On the Equilibrium Spectrum of Transient Waves in the Atmosphere

Wendell T. Welch

National Center for Atmospheric Research, Advanced Study Program, Boulder, CO*

Ka-Kit Tung

Department of Applied Mathematics, University of Washington, Seattle, WA

April 9, 1997

* The National Center for Atmospheric Research is sponsored by the National Science Foundation.

Abstract

The spectrum of transient waves in the mid-latitude atmosphere is explored. There are two maxima, one at planetary and one at synoptic scale. The very long wave maximum seems to be Rhines' wavenumber of cascade arrest, i.e. the scale at which a barotropic reverse energy cascade stops due to the prominence of linear Rossby wave dynamics. The synoptic peak is explained here by the theory of nonlinear baroclinic adjustment: a balance between quasi-linear energy extraction from the mean flow due to baroclinic instability and a resulting up-scale cascade. The mechanisms for the two energy maxima appear to operate independently of each other.

Computations with a quasi-geostrophic two-level model in a beta-plane channel corroborate the nonlinear baroclinic adjustment mechanism. It is demonstrated that the model and hence the theory, although simple, can simulate the dynamics of the real atmosphere both in the terms of the transient mid-latitude spectrum and the meridional temperature gradients achieved in summer and winter.

Nonlinear baroclinic adjustment is shown to agree with Salmon's theory of wave-wave equilibration from quasi-geostrophic turbulence. The former goes further, however; it is capable of explaining not only the equilibrium dynamics but also the mechanistic approach thereto, for different channel geometries and at different levels of forcing.

The mechanism is shown to be different from both stochastically forced transient growth and neutralization theories of the atmosphere.

1. Introduction

The behavior of transient eddies in the mid-latitudes yields a bimodal spectrum. This is shown in Fig. 1 (dashed line), which is the 500mb geopotential height variance for propagating disturbances in winter. The data is averaged over five winters at 50°N, from a study by Fraedrich and Böttger (1978). There are two maxima of power: one at the longest (planetary) scale and one in the synoptic (i.e. baroclinic) range. (Signs of similar behavior are implicit in Figs. 11 and 13 of Kao and Wendell (1970). See Hoskins and Pearce (1983) for a list of other studies.) Note that this differs from the spectrum of *total* variance, including propagating and stationary disturbances, which is fairly flat from the synoptic peak to the large scales (Rhines 1975; Leith 1972). Various theories have been proposed to explain the presence of the synoptic peak of transient variance.

The synoptic peak has been identified with baroclinic instability, for observations of the dominant eddies in the atmosphere have shown them to be of synoptic scale and baroclinic in nature (Charney 1971, p. 1088). The simplest explanation for their dominance would be, of course, linear: the most unstable wave from a linear analysis of the zonally averaged atmosphere would dominate the energy spectrum. Gall (1976) has shown, however, that the most unstable wave is shorter than the scale of either of the energy peaks. Zonal wavenumbers 12-15 have the highest growth rates, as calculated from a GCM linearized about the actual mean atmospheric state, whereas the transient eddies which dominate in the real (nonlinear) atmosphere are wavenumbers 4-7 (Randel and Held 1991). The latter is corroborated by Fig. 1. Thus a linear explanation is unable to account for the synoptic energy maximum. We note that this includes Stone's (1978) original proposal of "baroclinic adjustment", in which he reasoned that the most unstable wave would do all of the heat transport at equilibrium. Such reasoning would yield wavenumber 12 or so as the most energetic wave in the equilibrated state, but this does not agree with observations.

To reconcile for this discrepancy, a nonlinear up-scale cascade of energy, as presented in the geostrophic turbulence study of Charney (1971), is often cited. Rhines (1975) proposed that this cascade would dominate at small scales, whereas linear Rossby wave dynamics would be present at the largest scales. At the scale where these two mechanisms have roughly the same magnitude, Rhines said the cascade would stop, with the spectrum decreasing or level at lower wavenumbers. He offered a prediction for this dominant "wavenumber of cascade arrest". In model simulations here, to be detailed below, Rhines' calculation yields a long wave; indeed, it corresponds roughly to the energy maximum at the planetary wave scale in Fig. 1. Thus, his argument seems to explain the planetary scale maximum of transient variance. It does not, however, explain the synoptic peak.

Salmon (1980) united Charney's up-scale cascade with the concept of energy injection at synoptic scales due to baroclinic instability. He argued for a balance at the short scales between energy extraction from the mean flow and nonlinear transfer toward long scales. Such dynamics could yield an energy maximum at a synoptic scale that is longer than that of the most unstable wave. This he reproduced numerically with a fully nonlinear model run to statistical equilibrium. In this way he diagnosed the dynamics at equilibrium which could support a synoptic energy maximum.

Salmon's theory, however, does not explain why or how a particular wavenumber comes to dominate. Why does the cascade stop at a synoptic wave and not continue all the way to the longest scales? In a different climate would the synoptic peak occur at a different wavenumber? Can this be predicted? Can the wavenumber of maximum heat transfer also be predicted, for a different climate?

Here we will show that the theory of nonlinear baroclinic adjustment (Cehelsky and Tung 1991; Welch and Tung 1997) can answer the above questions. We will verify the theory using simulations from a high resolution two-level quasi-geostrophic baroclinic model in a β-plane channel with a flat bottom. Unlike the previous studies, the model now will be run in a parameter regime in which many zonal waves are linearly unstable, similar to the situation in the real atmosphere.

Of course there are many ways to make the model more realistic: using a spherical geometry, allowing the static stability to vary, including topography and hence stationary waves, and adding moisture, to name but a few. Here we use the simplest possible model, which still allows for baroclinic dry dynamics, in order to investigate most easily the qualitative features of meridional heat transport.

Although a two-level model cannot simulate properly the real atmosphere, there is a correspondence between linear stability analysis of a two-level model and tropospheric observations: the critical gradient in the former corresponds to the cutoff in the atmosphere between shallow waves, ineffective at transporting heat, and long deep waves which can efficiently flux heat poleward (Held 1978). Furthermore, it appears that adding more levels in the vertical may *not* change the qualitative results of the model. Pavan (1996) showed in her quasi-geostrophic Boussinesq model that increasing the layers from 3 to 20 had little effect on the main features of the solution. High vertical resolution was needed for qualitative convergence only when eddy momentum flux was crucial, e.g. in situations with a strong barotropic governor effect (James 1987).

We point out that the static stability has been held constant over time in this study. However, an important process in equilibrating baroclinic flows, in addition to the reduction of the horizontal temperature gradient, is the adjustment of the vertical temperature profile via vertical eddy heat fluxes (Gutowski et al. 1989; Zhou and Stone 1993). Here we neglect this effect in order to focus on the interaction of horizontal heat transport and the horizontal temperature profile, consistent with the quasi-geostrophic formulation adopted. In the future our results should be tested with a model which allows for variation of the static stability. In addition we have used quasi-geostrophic scaling, which is based on the Rossby number being small: $\mathrm{Ro} \equiv U/f_\circ L \ll 1$. However, as the driving increases the advective time scale, L/U, becomes small. The underlying expansion in Ro will not be valid, and hence neither will be our model output, at very high forcings. These issues are two of the main limitations of this model.

Two additional theories have been proposed to explain the equilibration of baroclinic flows. Lindzen (1993, 1994) has studied the possibility of neutralization, in which the atmosphere is always striving for a state which is linearly neutral or stable to perturbations of *all* wave scales. Can this be reconciled with the nonlinear baroclinic adjustment theory of Welch and Tung (1997)? At high enough forcing in that study with only two waves unstable, the longer wave was indeed at marginal stability at equilibrium, but the shorter wave was linearly unstable. Is this behavior still present in a realistic simulation with many waves unstable, or does Lindzen's theory apply? This issue will be investigated.

Also, Farrell and collaborators (Farrell and Ioannou 1994; DelSole and Farrell 1995; Farrell and Ioannou 1995; DelSole and Farrell 1996) have espoused yet another explanation. In the studies listed, the authors proposed a stochastically forced, linearized system to explain the dynamics of baroclinic equilibration in the atmosphere. The difference between their mechanism and that of nonlinear baroclinic adjustment must be pinpointed.

In section 2 the mathematical model is described briefly. Section 3 shows the model-simulated transient energy spectrum and compares with it Rhines' prediction of the wavenumber of cascade arrest. In this section the mechanism of nonlinear baroclinic adjustment also is proposed, and it is shown that both the spectral peaks of Fig. 1 can be explained. The agreement between this mechanism and the wave-wave equilibration theory of quasi-geostrophic turbulence is demonstrated. As a test of the model, its results are compared with observations in section 4. Finally, differences between the proposed theory and those of Lindzen and of Farrell are discussed in section 5, with conclusions in section 6.

2. Numerical Model

The numerical model used here is identical to that developed in Welch and Tung (1997) except for the values of a few parameters, which will be discussed below. The model solves the quasi-geostrophic equations on a β-plane, including Newtonian cooling to a radiative equilibrium temperature profile and Ekman friction at the Earth's surface. The domain approximates the mid-latitude troposphere: a channel of 45° width centered on 50°N and extending from the top of the Ekman layer to 200mb. We assume rigid walls, and a rigid (and flat) top and bottom. The model is finite-differenced into two levels[1] in the vertical. The non-dimensionalized, leveled equations in pressure coordinates are:

$$\frac{\partial}{\partial t}\nabla_\delta^2\Psi_1 = -\delta J(\Psi_1, \nabla_\delta^2\Psi_1) - \delta\beta\frac{\partial\Psi_1}{\partial x} + \omega_2 - \omega_0 \tag{2.1}$$

$$\frac{\partial}{\partial t}\nabla_\delta^2\Psi_3 = -\delta J(\Psi_3, \nabla_\delta^2\Psi_3) - \delta\beta\frac{\partial\Psi_3}{\partial x} + \omega_4 - \omega_2 \tag{2.2}$$

$$\begin{aligned}\frac{\partial}{\partial t}(\Psi_3 - \Psi_1) &= -\frac{1}{2}\delta J(\Psi_1 + \Psi_3, \Psi_3 - \Psi_1) - 2\sigma_\circ\omega_2 \\ &\quad -2h''\left[\Psi_3 - \Psi_1 - (\Psi_3 - \Psi_1)^\dagger\right]\end{aligned} \tag{2.3}$$

where $\Psi = \Phi/f_\circ$ is the geostrophic streamfunction and $\omega = dp/dt$ is the vertical velocity. Subscripts 1 and 3 indicate the upper and lower levels, 2 the interface, and 0 and 4 the top and bottom of the model, respectively. $\nabla_\delta = \delta^2(\partial^2/\partial x^2) + (\partial^2/\partial y^2)$ is the non-dimensionalized Laplacian, and a $\dagger$ identifies the radiative forcing. For details of the model, see Welch (1996).

