SURFACE DRAG AND THE CIRCULATION OF A BAROCLINICALLY UNSTABLE ATMOSPHERE

I.N. James
Department of Meteorology, University of Reading
READING RG6 2AU, U.K.

1. FEEDBACKS BETWEEN RESOLVED AND PARAMETRIZED PROCESSES

The theme of this paper is that feedbacks exist between the various diabatic and frictional processes, which are generally parametrized in numerical models, and also with the explicitly resolved large scale circulations. Models of various levels of simplicity will be presented to illustrate such feedback mechanisms.

The classical Lorenz energy cycle, exemplified by Fig. 1 taken from Kung and Tanaka (1983) makes this contention plain. Differential heating generates available potential energy. Conversions transform this available potential energy to eddy and eventually to zonal kinetic energy. Dissipation, especially at the surface, destroys kinetic energy and must eventually balance the energy input. Also noted on Fig. 1 are the residence times of energy in each of the boxes, calculated by dividing the energy content by the total flux through the box due to conversions and source/sink terms. The timescales are short; except for the zonal available potential energy, they do not exceed 10 days. Even the zonal available potential energy has a lifetime of 30-40 days, a value which is of course a global measure of the radiative time constant for the atmosphere. These short timescales mean that an error in, say, the dissipation rates will have a significant impact on the entire global energy budget during the course of a 10-day forecast. There would be least impact on the zonal available potential energy, but the zonal winds and steady and transient eddies could respond quite substantially.

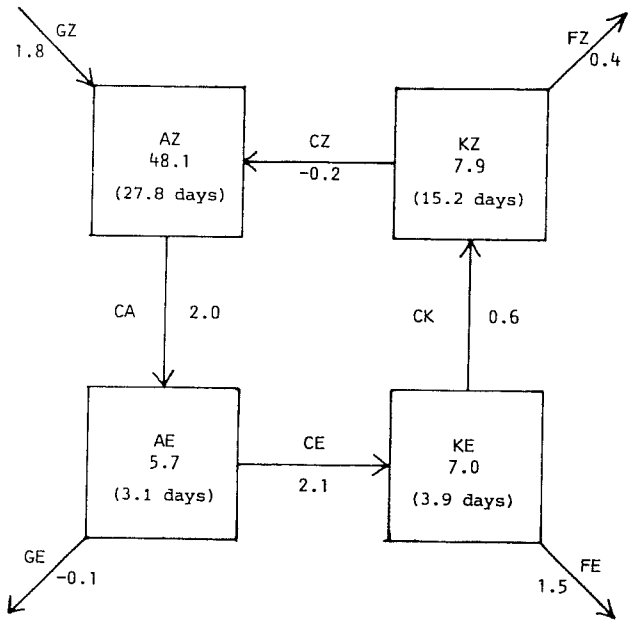

Fig. 1. A Lorenz energy diagram for the FGGE first special observing period, computed by Kung and Tanaka (1983) from ECMWF level IIIb analyses.

An alternative view is provided by potential vorticity considerations. I am indebted to Professor Brian Hoskins for this argument. Fig. 2 shows an isentrope S which intersects the surface of the Earth along the circuit L. The rate of change of potential vorticity is

$$\frac{D}{Dt} P = \frac{\dot{\theta}\zeta + \mathcal{F} \wedge \nabla\theta}{\rho} \quad . \tag{1}$$

Integrating over the surface S and using the circulation theorem gives

$$\frac{D}{Dt} \int_S \rho P dS = \oint_L \left(\mathcal{F} \frac{\partial\theta}{\partial n} - \dot{\theta}\zeta \right) d\ell \tag{2}$$

and so in the time average we must have

$$\overline{\oint \mathcal{F} \frac{\partial\theta}{\partial n} d\ell} = \overline{\oint_L \dot{\theta}\zeta \, d\ell} \quad . \tag{3}$$

Thus, if there is a net diabatic cooling around L the friction $\mathcal{F}$ must be negative, implying surface winds must be positive or westerly. Similarly, diabatic heating around L must imply easterly flow. The integral of P over S is closely related to the circulation around L, and this can be used to estimate a timescale for the modification of P by surface diabatic processes, and hence of the timescale needed for $\mathcal{F}$ and $\dot{\theta}$ to come into balance. Taking $\dot{\theta}$ as around 1 K day^{-1}, $\zeta = f = 10^{-4}$ s^{-1} and a typical surface wind as 5 m s^{-1}, the relaxation time for potential vorticity is again around 40 days, comparable to the radiative time constant.

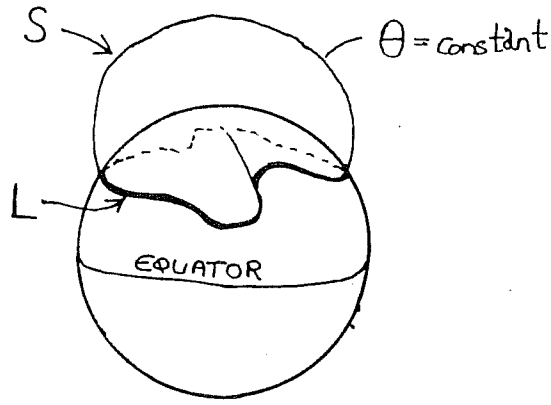

Fig. 2 Illustrating the isentropic surface S which intersects the Earth's surface around L.

Both these arguments serve to demonstrate the coupling between friction and diabatic processes on timescales no longer than that of the radiative time constant, and possibly a good deal shorter for certain specific aspects of the flow. In the succeeding sections, this coupling will be demonstrated using simple models of baroclinic flow on the sphere.

2. THE TROPICAL HADLEY CIRCULATION

Although described by Halley in 1685 and Hadley in 1735, simple quantitative accounts of axisymmetric circulations at low latitudes on a rotating planet have only been published relatively recently. The arguments of Held and Hou (1980) are particularly elegant and are surprisingly successful in accounting for the observed strength and extent of the tropical Hadley circulation.

Their model can be regarded as a two level model of the atmosphere. The surface layer experiences friction which reduces it to rest with respect to the ground on some unspecified but short timescale. The upper tropospheric flow, at height H, experiences no friction and so air conserves angular momentum as it moves away from the equator at upper levels. Conservation of angular momentum leads to a simple formula for the upper level zonal wind which, via the thermal wind equation and an integration with respect

to latitude leads to an expression for the variation of mid tropospheric temperature with latitude. Assuming that the temperature of the tropical atmosphere relaxes radiatively to some simple function symmetric about the equator, a condition of no net heating of parcels circulating in the Hadley cell leads to an expression for the width of the cell: the vertical and meridional velocities can also be calculated if a value for the radiative relaxation time is assumed. Fig. 3 illustrates their model.

