scaling in section 1. The nondivergent assumption is also justified by Charney's (1963) scale analysis of large-scale tropical motions. At any rate, it will be shown that it leads to the same type of singularity as in Eq. (16) but allows a much simpler initial value problem.

If the flow is horizontally nondivergent, a stream function can be defined such that:

$$u = -\frac{\partial \psi'}{\partial y}, \quad v = \frac{\partial \psi'}{\partial x} \quad . \tag{24}$$

Using these relations and Eqs. (1) and (2), we can derive the linearized barotropic vorticity equation valid for nondivergent flow at any latitude:

$$(\frac{\partial}{\partial t} + U\frac{\partial}{\partial x})\nabla^2\psi' + (\beta - \frac{d^2U}{dy^2})\frac{\partial\psi}{\partial x} = 0 \qquad (25)$$

One can see how this leads to an equation like (16) by assuming solutions of the form

$$\psi'(x, y, t) = Re[P(y)e^{i(kx - kct)}] .$$

The resulting equation is:

$$\frac{d^2P}{dy^2} + (-k^2 + \frac{1}{U-c}(\beta - \frac{d^2U}{dy^2}))P = 0 . \qquad (26)$$

Comparison with Eq. (16) indicates the formal assumption of complete non-divergence has eliminated the $-y^2$ and $y^2U/U-c$ terms in the coefficient of v in (16). Near a critical layer (where $\frac{1}{(U-c)}$ becomes dominant), the coefficient of v in (26) becomes singular in a manner similar to that in (16). Thus, the nondivergent model behaves like the divergent one near the critical latitude.

Since we wish to consider an initial value problem, the solutions are periodic in time only in an asymptotic sense. Thus, we assume periodic solutions only in x:

$$\psi'(x, y, t) = Re[\psi(y, t)e^{ikx}] .$$

This leads to the equation:

$$[\frac{\partial}{\partial t}(\frac{\partial^2}{\partial y^2} - k^2)]\psi = -Uik(\frac{\partial^2}{\partial y^2} - k^2)\psi - \beta^* ik\psi \quad \text{where} \quad \beta^* = \beta - \frac{d^2U}{dy^2}. \qquad (27)$$

The method adopted for the solution of this equation is that known as the modified Euler technique discussed by Lilly (1965) and Young (1968). Lilly found the method desirable for ordinary differential equations because it is neutrally stable and conserves mean square vorticity for spectral models. In extending its use to a partial differential equation, it retains computational stability. This difference scheme for the time dependence is:

$$[(\frac{\partial^2}{\partial y^2} - k^2)\psi]^{n+1} = [(\frac{\partial^2}{\partial y^2} - k^2)\psi]^n - \frac{ik\Delta t\beta^*}{2}[\psi^{n+1} + \psi^n]$$

$$- \frac{\Delta tUik}{2}\langle[(\frac{\partial^2}{\partial y^2} - k^2)\psi]^{n+1} + [(\frac{\partial^2}{\partial y^2} - k^2]\psi]^n\rangle \qquad (28)$$

where $t = n\Delta t$. Centered differences are used in y:

$$\frac{\partial^2 \psi}{\partial y^2} \doteq \frac{\psi_{i+1} + \psi_{i-1} - 2\psi_i}{(\Delta y)^2} \tag{29}$$

where $y = i\Delta y$.

Using (28) and (29), and rearranging to place all $n+1$ terms on the left hand side:

$$(1 + \frac{\Delta t U i k}{2})[\frac{\psi_{i+1}^{n+1} + \psi_{i-1}^{n+1} - 2\psi_i^{n+1}}{(\Delta y)^2} - k^2 \psi_i^{n+1}] + \frac{\Delta t \beta^* ik}{2} \psi_i^{n+1}$$

$$= (1 - \frac{\Delta t U i k}{2})[\frac{\psi_{i+1}^{n} + \psi_{i-1}^{n} - 2\psi_i^{n}}{(\Delta y)^2} - k^2 \psi_i^{n}] - \frac{\Delta t \beta^* ik}{2} \psi_i^{n} \ . \tag{30}$$

It is assumed that ψ is initially zero everywhere from $-D \leq y < +D$ but that a disturbance $\psi(D, t) = 1$ is raised along the northern boundary at $t = 0$ and thereafter kept constant with time. Alternatively, one could apply a periodic forcing in time to simulate a wave travelling along the boundary in the manner of Houghton and Jones (1969). However, since Eq. (25) is invariant under a Galilean transformation of the type

$$U(y) \rightarrow U(y) + U'$$

where U' is constant, the two approaches are equivalent. The initial and boundary conditions are therefore:

$$\psi(y, 0) = 0 \qquad \text{for } -D \leq y < +D$$

$$\psi(D, 0) = 1.0$$

$$\psi(D, t) = 1.0$$

$$\psi(-D, t) = 0.0.$$

The last equation implies a reflecting wall at $y = -D$ since $v = 0$ there; its effect is negligible here since very little energy propagates to that latitude in this solution.

With the boundary conditions there is a complete set of linear algebraic equations, one for each internal gridpoint, which determine ψ from its values for the preceding time step. The system can be written:

$$\underline{A}\, \underline{\psi}^{n+1} = \underline{b}^{n} \ .$$

If m is the number of internal gridpoints, $\underline{A}$ is an $m \times m$ matrix of coefficients which is constant with time, $\underline{\psi}^{n+1}$ is a column vector of m elements

and $\underline{b}^n$ is another column vector of m elements which depends on the values of ψ_i^n. Since ψ is complex, this is equivalent to the solution of 2m algebraic equations for real variables. In this work 49 internal gridpoints were used, resulting in 98 algebraic equations. The fact that $\underline{A}$ does not vary with time means that one may perform a Gaussian elimination on it only once and simply back substitute at each time step when there is a new term $\underline{b}^n$.

As an aid to the interpretation of the solutions to this problem, it is helpful to consider what classical wave theory tells us about the problem. It is well known that Eq. (25) allows only stable Rossby waves unless $(\beta - \frac{d^2U}{dy^2})$ is zero somewhere. Using WKB logic we may consider U to be locally constant. Solutions of the form

$$\psi = \mathrm{Re}[\psi^* e^{i(kx + \ell y - \omega t)}]$$

satisfy the dispersion relation:

$$\omega = Uk - \frac{(\beta - U'')k}{k^2 + \ell^2} \tag{31}$$

By setting $\psi = 1.0$ at the boundary, we have specified its phase to be constant in time. Thus the frequency should be zero there and (31) determines a value for ℓ^2:

$$\ell^2 = -k^2 + \frac{\beta - U''}{U(y)} . \tag{32}$$

In this manner, the two conditions at the boundary (that the wave be stationary with a given zonal wind) determine a wavelength in y. From the dispersion relation (31), one can compute a group velocity for the y direction:

$$C_{gy} = \frac{\partial \omega}{\partial \ell} = \frac{2k\ell(\beta - U'')}{(k^2 + \ell^2)^2} . \tag{33}$$

Thus the zonal wind at the boundary indirectly predicts an initial group velocity. With only qualitative accuracy we can assume that the disturbance excited at the northern boundary will propagate toward the equator with a local group velocity given by (33) and a structure $\ell^2(y)$ given by (32).

Since we are mainly concerned with the behavior near a critical layer, it will suffice to consider a linear shear flow. For convenience we will assume it to have a zero at y = 0. Thus

$$U(y) = y(\frac{dU}{dy}) \tag{34}$$

where $\frac{dU}{dy}$ = constant. The latitude y = 0 therefore is the critical latitude for the wave with speed c = 0.

Noticing that ℓ becomes large near $y = 0$, we can approximate Eqs. (32) and (33) there by:

$$\ell^2 \doteq \frac{\beta - U''}{y\frac{dU}{dy}}$$

$$C_{gy} \doteq (\beta - U'')^{-1/2} \left(\frac{dU}{dy}\right)^{3/2} 2k\, y^{3/2} \tag{35}$$

If a wave group approaches $y = 0$ from $y = D$ using this velocity, the time it takes to get there will be

$$t \doteq \frac{(\beta - U'')^{1/2} \left(\frac{dU}{dy}\right)^{-3/2}}{2k} \int_0^D y^{-3/2} dy \ . \tag{36}$$

The fact that this integral is not finite leads to the concept of absorption since the group is neither reflected nor transmitted—it is simply retarded. This explanation gives the right behavior for the phenomenon, but the WKB theory is wrong in principle. In order for a WKB approximation to be valid the local wave number ℓ should not vary much over a wavelength $(= 2\pi/\ell)$. In other words, the percentage change in wavenumber over a wavelength L_y :

$$\frac{\frac{\partial \ell}{\partial y} L_y}{\ell} = \frac{\partial \ell}{\partial y} \frac{2\pi}{\ell^2}$$

must be of order unity or less. In our case, (35) shows that ℓ is proportional to $y^{-1/2}$ near the critical level. Thus, near $y = 0$ the above expression gives

$$\frac{2\pi}{\ell^2} \frac{\partial \ell}{\partial y} = \frac{2\pi \frac{dU}{dy} y}{\beta - U''} \frac{\partial}{\partial y} \left(\frac{(\beta - U'')^{1/2}}{y^{1/2}}\right) = -\pi \frac{dU}{dy} (\beta - U'')^{-1/2} y^{-1/2}, \tag{37}$$

which goes to infinity as y goes to zero; thus the WKB approximation is invalidated near the singularity. However, it can still give us a first estimate of the rate of energy propagation away from the singularity.

Since we have chosen a linear wind profile and have chosen it to be zero at $y = 0$, the parameters $\ell(D)$, $U(D)$, $\frac{dU}{dy}$, and $C_{gy}(D)$ are all related. If any one of them is known along with the east-west wavenumber k, the other three can be computed. For convenience, however, they have all been calculated in the following table for the three numerical experiments we shall consider shortly:

Case	k	$\ell(D)$	U(D)	$(\frac{dU}{dy})$	C_{gy}
A	2.1	6.5	0.0225	0.015	1.33×10^{-2}
B	2.1	4.0	0.05	0.033	4.2×10^{-2}
C	0.7	1.7	0.3	0.2	2.0×10^{-1}

Cases A and B differ only in the mean shear $\frac{dU}{dy}$ (and hence U(D)). By our WKB logic we expect ψ in case B to have a larger wavelength in y and hence to propagate faster toward the equator. This is readily verified by inspection of Figs. 9 through 14. However, although the WKB method also predicts that the wavelength near the critical level becomes extremely small, it does not occur in the numerical solutions. In the quasi-steady solutions (the last figure of each series), no drastic shortening of wavelength near $y = 0$ is apparent.

One can scale this problem like the problems in sections 2 and 3 by choosing $\sqrt{g'H_0}$ to be 100 m/sec. This gives a length scale, $\hat{L}$, from (6) of approximately 2100 km and a corresponding time scale factor $\hat{T}$ of 5.85 hours. At this scaling, A and B correspond to waves whose east-west wavelength is approximately 6300 km and C corresponds to a wavelength of 19,000 km.

In the initial stages of development there is a rapid growth in the magnitude of ψ throughout the whole domain as a response to the impulsive start. This response dies off exponentially away from the forced boundary in a manner which is similar to potential flow. The initially irrotational motion satisfies:

$$(\frac{\partial^2}{\partial y^2} - k^2)\psi = 0 \ .$$

Solutions of this equation are

$$\psi = e^{\pm ky} \ .$$

This relation can be verified quantitatively for case C at $t = 3.0$. After this initial effect, the solution becomes oscillatory and the disturbance begins to propagate southward. This propagation slows down as the disturbance approaches $y = 0$ where the singularity exists in the asymptotic problem. Eventually the solution reaches a quasi-steady state in which it has negligible amplitude on the southern side of the critical latitude. This solution remains essentially unchanged even for times three to four times that needed for development.

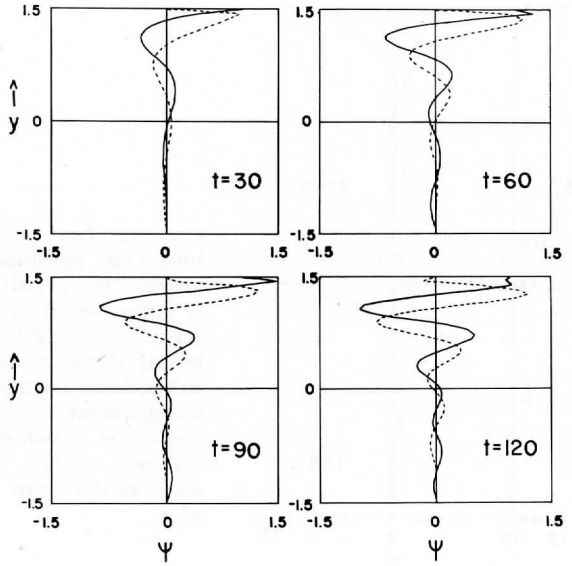


Fig. 9. Real (solid lines) and imaginary (dashed lines) parts of streamfunction, $\psi(y)$ for Case A for initial stages of development. Variables specified are k, an east-west wavenumber, and dU/dy, the shear of the mean wind.

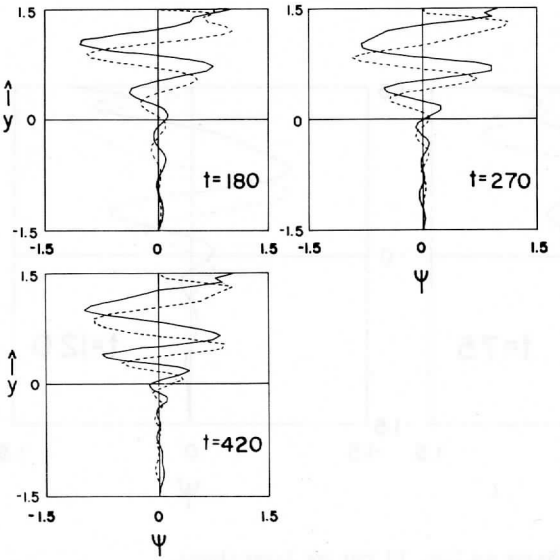


Fig. 10. Same as Figure 9 but for later times.

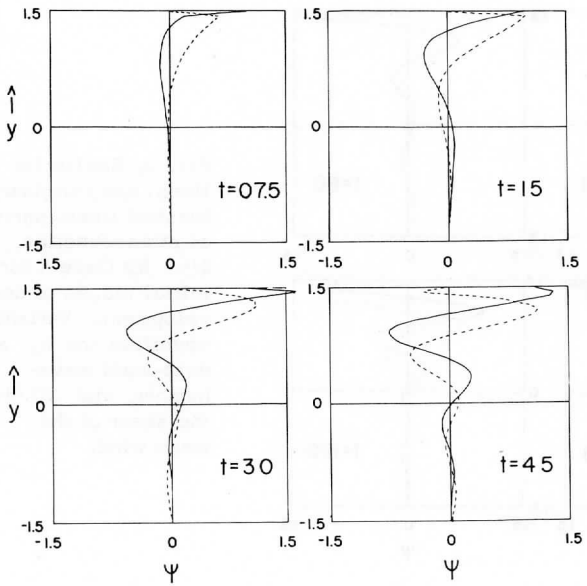


Fig. 11. Real (solid lines) and imaginary (dashed lines) parts of streamfunction, $\psi(y)$ for Case B for initial stages of development. Variables specified are k, an east-west wavenumber, and dU/dy, the shear of the mean wind.

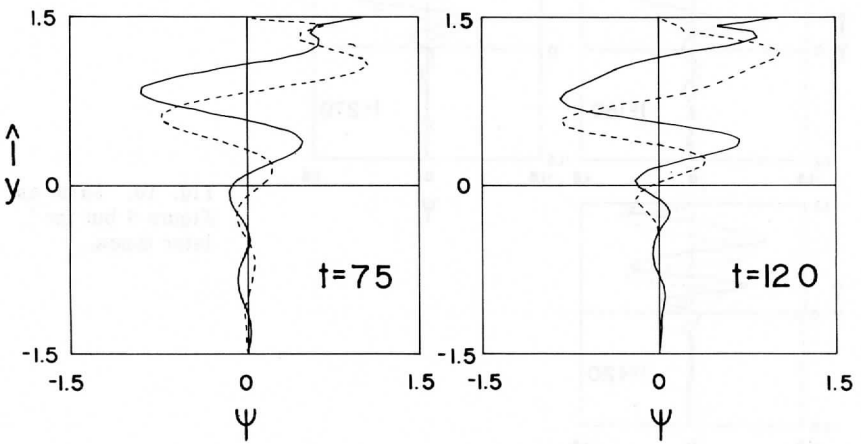


Fig. 12. Same as Fig. 11 but for later times.

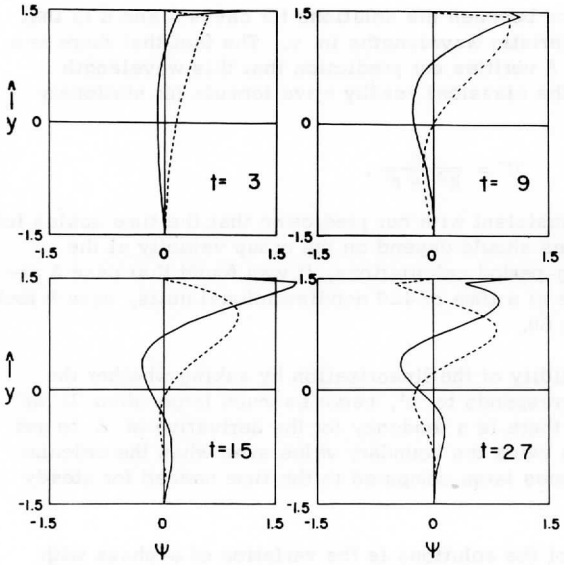


Fig. 13. Real (solid lines) and imaginary (dashed lines) parts of streamfunction, $\psi(y)$ for Case C for initial stages of development. Variables specified are k, an east-west wavenumber, and dU/dy, the shear of the mean wind.

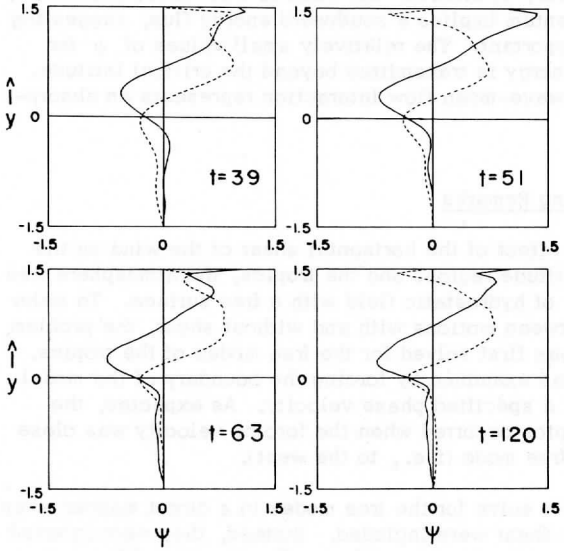


Fig. 14. Same as Fig. 13 but for later times.

The essential difference between the solutions for cases A and B is that they show different characteristic wavelengths in y. The fact that there is a shorter wavelength in case A verifies our prediction that this wavelength should be consistent with the classical Rossby wave formula for stationary disturbances:

$$U = \frac{\beta}{k^2 + \ell^2} .$$

The three cases are consistent with our prediction that the time scales for development to steady states should depend on the group velocity at the boundary. Based upon long-period calculations, it was found that case A attained a quasi-steady state at a time of 420 nondimensional units, case B took 120 units and case C about 60.

We can assess the validity of the linearization by asking whether the value of $\partial\psi/\partial y$, which corresponds to u', becomes much larger than U at the boundary. In case C there is a tendency for the derivative of ψ to get large, but it never exceeds twice the boundary value even when the calculations were carried out to times large compared to the time needed for steady development.

An important property of the solutions is the variation of ψ-phase with latitude, which corresponds to a tilt of the troughs and ridges from southwest to northeast. This tilt is indicative of a northward transport of zonal momentum. Application of the theory of Eliassen and Palm to this model shows that this northward flux of momentum implies a southward energy flux, suggesting that reflection is not too important. The relatively small values of ψ for $y < 0$ indicate that little energy is transmitted beyond the critical latitude. Thus we conclude that the wave-mean flow interaction represents an absorption of wave energy.

5. Summary and Concluding Remarks

In order to isolate the effect of the horizontal shear of the wind on the interaction between mid-latitude regions and the tropics, the atmosphere was idealized as a single layer of hydrostatic fluid with a free surface. To make a numerical comparison between motions with and without shear, the problem without basic zonal wind was first solved for the free modes of the tropics. A resonance phenomenon was examined by forcing the boundary of the model with a wave disturbance of a specified phase velocity. As expected, the largest response in the tropics occurred when the forcing velocity was close to the phase velocity of a free mode (i.e., to the west).

It was not convenient to solve for the free modes in a direct manner when the effects of a zonal wind shear were included. Instead, they were inferred from the resonant response to the forced problem. The forced problem was

treated only for those phase velocities for which the governing equation was not singular. Because the equation was singular for all phase velocities which were neither large eastward nor large westward, it was left for a special analysis. The resonant responses to the nonsingular forced problem indicated that the main effect of the zonal wind was to doppler shift the phase speeds of the free modes in the direction of the wind averaged over all latitudes of the model.

Because of this doppler shift, the observed winds appeared to inhibit interaction between the tropics and mid-latitudes. If the westerlies of mid-latitudes extended to the equator, the free modes of the tropics would move with phase speeds closer to those of mid-latitude disturbances and hence be closer to resonance.

The important case which could not be treated by the above analysis occurred when the longitudinal phase velocity of a propagating wave became equal to the mean zonal flow at some "critical" latitude. This case was treated by simplifying the model slightly by assuming the flow to be nondivergent. This assumption made it feasible to pose a problem in which the northern boundary of the initially undisturbed tropics was forced and the motion computed as a function of time and latitude. The numerical solutions to this marching problem indicated that singular line absorption of wave energy can occur at "critical" latitudes. Time scales for this energy loss from mid-latitude motions range from 2 - 3 days for planetary scales to several weeks for small scales (1000 km).

The observed shear of the zonal wind implies that this mechanism serves to inhibit the propagation of wave disturbances, just as the doppler-shift of the free modes does. An important difference, however, is that the doppler shift mechanism tells us that there is little interaction between the eddy motions originating in mid-latitudes and the tropical zonal wind. Moreover, there are both observational and theoretical grounds which support such a mechanism.

The observed eddy motions transport eastward momentum into the latitudes where the zonal winds are a maximum (Mak (1969)). The observed energy flux (energy flux here is devined as $\overline{p'v'}$, the pressure work done on the fluid north of a latitude circle by the fluid south of it) is southward (see Mak (1969)), and thus consistent with the momentum flux for the model treated here once a steady state has developed. Thus, the propagation of wave disturbances from mid-latitudes into the tropics is consistent with the observed momentum and energy fluxes.

Theoretical models have exhibited this northward momentum flux and southward energy flux which is observed. Mak's (1969) stochastic model also shows a very strong conversion from eddy kinetic energy to zonal kinetic energy in the tropics implying that the motions serve to maintain the zonal winds.

However, Mak did not discuss in detail the mechanism by which this momentum is taken from the easterly flow. Although he used a two-layer model, his results indicate little vertical coupling and thus it is reasonable to assume can be interpreted in terms of the one-layer model presented here.

Mak found that the eddy motions were much smaller in his model tropics than in mid-latitudes. The dominant motions had phase speeds directed westward relative to the mean flow. These motions are precisely the ones one would expect to propagate from our analysis of the forced modes in the presence of shear since the resonant responses corresponding to the free modes were always westward moving. However, the forcing at the boundary of Mak's model included eastward moving waves. For these waves, our model predicts singular line absorption and conversion of wave energy to mean flow energy. Interestingly, Mak's model shows very little wave energy in the tropics for these phase velocities but shows a very strong conversion from eddy kinetic energy to zonal kinetic energy.

In conclusion, the shear of the zonal wind can act to limit the propagation of wave disturbances into the tropics to those waves travelling more westward than the mean flow at any latitude. Other disturbances, which include most of the eastward moving or slowly westward moving disturbances of mid-latitudes, can influence the tropics by being absorbed by the mean flow and maintaining the easterlies.

Acknowledgments

The writers profited from discussions with Professor David D. Houghton of the University of Wisconsin and Dr. Robert E. Dickinson of the National Center for Atmospheric Research.

Financial support was provided by the Environmental Science Services Administration under grant number E-230-68-(G) at the Space Science and Engineering Center of the University of Wisconsin and the National Science Foundation with the University of Colorado. Computing funds were provided by the Wisconsin Alumni Research Foundation at the University of Wisconsin Computing Center and the National Science Foundation at the National Center for Atmospheric Research.

References

Booker, John R., and Francis P. Bretherton, 1967: "The Critical Layer for Internal Gravity Waves in a Shear Flow," J. Fluid Mech., 27, 513-539.

Bretherton, F. P., 1966: "The Propagation of Groups of Internal Gravity Waves in a Shear Flow," Quart. J. Roy. Met. Soc., 92, 466-480.

Charney, J., 1969: "A Further Note on Large-Scale Motions in the Tropics," J. Atmos. Sci., 26, 182-185.

Charney, J., 1963: "A Note on Large Scale Motions in the Tropics," J. Atmos. Sci., 20, 607-609.

Charney, J. G. and A. Eliassen: "On the Growth of the Hurricane Depression," J. Atmos. Sci., Vol. 21, No. 1, Jan. 1964, 68-74.

Dickinson, Robert E., 1968: "Planetary Rossby Waves Propagating Vertically Through Weak Westerly Wind Guides," J. Atmos. Sci., 25, 984-1002.

Dickinson, Robert E., 1969: "Development of a Rossby Wave Critical Level." Submitted for publication (N. C. A. R. Ms. No. 69-193, Boulder, Dec. 1969).

Eliassen, A. and Enok Palm, 1960: "On the Transfer of Energy in Stationary Mountain Waves," Geofysiske Publikasjoner, 22, no. 3, 1-23.

Hines, C. O. and C. A. Reddy, 1967: "On the Propagation of Atmospheric Gravity Waves Through Regions of Wind Shear," Jour. Geophys. Res., 72, no. 3, 1015-1034.

Houghton, David D. and Walter L. Jones, 1969: "A Numerical Model for Linearized Gravity and Acoustic Waves," Jour. Computational Physics, 3, no. 3, 339-357.

Jones, Walter L., 1967: "Propagation of Internal Gravity Waves in Fluids with Shear Flow and Rotation," J. Fluid Mech., 30, part 3, 439-448.

Koss, W. J.: "Further Theoretical Considerations of Tropospheric Wave Motions on Equatorial Latitudes," Monthly Weather Review, 95, no. 5, May, 1967, 283-297.

Lilly, D. K., 1965: "On the Computational Stability of Numerical Solutions of Time Dependent Non-Linear Geophysical Fluid Dynamics Problems," Monthly Weather Review, 93, no. 1, 11-26.

Lindzen, R. S., 1967: "Planetary Waves on Beta-Planes," Monthly Weather Review, 95, 441-451.

Lipps, F., 1970: "Barotropic Stability and Tropical Disturbances," Monthly Weather Review, 98, no. 2, 122-131.

Mak, Man-Kin, 1969: "Laterally Driven Stochastic Motions in the Tropics," J. Atmos. Sci., 26, 41-64.

Marayuma, T., 1968: "Time Series of Power Spectra Disturbances in the Equatorial Lower Stratosphere in Relation to the Quasi-Bienniel Oscillation," J. Meteor. Soc. Japan, 45, 404-417.

Matsuno, T., 1966: "Quasi-Geostrophic Motions in the Equatorial Area," J. Meteor. Soc. Japan, 44, 25-42.

Miles, J. W., 1962: "On the Generation of Surface Waves by Shear Flows, part 4," Jour. Fluid Mech., 13(3), 433-448.

Nitta, T. and M. Yanai, 1969: "A Note on the Barotropic Instability of the Tropical Easterly Current," J. Me. Soc. of Japan, 47, No. 2, pp. 127-130.

Rosenthal, S. L., 1960: "Some Estimates of the Power Spectra of Large-Scale Disturbances in Low Latitudes," J. Meteor. 17, 259-263.

Rosenthal, S. L., 1965: "Some Preliminary Theoretical Considerations of Tropospheric Wave Motions in Equatorial Latitudes," Monthly Weather Review, 93, No. 10, 605-612.

Rosenthal, S. L., 1967: "On the Development of Synoptic-Scale Disturbances over the Subtropical Oceans," Monthly Weather Review, 95, No. 6, 341-346.

Wallace, John M., 1969: "Some Recent Developments in the Study of Tropical Wave Disturbances," Bull. Amer. Meteor. Soc., 50, 792-799.

Yanai, M., and T. Maruyama, 1966: "Stratospheric Wave Disturbances Propagating over the Equatorial Pacific," J. Meteor. Soc. Japan, 44, 291-294.

Young, John A., 1968: "Comparative Properties of Some Time Differencing Schemes for Linear and Non-linear Oscillations," Monthly Weather Review, 96, no. 6, 357-364.

A NUMERICAL STUDY OF THE THREE-DIMENSIONAL STRUCTURE AND ENERGETICS OF UNSTABLE DISTURBANCES IN ZONAL CURRENTS

Rak To Song

Abstract

The multiple modes of dynamic instability, their structure and energy conversions for various zonal currents (pure baroclinic, pure barotropic, and mixed jet type) are investigated through use of a linear quasi-geostrophic model. Numerical results are obtained from the complete set of complex eigenvalues and eigenvectors of the finite difference version of the governing equations.

Numerical results are obtained for pure baroclinic and pure barotropic zonal currents and are compared with the well-known theoretical results. Their accuracy is examined by varying the number of subdivisions. It is found that the numerical method gives sufficiently accurate results in describing the structure and behavior of the primary unstable disturbances of cyclone scale. New results include information on the relation between the growth energetics, amounts of available energy, and disturbance wavelengths.

For jet-type zonal currents containing absolute vorticity extrema at various latitudes, there exist two primary modes of instability, each corresponding to a unique category of available zonal energy sources. The first is mainly characterized by baroclinic processes (modified by horizontal shear) and the other by barotropic mechanisms (modified by vertical shear). The baroclinic mode dominates for shorter waves, while the barotropic mode dominates for longer waves. The baroclinic mode increases the kinetic energy of the zonal current; conversely, the barotropic mode decreases the zonal kinetic energy. The presence of both energy sources tends to reduce individual rates of energy conversion. However, the baroclinic instability mechanism dominates in the lower layer, while the barotropic instability is most pronounced near the tropopause, particularly for longer waves.

Zonal currents which are similar but contain no active barotropic instability mechanism are found to possess only one strongly unstable mode which withdraws potential energy from the zonal flow, yet feeds kinetic energy into the jet.

Contents

List of Symbols

a	A variable parameter characterizing the magnitude of latitudinal shear of the zonal current
A	Amplitude of the perturbation stream function; also, available potential energy
A_r and A_i	The real and imaginary parts of A
$\underline{A}$	Column vector consisting of A at every grid point
A^*	Available potential energy density
b	A variable parameter characterizing the magnitude of vertical shear of the zonal current
$\underline{B}$ and $\underline{D}$	Coefficient square matrices of $(M-1) \times N$ order
c	Phase velocity
c_r and c_i	The real and imaginary parts of c
$C(,)$	Energy conversion rate from the first form to the second form of energy in the parenthesis
D	Meridional width of the basic channel
$\underline{D}^{-1}$	The inverse of matrix $\underline{D}$
e	The base of Naperian logarithms
E_A and E_k	Available potential energy error and kinetic energy error, respectively
f_0	Coriolis parameter at 45°N, $f_0 = 1.03 \times 10^{-4}$ sec^{-1}

g	Acceleration of gravity
HK_eC	Horizontal flux term of the eddy kinetic energy
i	Unit imaginary number; also an integer index for vertical subdivisions
$\underline{I}$	Identity matrix of $(M-1) \times N$ order
j	An integer index for latitudinal subdivisions
k	Zonal wave number, $k = \frac{2\pi}{L}$
kc_i	Growth rate
$\vec{k}$	Unit vector directed upward
K	Kinetic energy
K^*	Available kinetic energy density
l	Latitudinal wave number, $l = \frac{n\pi}{D}$, where n is integer
L	Zonal wavelength
M	Maximum number of grid points for latitudinal subdivisions from the center to the northern boundary of the channel
N	Maximum number of grid points for vertical subdivisions
p	Pressure
p_0	Surface pressure, $p_0 = 1000$ mb
q	Quasi-geostrophic absolute potential vorticity
Re	Real part
t	Time
$\bar{U}$	Basic zonal current velocity, $\bar{U} = -\frac{\partial \bar{\psi}}{\partial y}$
$\bar{U}_m$	Vertical mean value of $\bar{U}$
$\bar{U}'(p)$	Zonal current deviation from $\bar{U}_m$
v	Meridional component of geostrophic wind, $v = \psi_x$

v_{div}	Divergent part of meridional wind component
$\vec{V}$	Geostrophic wind, $\vec{V} = \vec{k} \times \nabla\psi$
VK_eC	Vertical flux term of the eddy kinetic energy
x	Coordinate toward the east
y	Coordinate toward the north
z	Altitude
W	Amplitude of the perturbation vertical velocity in the pressure coordinates
W_r and W_i	The real and imaginary parts of W
$(\,)_Z$ and z	Zonal energy integrated over the entire mass (capital Z) and integrated over one or two dimensions (small z)
$(\,)_E$ and e	Eddy energy integrated over the entire mass (capital E) and integrated over one or two dimensions (small e)
$(\,)_{x,y,p}$	Differentiation with respect to x, y, and p
α	Phase angle of the unstable perturbation stream function
β	Rossby parameter at 45°N, $\beta = 1.618 \times 10^{-11}$ (sec m)$^{-1}$
θ	Potential temperature; also, phase angle of the unstable perturbation vertical velocity
π	$\pi = 3.14159$
ρ	Density
$\sigma(p)$	A measure of static stability, $\sigma(p) = -\frac{1}{\bar{\rho}} \frac{\partial \ln \bar{\theta}}{\partial p}$
ϕ	Geopotential height, $\phi = gz$
ψ	Stream function
ω	Vertical velocity in the pressure coordinates, $\omega = \frac{dp}{dt}$
∇	Horizontal gradient operator on a constant pressure surface

1. Introduction

Large-scale weather phenomena in the extratropical atmosphere are associated with slowly traveling waves in the prevailing westerlies. One of the classical problems in theoretical meteorology has been the explanation of the origin, development, and structure of these waves, based upon the study of hydrodynamically unstable wavelike disturbances in the zonal current.