Several non-dimensional parameters have been introduced: β is the meridional gradient of the Coriolis parameter f, $\sigma_\circ$ is a measure of static stability (specified and held constant in our simulations), h'' is the relaxation time scale for the radiative forcing, and $\delta = L_y/L_x$ is a horizontal aspect ratio, where L_x and L_y indicate the scale of the domain in the zonal and meridional directions, respectively.[2] Specifying δ is equivalent to setting the length scale of the gravest zonal mode allowed (for a fixed channel width): the larger δ, the shorter the fundamental mode's wavelength. Thus, by varying δ, the number of zonal modes which are unstable can be varied.

In this study we are attempting to simulate a case akin to the real atmosphere in which many waves are unstable and participate in the baroclinic equilibration. We choose the value $\delta = 0.28$, for

[1] In Welch and Tung (1997) we called this same setup a two *layer* model, because it follows from the original formulation of Lorenz (1960), which he termed the same. However, both our model and his are actually *level* models, for the height at which the variables are to be evaluated is specified. See Pedlosky (1987).

[2] Note that we let the non-dimensional variable $\hat{x}$ range over $[0, 2\pi]$ whereas $\hat{y}$ only ranges over $[0, \pi]$. This is motivated by the no flow side wall boundary conditions and our choice of $\sin ny$ as the meridional form of the eigenfunctions. See Welch 1996.

it corresponds to a channel which circumscribes the Earth, i.e. to the real mid-latitudes.[3] Notice that such a small value of δ indeed allows many zonal waves to be linearly unstable. This can be seen in Fig. 2, which displays the marginal stability (dashed) curve for our choice of parameters. Finally, we use the value $\sigma_\circ = 0.06$ for the static stability parameter, which is in the range 0.05-0.09 calculated from tropospheric observations (Welch 1996). (In section 4 we will refine this value even further for winter vs. summer.)

Fig. 2 (dashed curve) shows that there are approximately 14 waves unstable, with the most unstable being $m = 11$. This is close to the scale of the most unstable wave in the real atmosphere ($m = 12 - 15$ from Gall 1976). As we shall see in the model results, this geometry with a "realistic" forcing of $\Delta T^\dagger = 90$K yields $m = 5$ as the dominant heat transporting wave at equilibrium, which is in the range found in observations by Randel and Held (1991). Thus with our chosen parameter values the model can simulate the key baroclinic wave scales of the current atmosphere.

The solution of the model is the same as in the previous two wave study. A ∇_δ^4-type sub-grid damping term is added to the vorticity equation at each level to represent the frictional effect of small scales. The dependent variables are expanded in the eigenfunctions of the Laplacian operator in the horizontal, and an ODE solver based on the Runge-Kutta method is used to calculate the expansion coefficients over time. Tests at various resolutions indicate that, for this case of many waves unstable, 26 modes must be retained in both the meridional and zonal directions to yield a converged solution. Also, the simulations must be carried out for approximately one year (3000 non-dimensional time steps) and averaged over the last four months (1000 time steps). All experiments in the subsequent discussions were run in this way, starting from the wave-free Hadley state, with all zonal and meridional wavenumbers perturbed with random but small magnitude. For further details see Welch (1996).

We will measure the modeled climate in several ways. One simple measure is the zonally averaged temperature difference (or "gradient") across the mid-latitudes in the middle troposphere, which we label $\Delta\overline{T}_2$. An expression for this can be derived by zonally averaging the thermodynamic energy equation (2.3), and using the hydrostatic relation $T \sim -\partial\Psi/\partial p$, to yield:

$$\frac{\partial \overline{T}_2}{\partial t} = -\delta\frac{\partial}{\partial y}\overline{(v_2' T_2')} + 2\sigma_\circ\overline{\omega}_2 + 2h''\left(T^\dagger - \overline{T}_2\right). \quad (2.4)$$

At equilibrium, then, the temperature gradient across the channel, $\Delta\overline{T}_{2,\mathrm{eq}}$, can be approximated by (dropping the "2" subscript):

$$\Delta\overline{T}_{\mathrm{eq}} \approx \Delta T^\dagger + \frac{\sigma_\circ}{h''}\,\Delta\overline{\omega}_2 - \frac{\delta}{2h''}\,\Delta\frac{\partial}{\partial y}\left(\overline{v_2' T_2'}\right), \quad (2.5)$$

where the last term is the differential eddy heat flux convergence or "heat transport". We will vary the magnitude of the radiative equilibrium forcing $\Delta T^\dagger$ and measure the resultant $\Delta\overline{T}$ and the heat

[3]First recognize that L_y is determined by the zonal jet, for the latter acts as a waveguide, confining baroclinic disturbances to a meridional range approximately 30 degrees in latitude. Then from (2.8) of Welch and Tung (1997) we have $m_E = 1.08\,m$, so that the wavenumber in an "Earthly" channel and that in our chosen channel are effectively the same. This confirms the realism of our channel.

transport by various waves at equilibrium.[4] The contribution by the vertical velocity is small in all cases we studied.

3. Explaining the Transient Energy Spectrum

In addition to the observed transient variance of geopotential height, Fig. 1 shows the energy spectrum generated by the model at $\Delta T^{\dagger} = 90$K, in simulation of winter. For easier comparison with observations, only the first ten wavenumbers are included.[5] Like the observed geopotential data, the model energy has two local maxima: one at the planetary scale and one at synoptic. The peak at $m = 1$ is larger in the simulated results (relative to the synoptic scale peak). This could be because the observed data is variance of the geopotential height only, whereas the model is tracking the perturbation energy, including kinetic energy and quasi-geostrophic potential energy, the last of which is proportional to geopotential height variance. (See Appendix for detail of the energy calculation.) The overall shapes of the curves are strikingly close, in spite of this difference, an indication that the model seems to be replicating the situation in the real atmosphere. Spectra at higher and lower forcings (not shown) have similar shape, with the synoptic peak shifting to lower wavenumbers for higher forcing.

We note that for geostrophic turbulence arguments to apply here, the model must include enough unstable waves to yield the nonlinear triad interactions expected in a turbulent regime. One way to investigate this is to look for a so-called "enstrophy-flux subrange" of two-dimensional turbulence (Holloway 1986). In such a spectral region, the flux of energy from wavenumber to wavenumber is zero while the flux of enstrophy is constant. From dimensional analysis and energetics arguments, these constraints imply that the perturbation energy should obey a formula of the form:

$$E'_m \sim m^{-\alpha} \tag{3.1}$$

for some constant $\alpha > 0$, regardless of the level of forcing (Pedlosky 1987). In Fig. 3 we have plotted the perturbation energy E' vs. zonal wavenumber for three different forcings in our model. By using a log-log plot, we can easily identify an enstrophy-flux subrange in the wavenumbers $m \geq 8$, for which the spectra have slope $\alpha \approx 3.5$. This value is close to that predicted ($\alpha = 3$) for quasi-geostrophic turbulence in the shortest wave scales (Charney 1971). The fact that it is slightly larger in magnitude might be due to the fact that the beta effect has some bearing on the smaller scales (Pedlosky 1987, pp. 252-253.) Here we simply note that there does seem to be an enstrophy-flux subrange in our energy spectra and hence the signs of fluid turbulence in the largest wavenumbers.

[4] As in Welch and Tung (1997), all quantities are projected onto the $\cos y$ mode, which gives a good and easy approximation to the full value.

[5] The scale of the geopotential variance on Fraedrich and Bottger's (1978) Fig. 4 lower left is inconsistent with their Fig. 2 bottom row. Thus we do not include their scale in our Fig. 1 but only compare the shapes of the two curves.

a. Rhines' Wavenumber

Rhines (1975) developed a prediction of the wave scale which would dominate in a turbulent flow subject to Rossby-wave dynamics, which is relevant to our present parameter regime. He reasoned that nonlinear triad interactions amongst short waves would result in a reverse cascade of energy up-scale (Charney 1971), but that such interactions would diminish at the large scales, which are dominated by the linear dynamics of Rossby waves due to the large Coriolis force. Thus the up-scale cascade would terminate at some wavenumber for which the Coriolis force and the nonlinear terms have the same magnitude. Rhines calculated this wavenumber of cascade arrest as $k_\beta \equiv (\beta/2U_{\rm rms})^{1/2}$, which can be expressed non-dimensionally for our study as:

$$m_\beta \equiv \sqrt{\frac{\beta}{2\delta U_{\rm rms}}}. \tag{3.2}$$

Here $U_{\rm rms}$ is the non-dimensional magnitude of the root mean square ("rms") eddy velocity. We have applied the above formula to our model to see if this prediction accords with our results. Table 1 lists $U_{\rm rms}$ for a range of forcings and the resultant Rhines wavenumber, m_β, calculated using (3.2). The rms velocity is a time-average at equilibrium, averaged over the channel and over the two levels; thus it is a measure of the barotropic energy in the channel, as in the study by Haidvogel and Held (1980). For comparison, the table includes the wavenumbers of the long wave energy maximum, $(m_{\rm Emax})_{\rm long}$, and of the synoptic energy maximum, $(m_{\rm Emax})_{\rm syn}$, from model output (as in the solid curve of Fig. 1).

$\Delta T^\dagger$	$U_{\rm rms}$	m_β	$(m_{\rm maxE})_{\rm long}$	$(m_{\rm maxE})_{\rm syn}$	$m_{\rm max\Pi}$
20	0.039	3.0	2	10	10
50	0.098	1.9	1	7	6
80	0.140	1.6	1	5	6
110	0.180	1.4	1	5	5
140	0.217	1.3	2	4	6

Table 1: Rhines wavenumber, m_β, for various forcings as calculated from the root mean square eddy velocity, $U_{\rm rms}$, according to (3.2). The two waves at which the transient energy is a local maximum are shown as $(m_{\rm maxE})_{\rm long}$ for a long wave and $(m_{\rm maxE})_{\rm syn}$ for a synoptic wave. The wavenumber of dominant heat transport is indicated by $m_{\rm max\Pi}$. All calculations are at equilibrium.