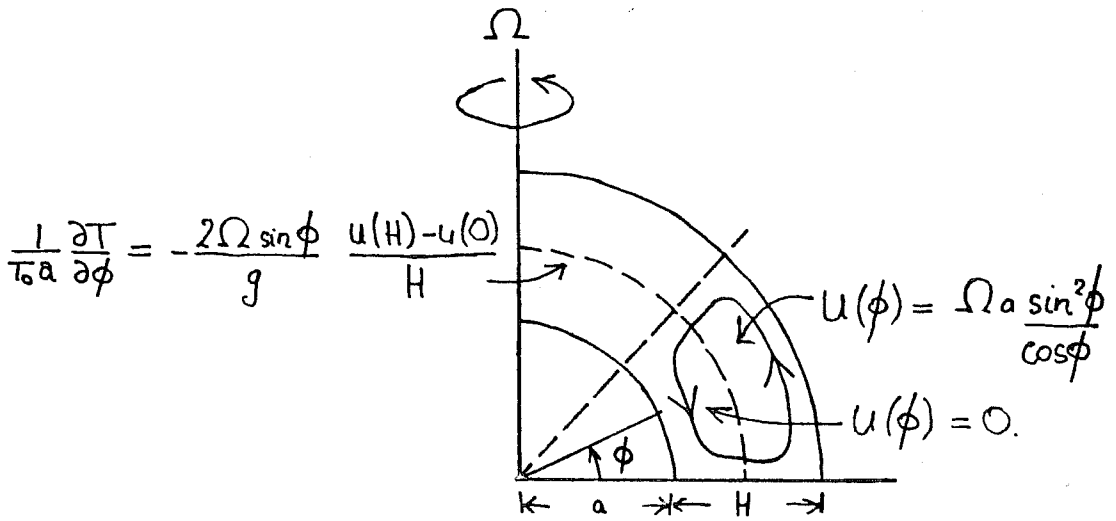

Fig. 3. The Held and Hou (1980) model of the Hadley circulation.

The model illustrates how the friction controls the large scale circulation, though, in this case, the upper level dynamics depend only weakly on the surface friction. The essential Held and Hou results are confirmed by a more elaborate axisymmetric primitive equation model. The thermodynamic equation includes relaxation on a 15 day timescale to a stably stratified basic state. The vorticity and divergence equations include vertical diffusion of the form $K\,\partial^2(\ )/\partial z^2$, which will set up a simple Ekman boundary layer in the presence of interior relative vorticity. The model can be integrated to a steady state, which is achieved after some 20-40 days of

integration, starting from a state of no motion. However, if the Ansatz coefficient K is reduced below some critical value (on the order of 1 m^2 s^{-1}), the flow remains unsteady. In full 3-dimensional models, this unsteadiness is much less pronounced, and it has been suggested that the eddies transport momentum at a rate which is equivalent to an eddy viscosity in excess of 1 m^2 s^{-1} in these integrations. But down to this critical diffusion, the flow changes only rather slowly as K is varied.

The simple arguments of Held and Hou assumed that the radiative equilibrium distribution of temperature was symmetric about the equator. It is interesting to extend the theory to the case when the radiative equilibrium temperature distribution reaches a maximum away from the equator, a situation which is more relevant to the solsticial seasonal mean. Such a modification has been made by Lindzen and Hou (personal communication) and reproduced by Paul James and the author at Reading University. As the maximum is moved away from the equator, the cells become highly asymmetric. The circulation is dominated by a major cell with ascent in the summer hemisphere and descent in the winter hemisphere, with only an extremely weak cell in the summer hemisphere. Paul James has prepared Figs, 4a and b using this modification of the theory; the results are confirmed using the same primitive equation model mentioned above (see Fig. 4c).

An important point of considerable generality is revealed by these calculations. The response, in the form of the Hadley cells, to the diabatic forcing, represented by the radiative equilibrium temperature field, is highly nonlinear. It follows that the mean meridional circulation through an annual cycle is much stronger than the response to the annual mean diabatic forcing. This latter is of course symmetric about the equator. Similar effects arise whenever the response to diabatic forcing is a nonlinear function of the forcing: thus any parametrization of such forcing must necessarily depend upon the space and time scales implicit in the parametrization.

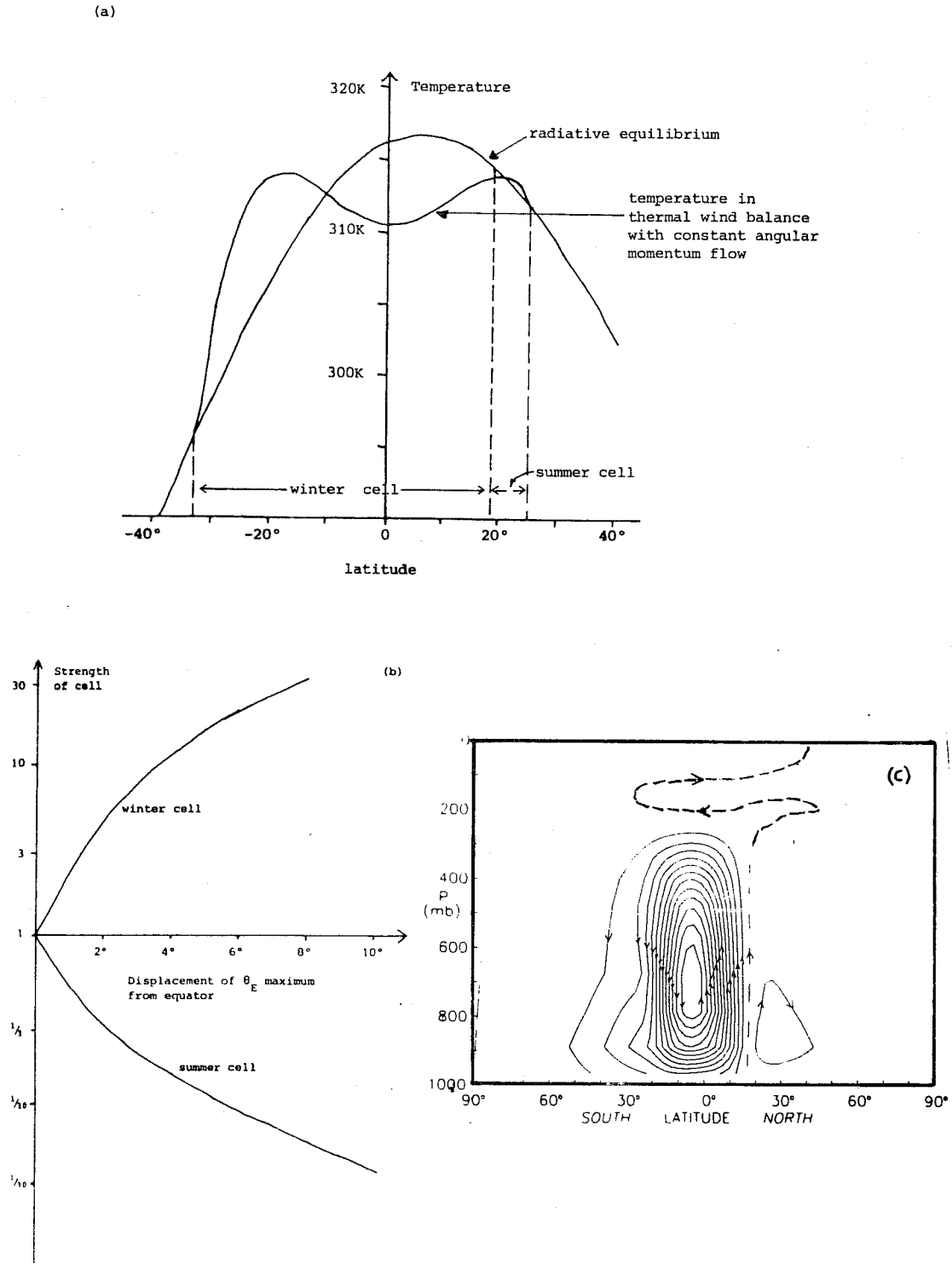

Fig. 4. The results of applying the Held and Hou model. (a) The thermal balance and width of the cell. (b) The strengths of the cells are a function of the latitude of the heating. (c) Axisymmetric primitive equation results for heating centred at 9°N.