Helmholtz (1886) first studied dynamic instability by considering wave motions along a surface of discontinuity with an abrupt change of wind and density (shearing instability) (see Lamb, 1932). Later Rayleigh (1913) showed that the stability of horizontal parallel flows would depend on the shape of the flow profile and that the current was unstable if the curvature of the current profile changed sign somewhere within the region. Further advances in the study of instability theory of an inviscid fluid was made by Lin (1945, 1955). However, these studies were restricted to nonrotating systems.

By introducing the uniform latitudinal variation of the Coriolis parameter, Rossby (1939) discovered the very important role of the earth's rotation and sphericity in the dynamics of atmosphere and obtained a law governing the propagation of large-scale waves in uniform westerlies. These waves were completely neutral or energetically inactive. The assumption of a uniform current in Rossby waves eliminated any available sources of energy and so could not explain the existence and development of the waves.

From one point of view, cyclogenesis may be regarded as a result of dynamic instability, whereby the amplitudes of small wavelike disturbances superimposed on the basic steady current will grow exponentially with time. Results obtained from this approach are far from complete if they are not supported by detailed information regarding the energetically active nature of the disturbances. It is now generally recognized that the large-scale (cyclone scale) waves possess both potential-energy-converting and kinetic-energy-converting mechanisms which are important factors in maintaining the general circulation against frictional dissipation. (See, for example, Starr, 1954, and Dutton and Johnson, 1967.)

To understand the processes, the structure and energetics of the three-dimensional disturbances superimposed on the basic zonal current containing both horizontal and vertical shear must be considered. In this type of basic zonal current, there are essentially two categories of energy sources: the first is the available potential energy associated with the vertical shear which is related to the horizontal temperature gradient caused by the latitudinal variation of solar heating (Lorenz, 1955); the second is the available kinetic energy associated with the horizontal shear which is related to the jet-like wind structure of middle latitudes.

Since the primary mathematical difficulty in the stability problem for such

zonal currents containing both energy sources is the occurrence of singularities and nonseparability in the governing higher order partial differential equation, numerous past studies have proceeded in two different directions by assuming that only a single energy source is present, which in turn leads to an ordinary differential equation. For convenience, these two different studies may be classified as "pure barotropic" and "pure baroclinic" instability studies, respectively.

The "pure barotropic" case is concerned with the inviscid zonal current possessing horizontal shear only, contained between two rigid walls. Kuo (1949) showed that if the meridional gradient of absolute vorticity of the basic barotropic current $(\beta - \frac{d^2\overline{U}}{dy^2}(y))$ changed its sign somewhere within the region, the current would be unstable with respect to nondivergent barotropic disturbances possessing horizontal scales of a few thousand km. In this case, the energy source was the "available kinetic energy" associated with latitudinal variations of the zonal current. Kuo also showed that the unstable waves could be characterized by troughs and ridges tilted from SW to NE north of the maximum current velocity and SE to NW south of the maximum current; such tilts appear to be in the opposite direction with respect to typical observed waves in the westerlies. The more recent results of Jacobs and Wiin-Nielsen (1966) show that these qualitative features of barotropic instabilities are not significantly altered when the presence of the vertical stratification of the atmosphere is taken into account.

Charney (1947) provided the first analysis of the stability of "pure baroclinic" zonal currents (possessing only uniform vertical shear supported by a latitudinally uniform temperature gradient). Since then this problem has been studied much more than the "pure barotropic" case by many investigators, notably by Eady (1949), Kuo (1952, 1953), Burger (1962). One common result of these investigations was that the pure baroclinic currents are unstable to quasi-geostrophic wave disturbances of cyclone scale $(O(10^3$ km)) for sufficiently large vertical shear. The waves produced by this instability mechanism are quite similar to observed waves in the mid-latitudes which tilt westward with height and convert the available potential energy of the zonal current to the eddy kinetic energy through the available eddy potential energy.

After these simplifications have been invoked in the baroclinic problem, formidable mathematical difficulties associated with the analytical solution of this singular characteristic-value problem still exist, even though methods for finding infinite series solutions are well known. For example, Green (1960) modified the "Eady problem" (vertically linear basic current without north-south dependence, $\beta = 0$, and constant stability parameter) so that there was a north-south gradient of the basic potential vorticity by introducing the "beta effect" $(\beta > 0)$ and discovered that the flow was unstable to disturbances of all wave lengths, even for small values of β. However, the numerically computed growth rates were very small in the short and long wave regions.

The differences between Eady's and Green's problems were illustrated by Bretherton (1966), based upon the concept of "critical layer instability" for quasi-geostrophic flows. This effect is associated with a critical layer (defined as a layer in which the phase speed of disturbances (c_r) is locally equal to the speed of the basic zonal current in the limit of no instability ($c_i \to 0$) where the disturbance structure changes in a singular manner. This instability, also investigated by Miles (1964), showed that it was significant in the stable or small growth rate region but insignificant in the region of large growth rates.

One way to avoid such difficulties is to utilize the approximate methods which allow more general zonal wind profiles to be considered. For example, Phillips (1951, 1954) and Thompson (1953) investigated the pure baroclinic stability problem by replacing the vertically continuous model with the discrete layer model. However, the reduced number of degrees of freedom in the vertical admits only disturbances whose vertical scale is the order of the depth of layer, resulting in a spurious stability for smaller scale disturbances.

Eliassen (1954) and Wiin-Nielsen (1961) utilized a finite Fourier series for the amplitude of stream function to determine the stability properties of the nondivergent and the divergent quasi-barotropic model, respectively. Haltiner and Song (1962) made some comparisons between the finite difference method and the Fourier series method for several quasi-barotropic models with both single-jet and double-jet currents and showed that the introduction of divergence into the model in the form of a Helmholtz term would tend to have a stabilizing influence for the single-jet current.

All of the above studies suffered limitations in their applicability to the real atmosphere because of the restrictive forms of the assumed zonal wind profiles. Typical observed zonal flow profiles are "jet-like," possessing both horizontal and vertical shear. Such flows will be referred to as "mixed" zonal currents, in view of the fact that they contain the energy sources characteristic of both barotropic and baroclinic basic flows. The increased physical complexity is reflected in added degrees of mathematical difficulty. Despite this, Pedlosky (1964) recently determined several necessary conditions for instability and found bounds on the growth rates and phase speeds of the quasi-geostrophic disturbances superimposed on a "mixed" zonal current in both a continuously-stratified and a two-layer model. In his following paper (1964), he investigated some energetic properties of unstable disturbances for currents with both parabolic-jet and cosine-jet latitudinal profiles by circumventing the nonseparability of the governing perturbation partial differential equation by means of replacing the continuous model with a two-layer model. Haltiner's (1963) work represents an improvement of this approach to a model with greater vertical resolution; by utilizing finite differences in the meridional plane, he was able to study the stability characteristics for realistic jet-like zonal currents. As in most such studies, these results were restricted mainly to the determination of quantitative stability characteristics, thereby resulting in relatively little information concerning the detailed structure and energetic

properties of unstable disturbances in the realistic model atmosphere.

The primary purpose of this study is to obtain information for representative zonal currents containing both horizontal and vertical shear of the following sort: (1) dynamic instability properties; (2) detailed structure of the three-dimensional unstable disturbances; (3) energetic properties of these disturbances; and (4) their relationships to the available sources of energy in the basic current. Hence, it is attempted to increase our understanding of the role of unstable cyclone waves in the maintenance of general circulation when both kinetic and potential energy converting mechanisms are included in the model.

It is natural to try to isolate these mechanisms as much as possible, at the same time trying to avoid artificially suppressing any of the natural coupling of the various physical processes. The technique which immediately suggests itself is to study linearized equations for the disturbances; this allows the unstable modes to interact realistically with the zonal flow (whose characteristics are known) while the mutual interactions of the unstable modes are suppressed.

Two criticisms of this approach may be mentioned. First, disturbances in the real atmosphere are of finite, rather than infinitesimal, amplitude. However, based upon comparisons of other linear theories with observation (see, for example, Charney, 1947), we envisage the present results as correctly predicting the essential features relating observed moving waves to the zonal flow. Second, it is sometimes argued that attention should be given only to the mode of maximum growth rate. This suggestion is certainly valid when the perturbations are all initially small. However, the presence of an entire spectrum of developed waves at any time in the real atmosphere changes the solution drastically. For example, a given wave mode characterized by relatively slow growth rates may be excited to large amplitude by the ever-present effects of surface inhomogeneities of interactions with other waves, while the most unstable wave may not be so well developed at the same time. In such a case, the energetics would at least initially be dominated by processes corresponding to the slowly growing mode.

Even in linearized form the governing equations must be solved numerically by following Haltiner (1963) and utilizing finite-difference approximations for both vertical and latitudinal derivatives. The time derivatives that remain may be handled in one of two ways. In the present study (beginning 1967) we assume oscillatory behavior of the disturbance variables in time, which casts the problem into eigenvalue form.

The alternative method is to follow Brown (1969) and employ an initial value approach by integrating the perturbation equations in time. This technique is inefficient when two modes have comparable growth rates, and ultimately allows the study of only the most unstable mode. Brown found two wavelengths of maximum instability for zonal flows which would be expected

to support both baroclinic and barotropic instability, suggesting that there were at least two modes of instability present. This conjecture was further supported by the fact that the shorter unstable wave strengthened the kinetic energy of the zonal jet, while the longer wave caused kinetic energy transfer in the opposite direction.

Thus, from one point of view the present work may be regarded as adding substantially to Brown's conclusions and to delineate the character of zonal flow instability more completely. The eigenvalue approach used to accomplish this possesses a peculiar disadvantage; namely, all eigenmodes satisfying the equations are found, leaving the investigator with an abundance of information. When attention is focused upon the unstable modes alone, we find in some cases a number of "secondary" unstable modes whose origin is uncertain, if not controversial (Arnason, 1967). These modes appear to be related to the finite difference version of the "critical layer instabilities" mentioned earlier, or they may be the result of truncation errors associated with singularities in the governing equations.

In analytical studies, the way of circumventing the singular solution of the inviscid equation for a simple barotropic model is to apply the asymptotic properties of the solution of the viscous counterpart in the limit of vanishing viscosity (see Lin, 1945; Wasow, 1948). The recent study of Yanai and Nitta (1968) sheds light on the numerical problem for inviscid barotropic flow. These authors showed that spurious unstable modes resulted from the effects of truncation errors acting on antisymmetric singular modes which were normally stable. These singular modes, which possessed "critical points" as a result of the shearing influence of the zonal flow, were found to become accurately stable as the grid resolution was increased.

Arnason (1967), in his study on the stability of a simple "primitive" baroclinic model, found that such spurious instabilities could not be removed by improving grid resolution. It would thus appear that the relevance of the slowly growing modes in a baroclinic numerical model is somewhat mathematically suspect. However, should these modes be physically real "critical layer instabilities," their small growth rates and singular structure suggest their relative unimportance in the real atmosphere.

Thus, in this study, an attempt will be made to pay attention to the physically most important modes associated with the two categories of available energy sources. To accomplish this, we will first study the "pure baroclinic" and the "pure barotropic" cases (which have been analytically studied extensively in the past) and make comparisons with the well known analytical results. This will allow an assessment of the effects and extent of truncation errors on the solutions. The following sections will then try to relate the primary modes of instability for the "mixed" zonal current to the unstable modes of the "pure baroclinic" and "pure barotropic" cases. This will be accomplished by continuously varying the structure of the basic current type (and

hence the forms of available energy) from that of the "pure baroclinic" to the "mixed" to the "pure barotropic" type.

In Chapter 2 of this paper, the eigenvalue problem for the numerical model is formulated. The derivation of the energy equations in a convenient form and discussion of the terms are presented in Chapter 3. Chapters 4 and 5 contain numerical results for the "pure baroclinic" and "pure barotropic" cases, respectively. Chapter 6 presents the detailed results for all perturbations of symmetric jet-like zonal flows with two different latitudinal wind profiles.

2. The Eigenvalue Problem for Linearized Perturbations in Finite-Difference Form

The large-scale atmospheric motions in extratropical regions can be considered quasi-geostrophic, convectively stable, slowly moving waves, and inviscid above the planetary boundary layer. In terms of the nondimensional parameters which play a central role in the scale analysis, their Rossby number is small ($0(10^{-1})$) and their Richardson number is large ($0(10^{2})$). Starting with this assumed information, scale analysis, a methodical procedure for obtaining approximate forms of the governing equations appropriate for these motions, may be applied. Good examples of this approach for quasi-geostrophic motions appear in the works of Charney (1948, 1962), Burger (1958), and Pedlosky (1964). The most complete exposition on the subtle points and implications is that of Phillips (1963).

If we are concerned with the relatively rapid development of these motions compared to the time scale for diabatic processes such as radiation, the latter may be neglected. Then the vorticity and the thermal equations for these quasi-geostrophic, adiabatic motions in the β-plane system may be written, respectively, in the following form:

$$\left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla\right) \nabla^2 \psi + \beta \psi_x = f_0 \omega_p , \tag{2.1}$$

$$\left(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla\right) \psi_p + \frac{\sigma(p)}{f_0} \omega = 0 . \tag{2.2}$$

In these equations, the independent variables x, y, p, and t represent the west to east coordinate, south to north coordinate, pressure, and time, respectively. The subscripts x, y, p, and t denote partial differentiation. ψ is the stream function for the geostrophic wind, $\vec{V} = \vec{k} \times \nabla\psi$ the nondivergent horizontal geostrophic wind vector, f_0 and β the Coriolis and the beta parameters, respectively (assuming constant values appropriate for a standard latitude), ∇ the "horizontal" del operator on constant pressure surfaces, $\omega = \frac{dp}{dt}$, and the measure of static stability is $\sigma(p) = -\frac{1}{\bar{\rho}} \frac{\partial \ln \bar{\theta}}{\partial p}$, where $\bar{\rho}$ and $\bar{\theta}$ are the density and potential temperature of the standard atmosphere,

respectively. The geometrical distortion is considered negligible on these large-scale motions of limited latitudinal extent, as has been pointed out by many former investigators, for example Burger (1958).

In Eq. (2.1), we have neglected the vertical advection of vorticity and twisting terms which are normally smaller than the other terms. This implies that the vertical advection of horizontal momentum in the horizontal equation of motion, from which the vorticity Eq. (2.1) is derived, has been ignored. Hence the vertical Reynolds stresses are an order of magnitude less than the horizontal stresses. Furthermore, the absolute vorticity in the divergence term of Eq. (2.1) has been replaced by an estimate involving the mean Coriolis parameter to avoid a spurious vorticity source (see Wiin-Nielsen, 1959).

In Eq. (2.2), the hydrostatic and geostrophic approximations have been used:

$$\frac{\partial \phi}{\partial p} = f_0 \frac{\partial \psi}{\partial p} = -\frac{1}{\rho} , \tag{2.3}$$

where ϕ is the geopotential height. The stability parameter σ is treated as a function of p only to have a sufficient energy invariant which has been shown by Lorenz (1960).

Now we can see that Eqs. (2.1) and (2.2) form a complete system of equations with two unknown dependent variables, ψ and ω. Differentiating (2.2) with respect to p and then eliminating ω_p from (2.1) results in the following quasi-geostrophic potential vorticity equation:

$$\frac{\partial q}{\partial t} + \vec{V} \cdot \nabla q = 0 , \tag{2.4}$$

where q is the absolute potential vorticity, defined as

$$q = \nabla^2 \psi + (f_0 + \beta y) + \left(\frac{f_0}{\sigma(p)} \psi_p\right)_p . \tag{2.5}$$

A similar expression in z-coordinates was derived by Pedlosky (1964).

For horizontal boundary conditions, rigid vertical walls will be assumed to exist at $y = 0$ and D to restrict the latitudinal extent of the disturbances, which is required in order to neglect the geometrical distortions. Then, by the kinematic boundary condition, these can be written as

$$v = \psi_x = 0 \qquad \text{at } y = 0 \text{ and } y = D . \tag{2.6}$$

For vertical boundary conditions, similarly, it will make the standard assumptions that

$$\omega = 0 \qquad \text{at } p = 0 \text{ and } p = p_0 = 1000 \text{ mb.} \tag{2.7}$$

Here the lower boundary condition is obtained from the assumption of a rigid level earth's surface having an approximately constant pressure (p_0 = 1000 mb); it may be noted that Lorenz (1957) has pointed out that the total energy will still be conserved even when $p = p_0$ is not a constant value. When substituted into (2.2), the vertical conditions (2.7) are expressed in terms of ψ alone:

$$(\frac{\partial}{\partial t} + \vec{V} \cdot \nabla)\psi_p = 0 \quad \text{at } p = 0 \text{ and } p = p_0 . \tag{2.8}$$

Now the linearization is accomplished by superimposing small perturbations on the time independent zonal solution of (2.4):

$$\psi = \bar{\psi}(y, p) + \psi'(x, y, p, t), \qquad \bar{\psi} >> \psi' . \tag{2.9}$$

Then it is obvious from (2.2) that

$$\omega = \bar{\omega}(y, p) + \omega'(x, y, p, t), \tag{2.10}$$

where $\bar{\omega}(y, p) = 0$ when the mean motions are frictionless and adiabatic. When (2.9) and (2.10) are substituted into Eqs. (2.1), (2.2), (2.4), (2.6), and (2.8), linearization of the advection terms by omission of products of perturbation variables produces the following system of equations and associated boundary conditions:

$$(\frac{\partial}{\partial t} + \bar{U}\frac{\partial}{\partial x})\nabla^2\psi' + (\beta - \bar{U}_{yy})\psi'_x = f_0\omega'_p , \tag{2.11}$$

$$(\frac{\partial}{\partial t} + \bar{U}\frac{\partial}{\partial x})\psi'_p - \bar{U}_p\psi'_x + \frac{\sigma(p)}{f_0}\omega' = 0 \tag{2.12}$$

$$(\frac{\partial}{\partial t} + \bar{U}\frac{\partial}{\partial x})[\nabla^2\psi' + (\frac{f_0^2}{\sigma(p)}\psi'_p)_p] + \psi'_x\bar{q}_y = 0, \tag{2.13}$$

$$\psi'_x = 0 \quad \text{at } y = 0, D, \tag{2.14}$$

$$(\frac{\partial}{\partial t} + \bar{U}\frac{\partial}{\partial x})\psi'_p - \bar{U}_p\psi'_x = 0 \quad \text{at } p = 0, p_0 . \tag{2.15}$$

$\bar{U} = -\bar{\psi}_y$ is the basic zonal current, and $\bar{q}_y = \beta - \bar{U}_{yy} - (\frac{f_0^2}{\sigma(p)}\bar{U}_p)_p$ the meridional gradient of potential vorticity of the zonal current.

Since the coefficients in these perturbation equations are functions of y and p, and the independent variables are x and t, separable solutions can be found of the harmonic wave form:

$$\psi' = \mathrm{Re}[A(y, p)e^{ik(x-ct)}], \tag{2.16}$$

where $A(y, p)$ is the complex amplitude in the meridional plane, $k = \frac{2\pi}{L}$ the zonal wave number, and c the eastward phase speed.

If this form is substituted into (2.13), (2.14) and (2.15), we obtain:

$$(\bar{U} - c)(A_{yy} + (\frac{f_0^2}{\sigma(p)} A_p)_p - k^2 A) + A\bar{q}_y = 0, \tag{2.17}$$

$$A = 0 \quad \text{at } y = 0, D, \tag{2.18}$$

$$(\bar{U} - c)A_p - \bar{U}_p A = 0 \quad \text{at } p = 0, p_0 . \tag{2.19}$$

Equation (2.17) together with the associated boundary conditions (2.18) and (2.19) constitutes an eigenvalue problem, or the stability problem, which contains several as yet undetermined parameters (such as c and k) in the equation and the boundary conditions. Since Eq. (2.17) and its boundary conditions are homogeneous, the solutions A, which are usually called eigenfunctions, will exist for only certain eigenvalues c, if all other parameters are fixed. The parameter k would be fixed, for example, by observation of the initial disturbance.

If the eigenvalue c is complex valued such that $c = c_r + ic_i$, from (2.16) we see that a wave disturbance with wave number k will travel at a speed c_r in the eastward direction with a time amplification (or damping) factor $e^{kc_i t}$. Since all the coefficients in (2.17) are real, the eigenfunction A for a corresponding complex eigenvalue c must be complex; the conjugates of c and A also satisfy the same fundamental equations. Therefore both of the amplified and damped disturbances exist simultaneously.

Due to the variable coefficients in Eq. (2.17), it is of a nonseparable mathematical form for most zonal flow profiles $\bar{U}(y, p)$. In addition, it may be singular, and possess multivalued solutions, as for example when we have a neutrally stable wave ($c_i = 0$) with $c = c_r = \bar{U}$ at some critical point.

If we omit the vertical variation of $\bar{U}$, σ, and A, Eq. (2.17) reduces to the familiar perturbation equation for nondivergent barotropic disturbances. On the other hand, eliminating latitudinal variations in $\bar{U}$ and A leads to the governing equation for pure baroclinic instability. In both cases the lateral boundary conditions (2.18) impose a finite lateral extent for the disturbances, and hence a changing latitudinal structure.

The eigenvalue problem in numerical form is formulated by transforming (2.17) to a system of algebraic homogeneous equations, by utilizing finite difference approximations for both y and p derivatives. Since the coefficients,

differential operators in (2.17), and the boundary conditions are latitudinally symmetric with respect to the center of the basic current (j = 1 in Fig. 1) when latitudinally symmetric basic currents are considered, we see that disturbances possessing pure latitudinal symmetry or antisymmetry each satisfy the system independently of the other. Therefore, only half of the channel needs to be considered; the boundary condition at $y = 0$ is replaced by the conditions $A_y = 0$ at $y = D/2$ for symmetric disturbances, $A = 0$ at $y = D/2$ for antisymmetric disturbances.

It may be noted here that most investigations in the past were limited to symmetric perturbations, following the logic that the antisymmetric perturbations interact less with the basic flow because they require vanishing northward motion at the center of the basic current. However, in this study, the antisymmetric case is also studied in an attempt to answer this question more precisely.

The meridional domain of the atmosphere is next subdivided vertically into N layers, and latitudinally into M bands from the center of the region ($y = D/2$) to the boundary line ($y = D$) as shown in Fig. 1.

Using centered difference approximations for the derivatives, the application of (2.17) at every interior grid point results in $(M-1) \times (N-2)$ homogeneous algebraic equations in $(M-1) \times N$ unknowns (the amplitudes of the perturbation stream function) when the lateral boundary conditions mentioned above are taken into account. At the vertical boundaries ($p = 0$ and $p = p_0$), Eq. (2.19) is applied by using forward and backward difference approximations, respectively. We then have a complete system of $(M-1) \times N$ homogeneous algebraic equations in $(M-1) \times N$ unknowns.

For convenience, in the matrix notation it can be written as:

$$(\underline{B} - c\underline{D})\underline{A} = 0, \tag{2.20}$$

where c is the phase speed, $\underline{B}$ and $\underline{D}$ are $(M-1) \times N$ order, square, real matrices involving the basic zonal current, wave number k, β, f_0, stability parameter, and mesh sizes of the grid. $\underline{A}$ is a column vector consisting of the complex amplitude of the stream function at every grid point. Since the matrix $\underline{D}$ is always nonsingular for this problem, Eq. (2.20) is equivalent to

$$(\underline{D}^{-1}\underline{B} - c\underline{I})\underline{A} = 0, \tag{2.21}$$

where $\underline{I}$ is the identity matrix and $\underline{D}^{-1}$ the inverse of $\underline{D}$. The condition for a nontrivial solution A of (2.21) is the singularity of the matrix $(\underline{D}^{-1}\underline{B} - c\underline{I})$ which results in the so-called "frequency equation" for c. Thus the phase speeds for this system are just the eigenvalues of the matrix $\underline{D}^{-1}\underline{B}$, and similarly its eigenvectors correspond to the column vector $\underline{A}$ (see Haltiner, 1963).

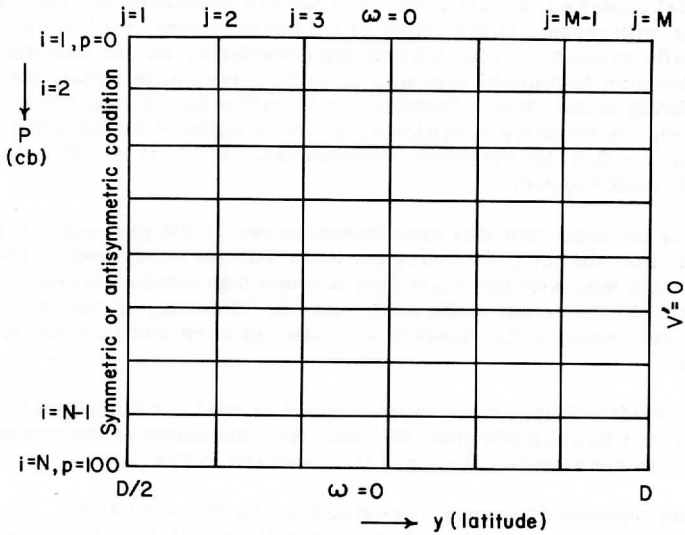


Fig. 1. Grid-point subdivisions and boundary conditions in the northern half of the meridional plane.

The problem now is to compute the eigenvalues and the corresponding eigenvectors for given distributions of the zonal current, stability parameter, zonal wavelength, etc. The problem of determining the eigenvalues and eigenvectors of a matrix is a well known subject and there are considerable references on it. Subroutines for large matrix operations suited to modern electronic computers are already available. The method used here is the "Q-R method" for a general square matrix (see the University of Wisconsin Computing Center, Users Manual, Vol. IV, Supplement D, pp. 19). Basically, the method consists of finding each eigenvalue by using equivalent orthogonal matrix transformations iteratively applied to the Hessenberg form (upper triangular, except for a single nonzero subdiagonal) of the original matrix.

In general it may be expected that the accuracy of the solutions to Eq. (2.21) will increase as M and N increase; naturally the computational effort increases as well. (For the larger matrices it was found that the computational time increased as about the 2.7^{th} power of the number of rows; for the 90×90 case, all eigenvalues and eigenvectors were obtained in $4\frac{1}{2}$ minutes on the CDC 3600.) Furthermore, it should be noted that singularities became more clear-cut as the number of grid points increase (see Bretherton, 1966).

For the mixed cases considered later, M and N will be restricted to the values 7 and 15, respectively, resulting in operations on a 90×90 matrix, a size which is about the maximum which can be treated by the CDC 3600 computer. For the pure baroclinic case, $M = 1$, and N will be varied up to 41 in order to judge the changing accuracy of the solutions. On the other hand, for the pure barotropic zonal current, M will vary from 5 to 13 while N varies from 5 to 15.

Having found the stream function values for a given eigenvalue, we relate the wind component

$$u' = \mathrm{Re}[\hat{U}(y, p)e^{ik(x-ct)}],$$
$$v' = \mathrm{Re}[\hat{V}(y, p)e^{ik(x-ct)}]$$

geostrophically to A by the following relations:

$$\hat{U} = -A_y,$$
$$\hat{V} = +ikA.$$

Similarly the temperature field

$$T' = \mathrm{Re}[\hat{T}(y, p)e^{ik(x-ct)}]$$

can be obtained from the hydrostatic relation (2.3) as

$$\hat{T} = -\frac{f_0 p}{R} A_p,$$

where R is the specific gas constant for air.

The vertical motion expressed in the form

$$\omega' = \mathrm{Re}[W(y, p)e^{ik(x-ct)}] \tag{2.22}$$

is obtained from the linearized form of the thermodynamic Eq. (2.12)

$$(\bar{U} - c)A_p - \bar{U}_p A - \frac{i\sigma(p)}{f_0 k} W = 0. \tag{2.23}$$

Now (2.16) and (2.22) can be written as follows:

$$\psi' = \mathrm{Re}[\{A_r(y, p) + iA_i(y, p)\}e^{ik(x-(c_r+ic_i)t)}], \tag{2.24}$$

$$\omega' = \mathrm{Re}[\{W_r(y, p) + iW_i(y, p)\}e^{ik(x-(c_r+ic_i)t)}]. \tag{2.25}$$

Using Euler's formula, these become

$$\psi' = e^{kc_i t}|A| \cos (kx - c_r t - \alpha), \tag{2.26}$$

$$\omega' = e^{kc_i t}|W| \cos (kx - c_r t - \theta), \tag{2.27}$$

where

$|A| = \sqrt{A_r^2 + A_i^2}$ is the absolute magnitude of A,

$\alpha = \tan^{-1}(-\frac{A_i}{A_r})$ the phase angle of ψ',

$|W| = \sqrt{W_r^2 + W_i^2}$ the absolute magnitude of W,

$\theta = \tan^{-1}(-\frac{W_i}{W_r})$ the phase angle of ω'.

Substituting (2.24) and (2.25) into (2.23) yields the real and imaginary parts of W as the following expressions:

$$W_r = \frac{f_0 k}{\sigma}[(\bar{U} - c_r)(A_i)_p - c_i(A_r)_p - A_i \bar{U}_p], \tag{2.28}$$

$$W_i = \frac{f_0 k}{\sigma}[-(\bar{U} - c_r)(A_r)_p - c_i(A_i)_p + A_r \bar{U}_p]. \tag{2.29}$$

From these equations, the distribution of vertical motion in the meridional cross section can be computed numerically if the eigenvector for an eigenvalue is computed in (2.21).

3. The Energy Equations

In the last section, it was shown that if an eigenvalue c in the frequency equation is a complex number, the disturbance would grow with time. Therefore there should be an increased intensity of energy conversions supported by the two available energy sources mentioned in the first chapter.

The purpose of this section is to derive all energy forms and conversion terms for quasi-geostrophic unstable disturbances in a manner similar to that introduced by Phillips (1954, 1956), and to discuss the exchanges of energy between the basic and disturbance flows, and between eddy available potential and eddy kinetic energy.

In the real atmosphere, wave disturbances of different wavelengths can exchange energy among themselves by nonlinear interactions, but in this model such exchanges have been precluded by the linearization procedure.

Multiplying (2.11) by ψ' and integrating it over one zonal wavelength by using the cyclic condition in x yields:

$$\left(\frac{1}{2}|\nabla\psi'|^2\right)_t = -\overline{U}\left(\psi'_x\psi'_y\right)_y + \left(\psi'\psi'_{ty}\right)_y + f_0\,\omega'\psi'_p - f_0\left(\omega'\psi'\right)_p. \qquad (3.1)$$

Applying a similar method to Eq. (2.12) by multiplying by $(f_0^2/\sigma(p))\,\psi'_p$ yields:

$$\frac{f_0^2}{2\sigma(p)}\left(\psi'_p\right)^2_t = \frac{f_0^2}{\sigma(p)}\overline{U}_p\,\psi'_x\psi'_p - f_0\,\omega'\psi'_p \qquad (3.2)$$

Here $0\!-\!\!-\!\!-L$ represents the x-integration over one zonal wavelength.

For convenience, Eqs. (3.1) and (3.2) can be written symbolically with the same order of appearance:

$$\frac{\partial K_e}{\partial t} = C(K_z, K_e) + HK_eC + C(A_e, K_e) + VK_eC, \qquad (3.3)$$

$$\frac{\partial A_e}{\partial t} = C(A_z, A_e) - C(A_e, K_e). \qquad (3.4)$$

The conversion terms $C(E_1, E_2)$ represent energy conversion from the energy form E_1 to the form E_2.

The physical interpretation of each term in these equations is quite clear. The available perturbation potential energy, A_e, is an expression consisting of the square of the temperature deviation from the zonal mean value and of the static stability parameter. It may be noted here that for a given temperature deviation, A_e is inversely proportional to the stability parameter. Lorenz (1955) has shown that this kind of expression represents an approximation to the exact definition of the portion of total potential energy of the hydrostatic atmosphere which is available for conversion to kinetic energy. Recently, Dutton and Johnson (1967) obtained the more general expression in isentropic coordinates and noted the kinds of errors inherent in the approximate form.

K_e represents the perturbation kinetic energy of the geostrophic horizontal wind discussed by Phillips (1954). It is important to note that only two terms, $C(K_z, K_e)$ and $C(A_z, A_e)$, contain both perturbation and basic state quantities and hence represent an interaction between these two quantities. More precisely, $C(K_z, K_e)$ represents the energy conversion from the zonal kinetic energy K_z to the perturbation or eddy kinetic energy K_e ("kinetic energy conversion"). This is accomplished by the horizontal Reynolds stresses and is significant when the zonal mean eddy momentum transport varies with latitude. The conversion from zonal to eddy available potential energy ("potential energy conversion"), denoted by $C(A_z, A_e)$, is positive when the perturbations transport

warm air northward and cold air southward, implying the northward transport of sensible heat by the perturbation meridional velocity v'.