It can be seen that Rhines' calculation does approximately predict the wavenumber of the long scale energy maximum. This is found at other forcings as well. An energetics analysis corroborates Rhines' idea, as displayed in Fig. 8. This figure is the case $\Delta T^\dagger = 90$K, for which we can interpolate from Table 1 that $m_\beta \approx 1-2$. The dashed line in the figure is the nonlinear gain or loss of energy by each wave. There is a strong nonlinear transfer of energy from short to long waves, peaking at $m = 1$. Thus a very long wave does seem to accumulate energy due to a cascade caused by nonlinear interactions of short waves. (This figure will be discussed further below.)

Table 1 demonstrates, however, that Rhines' argument does not explain the presence of a synoptic scale energy peak, because it predicts a wave scale which is much longer. We note that the derivation of m_β assumed a barotropic fluid with no forcing and no viscosity; Rhines' nonlinear transfer does not depend on the degree of baroclinicity of the fluid system. Now, in our case baroclinic instability determines both the amount of energy injected into the eddies and the scale of such energy injection. Since Rhines' upscale cascade does not depend on baroclinic stability *per se* (hence it should work for the baroclinic atmosphere as well) it should explain how that injected energy is distributed among different scales. It is surprising, then, that Rhines' mechanism does not determine the synoptic energy peak. The processes which do select this shorter scale peak must directly involve the baroclinicity of the system.

Also shown in the table is the wavenumber of the dominant heat transporting wave at equilibrium, $m_{\max\Pi}$ (determined from Fig. 6 below). The scale of this wave is always close to that of the synoptic scale energy maximum. Thus the explanation for the selection of the synoptic energy maximum may be tied to the explanation for the dominance of the main heat-transporting wave. We propose the mechanism of nonlinear baroclinic adjustment to explain the latter, and hence the former.

b. Nonlinear Baroclinic Adjustment

The theory of nonlinear baroclinic adjustment was presented in the study of Cehelsky and Tung (1991) and elaborated by Welch and Tung (1997). Here we will review the concept, corroborating it with model simulations, and we show how it can explain the synoptic peak in the energy spectrum.

Nonlinear baroclinic adjustment is motivated by the recognition of existence of an "atmospheric thermostat". There seems to be a mechanism which maintains the mid-latitude temperature gradient in the middle troposphere at a fairly constant value, even while the forced radiative equilibrium temperature gradient varies considerably with the seasons (Stone 1978). The theory posits that, given a certain forcing, there is a certain amount of heat which the sum of *all* baroclinically unstable waves attempt to transport poleward. The amount of heat increases linearly with the forcing, so that the resultant temperature gradient attains its robust value.

Both of these features are found in our model simulations, as shown in Fig. 4: the solid line rises linearly with forcing while the crosses are fairly constant. Note that the theory does not assume that the heat is transported by the most unstable wave alone, which was Stone's (1978) original proposal of *linear* baroclinic adjustment. He argued that no other waves would participate because they would be linearly stabilized by the action of the most unstable wave, which would reduce the flow to the minimum critical shear or temperature gradient. This, however, would yield the most unstable wave ($m \approx 12 - 15$) as the dominant heat transporter in the real atmosphere, whereas observations show it to be $m = 4 - 7$ (Randel and Held 1991). In the theory of *nonlinear* baroclinic adjustment, the heat transport is spread over many scales, and the wave which ends up dominating this process at equilibrium is longer than that most unstable.

The rule which determines how heat is transported by the various wavelengths is based on the concept of a nonlinear threshold for each wave, above which it can no longer grow. At this limiting magnitude, the wave will break and saturate, shedding excess energy to other scales. This is demonstrated in Fig. 5 with a model run at $\Delta T^{\dagger} = 120$K. Each frame shows contours of upper layer PV ("Q") vs. x and y at a certain point in time in the evolution to equilibrium, with solid (dotted) contours indicating total PV greater (less) than the average planetary contribution, $f_{\circ}$. The run is started from the (perturbed) Hadley solution, which at this forcing is supercritical to wavenumbers 1-14 (Fig. 2, dashed curve). Early on at 1.2 days, the most unstable wave (wavenumber 11) is the primary perturbation to the PV contours, as can be seen from the number of peaks in the perturbed $f_{\circ}$-contour in panel a. This wave does not maintain dominance, however. As the evolution proceeds, the dominant wave observed becomes longer and longer: from $m = 11$ initially to $m = 10$ (not shown) to $m \approx 9$ at 2.3 days (panel b) to $m \approx 8$ at 3.5 days (panel c), etc. Finally at equilibrium, it is $m \approx 6$ which is dominant, as shown at 92.6 days in panel d. This agrees with the data of Fig. 6 below.

Dominance shifts to a longer scale when the PV contours of the currently dominating wave curl up such that regions of negative PV gradient are created. Initially the gradients are dominated by β and hence are everywhere positive, as in panel a. Waves then begin to grow, distorting the contours. If a wave becomes very large, long thin fingers of PV start to form, as in panel b for wavenumber 9. Eventually these fingers begin to curl over onto themselves, creating regions of negative PV gradient; for example, see the top left section of panel b or the top of panel c. Where this occurs, small scale instabilities begin to grow. Panel c shows this, and the PV fingers that curl over are being "pinched off" by instabilities which have arisen on the side(s). Small "blobs" of PV thus are created, and these are in turn broken up, as they have negative PV gradients due to their closed shape. The blobs become smaller and smaller until they are dissipated by (sub-grid) friction. In this way, further growth of wavenumber 8 in panel c is prevented. The next longer wave begins to emerge in the PV contours as it grows larger than the breaking wave.

This idea is consistent with the criterion for instability given by the Charney-Stern Theorem: if the background potential vorticity gradient is zero or negative somewhere in the fluid, then the fluid may be unstable to small scale secondary perturbations (Charney and Stern 1962). That theorem was derived for the case of a zonally uniform flow, and hence the background flow was just the zonal mean. Here we are concerned with the stability of a zonal mean flow plus a heat transporting wave or waves. These waves are of large scale, however; from the point of view of small-scale secondary perturbations, such planetary or synoptic waves vary so slowly in the zonal direction as to appear locally constant. Thus we generalize Charney and Stern's idea by applying it to the *total* background flow, including zonal mean and wavy contributions. We claim that when the *total* PV gradient $Q_y = \overline{Q}_y + Q'_y$ becomes negative, the flow becomes potentially unstable. The dominant wave then breaks, shedding blobs of PV, and it cannot grow further. This was described above by the panels of Fig. 5.

This limit on the total PV gradient is in fact a limit on the wavy part, for the zonal mean gradient has a roughly constant (positive) value independent of forcing. To demonstrate this, consider the non-dimensional PV on each level, defined by:

$$Q_j \equiv 1 + \beta\left(y - \frac{\pi}{2}\right) + \nabla_\delta^2 \Psi_j + \frac{j-2}{2\sigma_\circ}\left(\Psi_1 - \Psi_3\right), \quad j = 1,3 \tag{3.3}$$

where the "1" is a non-dimensionalized $f_\circ$. The equilibrated zonal average can be approximated:

$$\overline{Q}_j \approx 1 + \beta\left(y - \frac{\pi}{2}\right) + \gamma_j \overline{T}_{2,\mathrm{dim'l}}(y), \quad j = 1,3 \tag{3.4}$$

where the relative vorticity is small compared to the planetary, as confirmed by model output, and hence has been neglected. The γ_j are non-dimensionalizing constants. (Note that $\gamma_1 < 0$.) The zonal mean meridional PV gradient on the upper level is then:

$$\overline{Q}_{1,y} = \frac{\partial \overline{Q}_1}{\partial y} \approx \beta + \gamma_1 \left(\frac{\partial \overline{T}_2}{\partial y}\right)_{\mathrm{dim'l}}. \tag{3.5}$$

We know that the cross-channel temperature gradient, $\Delta\overline{T}_2$, is robust at equilibrium from Fig. 4. In addition, the temperature approximately retains its forced $\cos y$ profile throughout the evolution (not shown). These two facts combined mean that $\partial\overline{T}_2/\partial y$ (at any y) is also fairly constant with forcing at equilibrium. Hence from (3.5) it must follow that the zonal mean PV gradient on the upper level retains roughly the same value regardless of forcing. (The same is true in the lower level.)

Now the wavy component of PV ($Q'_{y,m}$) oscillates in sign by definition. Thus the total PV gradient will be negative when the magnitude of the wavy component becomes larger than $\overline{Q}_y$. This puts a limit on how large the wavy PV gradient can be before it will render the flow unstable by creating regions of negative total Q_y.

Finally, the size of the wavy PV gradient can be related to the streamfunction magnitude, and through that to the amount of the heat flux. As derived in Welch and Tung (1997), we have:

$$< \left|Q'_{1,y}\right|_m > \sim m^{3/2} \sqrt{< |\Pi_m| >}. \tag{3.6}$$

Here Π_m is zonal average eddy heat flux convergence by wave m. The vertical lines indicate absolute values, and the angle brackets a meridional average. This relation shows that the same heat transported by a longer wave (smaller m) will yield a lower wavy PV gradient than if transported by a shorter wave (larger m). According to the generalized Charney-Stern theorem, then, we expect shorter waves to yield larger wavy PV gradients and render the flow unstable, i.e. to reach their nonlinear saturation level, at lower forcings (i.e. lower heat fluxes) than will longer waves.