3. BAROCLINIC MODELS WITH DRAG

The dynamics of an axysimmetric atmosphere are relatively straightforward, at least in as far as the flow will, with moderate dissipation rates, settle down to a unique steady state. When the axisymmetric constraint is relaxed, much greater complexity is possible. James and Gray (1987) carried out such integrations and this section is based on their paper.

The model used was similar to that described in the previous section, with Newtonian cooling towards some prescribed simple "radiative equilibrium" thermal state on a timescale of 15 days. Mechanical dissipation was introduced in the form of a simple Rayleigh friction (i.e., a term of the form $\underset{\sim}{u}/\tau_D$ on the right-hand side of the momentum equation) at the lowest model level. A rather coarse resolution with triangular truncation to wavenumber 21 and 5 equispaced levels in the vertical was employed, enabling long runs (of order 500 days) to be repeated with different imposed parameters. The drag timescale τ_D was varied between these runs.

When the drag timescale was short (i.e., large mechanical dissipation), the flow was baroclinically unstable and highly unsteady. In the midlatitudes, baroclinic waves developed and collapsed in an irregular fashion, resulting in large fluctuation of the poleward heat transport and the eddy kinetic energy on timescales up to 10-20 days. Fig. 5 shows the time mean zonal flow and meridional circulation for this case. It is a surprisingly faithful simulation of the gross aspects of the observed general circulation, with a low latitude Hadley cell and mid-latitude Ferrel cell driven by transient eddies associated with baroclinic instability of the strongly sheared subtropical jet.

The middle latitudes of this high drag run are filled with vigorous baroclinic waves; the fields of stream function at $\sigma = 0.3$ and temperature at $\sigma = 0.9$ are shown in Fig. 6. The dominant disturbances have zonal wavenumbers less than 8 and the characteristic baroclinic phase shift with height can be discerned.

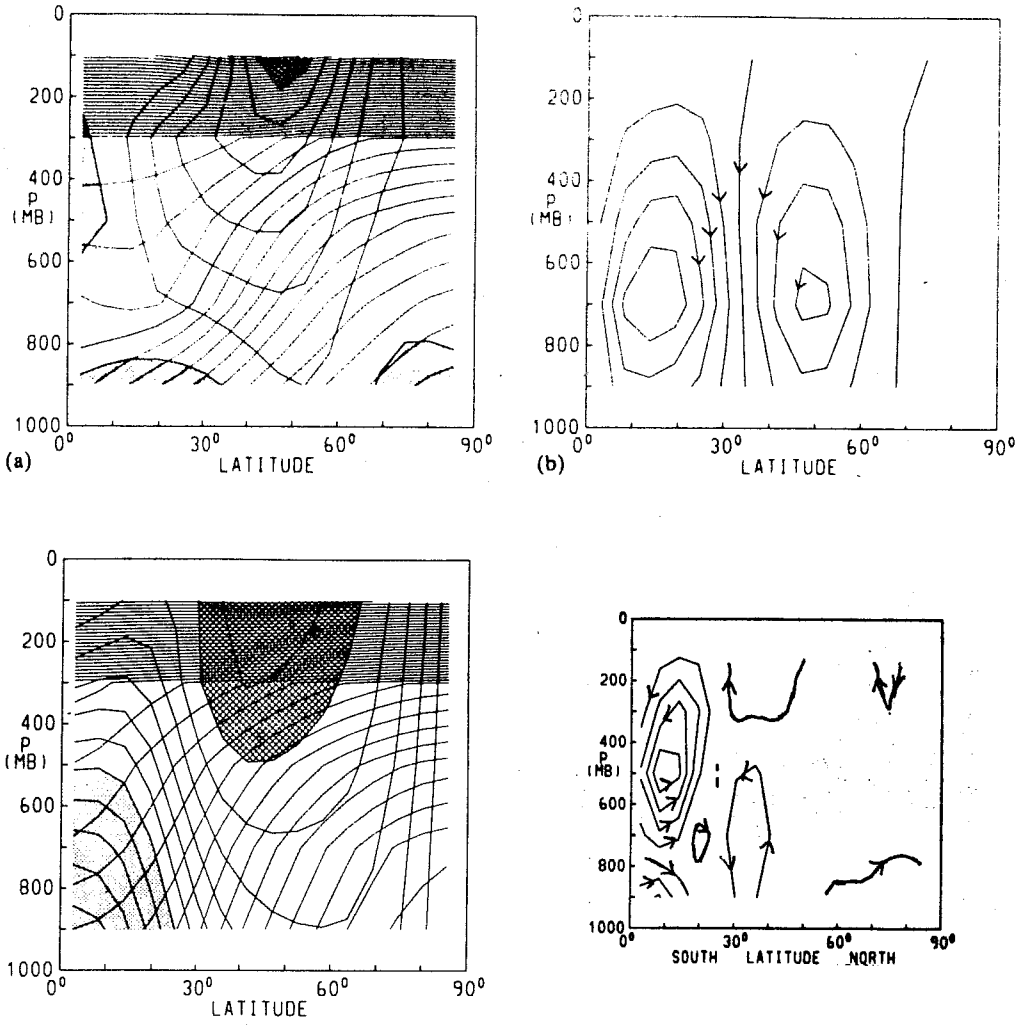

Fig. 5. The zonal mean climate of the simple model for extreme values of the drag.

(a) High drag case, [u] and [θ] (contour intervals 5 m s^{-1} and 5K).

(b) High drag case, meridional stream function, contour interval 22.4 mb s^{-1}.

(c) Low drag case, [u] and [θ] (contour intervals 5 m s^{-1} and 5K).

(d) Low drag case, meridional stream function, contour interval 1.44 mb s^{-1}.