The conversion of potential into kinetic perturbation energy ("perturbation energy conversion"), $C(A_e, K_e)$, represents the effect of vertical circulations in the zonal plane; it is positive for the relative upward motion of warm air and downward motion of cold air. It should be noted here that the effect of mean vertical circulation in the meridional plane has been neglected since we assumed $\bar{\omega} = 0$. $\bar{\omega}$ has been estimated to be very small in numerous past studies (see Phillips, 1956; Oort, 1964). Thus the temperature gradient in the basic state of the atmosphere will decrease when $C(A_z, A_e)$ and $C(A_e, K_e)$ are positive, as is usually the case for observed large-scale motions in the extratropical atmosphere.

The boundary terms in (3.3) are HK_eC and VK_eC, representing the horizontal and vertical kinetic energy flux convergences. These become zero when the entire atmosphere is considered, since the boundary conditions do not allow transport through the sides of the domain. More precisely, VK_eC represents the net rate of work done vertically by the pressure. This term may be significant in determining the vertical distribution of K_e, but become zero when it is integrated vertically due to the vanishing energy flux at the boundaries.

On the other hand, the physical meaning of HK_eC is less clear. To gain more insight into HK_eC, it may be written in two parts as follows:

$$HK_eC = \overset{0\text{———}L}{(\psi'\psi'_{ty})_y} = \overset{0\text{———}L}{(\bar{U}\psi'\psi'_x)_y} - \overset{0\text{———}L}{f_0(\psi' v'_{div})_y}, \qquad (3.5)$$

where v'_{div} represents the divergent part of the perturbation meridional velocity.

This relation has been derived from the linearized first equation of motion and the geostrophic relation, in addition neglecting the vertical advection term and using the cyclic condition in x (see Phillips, 1963). It should be noted here that the wind components in the acceleration terms in the first momentum equation are nondivergent, but v' in the Coriolis force term consists of the divergent part alone. The second term on the right side of (3.5) represents the net latitudinal pressure work rate which is analogous to VK_eC.

The first term reduces to the following expression for kinetic energy conversion after having been combined with $C(K_z, K_e)$ (see Lorenz, 1955):

$$\overset{0\text{———}L}{(\bar{U}\psi'_x\psi'_y)_y} - \overset{0\text{———}L}{\bar{U}(\psi'_x\psi'_y)_y} = \overset{0\text{——}L}{\psi'_x\psi'_y\bar{U}_y}. \qquad (3.6)$$

Since integrating (3.6) over y between two rigid boundaries results in the vanishing of the first term, the remaining two terms are exactly identical in integrated form. Thus, in this study HK_eC will be treated as a single term representing the horizontal K_e flux convergence. This term may be important in determining the latitudinal distribution of K_e, but becomes zero when it is integrated over y domain.

Integrating (3.1) and (3.2) over the y-p plane (meridional cross section) by using the appropriate boundary conditions reduces to:

$$\frac{\partial K_E}{\partial t} = C(K_Z, K_E) + C(A_E, K_E), \tag{3.7}$$

$$\frac{\partial A_E}{\partial t} = C(A_Z, A_E) - C(A_E, K_E), \tag{3.8}$$

where the capital subscripts represent the total energy integrated over the entire mass of the atmosphere.

Now we consider the dependence of the energy conversions on the structure of unstable disturbances. Substituting (2.24) and (2.25) into (3.1) and (3.2) and integrating over the entire mass of the atmosphere yields:

(a) $$C(A_Z, A_E) = e^{2kc_it}\, g^{-1}\, k\frac{L}{2} \int \frac{f_0^2}{\sigma}\, \bar{U}_p [A_r(A_i)_p - A_i(A_r)_p]\,dydp\,,$$

(b) $$C(K_Z, K_E) = e^{2kc_it}\, g^{-1}\, k\frac{L}{2} \int \bar{U}[A_i(A_r)_{yy} - A_r(A_i)_{yy}]dydp,$$

(c) $$C(A_E, K_E) = e^{2kc_it}\, g^{-1}\, f_0\, \frac{L}{2} \int [W_r(A_r)_p + W_i(A_i)_p]dydp,$$

(d) $$A_E = e^{2kc_it}\, g^{-1}\, \frac{L}{4} \int \frac{f_0^2}{\sigma}[((A_r)_p)^2 + ((A_i)_p)^2]\,dydp, \tag{3.9}$$

(e) $$K_E = e^{2kc_it}\, g^{-1}\, \frac{L}{4} \int [((A_r)_y)^2 + ((A_i)_y)^2 + k^2(A_r^2 + A_i^2)]\,dydp,$$

(f) $$HK_eC = e^{2kc_it}\, k\frac{L}{2} [c_i(A_r(A_r)_y + A_i(A_i)_y)_y + c_r(A_r(A_i)_y - A_i(A_r)_y)_y],$$

(g) $$VK_eC = -e^{2kc_it}\, f_0\, \frac{L}{2} [(W_rA_r + W_iA_i)_p]\,,$$

where the subscripts y and p denote differentiation and the subscripts r and i represent the real and imaginary parts.

By using (2.26) and (2.27), each term of (3.9) can be written in terms of the absolute magnitudes and phase angles of the unstable disturbances as:

$$\text{(a)} \qquad C(A_Z, A_E) = e^{2kc_i t} g^{-1}(-k\frac{L}{2}) \int \frac{f_0^2}{\sigma} \bar{U}_p(|A|^2 \alpha_p)\, dydp,$$

$$\text{(b)} \qquad C(K_Z, K_E) = e^{2kc_i t} g^{-1} k \frac{L}{2} \int \bar{U}(|A|^2 \alpha_y)_y dydp,$$

$$\text{(c)} \qquad C(A_E, K_E) = e^{2kc_i t} g^{-1} f_0 \frac{L}{2} \int [|W||A|_p \cos(\theta-\alpha),$$

$$+ |W||A|\alpha_p \sin(\theta-\alpha)]dydp, \qquad (3.10)$$

$$\text{(d)} \qquad A_E = e^{2kc_i t} g^{-1} \frac{L}{4} \int \frac{f_0^2}{\sigma}[(|A|_p)^2 + (|A|\alpha_p)^2]dydp,$$

$$\text{(e)} \qquad K_E = e^{2kc_i t} g^{-1} \frac{L}{4} \int [(|A|_y)^2 + (|A|\alpha_y)^2 + |A|^2 k^2]dydp,$$

$$\text{(f)} \qquad HK_eC = e^{2kc_i t} k \frac{L}{2}[c_i(|A||A|_y)_y - c_r(|A|^2 \alpha_y)_y],$$

$$\text{(g)} \qquad VK_eC = -e^{2kc_i t} f_0 \frac{L}{2}[|W||A|)_p \cos(\theta-\alpha) - |W||A|(\theta-\alpha)_p \sin(\theta-\alpha)].$$

From these relations, it can be seen that when the troughs (or ridges) of the stream lines tilt westward with height ($\alpha_p > 0$), the potential energy conversion is positive ($A_Z \rightarrow A_E$) for the usual tropospheric case of a negative $\bar{U}_p$.

In the region of $\bar{U}_y < 0$, the eastward tilt of the troughs (or ridges) with latitude ($\alpha_y > 0$) will give a positive kinetic energy conversion ($K_Z \rightarrow K_E$). However, the perturbation energy conversion, $C(A_E, K_E)$, and VK_eC are related to the difference in phase between ψ' and ω'. If the phase angle of ω' leads that of ψ' (typically by about 90 degrees) (see Charney, 1947; Kuo, 1952), the second terms in $C(A_E, K_E)$ and VK_eC (3.10(c) and (g)) are dominant and have opposite signs. Thus, if both ψ' and ω' tilt westward with height, the perturbations convert A_E into K_E and give negative VK_eC.

The second term in HK_eC (3.10(f)) is normally the larger (since $c_r > c_i$ except for very long wavelengths) and is similar to $C(K_Z, K_E)$, but with the opposite sign. Furthermore, we note that all energy terms involving available potential energy and ω' are inversely related to the static stability parameter.

It is important here to note that for real eigenvalues c, there exist only real eigenvectors in (2.21), resulting in zero phase angle for the stream function and exactly 90 degrees for the ω' waves (see (2.28) and (2.29)). Hence it can be easily seen in (3.10) that all conversion terms should vanish for the stable modes ($c_i = 0$), as we expect. Moreover, damped waves possessing the conjugate complex eigenvalue $c = c_r - ic_i$ ($c_i > 0$) exhibit energy conversions which have the same magnitude but are in the opposite direction to those of the amplifying mode ($c = c_r + ic_i$). (This may be proven by noting that α and $(\theta - 90°)$ change sign for the conjugate mode.)

It may be interesting to note that the fractional time rate of change of the perturbation energies (A_E and K_E) integrated over x or the entire mass of the atmosphere depends on the growth rate as follows:

$$\frac{(K_E)_t}{K_E} = 2kc_i, \tag{3.11}$$

$$\frac{(A_E)_t}{A_E} = 2kc_i. \tag{3.12}$$

Hence, as a crude check on the accuracy of the numerical results, an error function (which will be referred to later) is hereby introduced:

$$E_K = \left(1 - \frac{C(K_Z, K_E) + C(A_E, K_E)}{2kc_i K_E}\right) \times 100\%, \tag{3.13}$$

$$E_A = \left(1 - \frac{C(A_Z, A_E) - C(A_E, K_E)}{2kc_i A_E}\right) \times 100\%, \tag{3.14}$$

where E_K and E_A are the kinetic energy and available potential energy errors, respectively. It is easy to see (using (3.9(a), (c), (d), (2.28) and (2.29)) that E_A is exactly zero in the numerical computations. Thus, only E_K will be used in later sections.

We have seen that the energy conversions from the available energy sources of the basic current to the corresponding perturbation energies are proportional to the shears of the basic current (see (3.2) and (3.6)). Hence the following definitions will be invoked to express the amount of available energy sources as a function of the shears:

$$A^* = \frac{2}{D} \int_0^{p_0} \int_0^{D/2} \frac{1}{2\sigma} (\bar{U}_p)^2 \, dy dp, \tag{3.15}$$

$$K^* = \frac{2}{D}\int_0^{p_0}\int_0^{D/2} \frac{1}{2f_0^2}(\bar{U}_y)^2 dy dp, \qquad (3.16)$$

where A^* and K^* are the available potential energy density and the available kinetic energy density, respectively. Note that the symmetric nature of the basic flow has been assumed. These densities will be obtained by the use of Simpson's method for the "p-integration" and analytically for the "y-integration" for given $\bar{U}$ and σ values.

Pedlosky (1964) has shown that an upper bound on the growth rate is proportional to the sum of the maximum integrands in (3.15) and (3.16). The utility of these forms is all the more apparent when "pure" zonal currents are considered. For example, (3.15) shows that A^* is proportional to the squared magnitude of the vertical wind shear; most past studies of the "pure baroclinic" problem have expressed the stability characteristics in terms of this quantity. In the "pure barotropic" problem it is known that strengthening a current of fixed profile (and thus increasing K^*) leads eventually to an unstable state (see Kuo, 1949).

Finally, it should be noted that K^* depends inversely upon D. Brown (1969) used this as the basis for a procedure for changing K^* while keeping the flow profile fixed. This seems to be somewhat artificial, and in addition forces the latitudinal width of allowable disturbances to change in a prescribed way. For a given zonal wavelength, this means that the horizontal disturbance shapes (and hence behavior) are forced to change. The procedure followed in the present work keeps D fixed, thus allowing the disturbances and their energetics to develop most naturally.

4. Wave disturbances in pure Baroclinic Currents

As explained earlier, one purpose of this work will be to study the character of wave instabilities on a symmetric zonal jet as A^* and K^* are changed. Two special cases may be expected to provide a proper starting point: the "pure baroclinic" case with $K^* = 0$ and the "pure barotropic" case with $A^* = 0$.

In this chapter we consider the first of these, the "pure baroclinic" problem, by studying the stability of a zonal current without horizontal shear whose vertical wind profile is representative of the observed wind at the center of the jet:

$$\bar{U}(p) = \bar{U}_m + b\bar{U}'(p), \qquad (4.1)$$

$$\bar{U}(p) = 2(\bar{U}_m + b\bar{U}'(p)). \qquad (4.2)$$

$\bar{U}_m = 23.725$ m/sec is the vertically integrated mean wind, $\bar{U}'(p)$ measures

the profile of the deviation from $\overline{U}_m$, and b is the variable shear parameter. The first profile is the one reduced from the cosine-jet mixed case (considered later) by neglecting the horizontal variation; the second is similarly reduced from the parabolic-jet mixed case, also to be considered later. The distributions of $\overline{U}'(p)$ and $\sigma(p)$ are given in Fig. 2. $\overline{U}_m$ and $\overline{U}'(p)$ are estimated from Lorenz (1967) for winter at the center of the mean jet. $\sigma(p)$ is obtained from the standard atmosphere.

The governing perturbation equation (second-order total differential equation) derived from (2.11) to (2.15) and the energy equations for the pure baroclinic case are shown in Appendix A. A series of computations was made for a number of different values of the shear parameter, b, which are shown by dots on the vertical line (a) in Fig. 3 (the case $a = 0$). The number of grid points used here is $M = 1$ and $N = 15$, leading to a 15×15 matrix. In addition $f_0 = 1.03 \times 10^{-4}$ sec^{-1} and $\beta = 1.618 \times 10^{-11}$ sec^{-1}m^{-1} (corresponding to 45 degrees latitude), and the width of the channel D is 3,000 km. It should be noted that the disturbances are of finite lateral extent rather than infinite extent (see Phillips, 1954). Furthermore, the lowest latitudinal wave number $l = \frac{\pi}{D}$ has been used, since it is expected to grow fastest. However, the higher latitudinal modes may be significant for long waves when the stability of more general currents is considered.

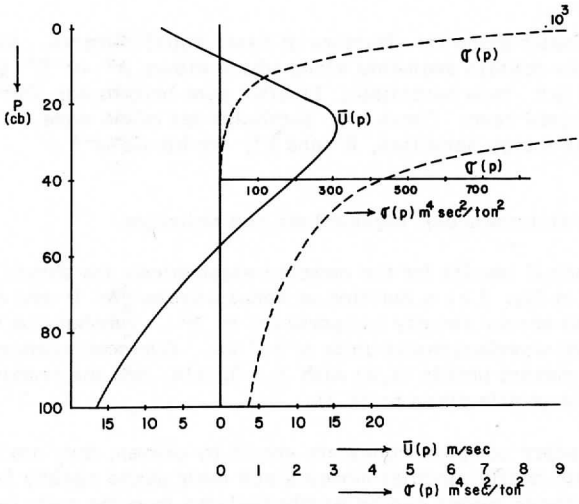


Fig. 2. Vertical profiles of the observed zonal wind deviation from its vertical mean and $\sigma(p)$ for the standard atmosphere.

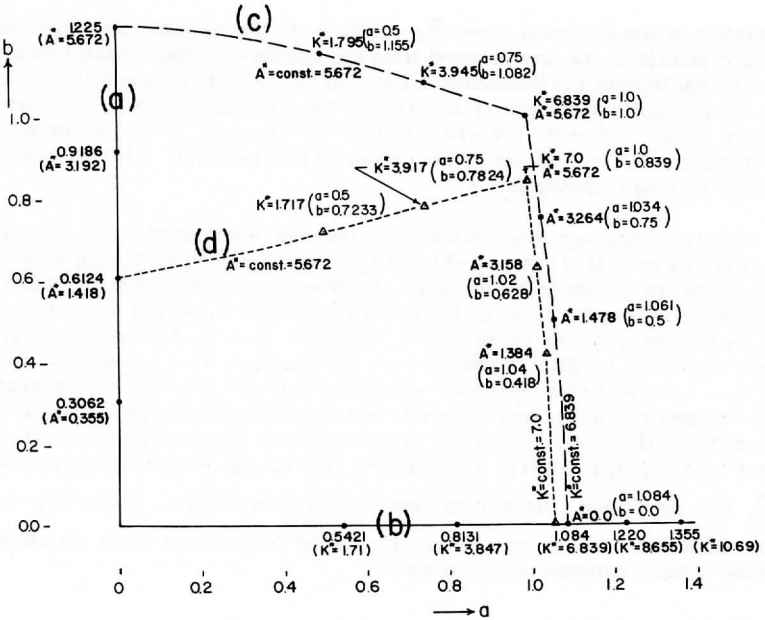


Fig. 3. Parameter plane for structure of basic zonal currents. All labelled curves contain segments along which either A^* or K^* is constant. Line (a): pure baroclinic; Line(b): pure barotropic; Curve (c): cosine-jet mixed case; Curve (d): parabolic-jet mixed case. The units of zonal energy densities, A^* and K^*, are kjoule m^{-3}.

4.1 Unstable disturbances: eigenvalues and structure

The numerical results for the complex eigenvalues are shown for the unstable modes in Fig. 4 as a function of zonal wavelength L and available zonal potential energy density (proportional to b^2). Numbers on the curves represent zonal wavelengths in units of 10^3 km. Curve (e) corresponds to the second zonal current profile (4.2) with b = 0.6124, and the remaining curves are for the first profile given by (4.1).

The secondary complex modes are shown by primes; they are apparently associated with nearly singular modes since their phase speeds fall in the range of $\bar{U}$. These modes can be easily isolated from the main (unprimed) mode, since their imaginary parts are very small and real parts are relatively large. Note that this mode does not appear for cases, (c) and (d), having smaller vertical shear.

The remaining eleven or thirteen real eigenvalues (among a total of fifteen values) apparently correspond to singular or "continuous spectrum" solutions (see Pedlosky, 1965), since they are ranged within the domain of basic current velocities.

From these results it can be seen that the value of c_i increases as the vertical shear increases, but the phase speed c_r decreases with increasing shear (b) for all wavelengths considered. These facts imply that the unstable disturbances have phase speeds equal to the basic current velocity somewhere below the vertical mean level (570 mb) due to the retrogressing effects of β and σ (see Derome and Wiin-Nielsen, 1966; Brown, 1969). For a fixed value of b or A^*, there is a particular wavelength for which c_i is a maximum (for example, 9.5×10^3 km for b = 1.225). This particular wavelength becomes longer as the vertical shear increases as shown by the maximum instability line (dashed line) in Fig. 5. This wavelength is much larger than would be found for the previously studied case of a disturbance with infinite lateral extent. The phase speed c_r for a fixed value of b also has a maximum value at an intermediate wavelength, in agreement with the approximate results of Kuo (1952) for the simple case of linear zonal currents.

Another interesting result can be seen by comparing case (a) and (e). The values of c_i are exactly equal for these two cases but the phase speeds c_r differ by 23.725 m sec^{-1} ($= \bar{U}_m$), because the basic current profiles have identical vertical shear (and hence A^*) but mean current velocities differing by $\bar{U}_m$. Hence we see that while c_i is related to the vertical shear, c_r depends mainly on the mean current velocity. Furthermore, it is important to note (see Appendix A) that the eigenvectors for case (a) and (e) are identical, since the values of $(\bar{U} - c)$ are equal for both cases. Thus, hereafter, only disturbances of case (a) will be considered; the same conclusions will apply to case (e).

Figure 5 shows the growth rates, kc_i, as a function of L and A^* (in units of kjoule m^{-3}) for group (I) in Fig. 4. Here, the dots represent the locations of numerical calculations. From these results, we see that intermediate wavelengths are more unstable than shorter and longer wavelengths. The maximum instability line shows that the wavelength of the most unstable disturbance increases but at a slower rate as A^* increases. For b = 1.225 (implying a shear about 20% too large) the most unstable wave has a wavelength L = 3,700 km and a growth rate $kc_i = 0.95$ day^{-1}; therefore, this wave will double its amplitude within 18 hours.

These results are generally in good agreement, except for very short wavelengths, with those for the more idealized linear current profile with constant fractional change of potential temperature with height ($\partial \ln \theta / \partial z$ = constant) discussed analytically by Kuo (1952). The main differences from Kuo are that the wavelengths of the most unstable disturbances are slightly smaller and their growth rates are larger in the realistic case. These differ-

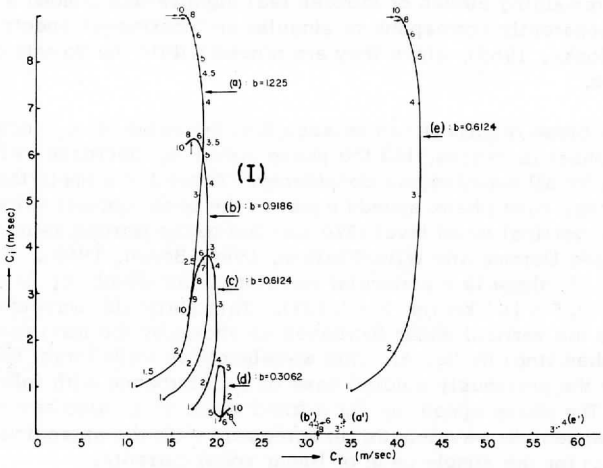


Fig. 4. Complex phase speed plane for pure baroclinic zonal currents. Numbers are zonal wavelengths L in 10^3 km. Group (I): $\bar{U} = \bar{U}_m + b\bar{U}'(p)$ for 4 values of b, M = 1, N = 15, D = 3,000 km, and $\bar{U}_m$ = 23.725 m sec^{-1}; (e): same as curve (c) of (I) except $\bar{U} = 2(\bar{U}_m + b\bar{U}'(p))$. The prime represents the corresponding secondary complex mode.

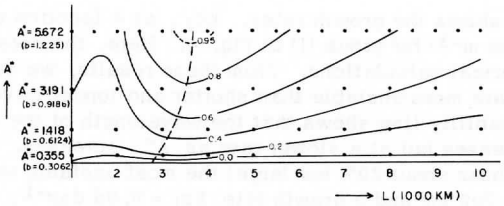


Fig. 5. Growth rate kc_i (day^{-1}) for pure baroclinic case as a function of L and A^* (kjoule m^{-3}) for the most unstable mode (group (I) and curve (e) in Fig. 4). Dots represent the locations of computed values. Dashed line shows maximum instability for fixed A^*.

ences seemed to be caused mainly by the smaller static stability parameter in the lower troposphere in this study. It is well known that the important effect of this parameter is to greatly stabilize the short wavelengths. Conversely, the effect of β is to stabilize the longer wavelengths and also to destabilize the short wavelengths (see Haltiner and Song, 1962; Brown, 1969).

It is interesting to see in Fig. 5 that a region of secondary maximum instability at very short wavelengths seems to have appeared, but we note that the truncation errors may be significant, since the value of c_i is much smaller (see the values of the error function in Appendix C).

The vertical distributions of the stream function (eigenvector) ψ' and the corresponding ω' profiles for case (a) in Fig. 4 are shown for several different wavelengths in Fig. 6. The upper part shows the normalized amplitudes (left) and phase angles (right) of ψ' and the lower part the absolute amplitudes (left) and phase angles of ω' obtained from Eq. (2.28) and (2.29) by using the computed normalized eigenvectors. It is seen that the amplitude of ψ' decreases rapidly with altitude for the shorter waves (2,000 km), but for the longer waves (6,000 km) there exists a maximum at the tropopause. Hence the long waves are pronounced at high levels, whereas the short waves are strictly a low level phenomenon.

The phases of both ψ' and ω' tilt westward with height in the troposphere and tilt eastward in the upper stratosphere; the latter leads the former by about 90°. As indicated earlier, these inclinations will give energy flow from A_Z to K_E through A_E. It should be noted that the westward tilt is larger for the longer waves, which is opposite to observed long waves in the atmosphere which are apparently forced waves, not free waves. These inclinations are also larger in the lower than the upper layer, especially in the ψ' field.

Another interesting result is that the magnitude of ω' decreases as the wavelength L increases, as expected from (2.28) and (2.29), and the level of horizontal nondivergence become slightly higher for longer waves. Note that these effects are most pronounced for the shorter waves.

Figure 7 shows the structure of the unstable perturbations (ψ' and ω') for case (a) in Fig. 4 with L = 4,000 km in the zonal cross section. Solid lines represent normalized stream function lines and the thinner dashed lines are ω' multiplied by the factor 10^{11}. It can be seen that the curves of $\omega' = 0$ lag slightly behind the troughs and ridges in the lower layer, and are a little ahead of them above 570 mb, indicating upward motions in front of ridges (southerly flow) and downward motions behind the ridges (northerly flow). This implies the upward flow of warm air, and downward motion of cold air, resulting in a positive perturbation energy conversion process. The maximum value of ω' occurs at about the 600 mb level. These results are strikingly similar to those of Kuo (1952) and Charney (1947).

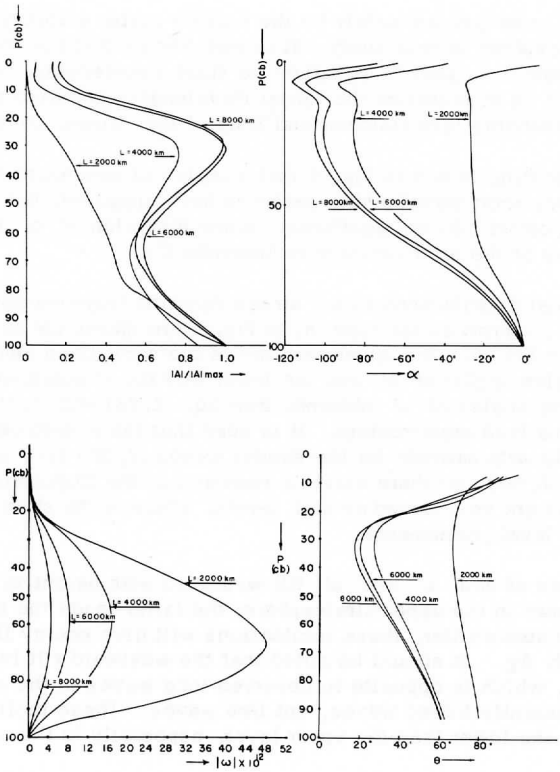


Fig. 6. Pure baroclinic case: normalized amplitudes (left) and phase angles (right) of the unstable ψ' and ω' waves as a function of pressure for several wavelengths of cases (a) and (e) in Fig. 4. Upper: stream function ψ'; lower: vertical velocity ω'.

4.2 Unstable disturbances: energy conversions

Numerical results of the total normalized energy conversion rates, integrated over the entire mass of the atmosphere, are presented in Fig. 8 as a function of wavelength L and A^*. Since the normalized eigenvectors have been used to compute the vertical velocity and energy conversions, the results show only the relative values. From these results it can be seen that two energy conversion rates, $C(A_Z, A_E)$(upper diagram) and $C(A_E, K_E)$(lower), are positive, indicating that the eddy kinetic energy of growing waves should be provided by the available zonal potential energy source through the eddy available poten-

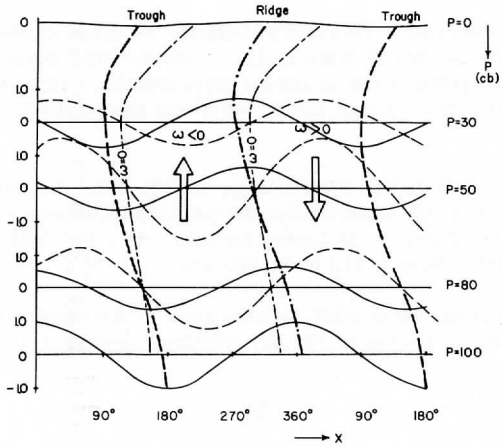


Fig. 7. Pure baroclinic case: structure of unstable disturbances in the zonal plane for cases (a) and (e) in Fig. 4 with L = 4,000 km. Solid lines are normalized stream function lines, dashed lines represent $\omega' \times 10^{11}$. Heaviest lines indicate trough and ridge lines. Lines of zero vertical motion are also indicated. Arrows indicate the sense of the vertical motions.

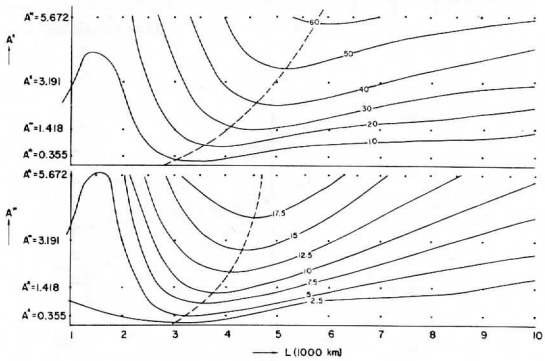


Fig. 8. Pure baroclinic case: energy conversions (in arbitrary units of 10^{-4}), integrated over the entire mass, as a function of L and A^* (kjoule m^{-3}) for the most unstable mode of Fig. 4. Upper: $C(A_Z, A_E)$; lower: $C(A_E, K_E)$. Dashed lines represent maximum conversion rates for fixed A^*.

tial energy. Furthermore, the general patterns of these conversions are similar to the distribution of growth rates in Fig. 5 except that individual maximum energy conversion rates occur at longer wavelengths. (It should be noted that $C(K_Z, K_E)$ becomes zero in this model for which the basic currents are independent of y.)

The reader is referred to Appendix C, which gives the exact numerical values of the energy conversions and the perturbation energy levels as well as the complex eigenvalues c and error function E_K, for case (a) in Fig. 4 with wavelengths ranging from 1,000 to 10,000 km.

Now we consider the spatial distributions of the energy quantities to find more useful information concerning the physical mechanism implied by each

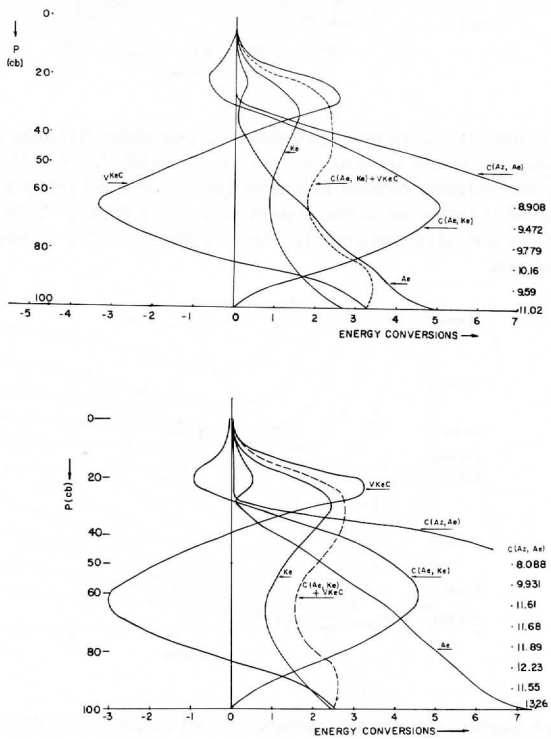


Fig. 9. Pure baroclinic case: vertical distributions of the perturbation energies (A_e and K_e, in units of 10^0) and energy conversions (10^{-5}), integrated over the horizontal plane, for cases (a) and (e) in Fig. 4. Upper: L = 4,000 km; lower: L = 6,000 km.

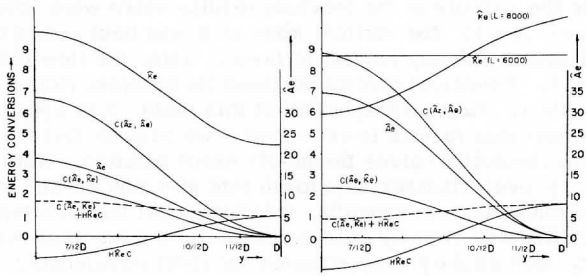


Fig. 10. Pure baroclinic case: latitudinal distributions of the perturbation energies ($\hat{A}_e$ and $\hat{K}_e$, in units of 10^{-5}) and energy conversions (10^{-9}), integrated over the zonal plane, for cases (a) and (e) in Fig. 4. Left: L = 4,000 km; right: L = 6,000 km. Also included in the right diagram is K_e for L = 8,000 km.

term. Figures 9 and 10 illustrate the vertical (integrated horizontally) and the latitudinal (integrated over x and p) distributions of the energy terms for two wavelengths (4,000 and 6,000 km) with the flow profile corresponding to case (2) in Fig. 4. Here 4,000 km corresponds to the wavelength of maximum instability and it will be seen in the next chapter that 6,000 km is about the wavelength of maximum instability for the pure barotropic case. From these results, the following conclusions are evident:

(1) The energy conversion terms for both waves, $C(A_z, A_e)$ and $C(A_e, K_e)$, are most pronounced in the lower layer near the center of the channel, as expected from the smaller values of $\sigma(p)$ in the lower layers, the larger values of the vertical shear at the center, and the form of the disturbance. Hence northward and upward heat transports are dominant in the lower layer and concentrated near the center, resulting in maximum temperature gradients below the tropopause with horizontal maxima at the center and near the boundary.

(2) The vertical and lateral flux terms, VK_eC and HK_eC, are important in determining the spatial distribution of K_e. VK_eC has two positive maxima at the ground and near the tropopause, apparently causing two maxima in the K_e distribution. The important result of the VK_eC process is to carry energy upward from the available energy source near the ground.

(3) The concentration of kinetic energy for the 6,000 km wave near the tropopause is weaker in the 4,000 km wave, which has a maximum K_e at the ground. Notice that the vertical distributions of all energy terms are slightly higher for the longer wave (6,000 km) and the values of VK_eC are somewhat larger at the tropopause for the 6,000 km wave.