With these thresholds in mind, let us return to the question of equilibration. For a given forcing, Fig. 4 shows the total amount of heat which must be transported to reach the robust equilibrium. If this exceeds the threshold of the most unstable wave (as it often does by a wide margin), then that wave will break and the excess heat will be taken up by the next longer wave, because it has a higher threshold. If this next longer wave reaches *its* threshold, then it will break and shed its energy. This process will continue until the total heat that must be transported has been accounted for, and thus the heat is spread over a spectrum of baroclinic waves. (Notice that relationship (3.6) above gives another argument for the *up-scale* cascade of Charney (1971). Energy cannot cascade to a shorter scale, even if that scale is linearly unstable, for it has a nonlinear threshold which is even smaller than that of the saturating wave.) Fig. 6 shows the simulated equilibrium heat transport by each of the most active waves over a range of forcings. At all but the lowest forcings there is significant contribution by more than just the most unstable wave ($m = 11$, from Fig. 2). Fig. 6 also demonstrates the presence of nonlinear thresholds. At $100 \leq \Delta T^{\dagger} \leq 150$K, wavenumbers 8-11 each have a heat transport that does not rise much with the forcing, indicating that these scales have reached saturation and have passed energy to the longer scale wavenumbers 4-7, which dominate at these forcings. Furthermore, the saturation level for the shorter waves are lower than for the longer waves: $(\Pi_{11})_{\rm satn} < (\Pi_{10})_{\rm satn} < (\Pi_9)_{\rm satn} < (\Pi_8)_{\rm satn}$, which corroborates the relation (3.6).

When an unstable wave reaches its saturation level, it is "nonlinearly stabilized" by sending energy to other scales. The longest wave to transport heat, however, will not have reached its threshold; hence it will not be stabilized nonlinearly at equilibrium. Rather, it stabilizes itself: through its heat transport, it reduces the zonal mean temperature gradient to a level at which it is linearly neutral, in the manner Stone envisioned for the most unstable wave. This is verified in Fig. 4. The squares mark the critical temperature gradient (from a linear stability analysis) of the dominant heat transporting wave at equilibrium for each forcing. The equilibrated temperature gradient falls near to the squares in each case, indicating that indeed the flow has been made approximately neutral with respect to the dominant wave. In the process, all longer waves have been rendered linearly stable. This can be seen from the marginal stability curve for the *equilibrated* flow at $\Delta T^{\dagger} = 90$K (solid line) in Fig. 2. Wavenumber 5 is the longest wave to transport heat in this case, and it reduces $\Delta \overline{T}$ down to near its critical gradient, $\Delta T_{\rm cr,5}$. This is lower than the critical gradient for longer waves, and thus the waves $m = 1 - 4$ are linearly stabilized at equilibrium.

It is this last wave, $m = 5$ for the example above, which acts as the atmospheric thermostat. It is the flexible component of the system, determining the equilibrium temperature gradient essentially by itself via a quasi-linear mechanism. In most cases this wave does the bulk of the heat transport and thus we call it the dominant heat transporting wave.[6] For all but the smallest forcings, this wave will be longer than the most unstable wave, the latter having been nonlinearly saturated.

The reason that the cascade does not continue further up-scale is that all the necessary heat transport has been borne by the shorter waves. This is a crucial difference between the synoptic peak achieved by nonlinear baroclinic adjustment and the planetary peak from the Rhines' mechanism of up-scale (barotropic) energy cascade. In the former, there appears to be an "objective": to transport just enough heat to maintain the atmospheric "thermostat". When that objective is achieved, no further up-scale cascade is needed, and hence it stops at the scale of the dominant heat transporting wave. The dominant wave then reduces the temperature gradient such that longer waves are linearly stabilized and do not participate in the dynamics. This equilibration process

[6] Depending on the level of forcing, this wave could actually be transporting *less* heat than its next shorter neighbor (which has saturated). However, it is still the *flexible* wave in the system, thus determining the dynamics, and in this sense it is still dominant.

depends on the radiative driving. For higher forcing, more heat flux will be required and hence more waves needed to transport heat to achieve equilibrium. The cascade will continue until a lower wavenumber is selected as dominant. This can be seen in Fig. 6: the wave which transports the most heat shifts to a longer and longer scale as the forcing $\Delta T^{\dagger}$ is raised.

Our theory is summarized in the schematic diagram of Fig. 7. Note that what determines when the temperature gradient shifts to a new regime is not the linear critical gradients of the waves, but rather their nonlinear saturation levels. Also, the schematic has an exaggerated ordinate for clarity; in the model atmosphere, the critical gradients of the various waves are very tightly packed. This is demonstrated for the equilibrated flow at $\Delta T^{\dagger} = 90\mathrm{K}$ in Fig. 2 by the flatness of the curve over many wavenumbers; the same shape is found at other forcings (not shown). Thus the levels of equilibration for the different regimes of Fig. 7 are actually very close in magnitude, producing a relatively constant temperature gradient over a range of forcings, as found in observations (and in Fig. 4).

We have used the theory of nonlinear baroclinic adjustment to explain the selection of the dominant heat transporting mode. This argument can also be used to explain the existence of a dominant synoptic energy mode, for the two scales are always close, as shown in Table 1. The energy cascades up-scale to a wavenumber that is lower than that most unstable, but the cascade does not continue to the longest scales as in the theory of Rhines (1975). It stops roughly at the dominant heat transporting wave. Now, there is indeed a difference between the dominant heat transporting mode and the (synoptic scale) mode with maximum perturbation energy. First, the wave of maximum energy is not necessarily the wave of maximum energy *extraction* from the mean flow. Model simulations do show them to be close, however (not shown). The wave of maximum energy often is just slightly longer than that of highest extraction, seemingly a byproduct of the continual cascade of energy to longer waves. Further, the wave of maximum energy extraction is not the same as the wave of maximum heat transport. However, both are quasi-linear quantities, measuring interaction between a wave and the mean flow, and model simulations do show that these two scales are also close. Thus one would expect that the (synoptic) wavelength of maximum energy and that of maximum heat transport would be close, but not always the same. This closeness implies that nonlinear baroclinic adjustment is the relevant mechanism to explain the occurrence of a synoptic scale peak in the energy spectrum.

An issue might be raised as to the relevance of the two-level model's energy spectrum to the real atmosphere. Due to its vertical simplicity, a two-level model has a critical temperature gradient separating stable and unstable waves, so that the longest waves are often stable. In contrast, in the continuous atmosphere there is no distinct cut-off and *all* waves are unstable. While the long waves have smaller growth rates than do the synoptic scales (Green 1960; Kuo 1973; Gall 1976), this does not necessarily prevent the long waves from achieving the largest energy (Schneider 1981). Thus, one might ask the question: Is the (solid line) energy spectrum of Fig. 1 wrong, because the longest waves in the two-level model are restricted from growing as much as they would in the real atmosphere?

This issue is not of consequence here, however. First, we showed above that the maximum of energy in the longest scale wave is due to a nonlinear cascade and not to the baroclinic instability of the fluid. The degree of instability (or stability) of the longest waves does not matter in the development of the argument; indeed, Rhines derived his wavenumber of cascade arrest based on a *barotropic* fluid. Furthermore, we have seen that the energy spectrum generated by our two-level model is similar to that in the real atmosphere. In the dashed line observations of Fig. 1 there is a local maximum of transient variance at the longest wave, on the same order of magnitude as the maximum at synoptic scales. In our model here there is a similar situation, even *without* the longest waves being unstable. Therefore, the vertical simplicity of the two-level model does not seem to compromise its ability to simulate the tropospheric energy spectrum.

c. *Relation to Quasi-Geostrophic Turbulence*

Our nonlinear baroclinic adjustment theory accords with the wave-wave equilibration theory of Salmon (1980) in his study of quasi-geostrophic turbulence. This can be seen in Fig. 8, which shows the perturbation energy growth, and components thereof, for each zonal mode at equilibrium for $\Delta T^{\dagger} = 90$K. (Calculations follow the method of Whitaker and Barcilon (1995); see Appendix.) The solid line represents quasi-linear extraction of energy from the mean flow, the dashed is nonlinear energy transfer between waves, and the dot-dashed line is the total of the linear processes: Ekman friction, sub-grid damping, and Newtonian cooling. Note that the components of energy change sum to approximately zero in our time-average, as indicated by the thin solid line, proof that the system is indeed at equilibrium.

The figure shows that there is nonlinear transfer of energy out of wavenumbers 5-12, while wavenumbers 1-4 gain energy nonlinearly. That is, the most unstable waves lose energy to longer, less unstable waves. This is due to the fact that these shorter heat transporting waves have reached their saturation levels and are breaking, as described above, and their energy is taken up always by waves of longer scale due to the relation (3.6). Fig. 8 also shows that the longer waves are losing energy to dissipation.[7] Because the most unstable waves continually lose energy nonlinearly, they can be maintained in a state in which they are linearly unstable. This is confirmed by the marginal stability curve at equilibrium for the case $\Delta T^{\dagger} = 90$K in Fig. 2 (solid line). The mean flow is supercritical to the most unstable waves at equilibrium.

The above energetics fit exactly with Salmon's wave-wave equilibration theory, which he developed to explain equilibration in a fluid with a forced mean baroclinic state, including Ekman friction at the Earth's surface (Salmon 1980). An equilibrium is achieved in which the most unstable waves gain energy quasi-linearly from the forcing (i.e. from the mean flow) and lose it nonlinearly to longer, less unstable waves on which dissipation acts. With such an energy balance, an equilibrium which is "supercritical" relative to the most unstable waves is possible.

It is now apparent that Salmon's turbulence theory can apply to the two wave case of Welch

[7] We note that our results are very similar to those of Whitaker and Barcilon (1995); compare to their Figs. 7 and 9.

and Tung (1997) (i.e. $\delta = 1.3$) as well. The dynamics for that case are exactly the same as in Fig. 8, except that all of the behavior is compressed into the two available modes and hence appears slightly different. In the two wave case, the single long mode must account for the dynamics of *all* the long waves in the theory above, and the single short mode must account for *all* the short waves, as we now show.

Consider the wavenumbers of the many wave case in Fig. 8 divided into the groups $m = 1 - 5$ and $m = 6 - 10$.[8] (We neglect wavenumbers $m > 10$ because they contribute little.) The energetics can be summarized in three points. First, modes $1 - 5$ have a net quasi-linear gain of energy from the mean flow, but this gain is smaller than that of modes $6 - 10$. Second, modes $1 - 5$ experience a net increase in energy due to nonlinear transfer, while the shorter wave group in net loses energy nonlinearly. Finally, the loss to dissipation is greater for $m = 1 - 5$ than for $m = 6 - 10$. In the two wave study, the roles of these two wave groups $m = 1 - 5$ and $m = 6 - 10$ are played by the long mode $m1$ and the short mode $m2$, respectively. Compare, for example, Fig. 6T in Welch and Tung (1997) with the wave groups of Fig. 8 here.