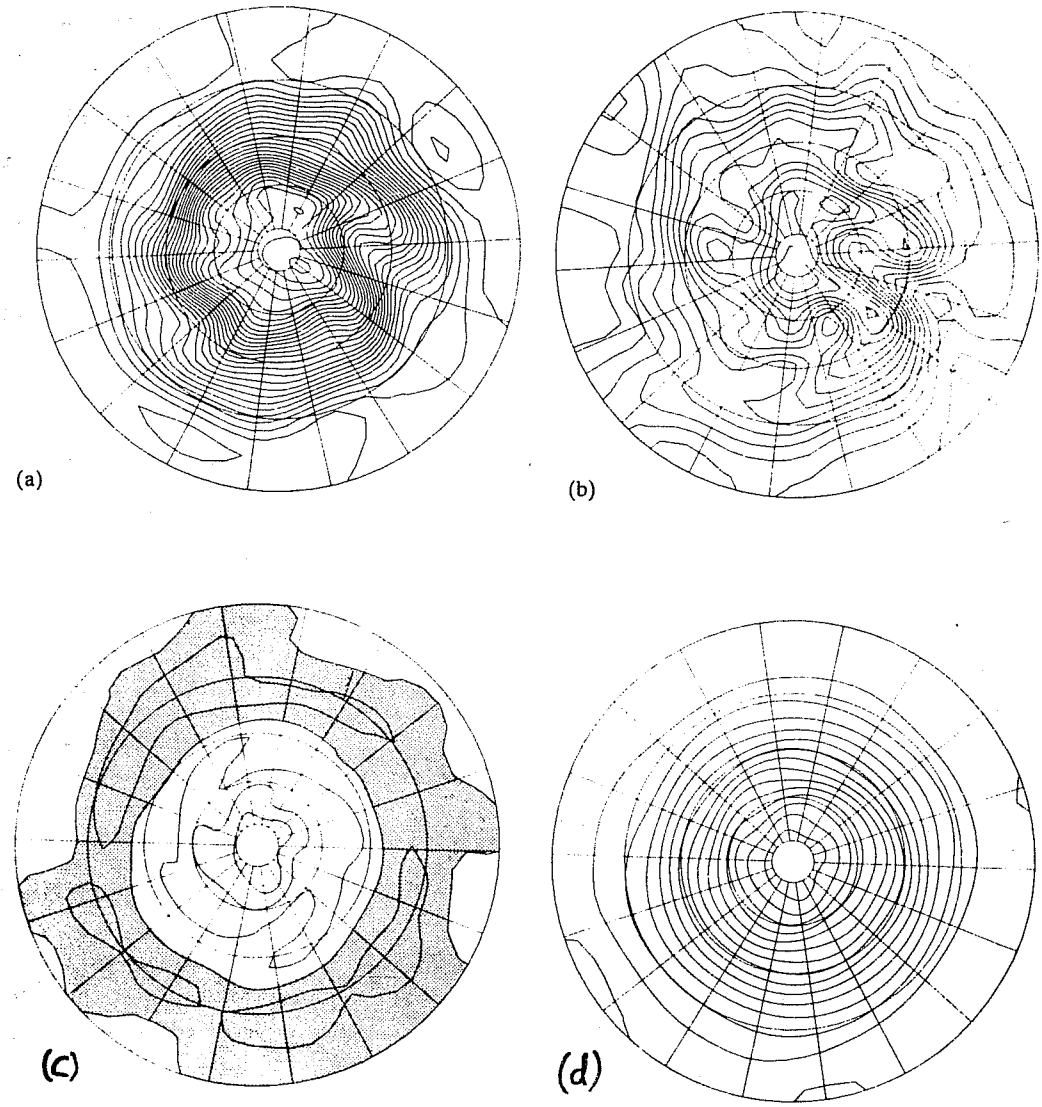

Fig. 6. Instantaneous flow in the simple model for extreme values of the drag.

(a) High drag case, stream function at $\sigma = 0.3$ (contour interval 1.5×10^7 m^2 s^{-1}).

(b) High drag case, temperature at $\sigma = 0.9$ (contour interval 4K).

(c) Low drag case, relative vorticity at $\sigma = 0.3$ (contour interval 7.29×10^{-6} s^{-1}).

(d) Low drag case, temperature at $\sigma = 0.9$ (contour interval 4K).

A dramatic change resulted from the reduction of the surface drag. The case when τ_D was 250 days, a very large value, is presented in this paper. Eddy activity is almost entirely absent and so the eddy transports are negligible. The temperature field differs only very slightly from the "radiative equilibrium" state, and the mean meridional circulation is weak. During the initial, transient phase of the integration, there was a period of 30-50 days when baroclinic activity was vigorous. During that phase, the temperature gradients were eroded and a strong barotropic component to the zonal flow developed. Once the eddies died away, the temperature relaxed back to the radiative equilibrium distribution on the 15 day radiative timescale, but the surface wind remained strong with only very limited dissipation by the small drag term. The final steady state contains a very low level of eddy activity which balances the slight dissipation of the surface wind.

The climate of the model changed smoothly between these extremes for intermediate values of the drag. The variation is most conveniently summarized in the change of the global energetics, shows in Fig. 7. As the drag is reduced, KE falls, rapidly at first, then more slowly. At the same time, AZ increases consistently with the reduced heat fluxes, and KZ increases as the surface winds build up at low drags. The baroclinic conversion CE is always in the direction of increasing KE, but becomes smaller as the drag is reduced. CK is related to the momentum fluxes and acts to increase KZ at the expense of KE. It increased slightly as drag was reduced at large drags but decreased steadily for lower values of drag, when the eddy activity became very small.

At first sight, these results appear paradoxical. The low drag runs have large baroclinicity and little stabilizing influence from dissipation. Yet they are nearly stable to baroclinic disturbances. The high drag runs are baroclinically very active, but have weaker temperature gradients, larger static stability and the large drag would be expected to reduce baroclinic growth rates. The usual criteria for baroclinic instability (Charney and Stern, 1962; see also the summary in James and Hoskins, 1985) are strengthened by the large temperature gradients and smaller static stability of the lower drag runs. The paradox is of course illusory; the Charney-Stern criterion is merely a <u>necessary</u> condition for instability which states that