To examine the validity of the previous results which were obtained by using $M = 1$ and $N = 15$, the vertical mesh size was next varied by dividing the atmosphere into a different number of layers, using the flow profile for case (a) in Fig. 4. Numerical results obtained for different values of N are presented in Table 1. From an inspection of this table, it is apparent that the accuracy of the previous results is very good if we assume that the case with $N = 41$ (maximum resolution) gives the nearly exact solution. Notice that the case with $N = 15$ overestimates the growth rate at 4,000 km and underestimates it for the remaining wavelengths, indicating that the wavelength of maximum instability was slightly underestimated in the previous results. The phase speed c_r was slightly overestimated for short wavelengths (less than 4,000 km) while the opposite was true for long waves. As indic earlier, the number of the singular modes increases as the value of N increases (to a maximum of 3 for $N = 41$) but their growth rates are very small and the eigenvalues are almost independent of wavelength.

Figure 11 shows the distribution of the eigenvector for different values of N, corresponding to the first mode for $L = 4,000$ km in Table 1. The major differences between cases of different N occur at higher levels, showing that the waves become more nearly of the shallow type as N increases. However, it is apparent that $N = 15$ produces high enough resolution to obtain the general features of the solutions to within less than 10% error. This conclusion will be applied to the numerical model in Chapter 6.

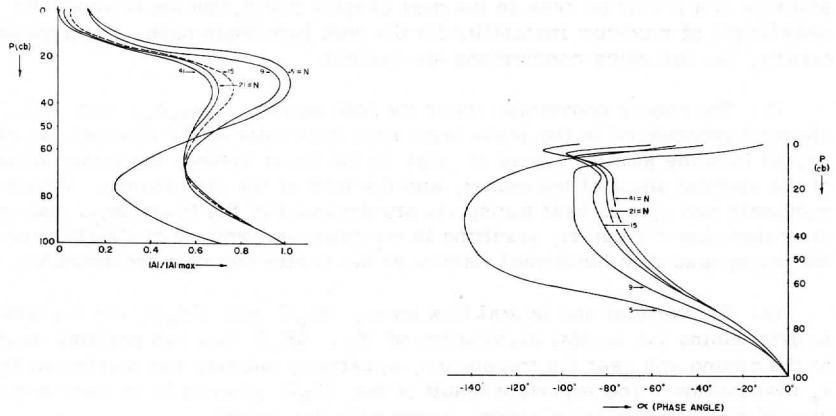


Fig. 11. Pure baroclinic case: vertical distributions of the normalized stream function amplitudes (upper) and phase angles (lower) obtained by varying the number of vertical subdivisions ($N = 5, 9, 15, 21$, and 41) for cases (a) and (e) in Fig. 4 with $L = 4,000$ km.

Table 1. A comparison of the numerical results for the pure baroclinic zonal current given by Eq. (4.1) for different values of N. Parameters: b =1.225, D = 3,000 km, and M = 1. Units: L in 10^3 km, energy conversions in 10^{-3} sec^{-1}, A_E and K_E are expressed in arbitrary nondimensional units, c and the secondary eigenvalues in m sec^{-1}, the kinetic energy error E_K in percentage, and kc_i in day^{-1}.

L (10^3 km)	N	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	(A_E, K_E) (10^{-3} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	E_K (%)	kc_i (day^{-1})	Secondary eigenvalue (m sec^{-1})
	5	3.73	1.38	99.7	127.6	17.37 ± i3.75	54.19	1.018	None
	9	1.88	0.94	58.7	83.4	17.10 ± i2.53	28.98	0.687	None
2	15	1.28	0.69	56.3	76.7	16.13 ± i1.69	16.09	0.460	None
	21	2.11	1.03	66.8	72.5	15.24 ± i2.58	12.44	0.700	39.33 ± i0.35
	41	2.12	1.09	65.6	74.0	14.90 ± i2.50	6.39	0.679	None
	5	2.86	0.40	207.5	128.4	14.93 ± i3.77	73.68	0.512	None
	9	5.69	1.74	174.2	129.2	17.68 ± i7.23	40.77	0.981	None
4	15	5.01	1.99	135.4	106.3	18.37 ± i7.13	17.06	0.969	34.01 ± i0.12
	21	4.72	1.93	127.8	99.3	18.41 ± i6.95	10.90	0.943	39.18 ± i0.44 36.35 ± i0.26
	41	4.616	1.99	122.0	96.8	18.31 ± i6.85	4.35	0.929	39.05 ± 0.056 41.35 ± 0.057
	5	1.06	0.09	218.5	104.5	13.32 ± i2.10	79.00	0.190	None
	9	4.47	0.86	210.0	102.8	16.25 ± i8.21	51.14	0.742	None
	15	6.61	1.88	261.0	131.2	17.52 ± i8.637	20.61	0.782	None
6	21	6.40	1.95	246.1	123.0	17.78 ± i8.644	12.56	0.782	39.17 ± i0.438 36.34 ± i0.27
	41	6.29	2.06	233.4	118.7	17.84 ± i8.661	4.40	0.784	39.04 ± i0.109 41.35 ± i0.077
	5					$c_i = 0$			None
	9	3.76	0.55	245.6	98.5	15.15 ± i8.31	57.05	0.564	None
	15	5.66	1.34	305.9	122.6	16.67 ± i9.00	22.91	0.611	None
8	21	6.39	1.62	333.7	131.8	17.02 ± i9.09	13.80	0.617	39.16 ± i0.436 36.35 ± i0.28
	41	6.82	1.88	343.3	136.7	17.16 ± i9.16	4.58	0.621	39.04 ± i0.13 41.34 ± i0.084
	5					$c_i = 0$			None
	9	3.38	0.42	287.1	103.5	14.39 ± i8.21	60.43	0.445	None
	15	5.15	1.09	357.0	126.3	16.06 ± i9.05	24.40	0.491	None
10	21	5.83	1.33	390.7	135.3	16.46 ± i9.17	14.65	0.498	39.16 ± i0.435 36.35 ± i0.27
	41	6.18	1.53	398.4	138.2	16.64 ± i9.27	4.74	0.503	39.04 ± i0.13 41.34 ± i0.088

5. Wave Disturbances in Pure Barotropic Zonal Currents

In this chapter the special case of the basic zonal current being independent of p, the pure barotropic case, is considered by assuming the current $\bar{U}$ possesses only horizontal shear and that the stratified atmosphere is characterized by the vertically dependent stability $\sigma(p)$. In such a model the conventional nondivergent mode of disturbances exists in barotropic form. In addition there exist baroclinic disturbances possessing horizontal divergence. These internal modes necessarily possess a vertical structure which is influenced by the stability $\sigma(p)$ (see Jacobs and Wiin-Nielsen, 1966).

The zonal current considered here is a cosine-jet profile which has been studied in the past:

$$\bar{U}(y) = \bar{U}_m(1 - a \cos \frac{2\pi y}{D}), \qquad 0 \leq y \leq D \ . \qquad (5.1)$$

Here $\bar{U}_m = 23.725$ m sec^{-1}, D is the latitudinal width of the channel and a is the variable horizontal shear parameter. Since this profile satisfies the necessary instability condition determined by Kuo (1949), it is likely that there exists an active barotropic instability mechanism in which the kinetic energy of some growing waves is converted from the available energy source, K_Z. To test these ideas, a series of numerical computations was made using the same values for most parameters as in the previous chapter (except for the basic current) in the governing Eqs. (2.17) - (2.19). Since $\bar{U}$ in (5.1) is symmetric with respect to the center of the channel, both symmetric or antisymmetric perturbations were considered. Also, results obtained for the case of symmetric perturbations in the symmetric parabolic-jet current are discussed at the end of the chapter.

5.1 Unstable symmetric disturbances: eigenvalues and structure for the "cosine jet"

The complex eigenvalues (with $c_i > 0$) obtained for several values of the shear parameter a (corresponding to dots on line (b) in Fig. 3) are shown in Fig. 12. Case (c) (a = 1.084) is the most realistic case; it possesses the same value of K^* as the cosine-jet mixed case with a = 1.0, b = 1.0 to be considered later. For this case, and using M = 7 and N = 15, there exist two pairs of complex modes: one corresponds to the nondivergent, one-layer, barotropic mode, the other corresponds to the lowest internal divergent mode and is denoted by triangles in Fig. 12. For short and intermediate wavelengths we see that the growth rate of the nondivergent mode exceeds that of the divergent mode.

Since this internal mode occurs for a prescribed stratification of the basic state of atmosphere, its structure is dependent on the vertical coordinate p (see Fig. 15). Clearly, in the finite difference system, only a certain number

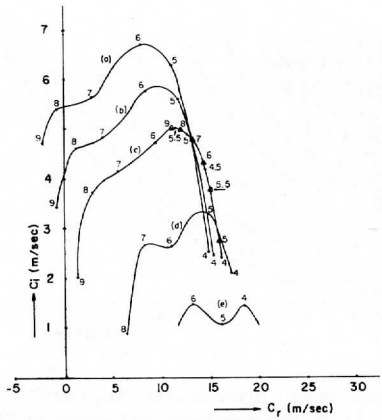


Fig. 12. Complex phase speed plane for the pure barotropic cosine-jet zonal currents given by Eq. (5.1), D = 3,000 km, $\bar{U}_m$ = 23.725 m sec^{-1}. Numbers are zonal wavelengths L in 10^3 km. Curve (a): a = 1.335; (b): a = 1.220; (c): a = 1.084; (d): a = 0.8131; (e): a = 0.5421. Dots in the triangle mark represent the lowest divergent or internal mode. M = 7 and N = 15 for curve (c); M = 7 and N = 5 for remaining curves.

of these internal modes (less than the number of grid points) can be resolved; they can be eliminated by decreasing the vertical subdivisions.

Hence, restricting ourselves to the nondivergent mode, the number of vertical subdivisions was decreased to N = 5 while the number of lateral subdivisions was held constant (M = 7) to save in computational effort. The results appear as cases (a), (b), (d) and (e). The appearance of two humps, particularly for small shears, was apparently due to truncation errors, for they were found to disappear as M was increased.

Other eigenvalues, which are not shown here, were real and corresponded to singular solutions having phase speeds equal to the current velocities somewhere in the channel. The lack of spurious unstable modes was apparently due to the use of finite differences for the computation of the variable coefficients in (2.17) (Yanai and Nitta, 1968).

Figure 13 shows the growth rate kc_i for the nondivergent mode as a function of L and K^*. The marginal instability curve ($kc_i = 0$) was determined analytically in a manner similar to Kuo (1949). See details in Appendix B. The numerical results for the position of the marginal curve are very close to the exact values, particularly for the critical shortwave cut-off at 3,470 km (see

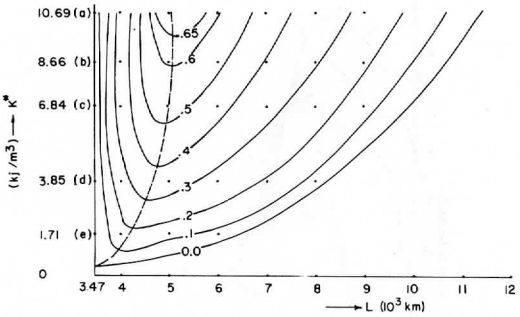


Fig. 13. Pure barotropic case: growth rate kc_i (day^{-1}) as a function of L and K^* (kjoule m^{-3}) for the nondivergent mode in Fig. 12. Dots represent the locations of computations. Dashed line shows maximum instability for fixed K^*. Marginal stability curve ($kc_i = 0$) is obtained theoretically.

Yanai and Nitta, 1968). The results in Figs. 12 and 13 are similar in appearance to those for the case of "pure baroclinic" instability, but the following differences are noticeable:

(1) the growth rates are much larger for the pure baroclinic current than for the pure barotropic one;

(2) the wavelength of maximum instability is shorter for the pure baroclinic case;

(3) for the pure baroclinic case, the phase speed is larger, especially for longer waves, and its zonal wavelength dispersion is less than for the barotropic case;

(4) the short waves ($L \leq 3,470$ km for $D = 3,000$ km) are barotropically stable ($c_i = 0$).

Figure 14 shows the latitudinal distribution of the normalized amplitudes (curves (a)) and phase angles (curves (b)) of ψ' for the nondivergent mode of case (c) in Fig. 12 with several wavelengths. It is seen that the amplitudes monotonically decrease and the waves tilt eastward with increasing latitude. This tilt becomes larger near the boundary as L increases.

The structure of the lowest internal mode (at the jet center $y = D/2$) for the same case as Fig. 14 (for $L = 6,000$ km) is indicated in Fig. 15. It can

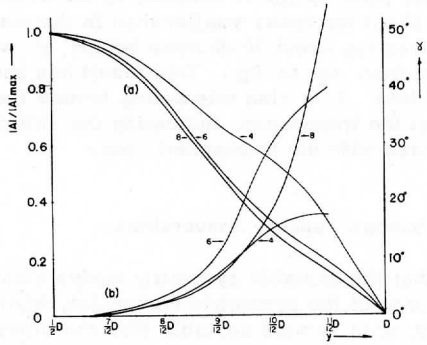


Fig. 14. Pure barotropic case: normalized amplitudes (curves (a)) and phase angles (curves (b)) of the stream function as a function of latitude for the unstable nondivergent mode of case (c) in Fig. 12. Numbers represent zonal wavelengths L in 10^3 km.

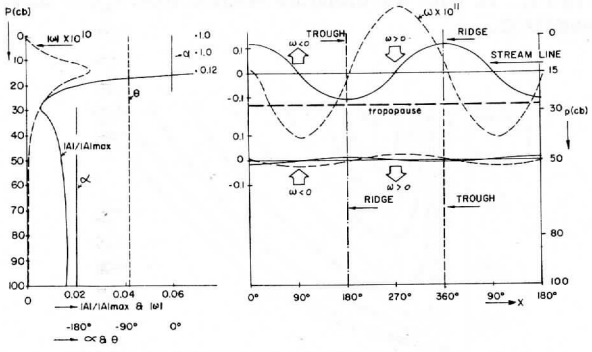


Fig. 15. Pure barotropic zonal current: structure of the unstable disturbances for the divergent mode of case (c) in Fig. 12 with L = 6,000 km. Left: vertical distributions of ψ' and ω'; right: structure in the zonal plane. Arrows indicate the sense of the vertical motions.

be seen that the major perturbation is confined to the stratosphere and the magnitude of ω' is about ten times smaller than in the pure baroclinic case. Furthermore, ω' waves lag about 90 degrees behind ψ' waves, resulting in an energy conversion from K_E to A_E. This result has been shown by Jacobs and Wiin-Nielsen (1966). It is also interesting to note the discontinuity in phase angle of ψ' at the tropopause, indicating the stratospheric wave is completely out of phase with the tropospheric one.

5.2 Unstable disturbances: energy conversions

We have seen that the unstable symmetric modes exist for intermediate wavelengths in the case of the symmetric cosine-jet, barotropic current and that the nondivergent mode is more unstable than the internal modes. Hence it is of special interest to study the energetics of the nondivergent unstable disturbances.

Figure 16 shows the total energy conversion $C(K_Z, K_E)$ as a function of L and K*. It is strikingly similar to the growth rates shown in Fig. 13 and shows that the perturbations grow at the expense of K_Z, as expected from the eastward tilt of ψ'. This implies northward (southward) eddy momentum transport north (south) of the center of the channel (y = D/2).

The exact numerical values of the energy conversions and the perturbation energy levels as well as the complex eigenvalues c and error function E_K, for case (c) in Fig. 12 with wavelengths ranging from 2,000 to 9,000 km, are given in Appendix C.

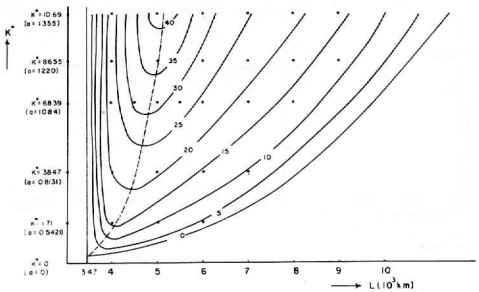


Fig. 16. Pure barotropic case: kinetic energy conversion, $C(K_Z, K_E)$ (in units of 10^{-4}), integrated over the entire mass, as a function of L and K* (kjoule m^{-3}) for the nondivergent mode given in Fig. 12. Dashed line shows locus of maximum conversion rates for fixed K*.

More information can be seen from Fig. 17, which presents latitudinal distributions (integrated over the zonal plane) of the energy terms for two wavelengths (4,000, 6,000 km) of case (c) in Fig. 12. From these results it can be seen that:

(1) $C(K_z, K_e)$ is positive near the boundary. Therefore, an initial single-jet current tends to be transformed into a double-jet current as a result of the meridional convergence of the Reynolds stresses (see Wiin-Nielsen, 1961; Haltiner and Song, 1962).

(2) HK_eC is important in determining the distribution of K_e and opposes the process $C(K_z, K_e)$.

(3) K_e has a maximum at the center for the shortwave (4,000 km); for the longwave (6,000 km) the maximum is shifted away from the center.

Decreasing the lateral mesh size and keeping $N = 5$, a numerical test of the validity of the previous results was made by considering one case (case (c) in Fig. 12) for several wavelengths. Table 2 and Fig. 18 show these results. From an inspection of these results it is apparent that the results with $M = 7$ are quite good. However, it can be noted that the case with $M = 7$ slightly underestimates the growth rates for short wavelengths (less than 6,000 km) and slightly overestimates them for the longwave (8,000 km), eliminating the secondary "hump." Very recently, Yanai and Nitta (1968) showed that in order to obtain sufficiently accurate eigenvalues, at least 20 subdivisions from $y = 0$ to $Y = d$ (corresponding to $M = 11$ in this study) were required.

5.3 Stable disturbances: antisymmetric perturbations of the cosine jet and symmetric perturbations of the parabolic jet

For antisymmetric perturbations with respect to the center of the channel, with the cosine-jet current given by (5.1) and the shear parameter $a = 1.084$, it is found numerically ($M = 7$, $N = 15$) that all eigenvalues for $L = 4,000$ and 6,000 km are real and lie within the range of $\bar{U}$, indicating stable perturbations. Eliassen (1954) obtained a similar result for the nondivergent barotropic instability problem by utilizing the truncated Fourier series method.

As a special case of the parabolic-jet mixed case, the following current profile is considered:

$$\bar{U} = 2\bar{U}_m\left(1 - a\left(\frac{2y}{D}\right)^2\right), \qquad |y| \leq \frac{D}{2} \tag{5.2}$$

where $\bar{U}_m = 23.725$ m/sec, $D = 3,000$ km. The value of the shear parameter a is chosen as 1.06 to give the same amount of K^* (7.0 kjoule m^{-3}) as in the corresponding mixed case. Notice that this current profile has a constant second derivative and hence no instability can occur for symmetric perturbations,

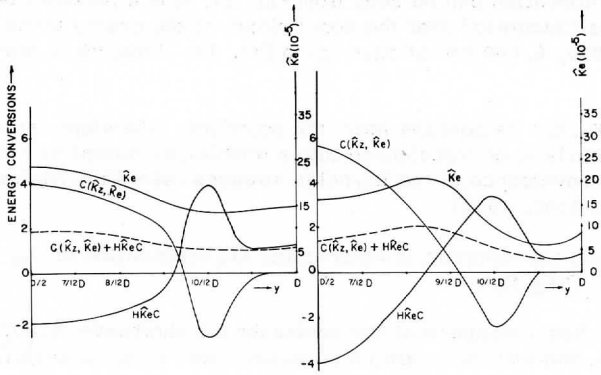


Fig. 17. Pure barotropic case: latitudinal distributions of the perturbation kinetic energy (in units of 10^{-5}) and the energy conversions ($C(\hat{K}_z, \hat{K}_e)$ and $H\hat{K}_eC$ in 10^{-9}), integrated over x and p, for the nondivergent mode of case (c) in Fig. 12. Left: L = 4,000 km; right, L = 6,000 km.

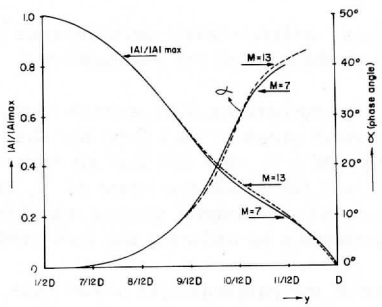


Fig. 18. Pure barotropic case: comparison of the normalized stream function amplitudes and phase angles obtained with M = 7 and M = 13 for the nondivergent mode of case (c) in Fig. 12.

Table 2. A comparison of the numerical results for the pure barotropic cosine-jet current given by equation (5.1) for different values of M. Parameters: a = 1.084, D = 3,000 km, and N = 5. Units: See Table 1.

l (10^3 km)	M	$C(K_Z, K_E)$ (10^{-3} sec^{-1})	K_E	$c = c_r + ic_i$ (m sec^{-1})	E_K (%)	kc_i (day^{-1})
4	5			$c_i = 0$		
	7	2.37	271.4	16.02 ± i2.386	-16.48	0.324
	9	1.93	260.8	16.82 ± i2.814	16.21	0.382
	13	1.60	250.1	17.45 ± i2.435	16.34	0.330
6	5	2.75	215.5	9.12 ± i5.635	- 8.04	0.510
	7	2.42	218.8	9.64 ± i4.702	-12.26	0.426
	9	2.35	225.0	8.76 ± i4.908	- 1.59	0.444
	13	2.39	226.8	8.86 ± i4.920	- 2.56	0.445
8	5	1.36	213.5	4.74 ± i3.625	-11.84	0.246
	7	1.43	224.0	2.97 ± i3.725	- 7.31	0.253
	9	1.26	233.0	3.04 ± i3.283	- 4.74	0.223
	13	1.31	237.5	2.88 ± i3.496	- 0.73	0.237

according to Kuo (1949). Numerical eigenvalues for $L = 4,000$ and $6,000$ km, with $M = 7$ and $N = 15$, considering symmetric perturbations only, are all found to be real and lie within the range of $\bar{U}$.

Results for the case of antisymmetric disturbances of the parabolic jet have not been obtained. However, we may conjecture that since they correspond to shorter wavelengths in y, they should be stable when the symmetric disturbances of all wavelengths are stable; i.e., in view of the above numerical results, they should always be stable.

6. Wave Disturbances in Zonal Jets: The Mixed Case

In the previous chapters, two idealized cases were considered, each in which the basic current varied in only one space dimension. From these results we tentatively concluded that the development of disturbances through baroclinic processes is more important than through barotropic processes in the troposphere. However, it may be expected that some changes will occur when we consider the stability problem for more realistic "jet-like" zonal currents. In this chapter, we consider currents containing both vertical and latitudinal shear which are symmetric with respect to the center of the channel. To examine the modification of the barotropic instability mechanism, we consider two different horizontal flow profiles: the cosine-jet (barotropically conditionally unstable) and the parabolic-jet (barotropically stable). Moreover, both symmetric and antisymmetric perturbations for cosine-jet currents (but only symmetric perturbations for parabolic-jets) will be studied.

6.1 Unstable symmetric disturbances: eigenvalues and structure for the "cosine jet"

The basic zonal currents considered here are latitudinally the cosine, single-jet profiles and vertically representative of those observed:

$$\bar{U}(y, p) = (1 - a \cos \frac{2\pi}{D} y)(\bar{U}_m + b\, \bar{U}'(p)), \qquad (6.1)$$

where a and b are the variable shear parameters. This profile is seen to be a combined form of the pure baroclinic and the pure barotropic cases discussed previously. Note that the necessary condition for barotropic instability is fulfilled in these profiles when a is large, since then the meridional gradient of absolute vorticity vanishes somewhere in the channel. Hence we may expect two primary complex modes, one related to each category of zonal energy sources. Similar profiles were used by Brown (1969).

As indicated earlier, the parameters a and b are carefully varied to give the following sequence of "experiments." First, a fixed value of A^* (5.672

kjoule m^{-3}) is chosen and K^* is increased from the pure baroclinic case to the "mixed case" (a = 1.0, b = 1.0 in Fig. 3). A fixed value of K^* (6.839 kjoule m^{-3}) is then used with decreasing values of A^* from this point to the pure barotropic case (see curve (c) in Fig. 3). In this manner results are obtained which are analogous to the "pure" cases discussed earlier, the difference being that both forms of available zonal energy are generally present. The relation between these cases is shown in Fig. 3.

The numerical results were obtained by using the grid resolution corresponding to M = 7 and N = 15; these values were found to give sufficiently accurate results in Chapters 4 and 5. Other parameters assumed the same values as in previous cases.

Numerical results of the complex eigenvalues for several values of a and b are shown in Figs. 19, 20 and 21. The most interesting feature of these results is that there exist two maximally unstable modes (curves (a) and (b)) which are individually related to each available energy source type. c_i for mode (a) decreases sharply as A^* decreases. On the other hand, c_i for mode (b) increases as K^* increases, particularly for relatively short waves; also its pattern is quite similar to that of the nondivergent pure barotropic mode (see Fig. 12). As shown before, mode (a) has a larger phase speed (and smaller dispersion with respect to L) than mode (b). This dispersion decreases as A^* decreases for mode (a); conversely, for mode (b), the dispersion increases as K^* increases. Therefore, it seems natural to name modes (a) and (b) as the baroclinic and barotropic modes, respectively, keeping in mind that each mode is modified by both kinds of shear in the zonal current.

Among the other (less unstable) modes shown in these figures, modes (d) and (d') correspond to the divergent barotropic (internal) modes and their c_i values decrease sharply with increasing A^*. Modes (c) and (c') correspond to the higher baroclinic internal modes which were eliminated in the pure baroclinic case by assuming a single lateral wave number ($l = \pi/D$).

The remaining unstable modes (which vary in number from zero to seven) possess very small growth rates and phase speeds which lie within the range of $\bar{U}$. It thus seems appropriate to label these as being "quasi-singular" modes; in some cases they appear to be highly internal but physically real modes. Otherwise their origin would appear to be of the same uncertain nature as the analogous modes found in the "pure baroclinic" case.

The growth rates kc_i of the two primary modes are presented in Fig. 22 as a function of L and K^* or A^*. The dots represent the locations of computations from which the isolines of kc_i have been drawn. The upper diagram represents the baroclinic mode and the lower represents the barotropic mode. From these results, the following features are evident:

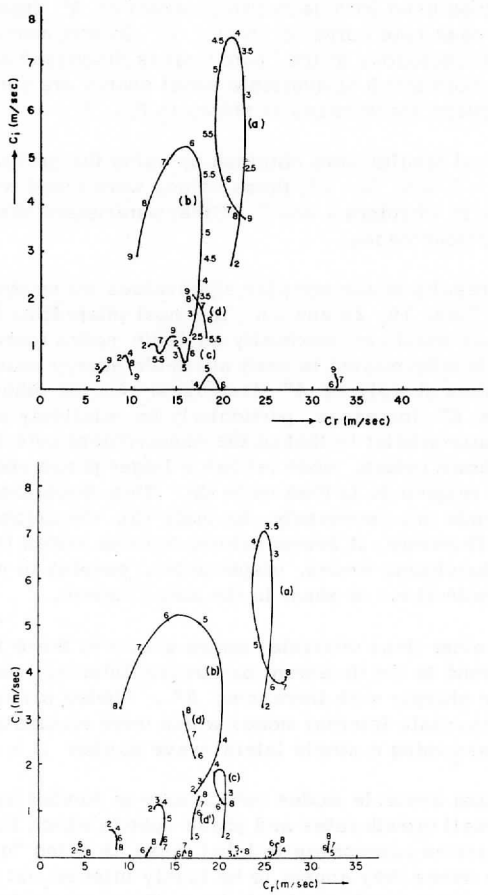


Fig. 19. Complex phase speed plane for symmetric perturbations in the cosine-jet mixed zonal current given by Eq. (6.1), M = 7, N = 15. Upper: a = 0.5 and b = 1.155; lower: a = 0.75 and b = 1.082. Numbers are L in 10^3 km. Curve (a): baroclinic mode; (b): nondivergent barotropic mode; (c) second internal baroclinic mode; (d) and (d'): first and second internal barotropic modes, respectively.

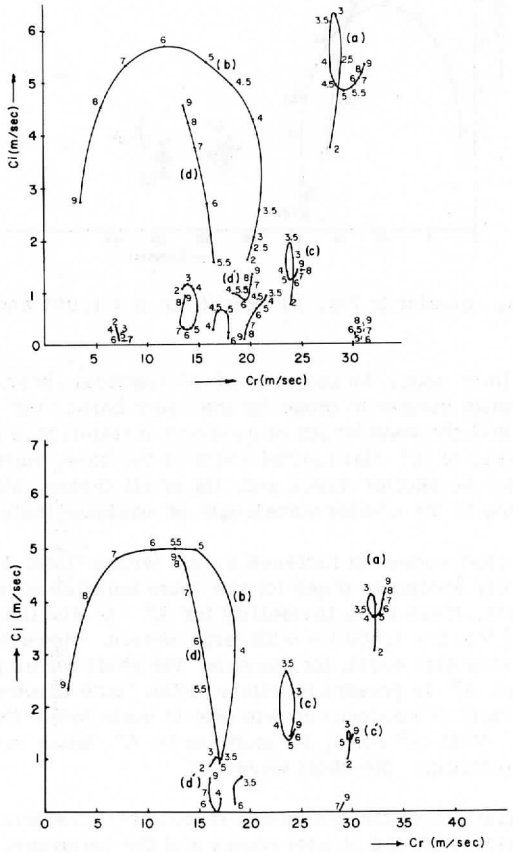


Fig. 20. Similar to Fig. 19 except for upper: a = 1.0 and b = 1.0; lower: a = 1.034 and b = 0.75. Curve (c') is a higher internal baroclinic mode.

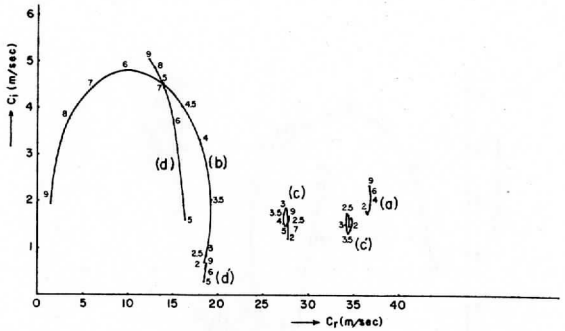


Fig. 21. Similar to Fig. 19 except for a = 1.061 and b = 0.5.

(1) Baroclinic mode: An increase of A^* (vertical shear) with a fixed K^* gives results quite similar to those for the "pure baroclinic" case shown in Fig. 5 except that the wavelength of maximum instability is shorter by about 50%. An increase of K^* (latitudinal shear of the basic current) with a fixed A^* destabilizes the shorter waves and, for small shears, stabilizes the longer waves, resulting in the shorter wavelength of maximum instability.

(2) Barotropic mode: An increase in K^* with a fixed A^* gives results which are broadly similar to those for the "pure barotropic" case shown in Fig. 13. For example, there is no instability for $K^* = 0$ and the maximum growth rates are found for L = 4,800 km with large shears. However, differences between these cases also exist; for example, the short waves possess significant instability when A^* is present (contrary to the "pure barotropic" case $A^* = 0$) and the wavelength of maximum growth rate is quite large (5,800 km) for small values of K^*. With K^* fixed, an increase in A^* tends to destabilize all wavelengths, including the short waves.

(3) A comparison of the two modes reveals that the baroclinic mode is dominantly unstable for the shorter waves and the barotropic mode for the longer waves. The region enclosed by the thick dashed lines in the lower diagram denotes cases where the barotropic mode grows faster than the baroclinic mode.

To obtain more information concerning the differences between the two primary modes, the structures of the perturbations (ψ' and ω') for two wavelengths (4,000 and 6,000 km) of the "jet" case with a = 1.0 and b = 1.0 are considered. The structures in the meridional plane for the 6,000 km wavelength are presented in Figs. 23, 24, 25 and 26. The results for the wavelength 4,000 km are shown in various zonal profiles in Figs. 27 and 28. From these results it can be seen that the two modes possess quite different structures, indicating

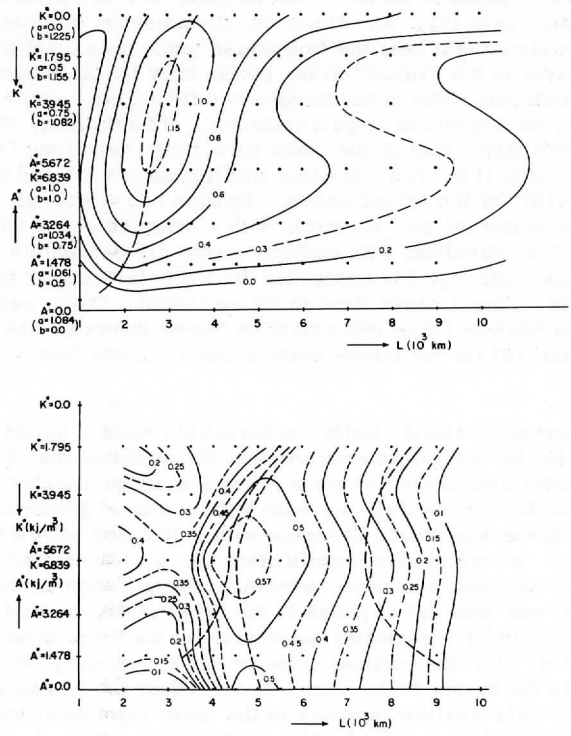


Fig. 22. Mixed case: growth rate kc_i (day^{-1}) as a function of L and K^* or A^* (kjoule m^{-3}). Upper: baroclinic mode; lower: barotropic mode. A^* has the constant value 5.672 in the upper half of the vertical axis; $K^* = 6.839$ in the lower half. Inside of thick dashed lines in the lower diagram, the growth rate of barotropic mode exceeds that of baroclinic mode.

different energetic properties as well. An inspection of these results and a comparison with Figs. 6 and 14 reveals the following interesting results:

(1) The stream function field: (i) baroclinic mode (Fig. 23): the primary maximum of amplitude is located at the ground (at $y = D/2$) with a secondary maximum at the tropopause for both wavelengths, in rough agreement with the "pure baroclinic" case (Fig. 6). However, the presence of large horizontal shear of the basic current near the tropopause tends to restrict this mode more to the lower layer in the "mixed" case, particularly for the longer waves. The vertical and horizontal tilts of the troughs and ridges are westward with height and latitude in the regions of large amplitudes. Interestingly, the vertical tilt is about 50% larger than in the "pure baroclinic" case (see Fig. 6). (ii) barotropic mode (Fig. 25): maximum amplitudes are located near the tropopause, especially for the longer waves. Troughs and ridges tilt westward with height, mainly below the 800 mb level, with a sharp decrease with increasing wavelength. The latitudinal tilts of these waves are eastward as in the "pure barotropic" case (see Fig. 14) except for the lower layers near the boundary where baroclinic effects cause them to tilt westward. These westward tilts in the lower layer become larger and extend to higher layers for the shorter wave. Also the vertical tilt for the longer wave is much smaller than for the former mode.