Thus the dynamics with two waves unstable seem to be just a compressed version of the situation in the many wave case.[9] With a slight adaptation, Salmon's wave-wave equilibration theory indeed can explain the energetics in the part of parameter space with few waves unstable. This removes the gap between his theory from quasi-geostrophic turbulence and the nonlinear baroclinic adjustment mechanism presented here.

4. Comparison of Winter vs. Summer and vs. Observations

As a test of the relevance of the model, the results have been compared with observations of the real atmosphere. First, recognizing that the static stability varies between the seasons, more specific values of $\sigma_\circ$ have been calculated for winter and summer separately, using data from Peixoto and Oort (1992). From their Fig. 7.5, and with $\rho = \rho_4 e^{-z/H}$ with $H \approx 7\text{km}$ and $\rho_4 \approx 1.275\text{kg/m}^3$ from Wallace and Hobbs (1977), we determine the following values for (non-dimensionalized) $\sigma_\circ$:

$$\begin{aligned} (\sigma_\circ)_{\text{winter}} &\approx 0.085 \\ (\sigma_\circ)_{\text{summer}} &\approx 0.077 \end{aligned} \tag{4.1}$$

We also choose the magnitude of forcing for each season separately, with

$$\begin{aligned} \left(\Delta T^\dagger\right)_{\text{winter}} &\approx 84\text{K} \\ \left(\Delta T^\dagger\right)_{\text{summer}} &\approx 56\text{K} \end{aligned} \tag{4.2}$$

[8]The groups could contain slightly different waves with the same effect, because the waves are "tightly packed" for this geometry. This is due to the smallness of δ, which causes the effective wavenumbers $m\delta$ of the waves to be close, and their critical gradients similar in magnitude (see Fig. 2). Hence the waves have similar behavior and can substitute for each other. This was not the case with only two waves unstable because of the relatively large $\delta = 1.3$.

[9]Recall that in the two wave study the long mode corresponds to wavenumber 5 on the Earth ($m_E = 5$), and the short wave to $m_E = 10$. Hence the grouping of waves above is even more compelling.

See Peixoto and Oort (1992), Fig. 6.14d.

Results are shown as crosses in Fig. 9. Notice first that the equilibrated temperature gradients are much smaller than the radiative equilibrium Hadley solution, agreeing with observations. Also note that $\Delta\overline{T}_{\mathrm{eq,wtr}}$ is slightly larger than $\Delta\overline{T}_{\mathrm{eq,sum}}$, which meets with our expectations.

Included in the figure are observed values of the temperature gradient in mid-channel, i.e. $-2\partial\overline{T}_{\mathrm{obs'd}}/\partial\hat{y}$ at $\hat{y} = \pi/2$, in winter and summer, taken from Peixoto and Oort (1992) Fig. 7.5. Fig. 9 shows that the results of the model are strikingly close to these observations. The model has approximated well the large reduction of the temperature gradient from the radiative equilibrium value. Moreover, the model results mimic the slight rise in equilibrium gradient from summer to winter. Thus we have some evidence that our model, albeit a simple two-level quasi-geostrophic model in a beta-plane channel, as well as the nonlinear baroclinic adjustment mechanism which it demonstrates, are relevant to equilibration in the real atmosphere.

5. Relation to Other Theories

The neutralization proposal by Lindzen (1993, 1994) is clearly different from our theory, because the former requires *all* waves to be linearly stable at equilibrium. Nonlinear baroclinic adjustment proposes that waves can be stabilized by different methods, in particular *non*linearly; not all waves must be linearly stabilized in order to achieve an equilibrium. This is evident in the marginal stability curve at equilibrium of Fig. 2 (dashed line); only the long waves have been linearly neutralized, while the short waves remain linearly unstable even at equilibrium. This was also found in the two wave study of Welch and Tung (1997), and thus appears to be true throughout parameter space.

Finally, we note that our nonlinear baroclinic adjustment theory, which is based on the presence of normal modes, differs from the proposals of Farrell and collaborators (Farrell and Ioannou 1994; DelSole and Farrell 1995; Farrell and Ioannou 1995; DelSole and Farrell 1996). Those studies used *linearized* quasi-geostrophic equations, showing how the non-normality of the linear operator can lead to transient growth of non-modal waves. Starting with a single initial value problem, Farrell (1989) demonstrated that an optimal perturbation in such a system can mimic the development of a single mid-latitude cyclone. The combined effects of many cyclones, i.e. the effect of baroclinic instability on climate, then can be simulated by continuously forcing the linearized system. This is done by modeling the omitted nonlinear interactions as a stochastic forcing in all wavelengths (plus a linear dissipation). Using this technique in the aforementioned works, the above authors were able to generate the eddy variance, i.e. heat and momentum fluxes (amongst other features), observed on synoptic scales in the real atmosphere. Therefore, they argued, it is transient growth due to non-normality of the linearized operator which causes baroclinic equilibration.

Here we do not make any assumption concerning the modal structure of the solution. Our calculations were done by stepping the fully nonlinear equations forward in time, allowing many waves to interact with each other and with the mean flow. Thus both normal modes and non-modal waves are allowed. The fact that the results here can be explained better with only normal mode thinking may be an issue of the function of nonlinearities, as we now discuss.

In a linearized problem for which the operator possesses normal modes but is non-normal, transient growth at small times gives way to normal mode (exponential) growth or decay at later times. This has been demonstrated by Farrell (1982) for the most unstable wave in the Eady problem. (See his Fig. 4.) Farrell and colleagues' non-normal linear theory of statistical baroclinic equilibration assumes that the system never proceeds past the transient stage to see the effects of growing (or decaying) normal modes. Energy is scattered by nonlinearities (parameterized as stochastic forcing), preventing evolution to an equilibrium state. On the other hand, our calculation of nonlinear baroclinic adjustment allows for short-term transiency, but it does not require that the system remain in this linear phase. The flow evolves past into a nonlinear regime, wherein the amplitudes of the normal modes are kept in check by nonlinear interactions among waves. In our case the discussion is particularly simple because the non-normal growth does not alter the mean flow significantly; hence normal modes of the original wave-free state are similar to those of the final state, and we effectively can follow the same normal mode through from initial linear instability to equilibration.

The key difference between the non-normal theory and ours, then, is the nonlinear interactions and the time scales over which they are presumed to have impact. Farrell and colleagues assume that nonlinearities act rapidly and catastrophically to scatter energy and interrupt the evolution to equilibrium. In our work here, we have made no such assumption, allowing the nonlinear interactions to occur at whatever pace the governing equations dictate. A more detailed inspection of the observed wave-wave (and wave-mean) interactions in the real atmosphere, and a comparison thereof with our equilibrium dynamics and with the stochastic forcing suggested by Farrell, could shed more light on this issue.

6. Summary and Conclusions

We have shown that the two maxima of energy in the transient spectrum of the mid-latitudes can be explained by two different theories. The long wave maximum is due to barotropic up-scale energy transfer and the cessation thereof at scales where linear Rossby wave dynamics take over, as proposed by Rhines (1975). The synoptic scale maximum is related to the amount of heat that needs to be transported at a given radiative forcing, as determined by the theory of nonlinear baroclinic adjustment. In the current climate this scale would have been the linearly most unstable baroclinic wave, except that that wave is unable (without reaching its saturation limit) to transport the large amount of heat required to neutralize the atmosphere; hence that most unstable wave breaks. The remaining heat is then transported by the next longer wave, and so on. By calculating the total amount of heat that must be tranported by all the waves and the saturation thresholds of each, one can then estimate the synoptic wave spectrum. Such an estimate is found to be near that of the observed spectrum. Note that the presence of baroclinic instability does not appear to interrupt the Rhines wavenumber cascade, nor does that cascade significantly affect the selection of the dominant heat transporting wave.

Combined with the results of Welch and Tung (1997), the nonlinear baroclinic adjustment mechanism has been shown to work throughout parameter space, explaining the equilibration of baroclinic flows with any number of waves unstable and at all forcings. This mechanism can account for the observed flexibility of meridional eddy heat transport in the atmosphere. An important part of the theory is that waves can be stabilized by different methods, linear and nonlinear, and that not all waves must be linearly stabilized in order to achieve an equilibrium. This is different from the neutralization theory proposed by Lindzen.

Finally, we have identified the similarity between the nonlinear baroclinic adjustment mechanism and the wave-wave equilibration theory of geostrophic turbulence. This latter theory applies to the subset of cases in which many waves are unstable and, with a slight reinterpretation, it can be seen to operate in the few-wave case as well. Wave-wave equilibration describes the overall dynamics of a baroclinic fluid state at equilibrium. Nonlinear baroclinic adjustment, on the other hand, also gives the dynamic mechanism for each wave and explains how the waves interact over time to yield a final equilibrium state. Its difference from Farrell's ideas of stochastically forced transient growth has been shown to lie in the time scale assumed for nonlinear interactions.

Although the model used here is quite simple, with (for example) quasi-geostrophic scaling and a fixed value of static stability, it seems able to represent qualitatively the observed dynamics of poleward atmospheric heat transport. It can simulate approximately the transient energy spectrum and the zonal mean meridional temperature gradient in winter vs. summer. Thus the mechanism proposed seems promising, but certainly testing it with more sophisticated models is an important next step.

Acknowledgments. WTW's research has been supported in part by NSF through its sponsorship of NCAR, in part by NASA under its Global Change Fellowship Program (grant 1812-GC92-0169), and in part by the following grant; KKT's research is supported by NSF's Climate Dynamics Program, under grant ATM-9526136.

Appendix

Energetics Calculation

To analyze the dynamics of our simulations, both at equilibrium and during the evolution, we have used an energetics analysis from the work of Whitaker and Barcilon (1995). Here we will derive the energy equation for our particular model. All quantities are non-dimensional unless otherwise indicated.