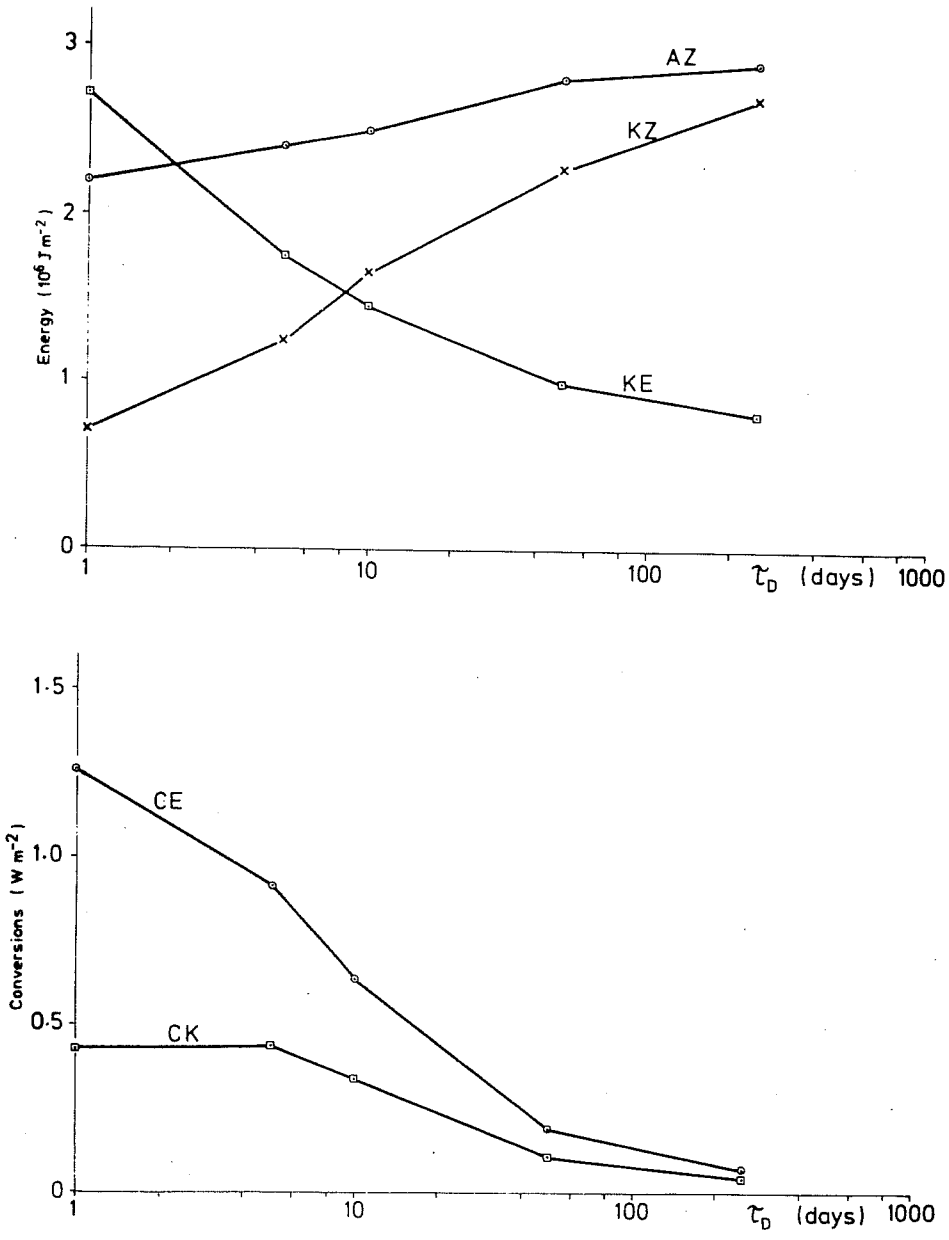

Fig. 7. Variation of energy and conversions in the simple model as a function of the drag timescale.

available potential energy exists. There is no guarantee that it is possible to sustain disturbances whose structure can exploit that available potential energy for substantial periods in any given flow. The classical model of baroclinic instability (those due to Eady, Charney and the two layer system) assume a basic flow with shear in the vertical but which is uniform in the horizontal. We may suspect that the greater baroclinic stability in the low drag runs reported in James and Gray (1986) is associated with the large horizontal shears which developed. I shall demonstrate that this was indeed the case in the next section.

4. BAROCLINIC INSTABILITY WITH HORIZONTAL SHEAR

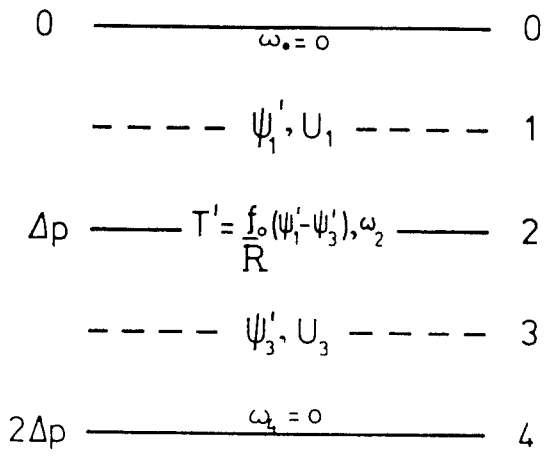

Fig. 8. Illustrating the nomenclature used to describe the two level model discussed in the text.

For simplicity, consider the two layer quasi-geostrophic system, linearized about zonal flow. This is described by the equations:

$$\frac{\partial}{\partial t} q_i' + ik\left[U_i q_i' + \left(\beta - \partial^2 U_i/\partial y^2\right)\psi_i' + \left(2 - i\right)\left(U_2 - U_1\right)\psi_i'\right] = 0 \quad (4)$$

where

$$q_i' = \frac{\partial^2}{\partial y^2}\psi_i' - \left(K_R^2 + k^2\right)\psi_i' + K_R^2\ \psi_{4-i}' \quad , \quad i = 1,3. \quad (5)$$

where q_i' is the perturbation potential vorticity, ψ_i' is the geostrophic stream function, and i = 1 refers to the upper level, i = 3 to the lower level. The parameter $K_R = (f/\sigma\Delta p)$, σ being a static stability parameter. Fig. 8 illustrates the notation. The mean zonal wind is given by

$$U_1 = \Delta U + Ay \quad , \quad U_3 = Ay \quad (6)$$

and for simplicity $\beta = 0$. A measures the horizontal, barotropic shear to which the growing instabilities are subject. Note that the addition of this simple linear horizontal shear does not modify the potential vorticity gradients or the vertical shear. It therefore has no impact on the criteria for the existence of baroclinic instability. James (1987) has solved the eigenvalue problem associated with Eqs. (4) and (5) to determine the structure and growth rates of unstable normal mode disturbances for various values of A. K_R was taken to be 500 km and ΔU to be 20 m s^{-1} for the calculations described here.