(2) The vertical velocity field: (i) baroclinic mode (Fig. 24): two maxima of vertical velocity, the primary maximum at the center of the channel ($y = D/2$), are located around the 650 mb level at a position slightly lower than in the "pure baroclinic" case (Fig. 6), with larger vertical gradients in the lower layer. Again these magnitudes decrease with increasing wavelength. The phase of the ω' waves is very complicated; the ω' waves lead the stream-function waves by about a quarter-wavelength at the center in the lower layer, but both waves are more out of phase in the upper layer, restricting the positive perturbation energy converting mechanism to the lower layer. (ii) barotropic mode (Fig. 26): the vertical velocity has a maximum, which is much smaller than in the former mode, near 600 mb and on the blanks of the jet for $L = 6,000$ km. This maximum appears in the lower layer and closer to the center of the channel ($y = D/2$) for the shorter wave. The ω' waves lead the stream-function wave by about 90 degrees near the location of the maximum vertical velocity.

From these results it may be concluded that the baroclinic processes for instability are dominant in the lower layers and the barotropic processes are most pronounced in the upper layers, particularly for the relatively longer waves. Summary profiles of the x, y, and p structure of the modes are contained in Figs. 27 and 28. The implications of these facts become clear by considering the energetics in the next section.

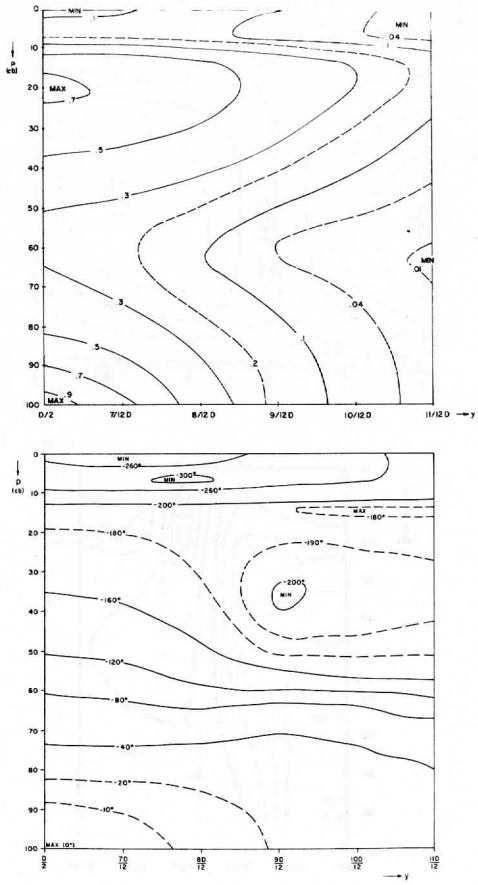


Fig. 23. Distributions of normalized amplitudes (upper) and phase angles (lower) of the stream function in the meridional cross section for the baroclinic mode (curve (a) in the upper diagram of Fig. 20) with L = 6, 000 km.

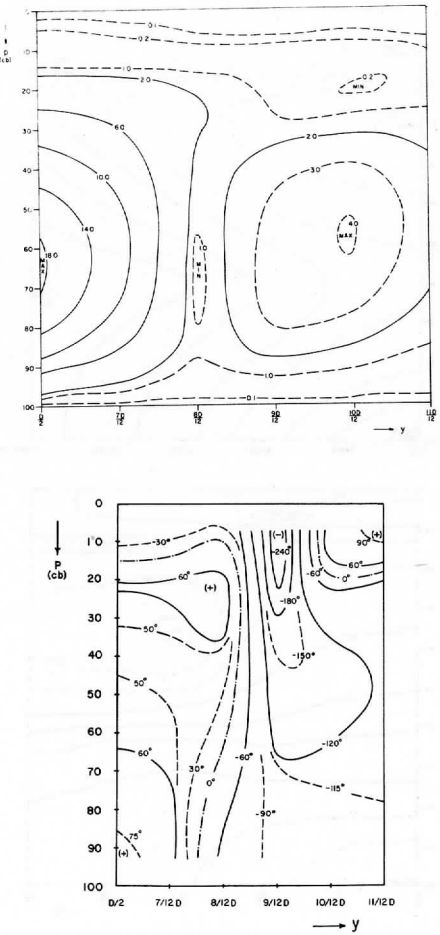


Fig. 24. Similar to Fig. 23 except for vertical velocity ω' (amplitude in arbitrary units of 10^{12}).

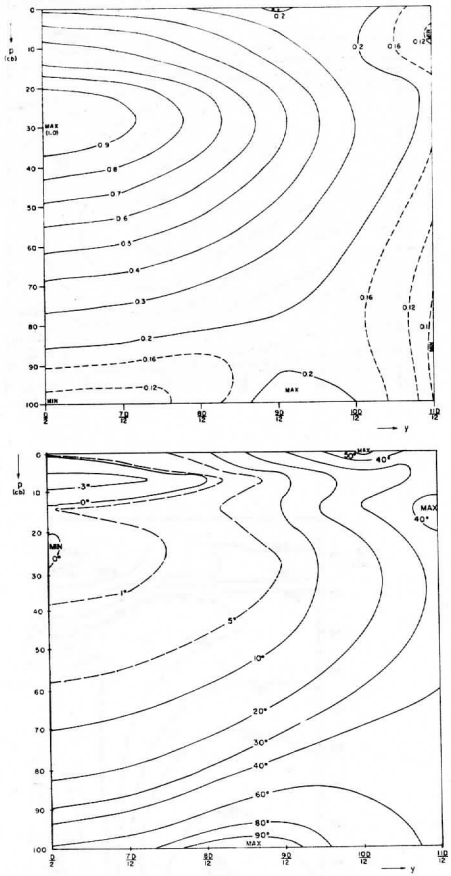


Fig. 25. Similar to Fig. 23 except for the barotropic mode.

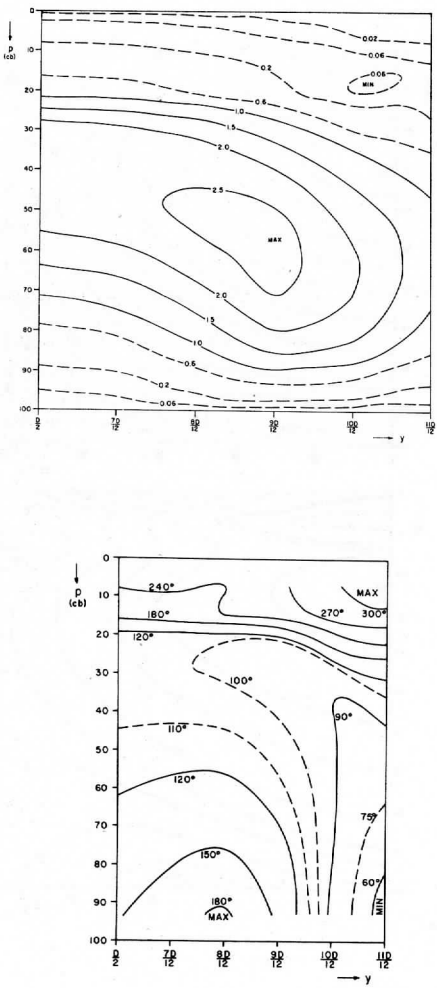


Fig. 26. Similar to Fig. 24 except for the barotropic mode.

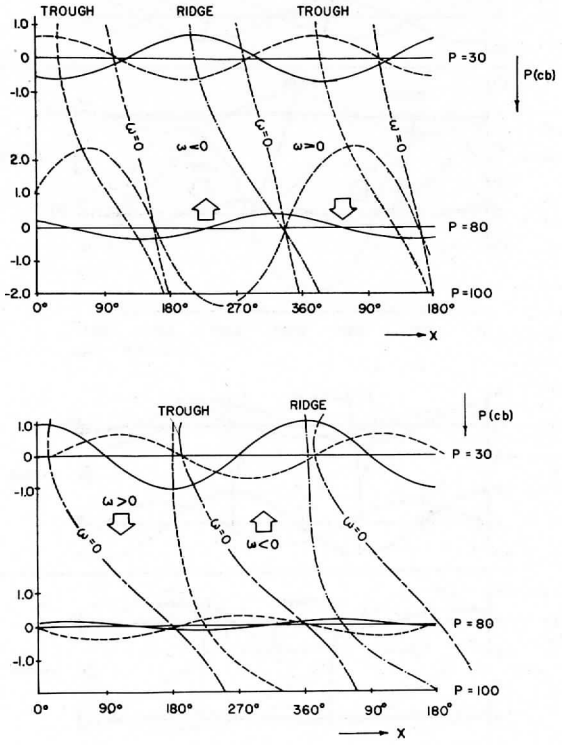


Fig. 27. Structure of unstable disturbances in the zonal plane at $y = D/2$ for cases (a) and (b) in the upper diagram of Fig. 20 with $L = 4,000$ km. Upper: baroclinic mode; lower: barotropic mode. Solid lines are normalized stream function lines and dashed wave lines are ω' waves (in units of 10^{11}).

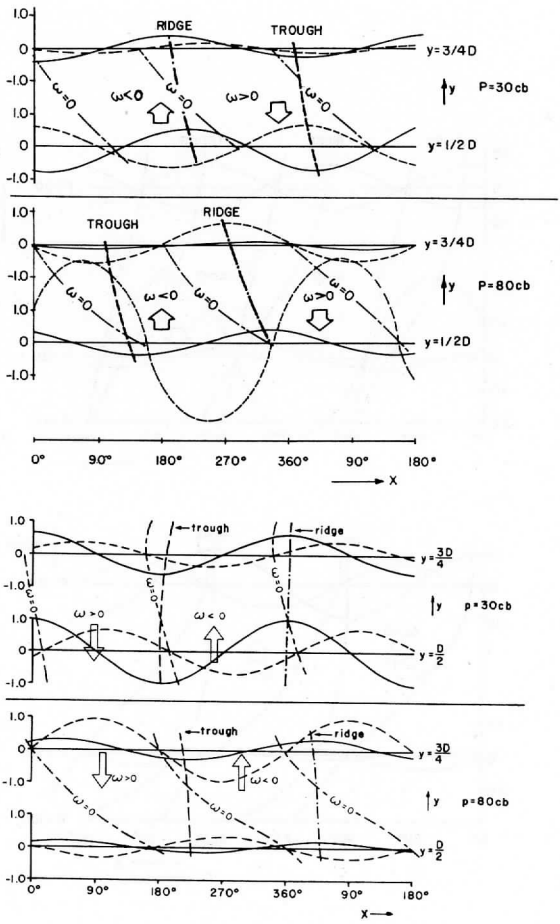


Fig. 28. Similar to Fig. 27 except for the horizontal plane at two different vertical levels (p = 30 and p =80cb). Upper: baroclinic mode; lower: barotropic mode.

6.2 Unstable symmetric disturbances: energy conversions

We now consider the energetic properties of the unstable disturbances, restricting ourselves to the aforementioned two primary modes (modes (a) and (b) in Figs. 19, 20, 21).

The potential energy conversion $C(A_Z, A_E)$ and the perturbation energy conversion $C(A_E, K_E)$ for the baroclinic and the barotropic modes are shown in Figs. 29 and 30, respectively, as a function of L and K^* or A^*. Curiously, both conversions in these figures have profiles qualitatively similar to the growth-rate distribution for the baroclinic mode (the upper diagram in Fig. 22). However, the wavelengths of maximum conversions are slightly larger than those of maximum growth rate and larger for the barotropic mode than the baroclinic mode. However, both modes have positive values for all cases considered here, with smaller values for the barotropic mode, indicating that in either mode these conversions are mainly determined by the pure baroclinic nature of the basic current. Another interesting result is that an increase of K^* with a fixed A^* tends to decrease the maximum conversion rate for the baroclinic mode, implying a damping effect by the horizontal shear of the current.

Similarly, the kinetic energy conversions $C(K_Z, K_E)$ are shown in Fig. 31. Here a positive sign indicates energy flow from K_Z to K_E. In contrast to $C(A_Z, A_E)$, this conversion changes its direction according to the relative magnitudes of A^* and K^* and the wavelength L. The baroclinic mode converts K_E into K_Z in a pronounced manner for relatively short waves, except for smaller values of A^* (1.478 kjoule m^{-3}), when the barotropic process dominates.

For the barotropic mode, a negative maximum induced by the baroclinicity is noticeable for relatively small values of K^* and wavelengths, where instability is absent in the "pure barotropic" case. It is important to note that the barotropic mode dominates this conversion for longer waves, particularly for larger values of K^*. On the other hand, the baroclinic mode dominates for short waves (see the thick dashed line in the lower diagram).

Energetic results were also obtained for the "divergent barotropic mode" (curve (d) in Figs. 19-21) whose growth rates were quite significant for the longer waves. This mode is characterized by a very small amount of available potential energy; the associated conversions are likewise small, with $C(A_E, K_E)$ being negative. $C(K_Z, K_E)$ is of course then positive and moderately large.

To gain more physical insight into these energy conversion processes, we now consider the vertical and latitudinal distributions for two representative wavelengths (4,000, 6,000 km) in the "mixed" case with $a = 1.0$, $b = 1.0$. These results are shown in Figs. 32 and 33 for the vertical distributions, and in Fig. 34 for the latitudinal distributions.

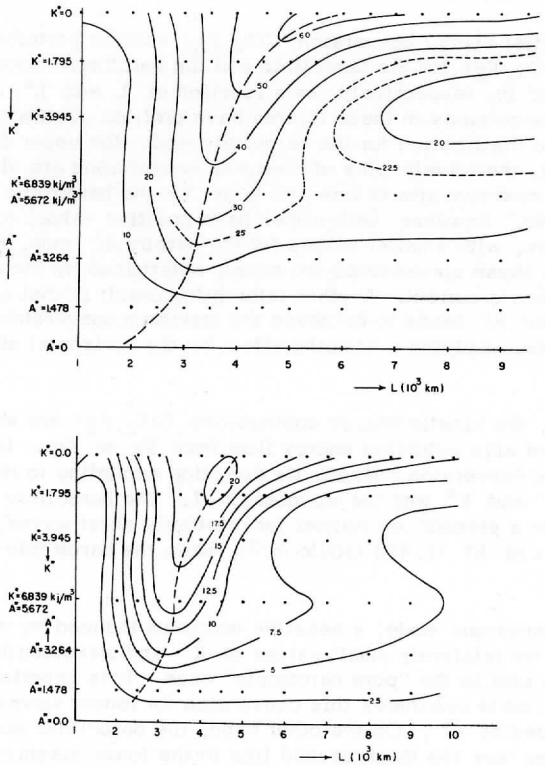


Fig. 29. Energy conversions integrated over the entire mass (in units of 10^{-4}) as a function of L and K^* or A^* for the baroclinic mode. Upper: $C(A_Z, A_E)$; lower: $C(A_E, K_E)$. $A^* = 5.672$ in the upper half of the vertical axis; $K^* = 6.839$ in the lower half.

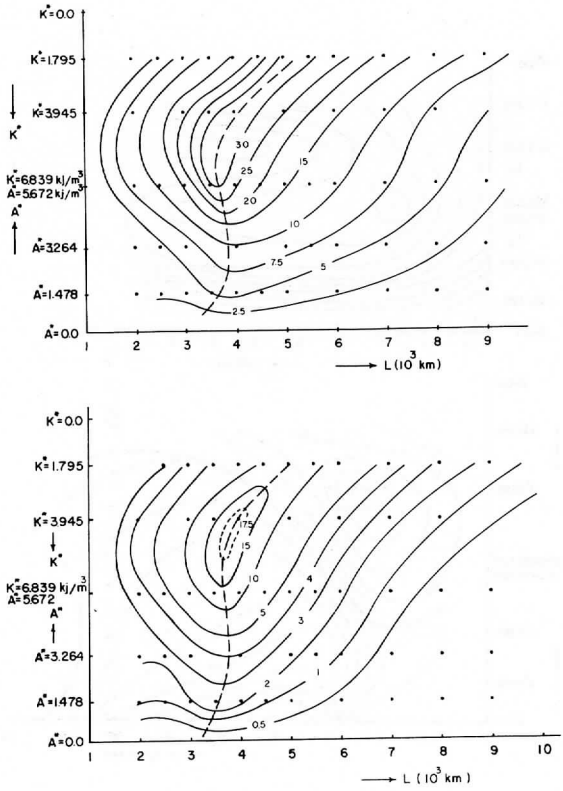


Fig. 30. Similar to Fig. 29 except for the barotropic mode.

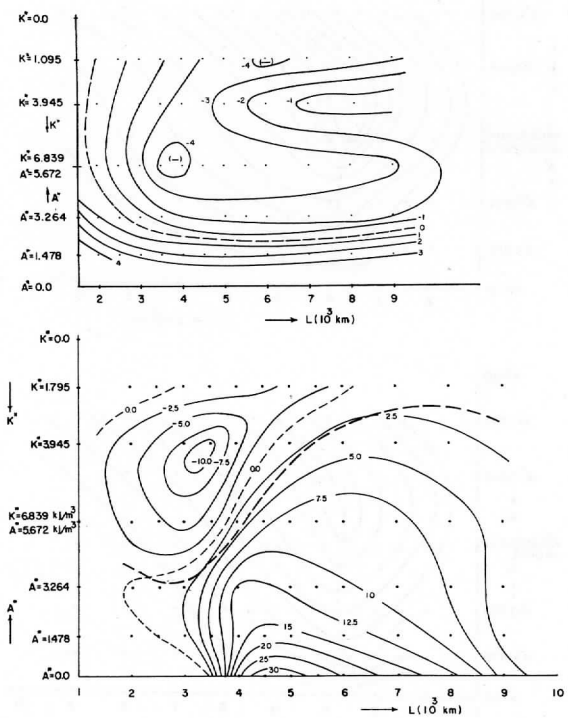


Fig. 31. Similar to Figs. 29 and 30 except for $C(K_Z, K_E)$. Upper: baroclinic mode; lower: barotropic mode. Positive sign denotes energy conversion from K_Z to K_E. Below the thick dashed line in the lower diagram, the conversion rate for the barotropic mode is greater than that of the baroclinic mode.

From these results it may be noted that the baroclinic mode has generally the same features for two different wavelengths. However, this is not true for the barotropic mode, indicating that this mode is significantly affected by zonal baroclinicity for the shorter waves. More interesting results are summarized as follows:

(1) Baroclinic mode: the vertical distributions shown in the upper parts of Figs. 32 and 33 are generally similar to those in Fig. 9 (pure baroclinic case), but the former become more of a shallow wave phenomena than the latter due to the larger damping effects of the horizontal shear in the upper layer. $C(K_z, K_e)$ has negative values with a maximum near the tropopause and thus causes a strengthening of the initial jet. VK_eC is important for both waves; its important role is to transmit the effects of the available energy sources in the lower layers to the upper layers and thus act as a main energy source for larger K_e values and negative $C(K_z, K_e)$ near the tropopause. The latitudinal distributions in the right part of Fig. 34 show negative values of $C(K_z, K_e)$ near the center and the boundary of the channel. This implies, from the convergence of eddy momentum transports, that the initial jet at the center will become stronger and two secondary jets will appear near the boundaries by unstable disturbances. It is also noted that both northward and upward sensible heat transports are concentrated near the center and HK_eC is less significant for the 6,000 km wave.

(2) Barotropic mode: the majority of K_e is located near the tropopause (due to larger values of $C(K_z, K_e)$) with a concentration near the center for both waves. However, for the 6,000 km wave, notice that $C(A_e, K_e)$ and VK_eC are negligible and A_e and $C(A_z, A_e)$ are evenly distributed below the 500 mb level, with much smaller values in the case of the 6,000 km wave. In contrast to the baroclinic mode, $C(K_z, K_e)$ has a positive maximum near the tropopause (thus decreasing the initial jet) but this term becomes negative near the ground for the 4,000 km wave. The latitudinal distributions for the 6,000 km wave in Fig. 34 show that $C(K_z, K_e)$ is positive near the center and negative near the boundary, indicating the tendency of the transformation of the initial single jet (latitudinally) into the double jet. It is also noticeable from Fig. 34 that the northward heat transports are concentrated near the center, and HK_eC is important in determining the latitudinal distribution of K_e for both waves.

Considering the change of the zonal current as a result of the meridional convergence of the Reynolds stresses, it is important to note that the baroclinic mode tends to increase zonal currents near the tropopause, with concentrations at the center of the channel and near the northern boundary, while the barotropic mode has a tendency to decrease the initial currents at the center and increase them in the region near the boundary, mainly in the upper layer. Hence a probable explanation for the maintenance of the zonal jet against frictional dissipation seems to center on the baroclinic mode. Similar results for the most unstable mode were found by Brown (1969).

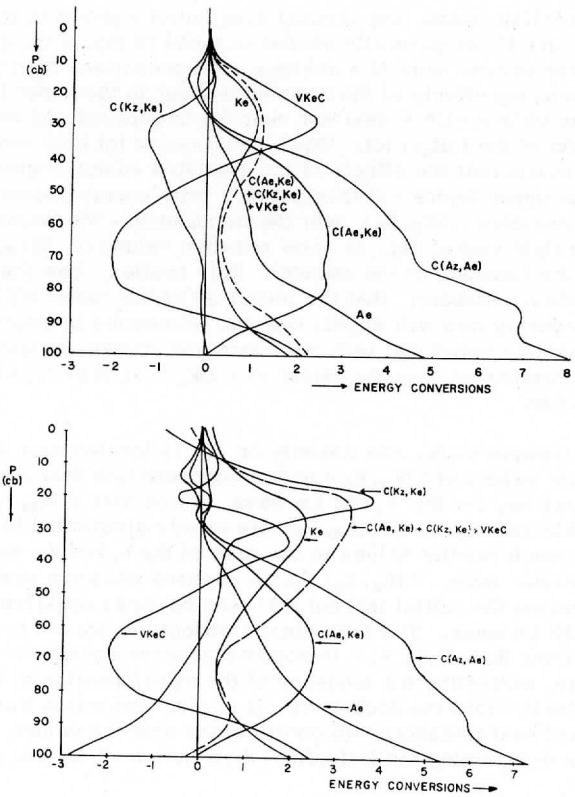


Fig. 32. Energy conversions (10^{-5}) and perturbation energies (10^{0}), integrated over x and y, as a function of pressure for the case with a = 1.0, b = 1.0, L = 4,000 km. Upper: baroclinic mode; lower: barotropic mode.

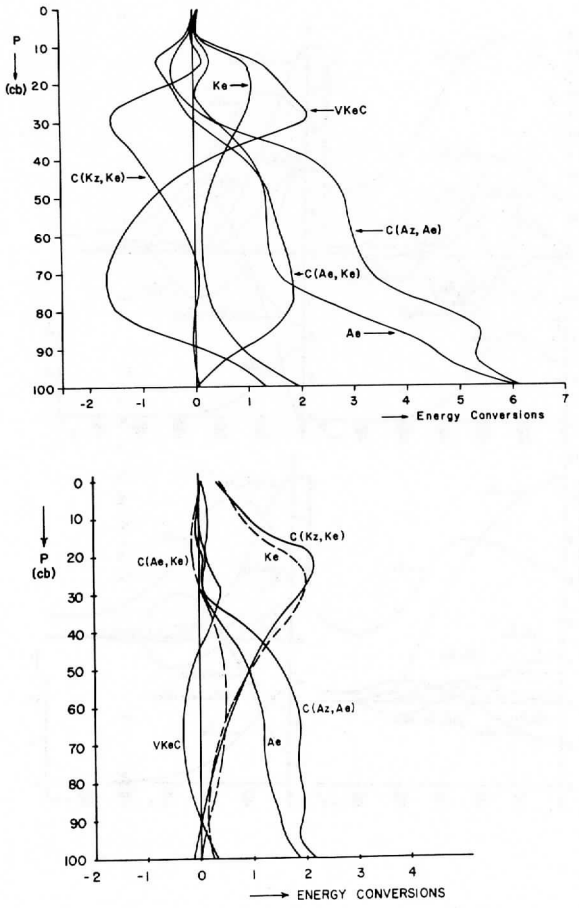


Fig. 33. Similar to Fig. 32 except for L = 6,000 km.

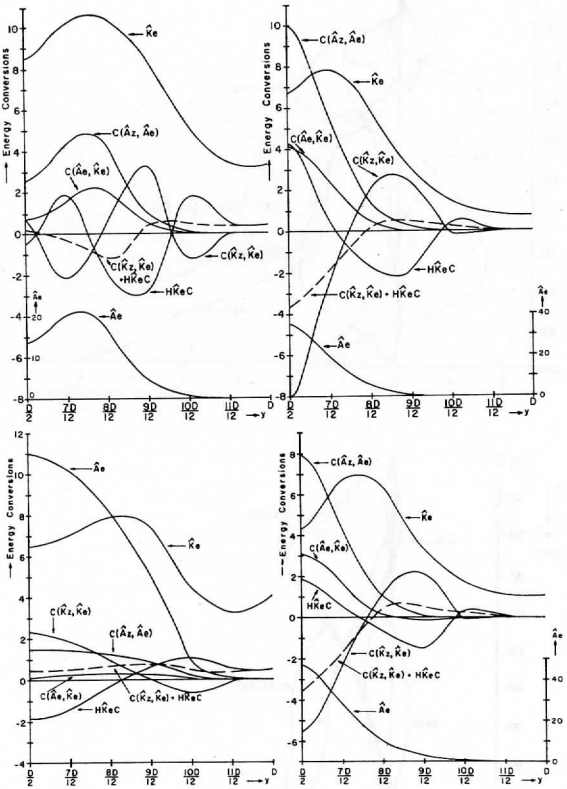


Fig. 34. Energy conversions (10^{-9}) and perturbation energies (10^{-5}), integrated over x and p, as a function of latitude (y) for the case with a = 1.0, b = 1.0, L = 4,000 km (upper), and L = 6,000 km (lower). Left: barotropic mode; right: baroclinic mode.

6.3 Unstable antisymmetric disturbances for the "cosine jet"

In Section 5.3 it was shown that the antisymmetric perturbations for the pure barotropic cosine jet current are stable in contrast to the case for symmetric perturbations. However, when only the antisymmetric perturbations of the lowest mode ($l = 2\pi/D$) between two rigid walls are allowed in the pure baroclinic current, the instability characteristics are unchanged if

$$k_{sym}^2 + (\frac{\pi}{D})^2 = k_{ant}^2 + (\frac{2\pi}{D})^2 , \qquad (6.2)$$

where k_{sym} is the zonal wavenumber for symmetric perturbations and k_{ant} the zonal wavenumber for antisymmetric perturbations. Hence the growth rate is lower for antisymmetric disturbances, but the instability still exists for wavelengths of O(1000 km).

In view of these facts, it is of interest to study the instability and energetic properties of antisymmetric perturbations for cosine-jet "mixed" currents. The complex eigenvalues obtained for different values of K^* with a constant A^* (5.672 kjoule m^{-3}) are shown in the upper part of Fig. 35. The most interesting result of this example is that there exists only one mode of primary instability, suggesting the barotropic mode is absent for these antisymmetric perturbations. Also present are some internal modes and "quasi-singular" modes with low growth rates.

The growth rates of the primary unstable mode are shown in the lower part of Fig. 35. A comparison of these results and those in Fig. 22 (symmetric perturbations-baroclinic mode) shows that the growth rates for antisymmetric perturbations are much lower than those for symmetric perturbations and also the wavelengths of maximum instability are slightly shorter for the antisymmetric waves.

The energy conversions and the eddy energies for this mode (with $a = 1.0$ and $B = 1.0$) are given in Table 3. The kinetic energy conversion $C(K_Z, K_E)$ is negative for all waves considered in this example, and the potential energy conversion and the perturbation energy conversion are positive. Thus, as will be seen in the parabolic-jet case, the energy source for the growing waves is A_Z, while K_Z is a sink.

6.4 Unstable symmetric disturbances for the "parabolic jet"

The basic zonal currents considered in this section are latitudinally parabolic, single-jet profiles given by

$$\bar{U}(y, p) = 2(1 - a(\frac{2}{D}y)^2)(\bar{U}_m + b\,\bar{U}'p)), \qquad |y| \leq \frac{D}{2}. \qquad (6.3)$$

This profile is a combined form of the pure baroclinic and the parabolic-jet

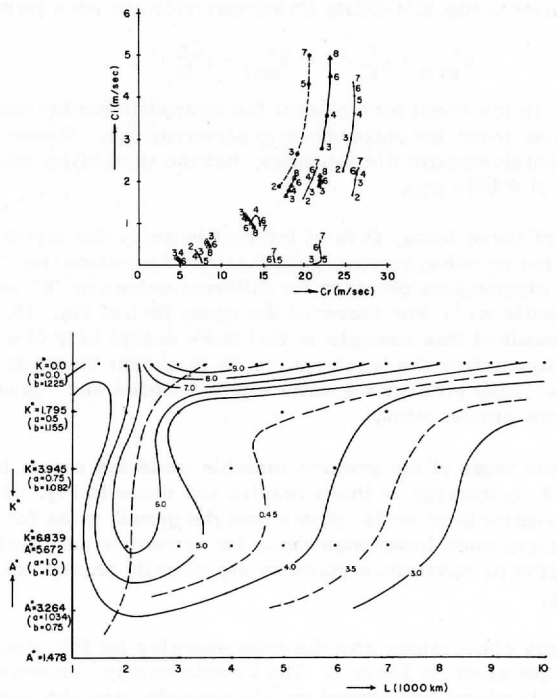


Fig. 35. Upper: complex phase speed plane for antisymmetric perturbations in the cosine-jet mixed current given by Eq. (6.1), M = 7, N = 15. Dots: a = 1.0, b = 1.0; triangles: a = 0.75, b = 1.082; circles: a = 0.5, b = 1.155. Lower: growth rate kc_i (day^{-1}) as a function of L and K* or A* (kjoule m^{-3}) for the most unstable mode in the upper diagram.

Table 3. Numerical results for antisymmetric perturbations in the cosine-jet mixed current given by Eq. (6.1) (for the most unstable mode). Parameters: a = 1.0, b = 1.0, M = 7, N = 15, and D = 3,000. Units: See Table 1.

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-3} sec^{-1})	$C(K_Z, K_e)$ (10^{-3} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	E_K (%)	kc_i (day^{-1})
2	1.43	0.64	-0.065	56.2	36.2	24.93 ± i2.25	-11.3	0.612
3	1.98	0.89	-0.23	86.4	54.0	25.69 ± i3.01	2.4	0.545
4	2.59	1.22	-0.42	121.4	79.0	25.98 ± i3.60	10.4	0.489
5	3.08	1.46	-0.53	160.9	105.6	26.03 ± i3.99	11.6	0.433
6	3.42	1.63	-0.58	202.5	132.7	25.99 ± i4.23	11.0	0.383
7	3.67	1.74	-0.61	245.0	159.8	25.95 ± i4.38	10.1	0.340

pure barotropic case discussed in the preceding chapters. Hence this is a flow for which the meridional gradient of absolute vorticity is always positive, suggesting that the barotropic instability mechanism is excluded.

Pedlosky (1965) has shown in the two-layer model that the disturbances occurring in such a basic current with no potential vorticity extrema within a layer are characterized by baroclinic instability alone and they increase the kinetic energy of the basic current (i.e., are barotropically stable).

The sequence of parameters a and b in (6.3) were chosen as in the cosine-jet profile and are shown by curve (d) in Fig. 3. The complex eigenvalues obtained by using the same parameters as in the preceding sections and the symmetric perturbation condition at the center of the channel are shown in Fig. 36. The upper diagram shows the results for a fixed value of A^* (5.672 kjoule m^{-3}) and three values of K^*, while the lower is for a fixed value of K^* (7.0 kjoule m^{-3}) and two values of A^*. The most interesting result of this figure is that there exists only one primary mode of instability; it is quite similar to the baroclinic mode for the cosine-jet profile (curve (a) in Figs. 19, 20, 21). Note that c_i for this mode decreases sharply as A^* decreases, but is not much changed for different values of K^* with a fixed value of A^*.