First we reformulate the vorticity and thermodynamic energy equations (2.1) - (2.3) into two equations for the conservation of potential vorticity ("PV") at each level. Solving the thermodynamic equation for ω_2 and substituting into the upper vorticity equation yields:

$$\begin{aligned}
\frac{\partial}{\partial t}\nabla_\delta^2\Psi_1 - \frac{1}{2\sigma_\circ}\frac{\partial}{\partial t}\left(\Psi_1 - \Psi_3\right) &= -\delta J\left(\Psi_1, \nabla_\delta^2\Psi_1 + \frac{1}{2\sigma_\circ}\Psi_3\right) - \delta\beta\frac{\partial\Psi_1}{\partial x} \\
&\quad -\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) + 2\nu_s\nabla_\delta^4\Psi_1 \\
&= -\delta J\left(\Psi_1, f + \nabla_\delta^2\Psi_1 + \frac{1}{2\sigma_\circ}\Psi_3 - \frac{1}{2\sigma_\circ}\Psi_1\right) \\
&\quad -\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) + 2\nu_s\nabla_\delta^4\Psi_1, \qquad \text{(A1)}
\end{aligned}$$

where we have used the fact that $J(\Psi_1, \Psi_1) = 0$. The above equation can be rewritten as:

$$\frac{DQ_1}{Dt} = \frac{\partial Q_1}{\partial t} + \delta J\left(\Psi_1, Q_1\right) = -\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) + 2\nu_s\nabla_\delta^4\Psi_1, \qquad \text{(A2)}$$

where the upper and lower level potential vorticities, Q_1 and Q_3, are defined by:

$$\begin{aligned}
Q_j &\equiv f + \nabla_\delta^2\Psi_j + \frac{2-j}{2\sigma_\circ}\left(\Psi_3 - \Psi_1\right), \quad j = 1, 3, \quad \text{with} \qquad \text{(A3)} \\
f &\equiv 1 + \beta\left(y - \frac{\pi}{2}\right). \qquad \text{(A4)}
\end{aligned}$$

(Note: the "1" in (A4) is a non-dimensionalized $f_\circ$.) Similarly we can develop a PV equation for the lower level:

$$\begin{aligned}
\frac{DQ_3}{Dt} = \frac{\partial Q_3}{\partial t} + \delta J\left(\Psi_3, Q_3\right) &= +\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) \\
&\quad +2\nu_s\nabla_\delta^4\Psi_3 - 2\nu\nabla_\delta^2\Psi_3, \qquad \text{(A5)}
\end{aligned}$$

For ease let us abbreviate the forcing terms by defining:

$$\begin{aligned}
D_1 &\equiv -\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) + 2\nu_s\nabla_\delta^4\Psi_1 \\
D_3 &\equiv +\frac{h''}{\sigma_\circ}\left(\Psi_3 - \Psi_1 - \Psi_3^\dagger + \Psi_1^\dagger\right) + 2\nu_s\nabla_\delta^4\Psi_3 - 2\nu\nabla_\delta^2\Psi_3. \qquad \text{(A6)}
\end{aligned}$$

This allows the PV equations to be written:

$$\frac{\partial Q_j}{\partial t} = -\delta J\left(\Psi_j, Q_j\right) + D_j, \quad j = 1, 3 \qquad \text{(A7)}$$

Now, divide each term in (A7) into a zonally and time averaged part, indicated by an overbar, and a perturbation to the time-zonal mean, indicated by an apostrophe, i.e. $g = \overline{g} + g'$, where

$$\overline{g} \equiv \frac{1}{\tau}\int_t^{t+\tau}\frac{1}{2\pi}\int_0^{2\pi} g\ dx\ dt \qquad \text{(A8)}$$

for any function $g(x,y,t)$. Here τ is the time averaging period, taken to be approximately four months (1000 time steps), which is much longer than several typical baroclinic life cycles (4-6 days). After dividing each term thus, (A7) becomes:

$$\begin{aligned}\frac{\partial Q_j'}{\partial t} &= -\delta J\left(\Psi_j', Q_j'\right) - \delta J\left(\overline{\Psi}_j, Q_j'\right) - \delta J\left(\Psi_j', \overline{Q_j}\right) + \overline{D}_j + D_j' \\ &= -\delta J\left(\Psi_j', Q_j'\right) - \delta \overline{u}_j \frac{\partial Q_j'}{\partial x} - \delta v_j' \frac{\partial \overline{Q}_j}{\partial y} + \overline{D}_j + D_j', \quad j = 1,3 \qquad \text{(A9)}\end{aligned}$$

where several terms have dropped out due to the definition (A8). This is the same as equation (A.1) in Whitaker and Barcilon (1995).

We can derive an energy equation for this fluid system by multiplying (A9) by $-\Psi_j'$ and manipulating the result, to yield:

$$\begin{aligned}\frac{\partial E'}{\partial t} &= \sum_j \left\{ \left[\delta \Psi_j' \left(J\left(\Psi_j', Q_j'\right) + \overline{u}_j \frac{\partial Q_j'}{\partial x} + v_j' \frac{\partial \overline{Q}_j}{\partial y} \right) \right.\right. \\ &\left. + \frac{\partial}{\partial x}\left(\Psi_j' \frac{\partial^2 \Psi_j'}{\partial x \partial t} \right) + \frac{\partial}{\partial y}\left(\Psi_j' \frac{\partial^2 \Psi_j'}{\partial y \partial t} \right) \right] \\ &\left. - \Psi_j' \overline{D}_j - \Psi_j' D_j' \right\}, \quad j = 1,3 \qquad \text{(A10)}\end{aligned}$$

where E' is the perturbation energy (to the zonal and time mean flow), defined by:

$$\begin{aligned}E' &\equiv \frac{1}{2}\left\{ \delta^2 \left(\frac{\partial \Psi_1'}{\partial x}\right)^2 + \left(\frac{\partial \Psi_1'}{\partial y}\right)^2 + \delta^2 \left(\frac{\partial \Psi_3'}{\partial x}\right)^2 + \left(\frac{\partial \Psi_3'}{\partial y}\right)^2 \right. \\ &\left. + \frac{1}{2\sigma_\circ}\left(\Psi_3' - \Psi_1'\right)^2 \right\} \qquad \text{(A11)}\end{aligned}$$

Note that E' includes the kinetic energy at both levels plus the potential energy due to the baroclinic shear.

We will call the term in the energy equation (A10) within square brackets the "nonlinear" term, or N. (Its last two time derivatives are actually from the left hand side of the PV equations (A9) and thus do not really contain nonlinear dynamics.) We can split N into a part that represents quasi-linear interaction of a wave with the mean flow and a part that represents wave-wave interactions, i.e. $N = N_q + N_w$, where:

$$N_w \equiv \delta \sum_j \Psi_j' J\left(\Psi_j', Q_j'\right) \qquad \text{(A12)}$$

and N_q is the remainder. (This will make sense when we take the zonal average, below.) Thus the perturbation energy equation (A10) can be written:

$$\frac{\partial E'}{\partial t} = N_q + N_w - \sum_j \left(\Psi_j' \overline{D} j + \Psi_j' D_j' \right), \quad j = 1, 3 \tag{A13}$$

Finally, we take the horizontal average of the above equation, where:

$$\overline{\overline{g}} \equiv \frac{1}{2\pi} \int_0^{2\pi} \frac{1}{\pi} \int_0^{\pi} g \; dy \; dx \tag{A14}$$

for any function $g(x, y, t)$.[10] Due to boundary conditions, including periodicity in x, all but one term of N_q equal zero and thus we have:

$$\frac{\partial \overline{\overline{E'}}}{\partial t} = \overline{\overline{G}} + \overline{\overline{N}}_w - \overline{\overline{\Psi_j' \overline{D}_j}} - \overline{\overline{\Psi_j' D_j'}}, \tag{A15}$$

where

$$\overline{\overline{G}} \equiv -\delta \sum_j \overline{\overline{v_j' Q_j' \overline{u}_j}}. \tag{A16}$$

$\overline{\overline{G}}$ is is the "generation" term of Whitaker and Barcilon (1995), representing the amount of energy generated in the waves by quasi-linear extraction from the mean flow.[11]

Now equation (A15) can be divided into contributions from each zonal wave. For example, terms which are $O(\Psi'^2)$, such as $\overline{\overline{G}}$ and those involving D_j, involve products of wavenumbers $+m$ and $-m$, summed up over all m. Rigorously speaking, this energy equation is only valid for the total of all zonal modes; it is not true for each wavenumber m individually. However, we tested our model and determined experimentally that (A15) *is* approximately true separately for each zonal wave.

Therefore, we have used the perturbation energy equation (A15), separated into zonal wavenumbers $m = 0 - M$, to calculate the various terms shown in Fig. 8. The quasi-linear extraction or wave-mean interaction term is equal to $\overline{\overline{G}}$, the nonlinear transfer or wave-wave interaction term is $\overline{\overline{N}}_w$, and the linear (dissipation plus forcing) term is $-\overline{\overline{\Psi_j' \overline{D}_j}} - \overline{\overline{\Psi_j' D_j'}}$. We have performed the calculation at equilibrium, time averaged in our usual way, and thus the derivative term with respect to time should be roughly zero. (This was shown to be true by the thin solid line in Fig. 8.)

The figures involving the energy in each mode, i.e. Figs. 3 and 1 (solid line), have been plotted using the definition (A11), after horizontally averaging and separating by zonal mode, and performing a time average at equilibrium.

[10] Note that the single bar is a zonal *and time* mean, while the double bar is a zonal mean plus a meridional mean. Thus a double bar over a zonal and time mean term does not remove the single bar representing the latter.

[11] Note that there is a minus sign missing in front of this term in equations (2.3), (2.5) and (A.6) in Whitaker and Barcilon (1995). Defined without the minus sign of (A16), this term would represent the amount of energy flowing *to* the mean flow *from* the eddies.

Note: to avoid lengthy calculations in our model, we did not determine the nonlinear wave-wave term $\overline{\overline{N}}_w$ directly. Instead, we calculated every other term in (A15), including the time rate of change (which is no more expensive than the others), and then solved for $\overline{\overline{N}}_w$ by subtraction. In addition, we used Fast Fourier Transforms to calculate the other $O({\Psi'}^2)$ terms.

References

Cehelsky, P. and K. Tung, 1991: Nonlinear baroclinic adjustment. *J. Atmos. Sci.,* **48**, 1930 – 1947.

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

Charney, J. G., 1971: Geostrophic turbulence. *J. Atmos. Sci.,* **28**, 1087 – 1095.

DelSole, T. and B. Farrell, 1995: A stochastically excited linear system as a model for quasi-geostrophic turbulence: analytical results for one- and two-layer fluids. *J. Atmos. Sci.,* **52**, 2531 – 2547.