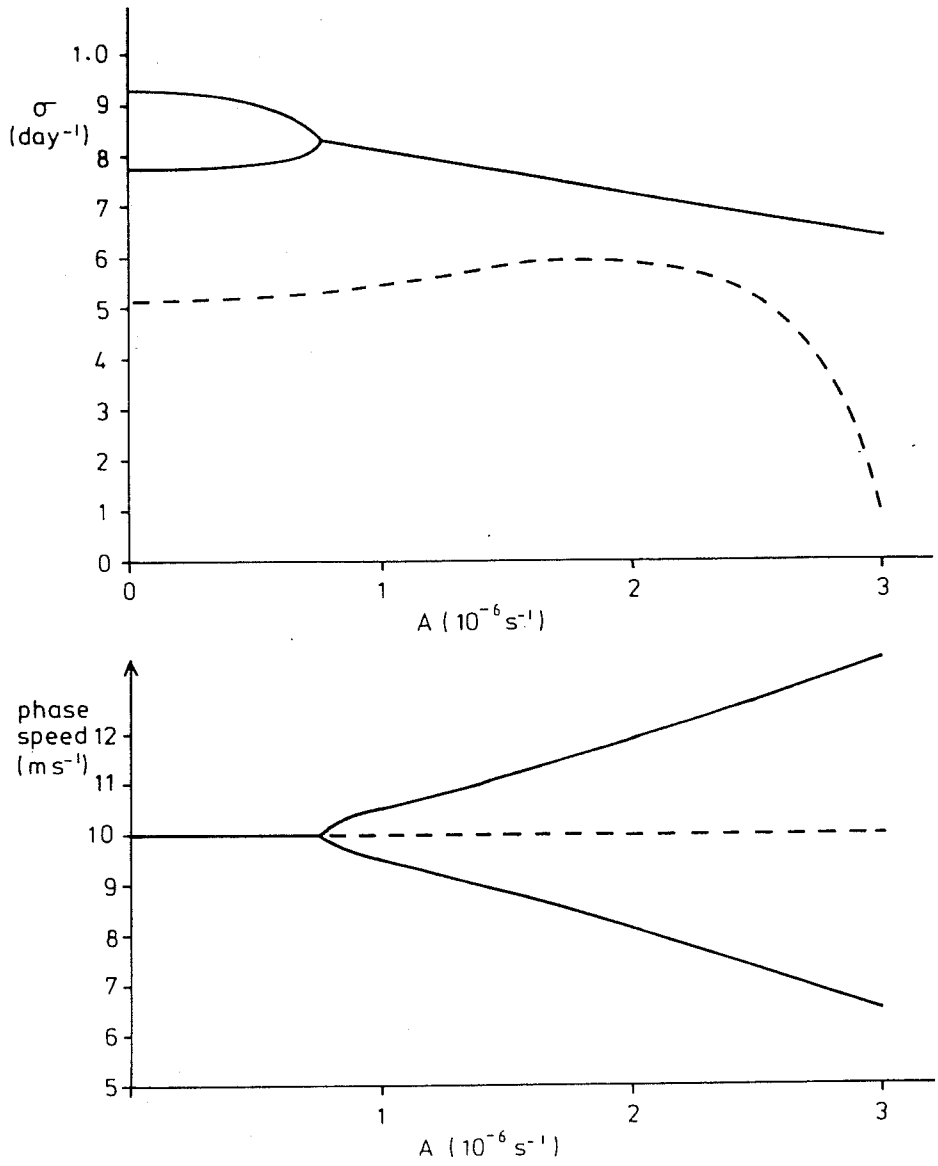

Fig. 9. Showing (a) growth rate and (b) phase speed of baroclinically unstable wavenumber 7 normal modes in the two layer system as a function of the horizontal shear A.

The results for A = 0 are of course the analytic text-book results, showing a most unstable mode with a growth rate of 0.92 day^{-1} and a phase speed of 10 m s^{-1} (i.e., the mean flow speed along the channel). As A approaches 0.7×10^{-6} s^{-1}, the growth rate of the most unstable mode drops slightly, while that of the next mode increases a little. The phase speed remains constant, at 10 m s^{-1}. But for larger values of A, the growth rates are reduced further and the phase speeds begin to change.

These changes are related to changes in the structure of the unstable normal modes, illustrated in Fig. 10. The essential effect of the shear is to

confine the normal modes in the meridional direction. At first, they are confined to the centre of the channel, with symmetry with respect to reversal of y and p; as might be expected from symmetry considerations, the steering level passes through the centre of the channel and so the phase speed is 10 m s^{-1}. For A in excess of 0.7×10^{-6} s^{-1}, the mode splits into two modes, each with the same growth rate, but confined to opposite sides of the channel. Their phase speed reflects the mean flow speed in that region. Fig. 10 shows just the modes on the southern boundary of the channel. The steering level intersects the lower layer at the point

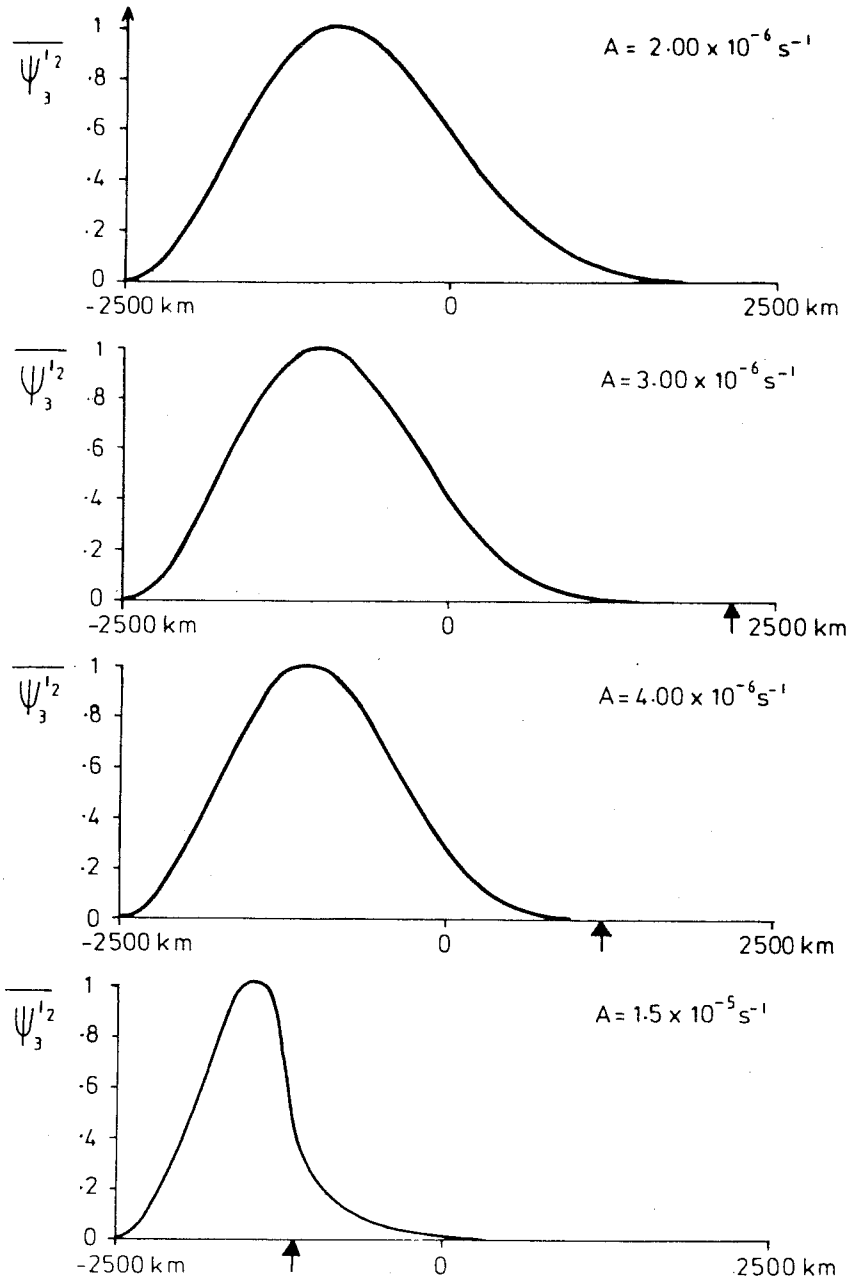

Fig. 10. Showing the streamfunction amplitude of the most unstable wavenumber 7 normal mode for various values of A.

indicated by the arrow; note that the mode has significant amplitude only at those latitudes where the steering level exists within the fluid.

It must be emphasised that the shears employed in the previous runs are not unrealistically large, yet growth rates are reduced by a considerable factor. Even larger reductions, or indeed complete stabilization, are possible if dissipative (Ekman pumping) terms are included. Furthermore, similar growth rate reductions are found when more realistic background states, including jets in which the artifical sidewall boundary conditions play no role, are used in place of the basic flow defined by Eq. (6). Details are given in James (1987).

Let us now return to the simulations carried out with different values of drag described in the previous section. The initial state in both the high and low drag cases was identical and had zero surface wind. As the eddies developed and the flow settled down towards its climatological state, the temperature gradients were weakened and a barotropic component of the wind was introduced. In the high drag runs, this surface wind was small while the reduction of the temperature gradients was substantial. In the low drag case, the climatological temperature field was rather close to the radiative equilibrium state, while the surface wind was very large. Following the ideas introduced by these calculations for the two layer system, let us see whether the strong horizontal shear in the low drag climate plays a significant role in reducing the level of eddy activity.