The remaining five unstable modes are significantly less unstable than the primary mode (these are not shown here for the cases $a = 0.5$, $b = 0.7233$ and $a = 1.04$, $b = 0.418$). It is also noted for this example that the phase speed (c_r) for the primary mode slightly increases as K^* or A^* decreases and is larger than that in the cosine-jet mixed case. The real eigenvalues (apparently singular modes) are ranged from $\bar{U}_{min}$ to $\bar{U}_{max}$, and are not shown here.

The growth rates, kc_i, of the most unstable mode are presented in Fig. 37 as a function of L and A^* or K^*. The results are quite similar to those of the baroclinic mode for the cosine-jet. A comparison of Figs. 37 with Fig. 22 reveals the following differences from the latter:

(1) the wavelength of maximum instability is larger.

(2) both the destabilizing effects of K^* (horizontal shear) for shorter waves and the stabilizing effects of a small amount of K^* for longer waves are smaller, especially for the shorter waves.

From these differences, it may be concluded that the presence of latitudinal shear as a parabolic profile has less influence on the baroclinic mode of instability than as a cosine-jet profile.

To find the energetic properties of the unstable disturbances for this mode, the case with $a = 1.0$ and $b = 0.839$ was considered. The numerical results

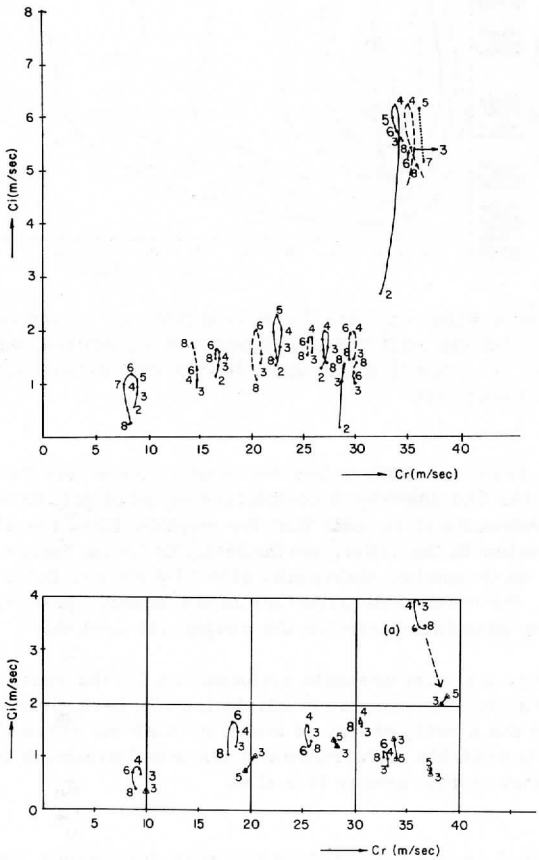


Fig. 36. Complex phase speed plane for symmetric perturbations in the parabolic-jet mixed current given by Eq. (6.3), M = 7, N = 15. In the upper, A^* has the constant value 5.672; solid line: a = 1.0, b = 0.839; dashed line: a = 0.75, b = 0.7824; dotted line: a = 0.5, b = 0.7233. In the lower, K^* = 7.0; dots: a = 1.02, b = 0.628; triangles: a = 1.04, b = 0.418. Numbers are L in 10^3 km.

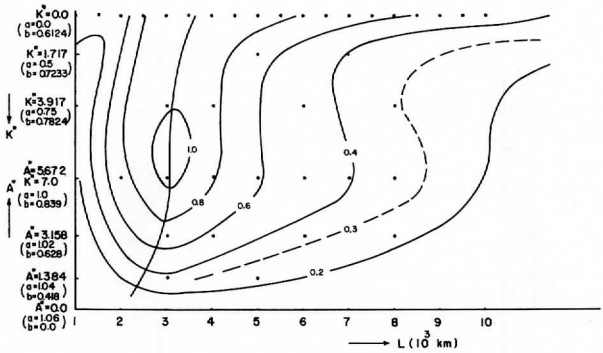


Fig. 37. Growth rates kc_i (day^{-1}) as a function of L and K^* or A^* (kjoule m^{-3}) for the most unstable mode of the parabolic jet (shown in Fig. 36). $A^* = 5.672$ in the upper half of the vertical axis, $K^* = 7.0$ in the lower half.

are given in Table 4, which shows the energy conversion from K_E to K_Z for all wavelengths considered. A comparison of these results with those in the cosine-jet (Appendix C) reveals that the magnitudes of the energy conversion terms are smaller in the latter, particularly for longer waves ($L \geq 4,000$ km), indicating a much smaller stabilizing effect by the parabolic lateral shear. Furthermore, the vertical distributions of the energy conversions are quite similar to the baroclinic mode for the cosine-jet current.

Therefore, the most unstable disturbances for the basic current without an absolute vorticity extremum are characterized by baroclinic mode processes and increase the kinetic energy of the basic current. Thus the only energy source for the unstable disturbances is the zonal available potential energy, while the zonal kinetic energy is a sink.

6.5 Summary of energetics for three unstable disturbance types

We now summarize the energetic properties of the unstable disturbances as a function of L for four different cases, by using the schematic energy flow diagram given in Fig. 38. The upper diagram shows the interactions between the zonal energies and the disturbances of two different wavelengths (4,000 and 6,000 km) for both the baroclinic mode (abbreviated as BC) and the barotropic mode (BT) for the symmetric perturbations on a cosine-jet mixed zonal current. Similarly, the lower diagram shows the baroclinic mode for the parabolic-jet mixed zonal current with symmetric perturbations (PB) and the antisymmetric perturbations on a cosine-jet mixed zonal current (AS). The flow

Table 4. Numerical results for the most unstable mode of the parabolic-jet mixed current given by Eq. (6.3). Parameters: a = 1.0, b = 0.839, M = 7, N = 15, and D = 3,000 km. Units: See Table 1.

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-3} sec^{-1})	$C(K_Z, K_E)$ (10^{-3} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	E_K (%)	kc_i (day^{-1})
2	1.11	0.51	-0.076	36.1	32.7	32.81 ± i2.65	20.4	0.720
3	2.82	1.33	-0.36	64.1	52.6	34.30 ± i5.55	20.4	1.003
4	3.68	1.53	-0.66	109.5	64.2	33.94 ± i6.23	30.6	0.845
5	3.66	1.41	-0.83	150.4	69.2	33.95 ± i5.97	44.3	0.649
6	3.48	1.26	-0.88	184.9	72.8	34.22 ± 5.73	56.5	0.518
7	3.34	1.16	-0.893	217.4	77.0	34.49 ± i5.59	65.8	0.433
8	3.25	1.09	-0.894	249.4	82.1	34.71 ± i5.51	72.8	0.374

of energy is in the direction of the arrow and the magnitude of the kinetic energy conversion in the upper diagram is shown by the length of the arrow.

It is evident from these results that the unstable symmetric perturbations superimposed on the mixed zonal current containing absolute vorticity extrema at various latitudes convert zonal kinetic energy into eddy kinetic energy, dominantly for the longer waves, while the unstable perturbations occurring in a current without the pure barotropic instability mechanism increase the zonal kinetic energy at the expense of eddy kinetic energy through the baroclinic mechanism. It is also interesting to note that the directions of $C(A_Z, A_E)$ and $C(A_E, K_E)$ are always positive for all cases considered in this study, but their magnitudes are significantly decreased for relatively large values of the zonal kinetic energy density.

7. Summarizing Remarks

The dynamic instability and the structure and energetic properties of the unstable disturbances in zonal currents, which are representative of those observed in the mid-latitude atmosphere, have been investigated through use of a quasi-geostrophic numerical model based upon a finite difference technique.

The numerical results obtained for several pure baroclinic and pure barotropic zonal currents were compared qualitatively with well known theoretical results and then the accuracy of the results was examined by varying the number of subdivisions. It was found that the numerical method, with vertically 14 subdivisions and latitudinally 6 subdivisions between the center and the rigid boundary, gives sufficiently accurate results for describing the structure and energetic properties of the unstable disturbances for intermediate wavelengths (order of a few thousand kilometers). Other interesting new results are summarized as follows:

Pure baroclinic case:

(1) The instability and structure of unstable disturbances are quite similar to those for the linear current with a constant stability parameter (Kuo, 1952), except for the slightly shorter wavelengths of maximum instability and larger growth rates, which seemed to be caused by the smaller values of the stability parameter in the lower troposphere.

(2) The number of singular complex modes increases as the number of subdivisions and the shear increases, but their growth rates are very small and almost independent of wavelength.

(3) The troughs and ridges tilt westward with altitude, mainly in the lower troposphere, and this inclination increases with increasing wavelength.

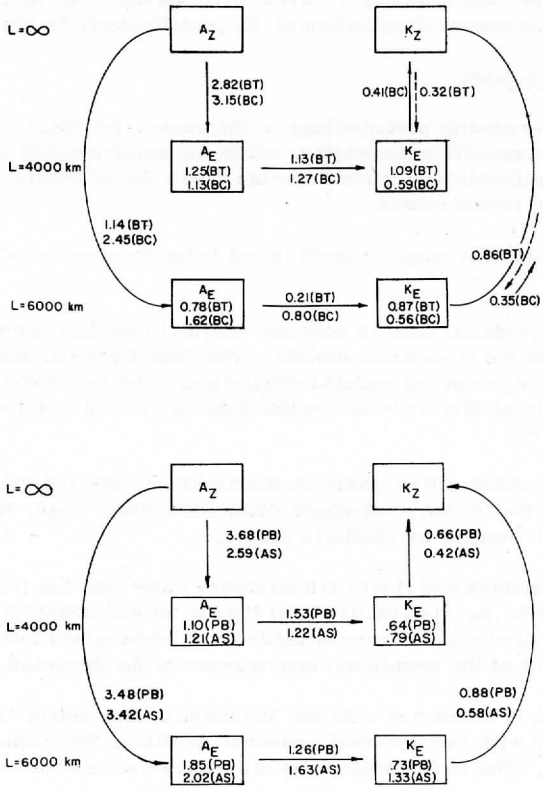


Fig. 38. Flow diagram of energy integrated over the entire mass. The flow of energy is in the direction of the arrow. Upper: cosine-jet current ($a = 1.0$, $b = 1.0$) with symmetric perturbations; BT: barotropic mode; BC: baroclinic mode. Lower: PB: parabolic-jet current ($a = 1.0$, $b = 0.839$) with symmetric perturbations; AS: antisymmetric perturbations in the cosine-jet current ($a = 1.0$, $b = 1.0$). Units: $A_E, K_E = 10^2$, energy conversions $= 10^{-3}$.

(4) The energy conversions, A_z to A_e and A_e to K_e, are dominant in the lower layer with concentrations at the center of the jet. From the vertical distribution of K_e, the shorter waves are characterized as shallow waves, while the longer waves as upper waves. The vertical flux term is important in determining the spatial distribution of K_e, particularly for the longer waves.

Pure barotropic case:

(1) The symmetric perturbations in the cosine-jet zonal currents are unstable for intermediate wavelengths, while the antisymmetric perturbations are stable. The parabolic-jet current is also stable for symmetric (and probably antisymmetric) perturbations.

(2) The singular complex mode is not found for wavelengths considered in this study.

(3) The numerical results near the marginal stability curve are in good agreement with the theoretical results. The first divergent unstable mode occurred when the number of vertical subdivisions has increased from 9 to 15. The growth rate of this mode is smaller than the nondivergent mode, but converts K_e to A_e.

(4) The wavelength of maximum instability for the symmetric perturbations is larger than that of the comparable "pure baroclinic" case, while the growth rates are much smaller for realistic shears.

(5) The troughs and ridges tilt eastward away from the jet. The energy conversion $C(K_z, K_e)$ is positive near the center and negative near the boundary. Thus, an initial single-jet current tends to be transformed into a double-jet current as a result of the meridional convergence of the Reynolds stresses.

Chapter 6 is concerned with the latitudinally symmetric "mixed" (jet-type) zonal currents with two different horizontal profiles, the cosine-jet and the parabolic-jet. The main conclusions are summarized as follows:

(1) For the zonal currents with absolute vorticity extrema at various latitudes (symmetric perturbations), there exist two most unstable modes corresponding to each category of the energy source; the baroclinic mode dominates the shorter waves, while the barotropic mode dominates the longer waves.

(2) Both the potential energy conversion and the perturbation energy conversion are positive for both modes. However, for a representative zonal current, the baroclinic mode increases the zonal kinetic energy, while the barotropic mode decreases it.

(3) The simultaneous presence of both baroclinic and the barotropic processes tend to reduce all energy conversions involving growing waves. How-

ever, the former dominates in the lower troposphere, while the latter is pronounced near the tropopause, particularly for longer waves.

(4) The presence of horizontal shear destabilizes shorter waves and stabilizes longer waves for the baroclinic mode, while the inclusion of the vertical shear destabilizes all wavelengths, especially for shorter waves, for the barotropic mode.

(5) The baroclinic mode tends to generate three jets near the tropopause—one at the center and the other two near the boundaries—while the barotropic mode tends to transform an initial single jet into a double jet near the tropopause.

(6) The internal complex modes and "quasi-singular" modes occurred, but their growth rates were significantly small.

(7) For the basic current without potential vorticity extrema (the parabolic jet), there exists only one strongly unstable mode; it is characterized by baroclinic processes. The same is also true for the antisymmetric perturbations in cosine-jet currents.

The results obtained in this thesis are necessarily of limited quantitative validity. Nevertheless, it is believed that many of the energetic properties of the solutions are accurate and useful when comparisons for changing available energy sources are considered. In particular, they seem to have illustrated the various peculiar modifications introduced into the disturbances as the shear-flow structure was altered.

It is the author's feeling that the qualitative results listed earlier in this section are very valuable. It appears that the relation between instabilities in the real atmosphere and those occurring in the theoretical "pure" cases is now much clearer. The results have allowed us to single out the "baroclinic" mode of instability whose energy conversions are in the same sense as those observed over great lengths of time in the real atmosphere. On the other hand, the method of analysis has allowed the other modes of instability to be examined. Since these modes may also be excited in the real atmosphere, these results appear to provide a possible explanation for the fluctuating character of disturbance growth atmospheric energetics over shorter periods of time.

It is thus hoped that the present results will benefit the synoptician, dynamicist, and scholar of the general circulation.

Acknowledgments:

I would like to extend my profound gratitude to Professor John A. Young for his guidance during the preparation of this thesis. In addition, I wish to express my sincere thanks and appreciation to the staff of the Meteorology Department of the University of Wisconsin for providing my graduate education.

I am also deeply indebted to Professor Donald R. Johnson for his guidance during my first year of graduate study, and to Professor David Houghton for his many helpful discussions.

Finally, I would like to express a special thanks to my wife, Miza, whose help and patience made the pursuit of this degree possible.

This research was supported on ESSA grant E-230-68G.

BIBLIOGRAPHY

Arnason, G., P. S. Brown, and E. A. Newburg, 1967: A Case Study of the Validity of Finite Difference Approximations in Solving Dynamic Stability Problems. J. Atmos. Sci., 24, 10-17.

Bretherton, F. P., 1966a: Critical Layer Instability in Baroclinic Flows. Quarterly J. of the Royal Meteorological Society, 92, 325-334.

Bretherton, F. P., 1966b: Baroclinic Instability and the Short Wavelength Cutoff in Terms of Potential Vorticity. Q. J. R. M. Soc., 92, 335-345.

Brown, J. A., Jr., 1969: A Numerical Investigation of Hydrodynamic Instability and Energy Conversions in the Quasi-Geostrophic Atmosphere, Part I and II. J. Atmos. Sci., 26, 352-365 and 366-375.

Burger, A. P., 1962: On the Non-existence of Critical Wavelengths in a Continuous Baroclinic Stability Problem. J. Atmos. Sci., 19, 30-38.

Burger, A. P., 1958: Scale Considerations of Planetary Motions of the Atmosphere. Tellus, 10, 195-205.

Charney, J. G., 1947: The Dynamics of Longwaves in a Baroclinic Westerly Current. J. Met., 4, 135-163.

Charney, J. G., 1948: On the Scale of Atmospheric Motions. Geofis. Publ. (Oslo), 17, 1-17.

Charney, J. G., 1962: Integration of the Primitive and Balance Equations. Proceedings of the International Symposium on Numerical Weather Prediction in Tokyo, November 7-13, pp. 131-151.

Dutton, J. and D. Johnson, 1967: The Theory of Available Potential Energy and a Variational Approach to Atmospheric Energetics. Advances in Geophysics, 12, 333-436.

Derome, J. F. and A. Wiin-Nielsen, 1966: On the Baroclinic Stability of Zonal Flow in Simple Model Atmospheres. The Univ. of Michigan, Tech. Report No. 2, 93 pp.

Eady, E. T., 1949: Long Waves and Cyclone Waves. Tellus, 1, 33-52.

Eliasen, E., 1954: Numerical Solution of the Perturbation Equation for Linear Flow. Tellus, 6, 183-193.

Green, J. S. A., 1960: A Problem in Baroclinic Stability. Quart. J. Roy. Met. Soc., 86, 237-251.

Haltiner, G. J., and R. T. Song, 1962: Dynamic Instability in Barotropic Flow. Tellus, 14, 383-393.

Haltiner, G. J., 1963: Finite Difference Approximations for the Determination of Dynamic Instability. Tellus, 15, 230-240.

Jacobs, S. J., and A. Wiin-Nielsen, 1966: On the Stability of a Barotropic Basic Flow in a Stratified Atmosphere. J. Atmos. Sci., 23, 682-687.

Kuo, H. L., 1949: Dynamic Instability of Two Dimensional Nondivergent Flow in a Barotropic Atmosphere. J. Met., 6, 105-122.

Kuo, H. L., 1952: Three Dimensional Disturbances in a Baroclinic Zonal Current, J. Met., 9, 260-278.

Kuo, H. L., 1953: The Stability Properties and Structure of Disturbances in a Baroclinic Atmosphere. J. Met., 10, 235-243.

Lamb, H., 1932: Hydrodynamics, 6th ed. Cambridge Univ. Press, 738 pp.

Lin, C. C., 1945: On the Stability of Two Dimensional Parallel Flows. Quart. Appl. Math., 3, 117-142, 218-234.

Lin, C. C., 1955: The Theory of Hydrodynamic Stability. Cambridge Univ. Press, 155 pp.

Lorenz, E. N., 1955: Available Potential Energy and the Maintenance of the General Circulation. Tellus, 17, 31-41.

Lorenz, E. N., 1960: Energy and Numerical Weather Prediction. Tellus, 12, 364-373.

Lorenz, E. N., 1957: Static Stability and Atmospheric Energy. Studies of the Atmospheric General Circulation II, Final Report under Contract No. AF 19(604)-1000, Department of Meteorology, MIT, 450-489.

Lorenz, E. N., 1967: The Nature and Theory of the General Circulation of the Atmosphere. W. M. O. Press, 161 pp.

Miles, J. W., 1964: Baroclinic Instability of the Zonal Wind. Parts II and III, J. Atmos. Sci., 21, 550-506 and 603-609.

Oort, A. H., 1964: On Estimates of the Atmospheric Energy Cycle. Monthly Weather Review, 92, 11, 483-493.

Pedlosky, J., 1964: The Stability of Currents in the Atmosphere and the Ocean. J. Atmos. Sci., 21, 201-219 and 342-353.

Pedlosky, J., 1965: On the Stability of Baroclinic Flows as a Functional of the Velocity Profile. J. Atm. Sci., 22, 137-145.

Phillips, N. A., 1951: A Simple Three-Dimensional Model for the Study of Large Scale Extratropical Flow Patterns. J. Meteor., 8, 381-394.

Phillips, N. A., 1954: Energy Transformations and Meridional Circulations Associated with Simple Baroclinic Waves in a Two Level Quasi-geostrophic Model. Tellus, 6, 273-286.

Phillips, N. A., 1956: The General Circulation of the Atmosphere: A Numerical Experiment. Quart. J. Roy. Met. Soc., 82, 123-164.

Phillips, N. A., 1963: Geostrophic Motion. Rev. of Geophysics, Vol. 1, 123-176.

Rayleigh, Lord, 1913: On the Stability of the Laminar Motion of an Inviscid Fluid, Phil. Mag., 26, 101-1010.

Rossby, C. G., et al., 1939: Relation between Variations in the Intensity of the Zonal Circulation and the Displacement of the Semi-Permanent Centers of Action. J. Marine Res., 2, 38-55.

Starr, V. P., 1954: Studies of the Atmospheric General Circulation, Final Report, Part 1, General Circulation Project Contract No. AF 19-122-53, Dept. of Meteor., MIT, 535 pp.

Thompson, P. D., 1953: On the Theory of Large-Scale Disturbances in a Two-Dimensional Baroclinic Equivalent of the Atmosphere. Quart. J. Roy. Met. Soc., 79, 51-69.

Wasow, W., 1948: The Complex Asymptotic Theory of a Fourth Order Differential Equation of Hydrodynamics. Ann. Math., 49, 852-871.

Wiin-Nielsen, A., 1959: On Certain Integral Constraints for the Time-Integration of Baroclinic Models. Tellus, 11, 45-59.

Wiin-Nielsen, A., 1961: On Short and Long Term Variations in Quasi-Barotropic Flow. Mon. Wea. Rev., 89, 461-476.

Wiin-Nielsen, A., J. A. Brown, and M. Drake, 1964: Further Studies of Energy Exchange between the Zonal Flow and the Eddies. Tellus, 16, 168-180.

Yanai, M., and T. Nitta, 1968: Finite Difference Approximations for the Barotropic Instability Problem. J. Meteor. Soc. Japan, 46, 389-403.

APPENDIX A

BASIC EQUATIONS FOR THE PURE BAROCLINIC ZONAL CURRENTS

1. The governing perturbation equation

When the basic zonal current in the system of perturbation equations (2.11)-(2.15) is independent of y, solutions can be found of the form

$$\psi' = \mathrm{Re}[A(p)e^{ik(x-ct)}e^{ily}], \tag{A.1}$$

where l is the latitudinal wave number. To satisfy the horizontal boundary conditions (2.14), Eq. (A.1) may be written as:

$$\psi' = \mathrm{Re}[A(p)e^{ik(x-ct)}]\cos ly, \qquad |y| \leq \frac{D}{2}, \tag{A.2}$$

where $l = \frac{n\pi}{D}$, $n = 1, 3, 5, \ldots$, and $D = 3,000$ km.

If this form is substituted into (2.13) and the vertical boundary conditions (2.15), we obtain

$$(\bar{U} - c)[(\frac{f_0^2}{\sigma(p)} A_p)_p - (k^2 + l^2)] + [\beta - (\frac{f_0^2}{\sigma(p)} \bar{U}_p)_p]A = 0 , \tag{A.3}$$

$$(\bar{U} - c)A_p - \bar{U}_p A = 0 \quad \text{at } p = 0, p_0 , \tag{A.4}$$

where the subscript p represents total vertical differentiation.

It may be noted that (A. 3) differs from the usual baroclinic instability problem with infinite lateral extent ($l^2 = 0$) which has been studied by most authors except Phillips (1954), in that the sum of the squares of zonal and of latitudinal wave number appears in place of the former alone. Thus the mathematical problem (A. 3, A. 4) is equivalent to that of the usual baroclinic stability problem. However, the presence of the latitudinal dependence of the perturbations shifts the curve c_i vs. L toward longer zonal wavelengths and decreases the growth rate, resulting in the slightly shorter wavelength of maximum instability. These differences increase as the value of l increases. Hence, by restricting ourselves to the most unstable internal mode, the lowest latitudinal wave number ($l = \pi/D$) is chosen in this study.

2. The energy equations

For the pure baroclinic zonal current and the perturbation stream function in the form of (A. 2), each term of the energy equations in (3.10) is replaced with:

$$\text{(a)} \qquad C(A_Z, A_E) = e^{2kc_i t} g^{-1}(-k\frac{L}{2})\frac{D}{4} \int \frac{f_0^2}{\sigma} \bar{U}_p(|A|^2\alpha_p)dp,$$

$$\text{(b)} \qquad C(K_Z, K_E) = 0$$

$$\text{(c)} \qquad C(A_E, K_E) = e^{2kc_i t} g^{-1}\frac{L}{2}\frac{D}{4} \int [|W||A|_p \cos(\theta-\alpha) + |W||A|\alpha_p \sin(\theta-\alpha)]dp, \qquad \text{(A. 5)}$$

$$\text{(d)} \qquad A_E = e^{2kc_i t} g^{-1}\frac{L}{4}\frac{D}{4} \int \frac{f_0^2}{\sigma}[(|A|_p)^2 + (|A|\alpha_p)^2]dp,$$

$$\text{(e)} \qquad K_E = e^{2kc_i t} g^{-1}\frac{L}{4}\frac{D}{4} \int (l^2 + k^2)|A|^2 dp,$$

$$\text{(f)} \qquad HK_e C = e^{2kc_i t}(-\frac{L}{2}kl^2) \cos 2\,ly(c_i|A|\,),$$

$$\text{(g)} \qquad VK_e C = -e^{2kc_i t} f_0 \frac{L}{2} \cos^2 ly[(|W||A|)_p \cos(\theta-\alpha) - |W||A|(\theta-\alpha)_p \sin(\theta-\alpha)].$$

APPENDIX B

ANALYTIC DETERMINATION OF CRITICAL WAVELENGTHS FOR COSINE-JET, NONDIVERGENT, PURE BAROTROPIC ZONAL CURRENTS

The zonal currents considered here are the cosine-jet pure barotropic profiles given by

$$\bar{U}(y) = \bar{U}_m(1 - a \cos \frac{2\pi}{D} y), \qquad 0 \leq y \leq D . \tag{5.1}$$

The linearized vorticity equation in the nondivergent model, after applying the harmonic form of the perturbation stream function, can be written as:

$$(\bar{U} - c)(A'' - k^2 A) + (\beta - \bar{U}'')A = 0, \tag{B.1}$$

where the prime denotes the differentiation with respect to latitude (y), A is the amplitude of the stream function.

Following Kuo (1949), we determine the critical wavelengths of the marginal stability curve for the basic profile given by (5.1).

1. The lower critical wavelength

The necessary condition for the existence of nonsingular neutral waves with $\bar{U}_{min} < c < \bar{U}_{max}$ is that $\beta - \bar{U}'' = 0$ at some critical latitude $(y = y_k)$ where $c = \bar{U}_k$ (the current velocity at $y = y_k$). Hence, if we substitute (5.1) into (B.1), and then use $c = \bar{U}_k$, we eventually obtain the simple relation

$$A_k'' = (k_k^2 - (\frac{2\pi}{D})^2) A_k , \tag{B.2}$$

where A_k represents the amplitude function of the neutral wave for the lower critical wavelength, k_k the critical wavenumber. Equation (B.2) admits the solution $A_k = C_1 \sin py + C_2 \cos py$, where C_1 and C_2 are constants, and

$$p^2 = (\frac{2\pi}{D})^2 - k_k^2 . \tag{B.3}$$

The boundary conditions $A_k(0) = A_k(D) = 0$ admit only the sine term, and the symmetric condition at $y = D/2$ $(A_k'(D/2) = 0)$ is satisfied by $p = (2n+1)\pi/D$, where n is any integer. However, since $k_k^2 > 0$, the only possible value of p is π/D. Thus we get

$$k_k^2 = 3(\frac{\pi}{D})^2 \text{ and } L_k = \frac{2}{\sqrt{3}} D . \tag{B.4}$$

Therefore, it can be seen that the lower critical wavelength depends on the latitudinal width of the channel, but not on the shear of the current.

2. The upper critical wavelength

Kuo (1949) further has shown that, for symmetric perturbations, the upper critical wavelength dividing unstable waves and neutral Rossby waves (for $c = \bar{U}_{min}$) is given by

$$k_0^2 = \frac{\beta \int_0^{D/2} A_0 \, dy}{\int_0^{D/2} (\bar{U} - \bar{U}_{min}) A_0 \, dy} , \qquad \text{(B. 5)}$$

where A_0 represents the amplitude function of the upper neutral wave. If we take $\bar{U}(y)$ as an approximate expression for $A_0(y)$, we obtain from (B. 5) after integration

$$L_0 = 2\pi \sqrt{\frac{a(1 + a/2)U_m}{\beta}} . \qquad \text{(B. 6)}$$

Hence, the upper critical wavelength is related to the shear and β. Also, it is interesting to note that the longer waves are stabilized by β, since L_0 decreases as β increases.

APPENDIX C

COMPLETE NUMERICAL RESULTS FOR THREE SELECTED CASES

1. The pure baroclinic zonal current:

$$\bar{U}(p) = \bar{U}_m + b\bar{U}'(p), \qquad (4.1)$$

where $\bar{U}_m$ = 23.725 m sec^{-1} the vertically integrated mean value, $\bar{U}'(p)$ the deviation from $\bar{U}_m$ (as shown in Fig. 2), and b = 1.225. The latitudinal wave number l = π/D, D = 3,000 km, M = 1, and N = 15. Units: See Table 1.

(a) Primary mode:

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-3} sec^{-1})	A_E	K_E	($c = c_r \pm ic_i$) (m sec^{-1})	kc_i (day^{-1})	E_k (%)	Secondary complex eigenvalue (m sec^{-1})
1	1.38	0.66	57.6	73.8	10.99 ± i0.99	0.54	28.4	None
2	1.28	0.68	56.2	76.7	16.13 ± i1.69	0.46	16.1	None
3	3.54	1.70	84.9	92.7	18.20 ± i5.16	0.93	15.0	33.88 ± i0.115
4	5.01	1.98	135.4	106.3	18.37 ± i7.14	0.97	17.1	34.01 ± i0.124
5	5.98	1.97	196.4	118.7	17.99 ± i8.13	0.88	19.0	None
6	6.61	1.88	261.0	131.2	17.52 ± i8.64	0.78	20.6	None
7	6.05	1.55	282.3	124.4	17.07 ± i8.89	0.69	21.9	None
8	5.66	1.34	305.9	122.6	16.67 ± i9.00	0.61	22.9	None
9	5.37	1.19	331.0	123.6	16.34 ± i9.047	0.55	23.7	None
10	5.15	1.09	357.0	126.3	16.06 ± i9.052	0.49	24.4	None
(b) Secondary mode:								
3	0.054	0.048	12.0	22.0	33.88 ± i0.115	0.021	-351.3	
4	0.051	0.041	25.2	16.7	34.01 ± i0.124	0.017	-527.7	

2. The cosine-jet pure barotropic current - symmetric perturbations:

$$\bar{U}(y) = \bar{U}_m(1 - a \cos \frac{2\pi}{D} y), \qquad 0 \leq y \leq D, \tag{5.1}$$

where $\bar{U}_m = 23.725$ m sec^{-1}, $a = 1.084$, and $D = 3,000$ km, $M = 7$, and $N = 15$. Units: See Table 1.

(a) Nondivergent mode:

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-3} sec^{-1})	$C(K_Z, K_E)$ (10^{-3} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	kc_i (day^{-1})	E_K (%)
2						all $c_i = 0$		
3						all $c_i = 0$		
3.5						all $c_i = 0$		
4	0	0	2.37	0	259.1	16.02 ± i2.39	0.32	-22.0
4.5	0	0	3.24	0	234.6	14.52 ± i4.16	0.50	-18.9
5	0	0	3.21	0	222.3	12.93 ± i4.89	0.53	-17.6
5.5	0	0	2.89	0	215.5	11.32 ± i5.02	0.50	-16.7
6	0	0	2.42	0	211.7	9.64 ± i4.70	0.43	-16.02
7	0	0	1.79	0	212.9	5.56 ± i4.17	0.32	-12.3
8	0	0	1.43	0	224.0	2.97 ± i3.73	0.25	- 9.3
9	0	0	0.71	0	237.2	1.07 ± i2.01	0.12	- 7.2

(b) Divergent mode:

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-5} sec^{-1})	$C(K_Z, K_E)$ (10^{-4} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	kc_i (day^{-1})	E_K (%)
5	0	-2.36	3.16	3.4	28.6	15.81 ± i2.78	0.29	-46.7
5.5	0	-3.01	3.74	3.5	27.5	15.04 ± i3.73	0.37	-46.7
6	0	-3.34	3.91	3.7	27.0	14.32 ± i4.30	0.39	-47.2
7	0	-3.58	3.86	4.1	27.0	13.05 ± i4.85	0.37	-48.9
8	0	-3.59	3.66	4.6	27.7	12.02 ± i5.02	0.34	-50.6
9	0	-3.51	3.42	5.0	28.8	11.18 ± i5.01	0.29	-52.1

3. The cosine-jet mixed current - symmetric perturbations:

$$\bar{U}(y, p) = (1 - a \cos \frac{2\pi}{D} y)(\bar{U}_m + b\bar{U}'(p)), \qquad (6.1)$$

where $\bar{U}_m = 23.725$ m sec^{-1}, $\bar{U}'(p)$ is the same as in Fig. 2, a = 1.0, b = 1.0, and D = 3,000 km. The number of subdivisions M = 7, and N = 15. Units: See Table 1.