DelSole, T. and B. Farrell, 1996: The quasi-linear equilibration of a thermally maintained, stochastically excited jet in a quasigeostrophic model. *J. Atmos. Sci.,* **53**, 1781 – 1797.

Farrell, B. and P. Ioannou, 1994: A theory for the statistical equilibrium energy spectrum and heat flux produced by transient baroclinic waves. *J. Atmos. Sci.,* **51**, 2685 – 2698.

Farrell, B. and P. Ioannou, 1995: Stochastic dynamics of the midlatitude atmospheric jet. *J. Atmos. Sci.,* **52**, 1642 – 1656.

Farrell, B. F., 1982: The initial growth of disturbances in a baroclinic flow. *J. Atmos. Sci.,* **39**, 1663 – 1686.

Farrell, B. F., 1989: Optimal excitation of baroclinic waves. *J. Atmos. Sci.,* **46**, 1193 – 1206.

Fraedrich, K. and H. Böttger, 1978: A wavenumber-frequency analysis of the 500 mb geopotential at 50°N. *J. Atmos. Sci.,* **35**, 745 – 750.

Gall, R., 1976: A comparison of linear baroclinic instability theory with the eddy statistics of a general circulation model. *J. Atmos. Sci.,* **33**, 349 – 373.

Green, J., 1960: A problem in baroclinic stability. *Quart. J. Roy. Meteor. Soc.,* **86**, 237 – 251.

Gutowski, W., L. E. Branscome, and D. Stewart, 1989: Mean flow adjustment during life cycles of baroclinic waves. *J. Atmos. Sci.,* **46**, 1724 – 1737.

Haidvogel, D. and I. Held, 1980: Homogeneous quasi-geostrophic turbulence driven by a uniform temperature gradient. *J. Atmos. Sci.,* **37**, 2644 – 2660.

Held, I., 1978: The vertical scale of an unstable wave and its importance for eddy heat flux parameterizations. *J. Atmos. Sci.,* **35**, 572 – 576.

Holloway, G., 1986: Eddies, waves, circulation, and mixing: statistical geofluid mechanics. *Ann. Rev. Fluid Mech.,* **18**, 91 – 147.

Hoskins, B. J. and R. P. Pearce, 1983: *Large-scale dynamical processes in the atmoshphere.* New York: Academic Press.

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

Kao, S.-K. and L. L. Wendell, 1970: The kinetic energy of the large-scale atmospheric motion in wavenumber-frequency space: I. Northern hemisphere. *J. Atmos. Sci.,* **27**, 359 – 375.

Kuo, H., 1973: Dynamics of quasi-geostrophic flows and instability theory. *Adv. Appl. Mech.,* **13**, 247 – 330.

Leith, C., 1972: Atmospheric predictability and two-dimensional turbulence. *J. Atmos. Sci.,* **28**, 145 – 161.

Lindzen, R., 1993: Baroclinic neutrality and the tropopause. *J. Atmos. Sci.,* **50**, 1148 – 1151.

Lindzen, R., 1994: The Eady problem for a basic state with zero pv gradients but $\beta \neq 0$. *J. Atmos. Sci.,* **51**, 3221 – 3226.

Lorenz, E., 1960: Energy and numerical weather prediction. *Tellus,* **XII**, 364 – 373.

Pavan, V., 1996: Sensitivity of a multi-layer quasi-geostrophic β-channel model to the vertical structure of the equilibrium meridional temperature gradient. *Quart. J. Roy. Meteor. Soc.,* **122**, 55 – 72.

Pedlosky, J., 1987: *Geophysical Fluid Dynamics* (Second ed.). New York: Springer-Verlag, 710 pp.

Peixoto, J. and A. Oort, 1992: *Physics of climate.* New York: American Institute of Physics, 520 pp.

Randel, W. and I. Held, 1991: Phase speed spectra of transient eddy fluxes and critical layer absorption. *J. Atmos. Sci.,* **48**, 688 – 697.

Rhines, P., 1975: Waves and turbulence on a beta-plane. *J. Fluid Mech.,* **69**, 417.

Salmon, R., 1980: Baroclinic instability and geostrophic turbulence. *Geo. Astro. Fluid Dyn.,* **15**, 167 – 211.

Schneider, E. K., 1981: On the amplitudes reached by baroclinically unstable disturbances. *J. Atmos. Sci.,* **38**, 2142 – 2149.

Stone, P., 1978: Baroclinic adjustment. *J. Atmos. Sci.,* **35**, 561 – 571.

Wallace, J. M. and P. Hobbs, 1977: *Atmospheric Science.* San Diego: Academic Press, 467 pp.

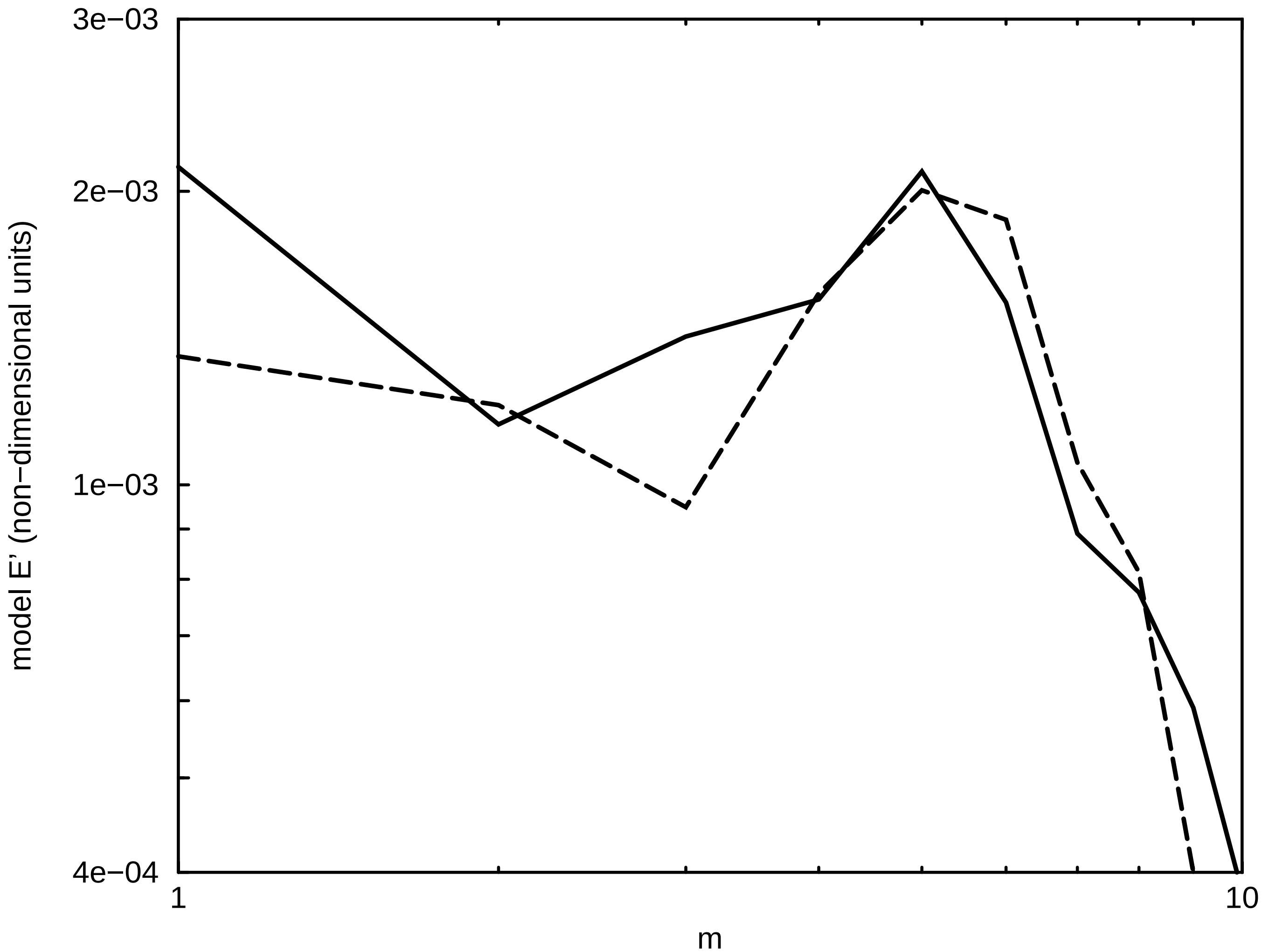
Welch, W. and K. Tung, 1997: Nonlinear baroclinic adjustment and wavenumber selection in a simple case. Submitted to *J. Atmos. Sci.*

Welch, W. T., 1996: *Nonlinear baroclinic adjustment and wavenumber selection as a mechanism for atmospheric heat transport.* Ph. D. thesis, University of Washington.