The linear normal modes which can grow on these zonal flows can be calculated using the spectral model to generate the elements of a large matrix of complex elements which describe the tendencies due to changes in each independent degree of freedom of the system. The eigenvectors of this matrix are the normal modes of the discretised model system and can be determined using standard methods. The solid curves in Fig. 11 show the growth of the most unstable normal mode as a function of zonal wavenumber for the initial states. These are identical, but of course the higher drag in the one case reduces the growth rates of the normal modes slightly. The dashed curves show the growth rates for the climatological state. In the low drag run, they are over an order of magnitude smaller than in the initial state. It is worth remarking that the most unstable normal modes shown in that curve had very noisy and small scale meridional structure;

it seems unlikely that they could grow to very large amplitude. The reduction of growth rates for the high drag case is less, consistent with the larger eddy activity in that run.

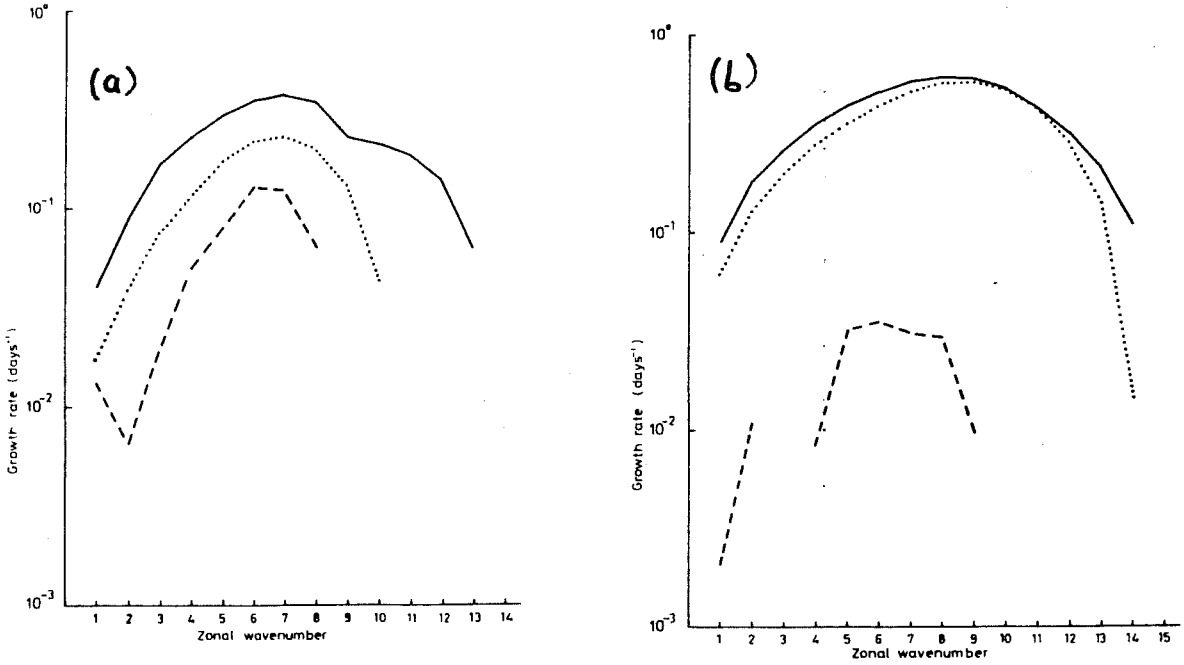

Fig. 11. Growth rate versus zonal wavenumber in (a) the high drag, (b) the low drag runs. The solid curve is for the initial (radiative equation) basic state and the dashed curve the climatological mean. The dotted curves are for the climatological state but with a barotropic correction setting the surface winds to zero.

Finally, the dotted curves show the importance of the horizontal shears in controlling the level of eddy instability in the two runs. The climatological state for both runs was modified by subtracting the surface wind from every level. This barotropic correction leaves the temperature field unaltered and has very little effect on the potential vorticity gradients (which, for large Richardson number, are dominated by vertical, not horizontal, curvature of the zonal wind field). But clearly the effect on the growth rates is dramatic. In the low drag run, growth rates return essentially to the values for the initial state, showing that horizontal shear was the dominant mechanism in stabilizing the flow. Even in the high drag case about half the stabilization must be attributed to the rather weak horizontal barotropic shears developed in the final climatology. It turns out that this stabilization by horizontal shear is a fundamental effect in the life cycles of baroclinic waves (described by, for example, Simmons and Hoskins, 1978); James (1987) showed that the collapse of the instability in the latter part of the life cycle is due principally to the development

of horizontal shear, not to the source of available potential energy becoming exhausted.

5. THE BAROTROPIC GOVERNOR

The momentum fluxes associated with unstable baroclinic waves are rather difficult to discuss in general terms, and require a full solution of the (generally inseparable) eigenvector problem for their determination. But certain general principles appear to hold, at least in a wide range of examples; these are conveniently summarized in terms of the familiar Lorenz energy cycle discussed in the first section.

The rate of change of total eddy energy can be written:

$$\frac{dE}{dt} = - \left\langle [u^*v^*][u]_y - \frac{R^*f}{\Delta p\ \sigma^2} [v^*T^*][T]_y \right\rangle \qquad (7)$$

where the quasi-geostrophic framework has been used. In balanced flow, horizontal temperature fluxes invariably imply vertical temperature fluxes, so that I do not distinguish between the eddy available potential and kinetic energies; E refers simply to the total eddy energy. The angle brackets denote integration over the entire atmosphere and other notation has its usual meaning. Baroclinic instability is characterised by a negative correlation between the poleward temperature flux and the temperature gradient, that is, by down gradient temperature fluxes. In contrast, the momentum fluxes tend to be positively correlated with the horizontal shears and so lead to a reduction in the rate of generation of eddy energy. This direct effect of momentum fluxes is not especially large, and is not the dominant process in accounting for the reduced growth rate in a horizontally sheared environment. Rather, the structural changes of the normal mode restrict the correlation between $[v^*T^*]$ and $[T]_y$ and so the growth rates are reduced.

However, the direction of the momentum fluxes is almost always such as to increase the horizontal shears in the vicinity of the developing disturbances. In the classical lifecycle experiments of Simmons and Hoskins (1978) a barotropic jet was added to the zonal flow as a result of the momentum fluxes associated with the evolving baroclinic disturbance. Similarly, zonal mean circulation statistics for the observed atmospheric circulation show that the transient momentum fluxes act to generate westerlies somewhat poleward of the latitudes of maximum eddy activity (see, for

example, Oort (1983)). Thus the depression belt tends to lie in a region of horizontal shear generated by the eddies themselves. James and Gray (1986) suggested that this effect is sufficiently ubiquitous in realistic situations to be idealised in the modified picture of the Lorenz energy cycle, shown in Fig. 12. We called this picture the "barotropic governor"; the eddies generate momentum fluxes which increase the horizontal shear in regions where baroclinic activity is concentrated. This shear tends to reduce the baroclinic activity until some climatological balance between the generation of energy by differential heating and its dissipation, essentially by friction, is achieved. The feedback between the horizontal shear and the generation of eddy energy governs the level of eddy energy and eventually the zonal kinetic energy, and plays a central role in ensuring that the sources and sinks remain in balance.