(a) Baroclinic mode:

L (10^3 km)	$C(A_Z, A_E)$ (10^{-3} sec^{-1})	$C(A_E, K_E)$ (10^{-3} sec^{-1})	$C(K_Z, K_E)$ (10^{-3} sec^{-1})	A_E	K_E	$c = c_r \pm ic_i$ (m sec^{-1})	kc_i (day^{-1})	E_K (%)
2	1.35	0.61	-0.058	32.0	28.5	28.51 ± i3.70	1.01	17.2
2.5	2.34	1.12	-0.16	45.7	40.8	29.30 ± i5.33	1.16	11.7
3	3.36	1.53	-0.29	69.3	53.3	28.90 ± i6.29	1.14	11.8
3.5	3.78	1.62	-0.41	98.4	62.4	28.26 ± i6.12	0.95	11.7
4	3.15	1.27	-0.406	113.3	58.7	28.43 ± i5.27	0.72	10.7
4.5	2.70	1.03	-0.38	122.9	54.5	29.11 ± i4.88	0.59	12.9
5	2.54	0.91	-0.37	135.1	53.7	29.71 ± i4.79	0.52	16.1
5.5	2.48	0.84	-0.36	148.4	54.6	30.18 ± i4.81	0.48	19.4
6	2.45	0.79	-0.35	162.2	56.3	30.56 ± i4.88	0.44	22.8
7	2.44	0.72	-0.33	190.7	60.6	31.09 ± i5.04	0.39	29.5
8	2.46	0.67	-0.32	219.7	65.5	31.44 ± i5.19	0.35	35.4
9	2.47	0.63	-0.31	248.9	70.8	31.69 ± i5.32	0.32	40.4

(b) Barotropic mode:

2	0.78	0.36	-0.31	41.7	29.8	20.07 ± i1.59	0.43	81.0
2.5	1.01	0.47	-0.41	59.0	38.2	20.41 ± i1.82	0.40	83.0
3	1.44	0.68	-0.50	88.1	58.2	20.76 ± i2.04	0.37	63.7
3.5	3.02	1.43	-0.43	173.9	134.1	21.32 ± i2.55	0.39	19.2
4	2.82	1.13	0.32	125.3	109.4	20.89 ± i4.29	0.58	1.4
4.5	2.07	0.69	0.74	99.1	94.4	18.74 ± i4.97	0.60	- 9.4
5	1.61	0.46	0.85	85.9	90.4	15.97 ± i5.34	0.58	- 7.4
5.5	1.38	0.32	0.87	80.9	88.4	13.73 ± i5.69	0.56	- 3.6
6	1.14	0.21	0.86	77.7	87.4	11.72 ± i5.68	0.51	- 3.7
7	0.79	0.087	0.82	73.5	90.3	8.21 ± i5.37	0.42	- 4.7
8	0.56	0.045	0.66	73.4	96.2	5.59 ± i4.47	0.30	- 4.4
9	0.31	0.025	0.38	74.9	102.3	3.38 ± i2.76	0.17	- 3.5

MASS CONVERGENCE IN A BAROCLINIC EKMAN LAYER

John A. Young
The University of Wisconsin, Madison

Abstract

An expression for the vertical mass outflow above a balanced Ekman boundary layer is derived for flow states possessing low level fields of baroclinity. It is found that the presence of horizontal temperature gradients may significantly alter the outflow from that expected when the pressure field is invariant throughout the boundary layer. The extent of this modification is greatest when the geostrophic wind varies substantially over the depth of the boundary layer, as in lower latitudes or over sloping terrain with thermal stratification near the surface.

The contribution of the baroclinic effect to the outflow produces upward motion over cold centers and downward motion over warm centers relative to that associated with the surface pressure field. Over sloping terrain with temperature lapse (increase) with height this corresponds to sinking (rising) motion in valleys and rising (sinking) motion over topographical maxima. The pressure work done by the boundary layer vertical motion usually drains energy from the free atmosphere above, but in sufficiently warm surface lows (or cold highs) this energy flux may actually be directed opositely.

1. Introduction

Despite the early theoretical work by Ekman (1905), our understanding of flow in the planetary boundary layer is still far from complete. Admittedly, steady progress is still being made on the problem of unaccelerated turbulent motion with constant pressure gradient force. These solutions have found useful application in the estimation of synoptic scale mass convergence in the boundary layer (see, e.g., Charney and Eliassen, 1949). However, these assumptions of spatial and temporal quasi-homogeneity are frequently violated in the real atmosphere. Unfortunately, the cases with large variations in the pressure forces are often those for which knowledge of the boundary layer motions is required, i.e., as in the hurricane or frontal zone. Some advances have recently been made along these lines using relatively crude representations of the

turbulent stresses. For example, Kuo (1969) has studied the steady flow in a viscous vortex, Businger and Ghing (1968) have examined the transient passage of a pressure disturbance, and Lettau (1967) and Holton (1967) have studied the development of jets over sloping, diurnally heated terrain.

As pointed out by Lettau (1967), the influence of surface height variations in conjunction with the thermal stratification of the atmosphere may be considered to be a kind of thermal wind effect in the boundary layer. Steady-state solutions along these lines have recently been exhibited by Mahrt and Schwerdtfeger (1970) for the special case of flow in the strong inversion layer over Antarctica. As Mendenhall (1967) has shown, the influence of boundary layer baroclinity need not be limited to sloping boundary layers, but may be statistically important even over the open ocean.

Figure 1 (taken from Lettau (1967)) schematically indicates the kinds of distortions in the wind field which can occur with various orientations of the surface pressure and temperature fields. Depending upon the orientation, we see that the magnitude and direction of the vertical wind shear can vary substantially if the vertical differences in the geostrophic wind vector are significant over the depth of the Ekman layer. Since mass convergence in a balanced boundary layer is known to depend upon the surface stress only, it is perhaps to be expected from Fig. 1 that the surface pressure distribution does not completely prescribe the boundary layer outflow.

The purpose of this paper is to explore, in the simplest relevant model, the kinds of modifications that can be induced in the boundary layer convergence by a field of baroclinity, and to discuss their possible significance in the real atmosphere. Section 2 contains derivation of expressions for mass outflow above a balanced boundary layer in the presence of a vertically varying geostrophic velocity field. Applications of these formulas are discussed in Section 3, where it is suggested that low-level temperature gradients in the tropics or sloping terrain at any latitude may induce systematic changes in the patterns of mass outflow.

2. Vertical motion expressions

We begin by considering the horizontal equations of motion for unaccelerated flow in the planetary boundary layer:

$$f\vec{k}\times\vec{V} = -\frac{1}{\rho_s(z)}\nabla p' + K\frac{\partial^2\vec{V}}{\partial z^2} \quad . \tag{2.1}$$

Here we have made a Boussinesq approximation in the pressure gradient force term by using $\rho_s(z)$, the slowly varying density in a motionless reference state. The deviation pressure $p' = p - p_s(z)$ will be assumed to obey the hydrostatic relation $\partial p'/\partial z = -\rho' g$ where $\rho' = \rho - \rho_s(z)$. The Coriolis parameter f is taken to be constant as is the coefficient of eddy viscosity K. $\vec{V} = u\,\vec{i} + v\,\vec{j}$

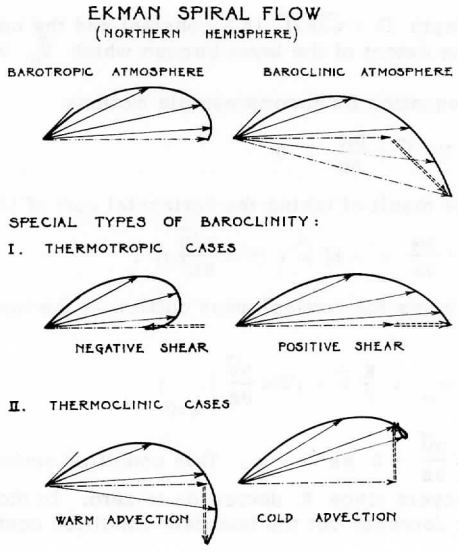


Fig. 1. Schematic Ekman spiral wind profiles (solid lines) for barotropic and various baroclinic flow situations, taken from Lettau (1967). The dashed vectors represent the thermal wind for the boundary layer.

is the horizontal velocity vector, where $\vec{i}$, $\vec{j}$ and $\vec{k}$ are the unit vectors in the x, y, and z directions. The horizontal del operator is

$$\nabla = \frac{\partial}{\partial x}\vec{i} + \frac{\partial}{\partial y}\vec{j} .$$

The geostrophic wind vector may be defined by

$$f\vec{k} \times \vec{V}_g = -\frac{1}{\rho_s(z)} \nabla p' . \qquad (2.2)$$

For the purposes of this study we allow it to vary with height in a layer near the earth's surface according to

$$\vec{V}_g(z) = \vec{V}_{g_0} \times \hat{\vec{V}}_g [1 - e^{-\alpha z/D}] . \qquad (2.3)$$

Here $\vec{V}_{g_0}$ is the surface geostrophic wind and $\vec{V}_g = \vec{V}_{g_\infty} - \vec{V}_{g_0}$ is the total variation of the geostrophic wind between the top and bottom of the boundary

layer. The Ekman depth $D = \sqrt{2k/f}$ is a constant and the nondimensional factor α determines the extent of the layer through which $\vec{V}_g$ varies significantly.

The continuity equation for incompressible motions

$$\nabla \cdot \vec{V} + \frac{\partial w}{\partial z} = 0 \tag{2.4}$$

may be used with the result of taking the horizontal curl of (2.1) to give

$$f\frac{\partial w}{\partial z} = -K[\vec{k} \cdot (\nabla \times \frac{\partial^2 \vec{V}}{\partial z^2})] \ . \tag{2.5}$$

Integration of (2.5) gives the vertical mass outflow at the top of the boundary layer

$$w_\infty = \frac{K}{f}\vec{k} \cdot (\nabla \times \frac{\partial \vec{V}}{\partial z}\Big|_{z=0}) \tag{2.6}$$

providing that $K\nabla \times \frac{\partial \vec{V}}{\partial z} \to 0$ as $z \to \infty$. This condition seems realistic for turbulent boundary layers since K decreases to zero. In the present idealized solutions K cannot decrease but the boundary condition continues to hold because we assume

$$\vec{V} = \vec{V}_g = \vec{V}_{g_\infty} \qquad \text{as } z \to \infty \tag{2.7}$$

where $\vec{V}_{g_\infty}$ is constant (see Eq. (2.3)). It is interesting to note that this simple model with constant K would not satisfy this upper condition if the Hesselberg-Sverdrup (1915) solution with constant geostrophic wind shear were used.

The solution to Eqs. (2.1) - (2.3) with boundary conditions given by Eq. (2.7) and the no-slip condition

$$\vec{V}(0) = 0 \tag{2.8}$$

is straightforward and has been given in Mahrt and Schwerdtfeger (1970). Using the complex notation $V = u + iv$, $i = \sqrt{-1}$, it may be written conveniently as

$$V = V_{g_0} + \hat{V}_g[1 - F(\alpha)e^{-\alpha z/D}] - [V_{g_0} + \hat{V}_g(1 - F(\alpha))]e^{-(1+i)z/D} \tag{2.9}$$

Here we define $F(\alpha) = \frac{4 - 2i\alpha^2}{4 + \alpha^4}$. It follows that

$$\frac{\partial V}{\partial z}\Big|_{z=0} = (1+i)V_{g_0} + \frac{1}{D}[\alpha F(\alpha) + (1+i)(1 - F(\alpha))]\hat{V}_g \tag{2.10}$$

depends upon $\hat{V}_g$ in a complicated way. Use of Eq. (2.6) then yields the simple relation

$$w_\infty = \frac{D}{2} \vec{k} \cdot \{\nabla \times (\vec{V}_{g_0} + L(\alpha) \hat{\vec{V}}_g)\} \tag{2.11}$$

where $L(\alpha) = \frac{\alpha[4 + \alpha(\alpha^2 - 2)]}{4 + \alpha^4}$ is shown in Fig. 2. The first term in (2.11) expresses the relation between vertical motion and vertically constant geostrophic wind; it agrees with the result of Charney and Eliassen (1949) when boundary condition (2.8) is utilized in their solution. The second term represents a correction due to the presence of vertical geostrophic wind shear. Roughly speaking, it is seen to be significant relative to the first term when the product $L(\alpha)|\hat{\vec{V}}_g|$ is comparable to $|\vec{V}_{g_0}|$.

In general, Eq. (2.11) is seen to relate the net mass outflow to the vorticity of the geostrophic wind at some intermediate height in the boundary layer where $\vec{V}_g = \vec{V}_{g_0} + L(\alpha)\hat{\vec{V}}_g$. This form of the relation is not, however, its most useful. The Boussinesq form for the thermal wind relation is

$$\frac{\partial \vec{V}_g}{\partial z} = \frac{1}{T_s} \frac{g}{f} \vec{k} \times (\nabla \times T) \tag{2.12}$$

where g is gravity and T absolute temperature. Combination of (2.12) with Eq. (2.3) yields

$$w_\infty = \frac{D}{2} \{\frac{1}{\rho_s f} \nabla^2 p_0' + M(\alpha) D \frac{g}{f T_s} \nabla^2 T_0\} \tag{2.13}$$

where $p_0' = p'(x, y, 0)$, $T_0' = T(x, y, 0)$ and $M(\alpha) = L(\alpha)/\alpha$. The first term contributes to rising motion near centers of low surface pressure in the usual manner, while the baroclinic correction term contributes to upward motion near centers of cold surface temperature.

Finally, application of (2.2), (2.3) and the hydrostatic relation allow (2.13) to be written in the form

$$w_\infty = \frac{D}{2} \{\vec{k} \cdot (\nabla + \vec{V}_{g_\infty}) + N(\alpha) D \frac{g}{f T_s} \nabla^2 T_0\} \tag{2.14}$$

where $N(\alpha) = (L(\alpha) - 1)/\alpha \leq 0$.

3. Applications

a. Middle latitudes

Contraty to the tropics, the atmosphere of middle latitudes is characterized by easily measured variations in pressure and temperature associated with synoptic-scale flow patterns. Equation (2.13) thus allows the estimation of w_∞ from measurements of p_0 and T_0 provided the depth of the thermal boundary layer D/α is known. Reference to the function $M(\alpha)$ in Fig. 2 suggests that for a given scale and magnitude of temperature variations, the baroclinic correction becomes small as the layer becomes thin ($\alpha \to \infty$).

The behavior of $M(\alpha)$ for $\alpha << 1$ also suggests that the baroclinic contribution to w_∞ is a maximum for thick baroclinic layers. However, in this case w_∞ would refer to the vertical motion at heights $z \sim \frac{D}{\alpha} >> D$ which are well above the region of turbulent influence in the real atmosphere. In addition, even with the constant K model, expansion of (2.3) for $Z = 0(D)$ shows that $\vec{V}_g(z)$ corresponds to the constant shear case studied by Hesselberg and Sverdrup (1915) for which it can be shown that the baroclinic correction is identically zero.

b. Tropics

The importance of the Coriolis parameter f in the derived results begins to appear when one considers the low latitude case where f is small. Equation (2.14) shows that an explicit dependence upon f appears in the baroclinic term. In addition both terms depend upon D which is implicitly dependent upon f in this model: $D = (2k/f)^{1/2}$.

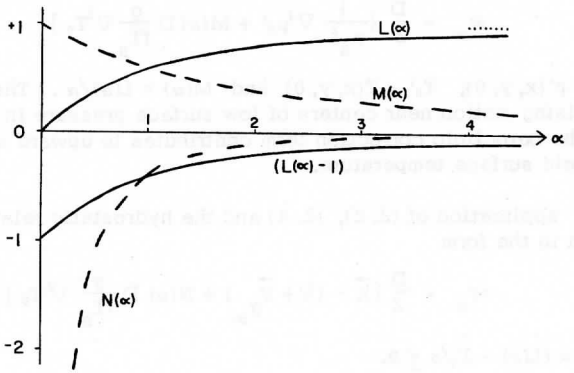


Fig. 2. Coefficients for baroclinic correction terms in expressions for vertical motions at top of boundary layer. α is inversely proportional to the relative depth of the surface baroclinic layer.

It is possible that D does not change with f(y) in the real atmosphere; (2.14) then shows that the "barotropic" term may vary only with the vorticity of the geostrophic wind above the boundary layer. On the other hand the baroclinic correction is seen to increase at low latitudes. This correction, written relative to the more easily observed wind above the boundary layer, contributes toward upward motion near warm centers of surface temperature. It would thus appear that the persistent gradients of sea surface temperature found in some tropical waters might, through their influence on low level baroclinity, cause systematic vertical motions in the lower atmosphere through this mechanism.

c. Rate of pressure work

From an energetic point of view a planetary boundary layer influences the free atmosphere above it through the mechanism of pressure work across the top of the boundary layer. This rate of energy change is given by

$$W = \overline{w'_\infty p'_\infty} \tag{3.1}$$

where p'_∞ is the deviation pressure at the outer edge of the boundary layer and the bar represents integration over the horizontal plane covering the area of interest. The sign in (3.1) has been chosen so that W reflects the sense of energy changes experienced by the free atmosphere. Thus negative W implies a flux of energy directed from the free atmosphere into the boundary layer, as would be found to occur in the simple "spin-down" process.

Substitution of w_∞ from Eq. (2.14) and the use of (2.2) allows us to write

$$W = \frac{D}{2}\left\{ \frac{1}{\rho_s f}\,\overline{p'_\infty \nabla^2 p'_\infty} + N(\alpha)\, D\, \frac{g}{fT_s}\,\overline{p'_\infty \nabla^2 T_0} \right\} \tag{3.2}$$

Integration of the first term by parts gives

$$\overline{p'_\infty \nabla^2 p'_\infty} = \overline{\nabla \cdot (p'_\infty \nabla p'_\infty)} - \overline{|\nabla p'_\infty|^2} \quad . \tag{3.3}$$

Gauss's Divergence Theorem states that the first term on the right-hand side is negligible if we consider a closed domain or if an open domain is chosen so that either p'_∞ or its normal derivative is small on the boundaries. The latter assumption is a reasonable approximation if the domain includes all of a disturbance field in its interior. In either case, only the second term will be considered in this analysis; use of (2.2) then gives

$$\overline{p'_\infty \nabla^2 p'_\infty} = -\rho_s^2 f^2 \overline{|\vec{V}_{g_\infty}|^2} \quad . \tag{3.4}$$

A similar procedure may be followed, with the additional use of Eq. (2.12), to transform the second term in (3.2). The result is

$$W = -\frac{D}{2}\rho_s f\left\{ |\vec{V}_{g_\infty}|^2 + (L(\alpha) - 1)\,\overline{\vec{V}_{g_\infty} \cdot \hat{\vec{V}}_g} \right\} \tag{3.5}$$

Using (2.3), this may be put in the more convenient forms

$$W = -\frac{D}{2}\rho_s f\left\{ \overline{|\vec{V}_{g_0}|^2} + L(\alpha)\overline{|\hat{\vec{V}}_g|^2} + (L(\alpha) + 1)\overline{\vec{V}_{g_0} \cdot \hat{\vec{V}}_g} \right\} \tag{3.6}$$

$$W = -\frac{D}{2}\rho_s f\left\{ L(\alpha)\overline{|\vec{V}_{g_\infty}|^2} - (L(\alpha) - 1)\overline{\vec{V}_{g_\infty} \cdot \vec{V}_{g_0}} \right\} . \tag{3.7}$$

We see from (3.6) that W is negative unless the bracketed terms sum to a negative quantity. The first term in brackets is that expected in the absence of baroclinity ($\hat{\vec{V}}_g = 0$) and is positive. This verifies that the "barotropic" boundary layer always extracts energy from the free atmosphere.

The remaining terms are corrections due to baroclinity. The first of these is also positive and gives a maximum contribution when the baroclinic layer is thin. The second correction term may be of either sign, depending upon the relation between the surface pressure and temperature fields. When the thermal wind vector opposes the surface geostrophic wind, $\vec{V}_{g_0} \cdot \hat{\vec{V}}_g$ is negative; depending upon its size, W may be made less negative or even positive. This kind of situation arises most naturally when the geostrophic flow patterns are of a special "equivalent barotropic" nature characteristic of warm surface lows or cold surface highs.

Equation (3.7) gives an idea of how large the baroclinity must be to reverse the energy flux W. Since $(L(\alpha) - 1)$ is negative, a necessary condition is that $\vec{V}_{g_\infty} \cdot \vec{V}_{g_0} < 0$; i.e., the direction of the geostrophic wind must essentially reverse itself over the depth of the baroclinic layer. Thus, it would appear that, for $\alpha = 0(1)$ and $|\vec{V}_{g_0}|/|\vec{V}_{g_\infty}| = 0(1)$, the energy flux will be directed from the boundary layer into the free atmosphere only in warm surface lows (or cold highs) which become highs (lows) at the top of the boundary layer. This kind of situation would not be found frequently in middle latitudes, but the strong influence of temperature gradients upon the weak pressure patterns of low latitudes suggests its possible occurrence there.

d. Sloping terrain

While the preceding discussion centered upon the effects of horizontal temperature gradients over a flat surface, it should be noted that such gradients may exist due to temperature variations in a direction normal to a sloping portion of the earth's surface. Figure 3 illustrates two coordinate systems which may describe such a situation. The (x, y, z) system is the conventional

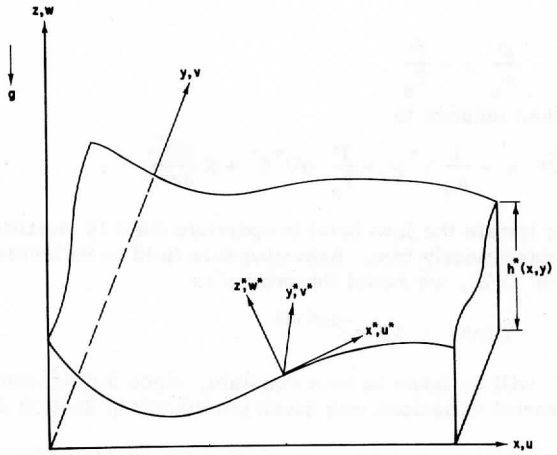


Fig. 3. Coordinate systems and notation for flow over sloping terrain

tangent plane system used earlier. The (x*, y*, z*) system is oriented so that z* is normal to the sloping ground, with the x* - y* plane tangent to the surface at the point shown. u*, v* and w* are the velocity components in the x*, y* and z* directions.

The equations for unaccelerated motion given in (2.1) for the (x, y, z) system must first be transformed into the (x*, y*, z*) system. The transformation

$$\nabla p' = \nabla^* p' - \frac{\partial p'}{\partial z} \nabla h^* \tag{3.8}$$

becomes, in the limit of small slopes ($|\nabla h^*| \ll 1$) and using the hydrostatic assumption

$$\nabla p' = \nabla^* p' + \rho' g \nabla^* h^* \quad . \tag{3.9}$$

Note h* is a function of x* and y*, in contrast to z*. Using the smallness of the ground slope to approximate $\vec{V}$ by $\vec{V}^*$, (2.1) becomes

$$f \vec{k}^* \times \vec{V}^* = -\frac{1}{\rho_s(z)} \nabla^* p' - \frac{\rho'}{\rho_s(z)} g \nabla^* h^* + K \frac{\partial^2 \vec{V}^*}{\partial z^{*2}} \quad . \tag{3.10}$$

As shown by Holton (1967), pressure variations in the equation of state $p = \rho RT$ may be ignored if the boundary layer is thin compared to the scale height,

so that we have

$$\frac{\rho'}{\rho_s} = -\frac{T'}{T_s} . \tag{3.11}$$

Equation (3.10) then reduces to

$$f\vec{k}^* \times \vec{V}^* = -\frac{1}{\rho_s}\nabla^* p' + \frac{T'}{T_s} g\nabla^* h^* + K\frac{\partial^2 \vec{V}^*}{\partial z^{*2}} . \tag{3.12}$$

Over sloping terrain the low-level temperature field is stratified such that $T' = T'(z^*)$ is approximately true. Assuming this field to be confined to a layer of e-folding depth D/α, we model the profile as

$$T'(z^*) = T_0' e^{-\alpha z^*/D} . \tag{3.13}$$

The quantity T_0' will be taken to be a constant, since a discussion of the effects of its horizontal variations was given previously in Section 2.

With the substitution of (3.13), we see that Eq. (3.12) reduces to the forms with variable geostrophic wind in (2.1 - 2.3) if we make the correspondences

$$-\frac{1}{\rho_s}\nabla^* p' = f\vec{k}^* \times \vec{V}_{g_\infty}^* \tag{3.14}$$

and

$$\frac{T_0'}{T_s} g\nabla^* h^* = -f\vec{k}^* \times \hat{\vec{V}}_g^* . \tag{3.15}$$

As shown by Holton, the continuity equation for the case of large horizontal scales and small slopes reduces to the form

$$\nabla^* \cdot \vec{V}^* + \frac{\partial w^*}{\partial z^*} - \frac{c_v}{c_p}\frac{1}{H_s}\vec{V}^* \cdot \nabla^* h^* = 0 . \tag{3.16}$$

The third term arises from density variations experienced by a parcel moving upslope or down-slope and can be shown to be small by the following analysis: assuming U^*, L^* as characteristic velocity and length scales, the first term has a maximum size of U^*/L^* in the boundary layer, while the third term has a maximum magnitude $(U^*/L^*)(\delta h^*/Hs)$, where δh^* is the ground height variation over a distance L^*. Thus, when the topographical variations are small compared with the scale height Hs (as must usually be the case for balanced motions to occur), the third term may be neglected and Eq. (3.16) becomes

$$\nabla^* \cdot \vec{V}^* + \frac{\partial w^*}{\partial z^*} = 0 . \tag{3.17}$$

Neglecting variations in Ts as it appears in (3.15) allows our approximate equations to reduce to the forms (2.1 - 2.4) given in Section 2. It follows that

the flow in an outwardly normal direction from the boundary layer is found approximately by replacing T_0 in Eq. (2.14) by $T_0'(h^*/(D/\alpha))$:

$$w_\infty = \frac{D}{2}\left\{ \vec{h}\cdot(\nabla^* \times \vec{V}^*_{g_\infty}) + (L(\alpha) - 1)\frac{g}{f}\frac{T_0'}{T_s}\nabla^{*2}h^* \right\} \tag{3.18}$$

The first term is the outflow expected in the absence of boundary slope or temperature gradients along the boundary. The second term represents the contribution from the slope variations when the temperature along the slope is constant. We see that if $\delta h^*/(D/\alpha) = 0(1)$ and $T_0'/T_0 = 0(1)$ then the slope correction may be comparable with that due to temperature differences along a flat terrain as found in Section 2.

The sense of this correction is such that, under lapse conditions $T_0' > 0$, outward motion ($w^* > 0$) occurs when h^* is a local maximum. Conversely, increasing termperature normal to the surface ($T_0' < 0$) yields outward motion where h^* is a local minimum. Thus with lapse conditions outward motion is found over topographical maxima, while it is found over the valleys with a strong inversion ($T_0' < 0$, such as is found in Antarctica). These results seem in qualitative agreement with the patterns recently given by Mahrt and Schwerdtfeger (1970).

Finally, the vertical motion w_∞ in the (x, y, z) system is given in this case simply by

$$w_\infty = w^*_\infty + \vec{V}_{g_\infty} \cdot \nabla h^* \tag{3.19}$$

where w^* is given in (3.18) and the second term represents the usual geometrical transformation factor evaluated at the top of the boundary layer.

4. Concluding remarks

The discussions of the previous sections suggest that low-level temperature fields may significantly alter the patterns of mass convergence in the boundary layer over sloping or flat ground. The derived formulas may of course be criticized quantitatively, since they have been obtained from an idealized model. Nevertheless, they appear to provide useful qualitative guidelines which can be considered to be the first steps in a more general theory.

The kinds of improvements to be made are several in number. First, the assumption of constant eddy viscosity K should be dropped and replaced by a more realistic formulation utilizing some sort of mixing length theory. However, an immediate consequence of this approach would be that the nonlinearity of the formulation would rule out the linear independence of the barotropic and baroclinic solutions.

A simpler alternative to this would be to retain the Ekman assumption in the bulk of the boundary layer and model the flow near the earth by a conventional Prandtl surface layer. However, closer examination reveals that the Prandtl constraint that the wind shear and velocity vectors be strictly parallel would be inconsistent with the changing pressure force vector characteristic of a general baroclinic flow.

The assumption of unaccelerated flow used in this analysis is a conventional one, but it is often a poor one in middle latitudes, where baroclinic frontal zones are usually accompanied by appreciable accelerations. At lower latitudes where the baroclinic effect is potentially larger, significant accelerations may be induced by small-scale changes in the baroclinity or by the rapid vanishing of f, the Coriolis parameter. Finally, the adiabatic expansions associated with the boundary layer vertical motions may represent a feedback mechanism linking the hydrodynamics and thermodynamics. This seems particularly true in view of the close relation between vertical motion and the temperature field in the solutions found here.

In closing, it would thus appear that final confirmation of the relevance of these ideas awaits the development of three-dimensional numerical models both in the boundary layer and the free atmosphere above it.

Acknowledgments

This research was supported by E. S. S. A. grant E-230-68-G and N. S. F. grant GA-12112 at the University of Wisconsin.

References

Businger, J. A. and Jason, K. S. Ching, 1968: The response of the planetary boundary layer to time varying pressure gradient force, Journal of Atmospheric Science, 25, 1021-1025.

Charney, J. G. and A. Eliassen, 1949: A numerical method for predicting the perturbations of the middle-latitude westerlies, Tellus 1 (2), 38-54.

Ekman, V. W., 1905: On the influence of the earth's rotation on ocean currents, Arkiv. Matem., Astr. Fysik Stockholm, 2 (11).

Hesselberg, T. and H. U. Sverdrup, 1915: Die Riebung in der Atmosphere, Veroeff. Geophys. Inst., Univ. Leipzig, 2nd Series, 1.

Holton, J. R., 1967: The diurnal boundary layer wind oscillation above sloping terrain. Tellus 19, 199-205.

Kuo, H., L., April, 1969: Axisymmetric flows in the boundary layer of a maintained vortex, Plan. Circul. Proj. Report No. 15, Dept. of Geophysical Sciences, U. of Chicago.

Lettau, H. L., 1967: Small to large-scale features of boundary-layer structure over mountain slopes, in "Proceedings of the Symposium on Mountain Meteorology," Atmospheric Science Paper No. 122, Colorado State University.

Mahrt, L. J. and W. Schwerdtfeger, 1970: Ekman spirals for exponential thermal wind, to be published in Boundary Layer Meteorology.

Mendenhall, B. R., December, 1967: A statistical study of frictional wind veering in the planetary boundary layer, Atmospheric Science Paper No. 116, Dept. of Atmospheric Science, Colorado State University.

A NOTE ON BOUNDARY LAYER DYNAMICS AND THE ITCZ

J. R. Holton and J. M. Wallace
Department of Atmospheric Sciences, University of Washington

and

John A. Young
Department of Meteorology, University of Wisconsin

1. Introduction

The intertropical convergence zone (ITCZ) is the narrow east-west band of vigorous cumulonimbus convection and heavy precipitation which separates the trade wind regimes of the northern and southern hemispheres. Within this region precipitation exceeds evaporation by a factor of 2 or more, the excess being provided by the moist, converging trade current. The convergent, low level flow in this region is marked by strong cyclonic vorticity.

The spatial and temporal variability of the ITCZ are well documented in the recent literature (Bjerknes et al., 1969; Hubert et al., 1969; Kornfield and Hasler, 1969). The ITCZ is particularly persistent and sharply defined over the North Pacific and Atlantic, between 5 and 10 N. It occasionally appears over the SE Pacific at 5 - 10 S. The influence of large land masses apparently inhibits the development of a well defined ITCZ at other longitudes. There are seasonal and year to year variations in position and intensity. For example, in the Western Atlantic the zone shifts from just north of the equator in winter and spring to around 8 - 10 N in summer. The Western Pacific ITCZ is well defined during some years and virtually absent during others.

The ITCZ is rarely, if ever, found at the equator, except in the vicinity of large land masses. The conspicuously cloud free "equatorial dry zone" has been viewed by Bjerknes et al. (1969) and others as a consequence of cool, upwelling ocean water associated with the equatorial current system. Pike (1970) has recently demonstrated the development of a low-latitude steady convergence zone in a coupled atmosphere-ocean model.

While the coincidence between equatorial upwelling and clear skies is quite remarkable over the eastern oceans, the equatorial day zone occurs in the absence of a well defined minimum of sea surface temperature over the western oceans. An effort to provide an ITCZ mechanism less dependent upon sea-surface temperature was made by Charney (1966). This theory characterized

the ITCZ as the result of a CISK process in which boundary-layer convergence on a large scale maintains the supply of latent energy which drives the system through organized convection. Charney concluded that boundary layer convergence (most efficient away from the equator) and the effects of conditional instability and available moisture (a maximum at the equator) can jointly produce a steady, zonally symmetric convergence zone at low latitudes.

However, when viewed at a particular instant in time, rather than in the usual climatological sense, the ITCZ exhibits a somewhat more complicated structure. Inspection of individual satellite photographs shows that it can only rarely be identified as a long, unbroken band of heavy cloudiness. Rather, it is usually made of a number of "cloud clusters," separated by large expanses of relatively clear skies. Chang (1970) and Reed (1970) have shown that cloud clusters are, in turn, manifestations of large scale westward propagating wave disturbances. They are marked by heavy precipitation (> 2 cm day^{-1}), which requires strong low level convergence ($\sim 10^{-5}$ cm sec^{-1}), strong ascent at middle levels (~ 2 cm sec^{-1}) and strong upper level divergence. Fujita et al. (1969) and Williams (1970) have reported that divergent motions of this size are typical in the vicinity of cloud clusters. They have also shown that the clusters represent concentration of cyclonic vorticity at low levels and anticyclonic vorticity at 200 mb. At low levels the convergence is mainly in the meridional wind component. The clear areas between clusters are marked by weak subsidence together with vertical distributions of vorticity and divergence opposite in sign to those within the clusters.

Thus in a climatological sense, the ITCZ may be considered to be the locus of the cloud clusters associated with a train of westward propagating wave disturbances. The low level convergence and cyclonic shear within the zone in the long term mean is an average of periods of strong convergence and cyclonic vorticity, coincident with the passage of cloud clusters, together with periods of weak divergence and anticyclonic vorticity which mark the passage of clear areas. The concentration of precipitation, divergence, and vorticity which we call the ITCZ is not simply a steady atmospheric state, but should be viewed as a consequence of the dynamical properties of large scale equatorial wave disturbances.

The purpose of this note is to propose a mechanism which can account for the strong latitudinal preference in travelling wave disturbances at low latitudes, and to qualitatively begin to answer the following questions: 1) Why is the ITCZ so well defined, even in the absence of a distinctive sea-surface temperature field? 2) How might we account for seasonal and yearly variations in the intensity and latitude of this band? Finally, we wish to comment upon the problem of relating the low-level vertical motion to the vorticity of the horizontal flow field at low latitudes.

2. Theoretical Background

Yamasaki (1970) and Hayashi (1970) have recently investigated the possibility that the CISK mechanism is the energy source for planetary waves in the tropics. Those studies each have certain weaknesses. Yamasaki treated only two-dimensional quasi-geostrophic flows, and Hayashi did not properly take into account the role of boundary layer friction. However, their analyses do provide at least some theoretical evidence that latent heat release is the driving mechanism for the observed easterly waves. Further evidence is provided by the diagnostic model of Holton (1970) in which it is shown that a specified heating distribution with cloud cluster scale can force an atmospheric response very similar to that of the observed waves.

An essential feature of the CISK mechanism is the existence of low level convergence to provide the necessary moisture to maintain the large scale convection pattern. For midlatitude quasi-geostrophic waves, Charney and Eliassen (1949) have shown that the convergence in the Ekman layer is proportional to the vorticity at the top of the boundary layer. However, such a simple relationship does not hold near the equator. In fact, it is shown below that for a propagating wave on the equatorial β-plane, the Ekman layer solution is singular along a critical latitude at which the angular frequency of the wave is equal to the Coriolis frequency. As a consequence of this singularity the convergence in the boundary layer will tend to be concentrated at the critical latitude.

To illustrate this singular behavior it is sufficient to consider the linearized equations for a barotropic disturbance in a resting atmosphere on the equatorial β-plane. If we nondimensionalize the equations using length scale $L \equiv (c/\beta)^{1/2}$, time scale $\tau \equiv (c\beta)^{-1/2}$, and depth scale H, where $c \equiv gH$ is the horizontal velocity scale, $\beta \equiv 2\Omega/a$ with Ω the angular velocity of the earth and a the radius of the earth, and H is the mean depth of the barotropic fluid,* we obtain,

$$\frac{\partial u}{\partial t} = -\frac{\partial \phi}{\partial x} + yv + D\frac{\partial^2 u}{\partial z^2} \tag{1}$$

$$\frac{\partial v}{\partial t} = -\frac{\partial \phi}{\partial y} - yu + D\frac{\partial^2 v}{\partial z^2} \tag{2}$$

$$\frac{\partial \phi}{\partial t} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \tag{3}$$

where $D = KH^{-2}(c\beta)^{-1/2}$ is the nondimensional vertical eddy viscosity, assumed constant.

*Matsuno (1966) has shown that the same equations apply in a stratified atmosphere provided that H is interpreted as the "equivalent" depth of tidal theory.

The system (1) - (3) may be analyzed by conventional boundary layer techniques. We assume that solutions exist of the form

$$\begin{Bmatrix} u \\ v \\ \phi \end{Bmatrix} = \mathrm{Re}\left[\begin{Bmatrix} u^I(y) + u^B(y, z') \\ v^I(y) + v^B(y, z') \\ \phi(y) \end{Bmatrix} e^{i(\omega t + kx)} \right] \tag{4}$$

where $z' = zD^{-1/2}$ is a stretched coordinate. Substituting (4) into (1) - (3) and taking the limit $D \to 0$ while holding z constant we obtain the interior equations

$$i\omega u^I = -ik\phi + yv^I \tag{5}$$

$$i\omega v^I = -\frac{\partial \phi}{\partial y} - yu^I \tag{6}$$

$$i\omega \phi = -iku^I - \frac{\partial v^I}{\partial y} \tag{7}$$

which is just the system discussed by Matsuno (1966).

We next substitute (4) into (1) - (3) and subtract out the interior variables by applying (5) - (7). The result is the boundary layer system

$$i\omega u^B = yv^B + \frac{\partial^2 u^B}{\partial z'^2} \tag{8}$$

$$i\omega v^B = -yu^B + \frac{\partial^2 v^B}{\partial z'^2} \tag{9}$$

To solve (8)-(9) we must apply appropriate boundary conditions. We require that $u, v = 0$ at $z' = 0$ and $u \to u^I$, $v \to v^I$ as $z' \to \infty$. Thus

$$\begin{array}{lll} u^B = -u^I & v^B = -v^I & z' = 0 \\ u^B \to 0 & v^B \to 0 & z' \to \infty \end{array} \tag{10}$$

The solution of (8)-(9(subject to the boundary conditions (10) is

$$v^B(z') = A(y)e^{-r_1 z'} + B(y)e^{-r_2 z'} \tag{11}$$

$$u^B(z') = iA(y)e^{-r_1 z'} - iB(y)e^{-r_2 z'} \tag{12}$$

where(assuming $\omega > 0$)

$$r_1 = \begin{cases} (1+i)\sqrt{\frac{\omega+y}{2}}, & y > -\omega \\ (1-i)\sqrt{\frac{\omega+y}{2}}, & y < -\omega \end{cases} \qquad r_2 = \begin{cases} (1+i)\sqrt{\frac{\omega-y}{2}}, & y < \omega \\ (1-i)\sqrt{\frac{y-\omega}{2}}, & y > \omega \end{cases}$$

and

$$A(y) = -\frac{1}{2}(v^I - iu^I)$$

$$B(y) = -\frac{1}{2}(v^I + iu^I)$$

or in terms of ϕ

$$A(y) = -\frac{i}{2}(k\phi + \frac{\partial\phi}{\partial y})/(y+\omega) \tag{13}$$

$$B(y) = -\frac{i}{2}(k\phi - \frac{\partial\phi}{\partial y})/(y-\omega) \quad . \tag{14}$$

Matsuno showed that the solutions of (5) - (7) can be expressed in terms of ϕ as

$$\phi = +\frac{1}{2}(\omega_n - k)\psi_{n+1} - n(\omega_n + k)\psi_{n-1} \tag{15}$$

where $\psi_n = e^{-1/2y^2} H_n(y)$, with $H_n(y)$ the Hermite polynomial of order n.

ω_n must satisfy

$$\omega_n^2 - k^2 + \frac{k}{\omega_n} = 2n+1 \qquad n = 0, 1, 2 \ldots \quad . \tag{16}$$

The vertical structure of the solutions (11) and (12) exhibits interesting behavior, particularly near the "resonance points" where $y = \pm\omega$. Unlike ordinary resonances, the solutions do not possess infinite amplitudes at these points. Instead, it may be verified that the coefficients A(y) and B(y) are everywhere finite and smoothly varying, in accord with the character of the flow above the boundary layer. *

However, since the depth of the boundary layer is seen to be proportional to $|y-\omega|^{1/2}$ as $y \to \omega$, the vertically integrated mass transport in the boundary

* Charney (private communication) has recently obtained a solution with similar properties by assuming that the flow above the boundary layer is in geostrophic balance, in contrast to the boundary conditions used in Eq. (10).

layer possesses a singularity of the order $|y - \omega|^{1/2}$ at $y = \omega$. (Similar arguments hold for $y = -\omega$.) This property of boundary layers in contained, rotating viscous fluids has been previously noted (Greenspan (1968). Its direct importance in the earth's atmosphere seems uncertain, since other factors may restrict the boundary layer thickness.

Another important feature of the boundary layer solution is the phase shift which occurs at the critical latitudes where $y = \pm \omega$.* This phase shift is capable of producing large meridional variations in the velocity components, and hence large values of horizontal convergence and vorticity.

The origin of this phase shift can be understood by consideration of (11) and (12). The first mode of each solution varies as $e^{-r_1(y)z'}$, and corresponds to the upward diffusion of momentum away from the surface at all latitudes. It may be thought of as a vertically damped viscous wave of the sort found above an oscillating plane in Stokes' problem of the second kind (Schlichting (1960)) which has been modified by Coriolis forces.

The second mode behaves in a similar way at low latitudes where Coriolis forces are small. However, at higher latitudes $|y| \geq |\omega|$ the Coriolis forces produce a downward phase propagation which accounts for a 90° phase shift for this mode across the critical latitude. In the low frequency limit $|\omega| \ll |y|$, the downward propagation matches the upward propagation in such a way that the classical Ekman structure of a vertically damped standing wave results.

Some characteristics of the solutions (11)-(12) are illustrated for the $n = 0$ and $n = 1$ modes in Figs. 1-3. According to Holton (1970), the $n = 1$ mode most closely resembles the maximum response mode for a moving ITCZ heat source. On the other hand, the $n = 0$ mode possesses strong cross-equatorial flow and is thus further from geostrophic balance.

Figure 1 illustrates the flow patterns which occur at low latitudes for the $n = 0$ mode. We see that deviations from the classical Ekman boundary flow are responsible for a circulation center lying between those associated with the actual and geostrophic flows above the boundary layer. Similar deviations for the $n = 1$ mode (not shown here) were noted only near the equator.

Figure 2 shows the pattern of upward motion

$$w^B(y, \infty) = -D^{1/2}\left[ik \int_0^\infty u^B dz' + \int_0^\infty \frac{\partial v^B}{\partial y} dz'\right]$$

---

*It is interesting to note that Lettau (1967) has noted that similar phase shifts seem to occur in diurnally varying flow over sloping terrain at the "critical latitude" of 30° N.

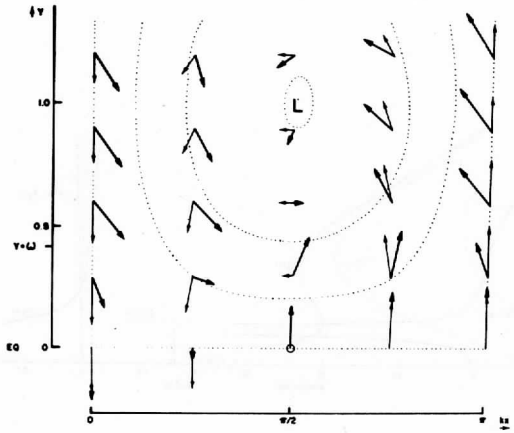


Fig. 1. Flow patterns for n = 0 node. Thick arrows represent boundary layer flow at $z' = 0.6$ exaggerated three times relative to flow above boundary layer (thin arrows). Dotted lines represent disturbance geopotential field.

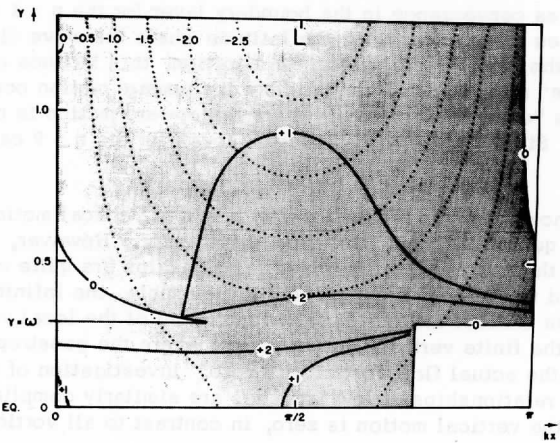


Fig. 2. Vertical motion at top of boundary layer for n = 1 mode. Shading denotes area of upward motion, with intensity (in nondimensional units of 10 $D^{1/2}$) indicated by solid lines. Dotted lines are disturbance geopotential.

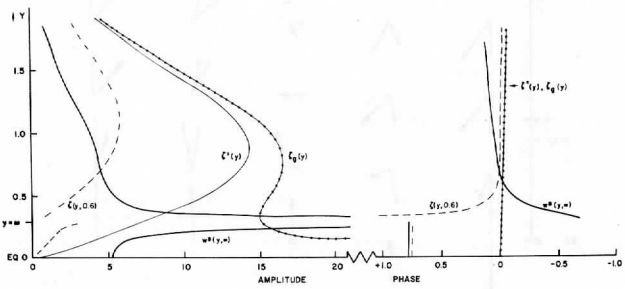


Fig. 3. Vorticity and boundary layer vertical motion relationships for n = 1 mode. Vertical motion amplitude is expressed in nondimensional units of $2.5D^{1/2}$. Phase is in radians.

produced by mass convergence in the boundary layer for the n = 1 mode. This motion possesses a maximum along the latitude where the wave disturbance frequency equals the Coriolis parameter. In addition, this latitude also separates the "midlatitude" regime to the north (where the upward motion occurs along the trough) from the "equatorial" regime (where the upward motion is not centered in the trough). Similar characteristics were found for the n = 0 case, not shown here.

Figure 3 shows that the proportionality between vertical motion and vorticity at any level is quite good well away from the equator. However, near and equatorward of the critical latitude these relationships are quite complicated, both with regard to amplitude and phase. For example, the infinite vertical motion found at the critical latitude is far equatorward of the local vorticity maxima. At the equator the finite vertical motion occurs where the geostrophic vorticity is infinite and the actual flow vorticity is zero. Investigation of the n = 0 case shows that the relationships at low latitudes are similarly complicated; e.g., at the equator the vertical motion is zero, in contrast to all vorticity quantities.

In closing, it should be emphasized that this idealized model treats the disturbance as a pure mode of single frequency, ω. In reality, atmospheric waves are transient disturbances whose energy is distributed over a continuous frequency spectrum. The boundary layer convergence should thus be distributed over a finite band of latitudes rather than restricted to a single latitude. Presumably, then, the maximum convergence would occur at the latitude where the

Coriolis frequency corresponded to the frequency of maximum power in the wave spectrum.

3. Discussion

From the foregoing discussion it is apparent that the boundary layer convergence associated with a spectrum of equatorial waves whose Doppler shifted frequencies (i.e., the frequencies relative to the mean zonal wind) lie between ω_1 and ω_2 will be mainly confined to the band of latitude in which the value of the Coriolis parameter lies between ω_1 and ω_2. If we associate boundary layer convergence with convection, it follows that the latitudinal distribution of rainfall produced by the passage of equatorial waves is to some extent determined by the frequency spectrum of the waves.

For equatorial waves which are forced by latent heat release, the above argument can be applied in reverse: If conditions favorable for convection (e.g., warm ocean temperatures, availability of moisture) are confined to a band of latitude in which the Coriolis parameter lies between f_1 and f_2, those waves whose Doppler shifted frequencies lie between f_1 and f_2 will tend to be favored. In other words, the frequency distribution of equatorial waves driven by latent heat release is to some extent determined by the latitudinal distribution of ocean temperature and moisture in the zonal mean.

Thus we conclude that there is a relationship between the frequency distribution of equatorial wave disturbances and the latitudinal distribution of precipitation in the waves. At this point we can make no statement as to which is cause and which is effect. In order to understand how this type of interaction is likely to result in a convergence zone in the long term mean, we must consider refinements to the linear theory discussed in the previous section.

The critical latitude represents a region of large velocity gradients and hence nonlinear effects. It is quite likely that the Reynolds stresses associated with the moving wave could force a mean circulation pattern near the critical latitude, in much the same manner that steady currents are produced by oscillating tidal flows in variable estuaries (Abbott, 1960).

However, in the tropical atmosphere, the dominant nonlinearity would seem to be associated with the release of latent heat, where temperature changes accompanying rising motion are quite different from those associated with sinking motion. As a result, it seems likely that the low level convergence associated with the waves will be more strongly concentrated near the "critical latitude" than the corresponding divergence. This means that in the long term mean there would be convergence at the latitude of maximum wave activity and divergence at other latitudes. Because of the earth's rotation, this steady state circulation may tend to produce a concentration of vorticity in the convergence zone.

Observational Evidence

Observational verification of the above mechanism would require, as a starting point, a large sample of simultaneous measurements of Doppler-shifted wave frequency and latitude of the ITCZ, preferably from a number of geographical regions, and during several seasons and years. The spectral data currently available are not adequate for this purpose. For the present, we can only indicate that the range of "critical latitudes" inferred from the frequency spectrum of tropical wave disturbances is not inconsistent with the observed position of the ITCZ.

As evidence of this we note that spectral peaks in the range .20 - .25 cpd have been observed at a large number of tropical stations and during a number of different time periods (Rosenthal, 1960; Yanai et al., 1968; Wallace and Chang, 1969; Chang, Morris and Wallace, 1970). In at least some of these cases the mean zonal wind speed in the boundary layer was sufficiently small compared to the wave speed so that the Doppler shifted frequency was approximately the same as the ground based frequency. The range of "critical latitudes" corresponding to this frequency band is 6 - 7°, which is close to the most frequent ITCZ position.

The observed lack of stationarity in time series of tropical data may be linked to variations in position and intensity of the ITCZ. There have been extended periods when the frequency spectrum of tropical disturbances in the western Pacific has failed to show any distinct peak in the .20 - .25 cpd frequency range. At such times most of the variance is contributed by waves with much lower frequencies, and the Doppler shifted spectrum resembles red noise. If the above arguments are correct, we should not expect to see a sharply defined ITCZ in such situations. It is then perhaps significant that the satellite maps of mean cloud amounts presented by Bjerknes et al. do not show a well defined ITCZ in the Western Pacific during the period January 1963 - July 1964, a time marked by predominantly low frequency waves in this region (Wallace and Chang, 1969, Fig. 9). The same maps show a well defined ITCZ during 1967, a period marked by a strong spectral peak in the .20 - .25 cpd range.*

4. Concluding Remarks

It should be emphasized that the proposed relationship between ITCZ position and the frequency spectrum of equatorial wave disturbances represents but one possible link in the complex chain of interactive mechanisms which control tropical climate. A complete explanation of the ITCZ and its spatial and temporal variations must also take account of proper modelling of the CISK process, air-

*A detailed study of this period is now in progress at the University of Washington.

sea interaction and topographical influences. In a similar manner, a complete explanation of the frequency spectrum of tropical wave disturbances must also take into account many factors not considered here. The primary value of the simple interactive scheme proposed above is to provide a conceptual framework which may be helpful in interpreting observational data and numerical models of the tropical general circulation.

Acknowledgments

This research was begun while one of us (J. Y.) enjoyed a stay at the University of Washington as a visiting scientist. Research was sponsored under NSF grant GA 629-X2 at the University of Washington and ESSA grant E-230-68-G and NSF grant GA-12112 at the University of Wisconsin.

References

Abbott, M. R. 1960: Boundary layer effects in estuaries, J. Mar. Res. 18, 2, 83-100.

Bjerknes, J., L. J. Allison, E. R. Kreins, F. A. Godshall, G. Warnecke, 1969: Satellite mapping of the Pacific tropical cloudiness. Bulletin of the A. M. S., 50, 313-322.

Chang, C. P., 1970: Westward propagating cloud patterns in the tropical Pacific as seen from time-composite satellite photographs. Journal of the Atmospheric Sciences, 27, 133-138.

Charney, J. G., 1966: Some remaining problems in numerical weather prediction. Advances in Numerical Weather Prediction, The Travelers Research Center, Inc., 61-70.

Charney, J. G., and A. Eliassen, 1949: A numerical method for predicting the perturbations of the middle-latitude westerlies. Tellus, 1, 38-54.

Fujita, T. T., K. Watanabe and T. Izawa, 1969: Formation and structure of equatorial anticyclones caused by large-scale cross-equatorial flows determined by ATS-I photographs. Journal of Applied Meteorology, 8, 649-677.

Greenspan, H. P., 1968: The Theory of Rotating Fluids. Cambridge University Press, 327 pp.

Hayashi, Y., 1970: Absorption of energy and momentum due to large-scale upward propagating waves generated by latent heat released in the tropical troposphere. Submitted to Journal of the Meteorological Society of Japan.

Holton, J. R., 1970: On the theory of easterly waves. Proceedings of the Symposium on Tropical Meteorology, Honolulu.

Hubert, L. F., A. F. Krueger and J. S. Winston, 1969: The double intertropical convergence zone—fact or fiction. Journal of the Atmospheric Sciences, 26, 771-773.

Kornfield, J. and A. F. Hasler, 1969: A photographic summary of the earth's cloud cover for the year 1967. Journal of Applied Meteorology, 8, 687-700.

Matsuno, T., 1966: Quasi-geostrophic motions in the equatorial area. Journal of the Meteorological Society of Japan, 44, 25-43.

Pike, A. C., 1970: The Inter-tropical convergence zone studied with an interacting atmosphere and ocean model. Scientific Report No. 2, Rosenstiel School of Marine and Atmospheric Sciences, Univ. of Miami, 25 pp.

Reed, R. J., 1970: Structure and characteristics of easterly waves in the equatorial western Pacific during July-August, 1967. Proceedings of the Symposium on Tropical Meteorology, Honolulu.

Schlichting, H., 1960: Boundary Layer Theory. New York, McGraw-Hill Book Co., Inc., 647 pp.

Williams, K., 1970: Characteristics of the wind, thermal and moisture fields surrounding the satellite observed mesoscale trade wind cloud clusters of the Western North Pacific. Proceedings of the Symposium on Tropical Meteorology, Honolulu.

Yamasaki, M., 1970: Large-scale disturbances in a conditionally unstable atmosphere in low latitudes. Papers in Meteor. Geophys., 20, (in press).

REAL TIME ATS DATA PROCESSING USING WIDE BANDWIDTH VIDEO STORAGE AND DISPLAY*

Robert Krauss

This report presents the results of a brief study of the advantages of a wide bandwidth video storage and display system for processing ATS picture data. The system can provide great speed advantages in obtaining winds and quality eye pleasing motion pictures, as well as serving to monitor ATS camera operation in real time. Its main disadvantages are its limited resolution elements per picture and lack of photometric fidelity and range. Compensation for the disadvantages is attainable with other components of the SSEC data processing system. Set-up cost will be high, although operating costs could be justified by output, given high system use. Thus, the major decision is whether the potential utility of such an addition to SSEC justifies the initial cost.

Before describing the display system and its operation, it is useful to define the elements separately:

Disk Memory—This is generally I/O bandwidth fixed by the bandwidth of the TV system to which the memory is attached. Because of the necessity for constant rotation speed and precision head alignment, the mechanical and electrical problems inherent in variable I/O rates prohibit connecting a megahertz disk into the nominal kilohertz ATS-SSEC data chain in any useful fashion. Thus, any use of wide bandwidth video presupposes an interface or transducer which, with a minimum of noise generation, can bandwidth expand or compress by two to three orders of magnitude.

Videotape Recorder—A standard TV recorder to make a permanent record of the data on the analog disk memory.

TV Camera—A standard TV camera suitable for introducing hard copy pictures, overlays, titles and annotations into the system.

---

*This work was supported on ESSA Grant E-230-68G.

Graphics Display—Adage display at the University of Wisconsin Computing Center. Essentially a line plotter like the Calcomp, with the added advantage of having a TV screen and an operator at the controls with a light pen to provide decision-making and feedback capability to the 1108 computer. Current SSEC plans call for including Adage display capabilities in conversion of the CDC 3600 plotting routines to the 1108.

Storage Tube—The interface for bandwidth compression and expansion will be a storage vidicon with 400 - 500 line storage capability extending for up to 5 minutes before requiring refreshment. This is where the resolution and photometry restrictions appear. The tube will store data on its face at one speed and read out at another.

Video Control Center—Serves to transfer the high bandwidth video between elements of the system. It contains the timing delays which shift the images up, down, or sideways and allows choosing to display disk picture tracks separately, superimposed, alternately for blink comparison, or sequentially for weather motion study and movie editing. One can simply look at the displays, use the videotape recorder to make permanent copies from disk playback or make movies or still pictures. Alignment and editing are quick and efficient as pictures are transmitted in seconds rather than minutes, and erasure of errors from the disk is a simple matter of pressing a button to delete a particular frame track.

Our present sources of data, aside from hard copy, are a) high resolution analog data recorded at the ground station from which precision digitizing may be done, b) low resolution analog data obtained in real time over telephone lines, which can be digitzed to reasonable precision if the satellite spin rate and sample rate are known, and c) digital data recorded directly from core at the ground station.

From the real time analog data, negatives can be exposed on the SSEC precision display and immediately developed. Cloud displacement measurements can be started on the original negatives within an hour after exposure to obtain winds, and the daily set of negatives can be aligned and exposed for movies on the next day. Thus, we can get winds the same day, and movies within one to two days after the actual weather. Project Weathernet will have stills ready within an hour, contact printed from negatives produced with real time data. These will be generated at three-hour intervals. Since the real time data is of lower quality and has no A/D reference frequency, we have at present no way of knowing how precisely winds can be determined by means of the several analog and digital procedures at our disposal. Studies presently underway should settle that problem.

The high resolution analog and digital data has the best geometric and photometric fidelity, but the highest accuracy is obtainable only through a computer at greater expense than analog methods. The results are presently

displayed in the form of dumps on the line printer or as contours on the Calcomp plotter. While very precise, the data are difficult to interpret in nonpictorial form. Conversion to an analog signal for generating a precision display transparency is time-consuming and really serves only as an end result.

The other advantage of the computer besides its photometric and geometric precision is its mapping or picture rectification ability. Any display of the raw data will contain geometric distortions due to satellite subpoint motion, elliptical orbit, equal angle scan, and digitizing rate changes. In any measurement of cloud displacement, these well-documented distortions generate a hierarchy of correction factors, producing, in sum, a transfer function that will allow correction of any given measurement of cloud displacement to the desired accuracy within the limits of ATS camera resolution. We also have means of processing and filtering the signal to enhance those characteristics of size or brightness we desire to measure. Thus, wind and cloud measurement is no serious problem. In making movies, however, we are in trouble, since analytical correction factors can't be used easily outside the computer. One can align small portions of successive pictures satisfactorily by the blink method to produce decent movies. This is because the distortion errors accumulated on the transparency are small over small distances. In the case of half or full earth pictures, however, any means of bending, tilting, twisting, or stretching the negative to obtain complete coincidence with the previous picture is doomed to failure since the subject (earth) has actually rotated in some fashion and presents a different aspect. Such distortions of the negative may minimize the sickening lurches of the earth's landmarks but will never keep all of them perfectly stable over a 12 hour period no matter how carefully the picture alignment is done. Thus, a really aesthetically good movie with all of the earth standing completely still is not possible without the aid of a computer to map the different aspects of the earth to a standard one, i.e., to rotate the earth the correct amount in each picture. Such a mapping requires random access to points in a two-dimensional array, a concept totally removed from the linear data streams in which the analog data exists. The computer, with a large random access digital drum memory of several million words appears the only solution. Unfortunately, mapping a picture of $2400 \times 8192 \approx 2 \times 10^7$ elements can be very costly. If one considers doing this for every frame of a movie, the cost would be prohibitive, even though the earth would no longer jump around. We can cheat a bit, using analytical mapping techniques on every 100th sample and linear interpolation on the others in between. This may cut down on computer time by a factor of 10 or 20, but still requires time-consuming handling of each picture element. It is certainly a technique worth looking into, however.

Another approach is to use the wide band TV system we have discussed. The low number of elements per picture necessitated by the storage vidicon can be turned to our advantage since now we have only $500 \times 900 = 4.5 \times 10^5$ elements to map—a factor of 40 - 50 difference in computer time. Resolution will be cut by a factor of 6 - 8, conveniently just about what one would expect to be the difference between direct ground station and real time telephone line data.

Major weather patterns should still be clearly visible on the earth's disk, and closeups which utilize more than every 5th or 6th line are always possible. Thus, the resolution limitation using TV is not necessarily absolute but would be better described in terms of trading off field of view against resolution.

Consider an average day's operation. The signal from telephone or tape is filtered and enhanced, passes through the transcriber where it is digitized, and then into the computer for rectification and navigation correction. If the A/D reference frequency is not present, an approximate value is chosen, based on the present satellite rotation speed. The most recent precision tapes have been used to locate landmarks and maintain a precise satellite ephemeris. Presumably enough accuracy can be maintained by prediction that landmarks have to be measured at most once a day. Several different courses of action are now available.

A. Navigation update program:

1. Put full earth picture onto 1108 drum and recall a portion with a navigation landmark, displaying on graphics display as brightness contours.

2. Recall one of several stored contour masks or landmark outlines and display alternately with actual picture contours on graphics display (blink method).

3. Twist dials on graphics display to move mask to position of actual landmark where mask and landmark contours coincide.

4. Press a button to fix line and element number of landmark point.

5. Repeat steps 1 - 4 for two or three other standard landmarks.

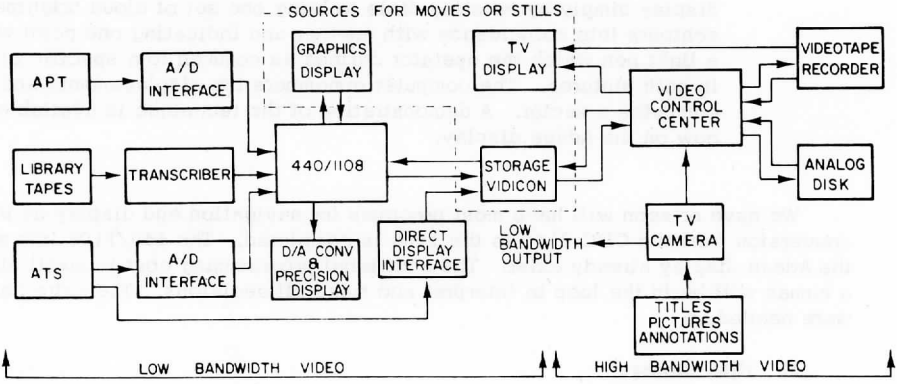
6. Program compares present satellite position with predicted position from previous day's navigation measurements. If no unexpected change in attitude, day's work can begin. Any change would require measuring the next 2 or 3 pictures to establish a new satellite ephemeris.

B. Movie making:

1. An enhanced and filtered picture comes in and is displayed in whole with most lines missing or in part at maximum resolution. The entire picture goes into the 1108 computer drum memory.

2. The picture line and element coordinates are adjusted using earth edge correlation techniques to center the disk for TV display.

PROPOSED TV PROCESSING SYSTEM



3. 3. For full earth display, every fifth line is placed on the storage vidicon, navigation and distortion corrected by the computer. Since points are being moved about on a virtual sphere, alignment and centering of the image should not be severely affected.

4. When storage vidicon is filled, it is read out onto a track on the analog disk at the wide bandwidth rate.

5. The track just entered is played back on the TV display alternating with the previous track (picture) and alignment is checked. Delay circuits in the control center are adjusted for perfect x - y positioning with respect to earlier picture. When delays are properly set, the storage vidicon is read through the delays onto the disk.

6. When disk is full, pictures can be transferred in sequence to videotape or played back on TV display and photographed with movie camera.

7. Closeups can be made by recalling every line of a portion of the full earth picture from the 1108 drum and editing and storing on the disk in a similar manner. In passing through the computer from drum, elements can be brightness enhanced or Fourier analyzed to bring out special details such as cumulonimbus shear in frontal systems or small convective cells.

C. Winds:

1. The blink method can be applied to wind measurement on the graphics display simply by twisting dials to bring one set of cloud brightness contours into coincidence with another and indicating one point with a light pen which the operator defines as common to a specific cloud in both pictures. The computer remembers the displacements and computes a vector. A demonstration of the technique is available now on the Adage display.

We have or soon will have most programs for navigation and display as the conversion from the CDC 3600 to the 1108 is completed. The 440/1108 link and the Adage display already exist. The additional programming cost is small since a human will be in the loop to interpret and make all decisions. The extra hardware needed is:

1. disk storage
2. videotape recorder
3. storage vidicon
4. TV display screen
5. video control electronics.

Items 2, 3 and 4 are relatively inexpensive. They can be added to the present SSEC hardware configuration with only minor alterations and would serve to help test picture display and storage techniques. New videotape recorders with stop motion capability might function as crude disk storage and playback. The control electronics could be tested in part and built in house in stages. The disk is expensive but might be rented for a starter. The sophistication of productions attainable with such a system is limited only by the ingenuity of the operator, since the system can be designed to do any of the tricks one sees on commercial TV. All editing, cropping, enhancing, filtering and aligning can be done in real time as the pictures arrive. Partial time sequences of the synoptic weather patterns or closeup detail exist on the analog disk, ready for playback at any time.

One can process up to 40 of the 500 × 900 element storage vidicon frames in a computer for the cost of a single high quality 2000 × 8000 element complete ATS frame since most of the cost of processing is not in enhancing or filtering, but simply the time involved in moving a certain quantity of picture elements from one spot to another. If the rectification and enhancement were done with the University of Wisconsin 1108 computer in nearly real time, one would also expect a 3-5% drop in system time sharing efficiency, an additional cost above actual elapsed computer time.

The problem facing the meteorologist today is that while observing tools like ATS exist to provide great amounts of data, the scientist is too much

involved in the process of observing rather than the interpreting of the observations. When the meteorologist is taking on the additional jobs of electronics engineer, mathematician and geometer, and computer programmer, he cannot concentrate on using the data most effectively. The situation proposed will combine both the accuracy of the computer and the speed and efficiency of TV to isolate and display in the most effective way for the meteorologist the changing aspects of the world's weather. Once again, the meteorologist can concentrate on the motion of the weather rather than the manipulation of bits.