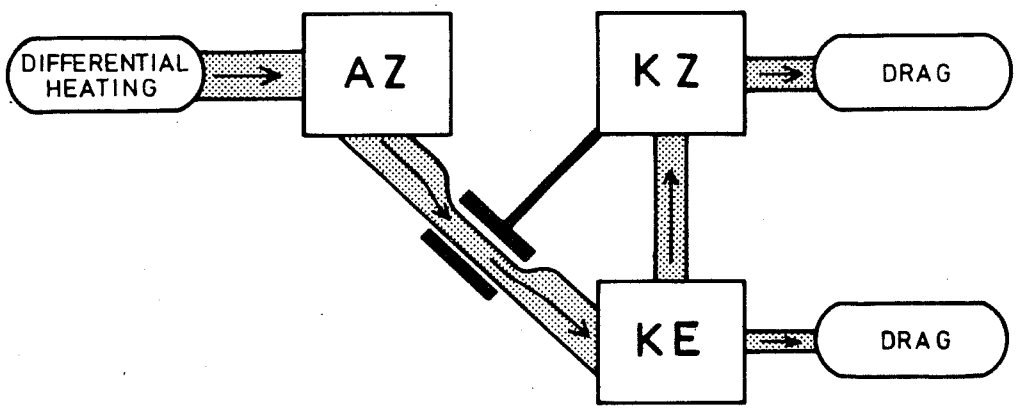

Fig. 12. Illustrating the "barotropic governor".

5. IMPLICATIONS

Fig. 13 shows the current ECMWF wind error for 10-day forecasts during the Northern Hemisphere winter. The error is broadly barotropic in structure, with westerly errors at 50°N and easterly errors around 30°N. The typical horizontal shear is as much as 2×10^{-6} s^{-1}, quite enough to cause significant reductions in the baroclinic growth rates and hence to the poleward temperature fluxes, according to the results discussed in the previous section.

It can easily be imagined that as a consequence of the error in the westerly

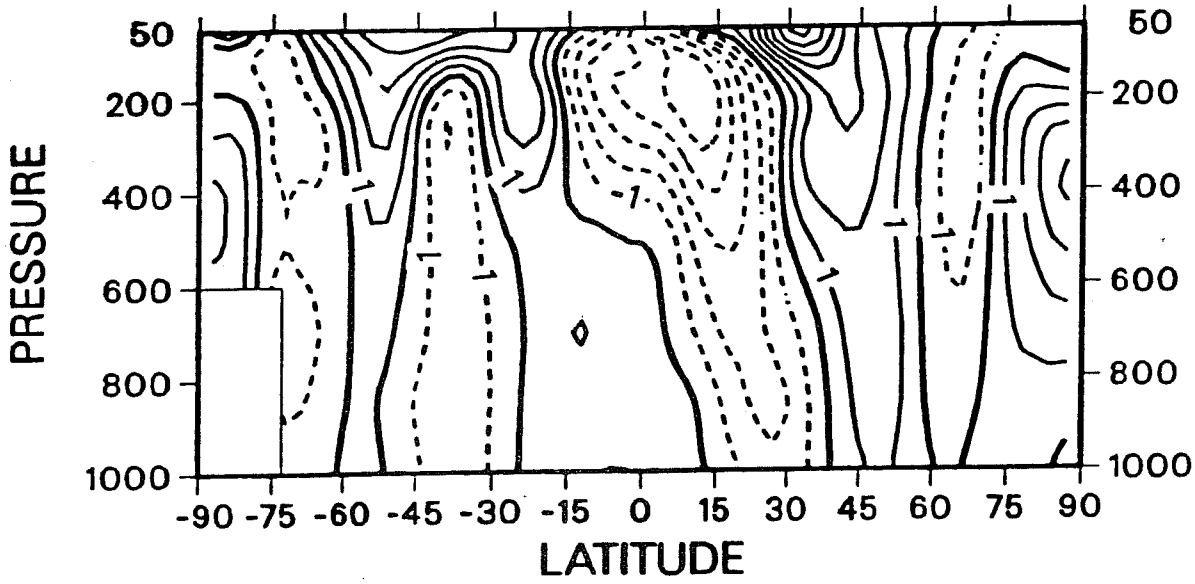

Fig. 13. The error in the 10-day forecast zonal wind at ECMWF for the period December 1986 to February 1987. Contour interval 1 m s^{-1}.

wind, there will be a tendency by day 10 of the forecast for the eddy temperature flux to be too small and consequently for the warming of the tropics to be exacerbated. The point is a fairly straightforward one which nevertheless needs emphasising periodically. An error in, for example, the tropic temperature field need not be the direct consequence of an error in the tropical diabatic heating. It could be the consequence of a global deficiency in the atmospheric dynamics which might have its origin in the parametrization of mechanical dissipation in midlatitudes. The barotropic governor mechanism suggests a particular route whereby this type of interaction might take place. It seems to have received rather little attention in the literature previously, but in view of the apparently ubiquitous nature of the governor mechanism (which is related to older ideas of "negative viscosity"), I suggest it could well be important when considering systematic errors in forecasting and general circulation models.

Most of the discussion here has been in the context of the zonal mean flow. The barotropic governor also apparently plays a role in the zonal variations of the flow, and in determining the length of the major storm tracks. The storm tracks of the northern hemisphere start over the western side of the major ocean basins. The change of surface drag from land to ocean may well play a crucial role in the extra generation of eddy kinetic energy in such regions (Paul Valdes, personal communication). As depressions go through

their lifecycle, a developing surface wind is generated which in turn helps to account for the greater stability of the flow over the eastern part of the ocean basin. A failure to dissipate these westerlies adequately could distort the length and shape of the storm track and lead to important systematic errors in the distribution of transient eddy activity by the end of a 10-day forecast.

The ideas discussed in this section are of course speculative. But they could easily be investigated using the kinds of simplified models and techniques mentioned in the previous sections. It is hoped to extend the work described in James and Gray (1986) and James (1987) in these directions.

References

Charney, J.G. and Stern, M., 1962: On the stability of internal baroclinic jets in a rotating atmosphere. J. Atmos. Sci., 19, 159-172.

Held, I.M. and Hou, A.Y., 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. J. Atmos. Sci., 37, 515-533.

James, I.N., 1987: Suppression of baroclinic instability in horizontally sheared flows. J. Atmos. Sci., 44, 3710-3720

James, I.N. and Gray, L.J., 1986: Concerning the effect of surface drag on the circulation of a baroclinic planetary atmosphere. Q.J. Roy. Met. Soc., 112, 1231-1250

James, I.N. and Hoskins, B.J., 1985: Some comparisons of atmospheric and internal baroclinic instability. J. Atmos. Sci., 42, 2142-2155.

Kung, E.C. and Tanaka, H., 1983: Energetic analysis of the global circulation during the special observation periods of FGGE. J. Atmos. Sci., 40, 2575-2592.

Oort, A.H., 1983: Global atmospheric circulation statistics, 1958-1973. NOAA Professional Paper 14 (U.S. Department of Commerce).

Simmons, A.J. and Hoskins, B.J., 1978: The life cycles of some nonlinear baroclinic waves. J. Atmos. Sci., 35, 414-432.