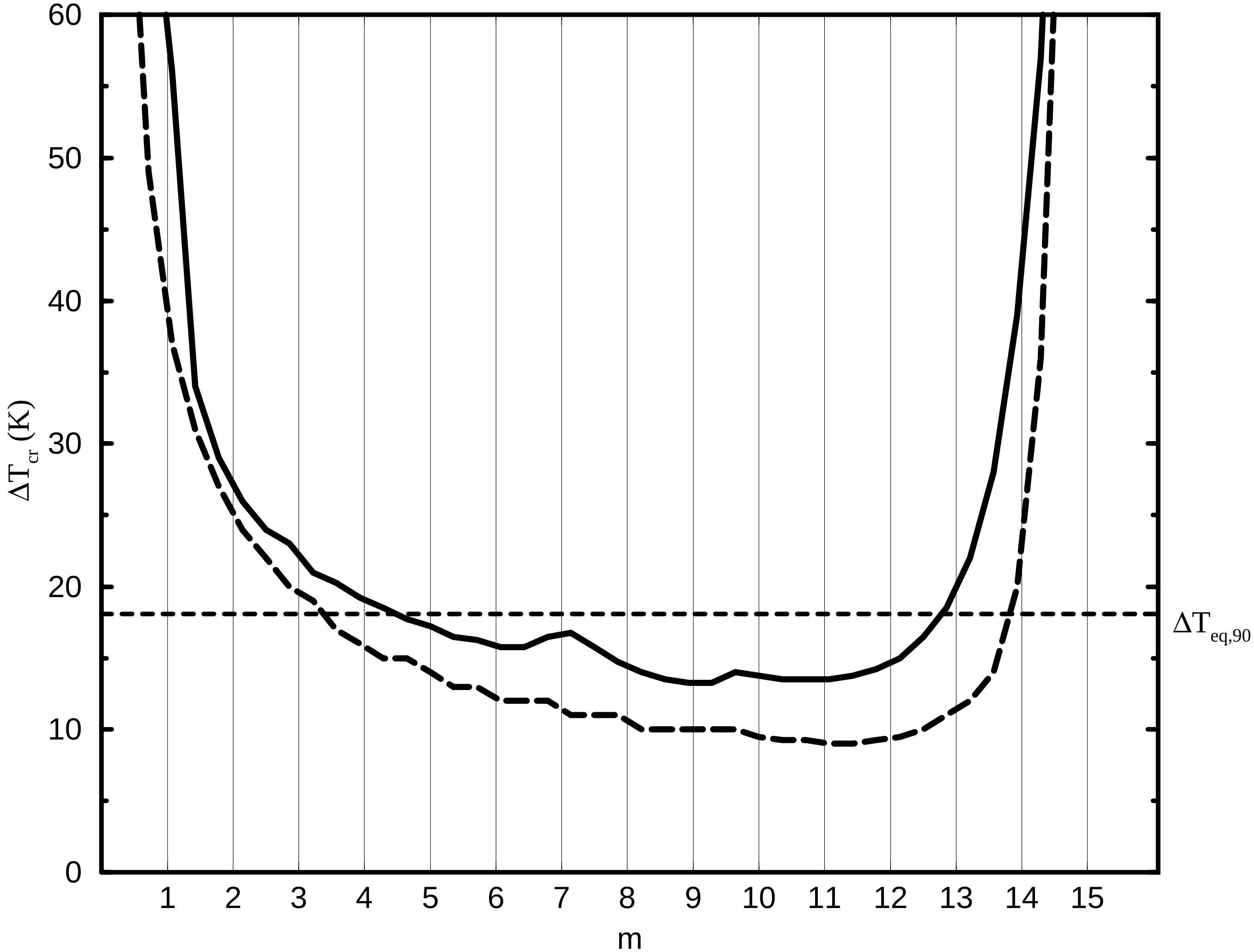
Whitaker, J. and A. Barcilon, 1995: Low frequency variability and wavenumber selection in models with zonally symmetric forcing. *J. Atmos. Sci.,* **52**, 491 – 503.

Zhou, S. and P. H. Stone, 1993: The role of large-scale eddies in the climate equilibrium. Part II: Variable static stability. *J. Climate,* **6**, 1871 – 1881.

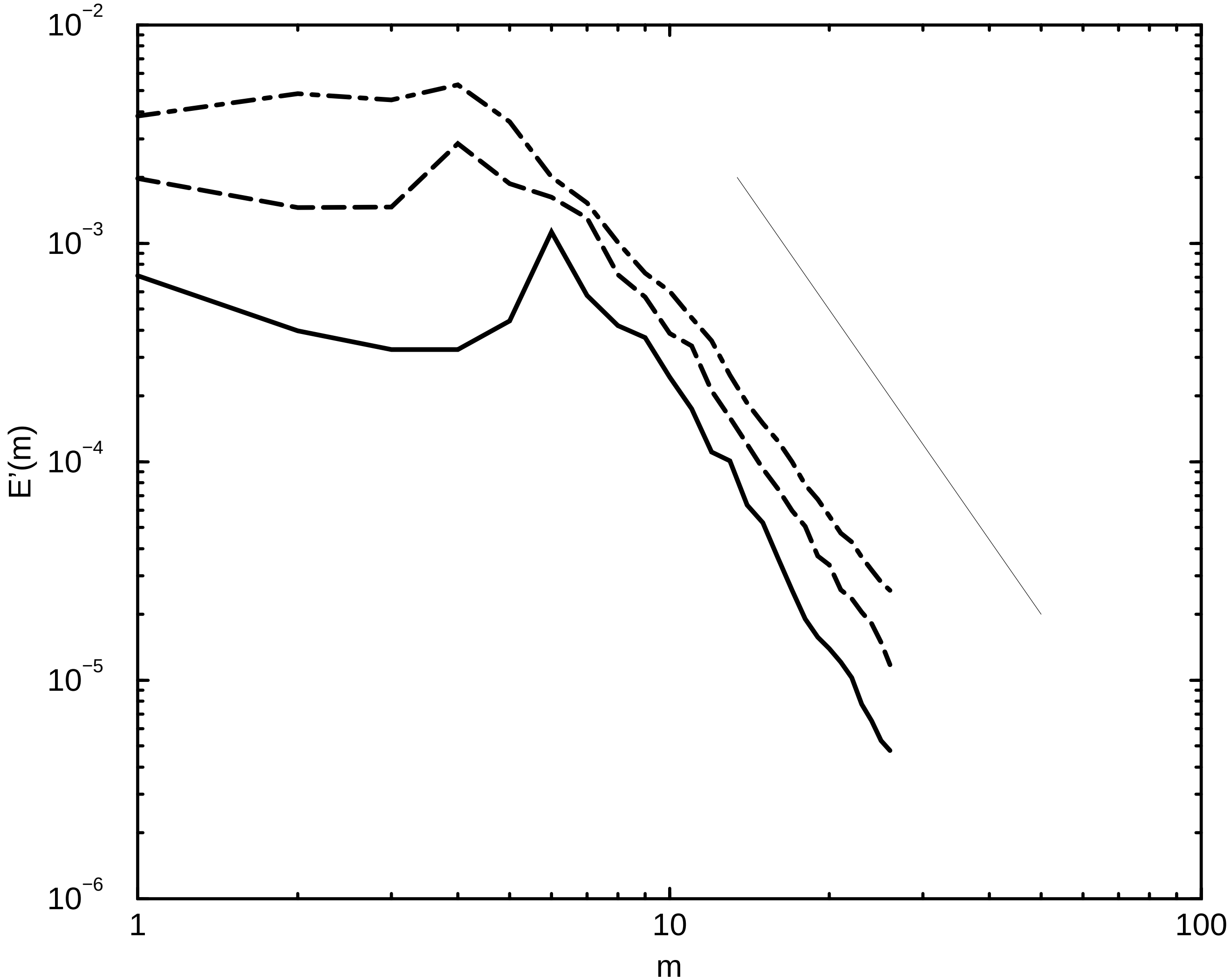
List of Figures



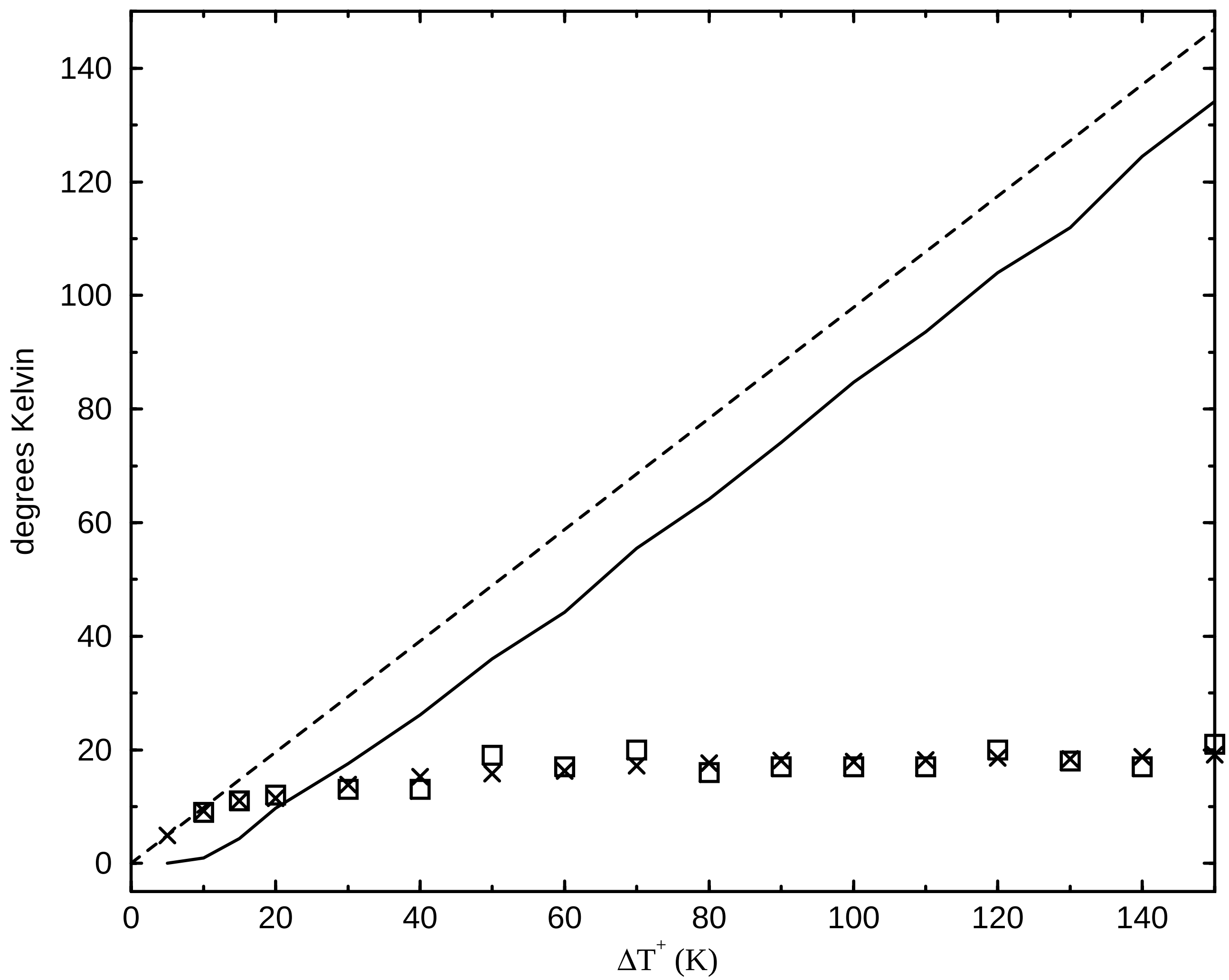
model E' (non−dimensional units)
3e−03
2e−03
1e−03
4e−04
1
10
m



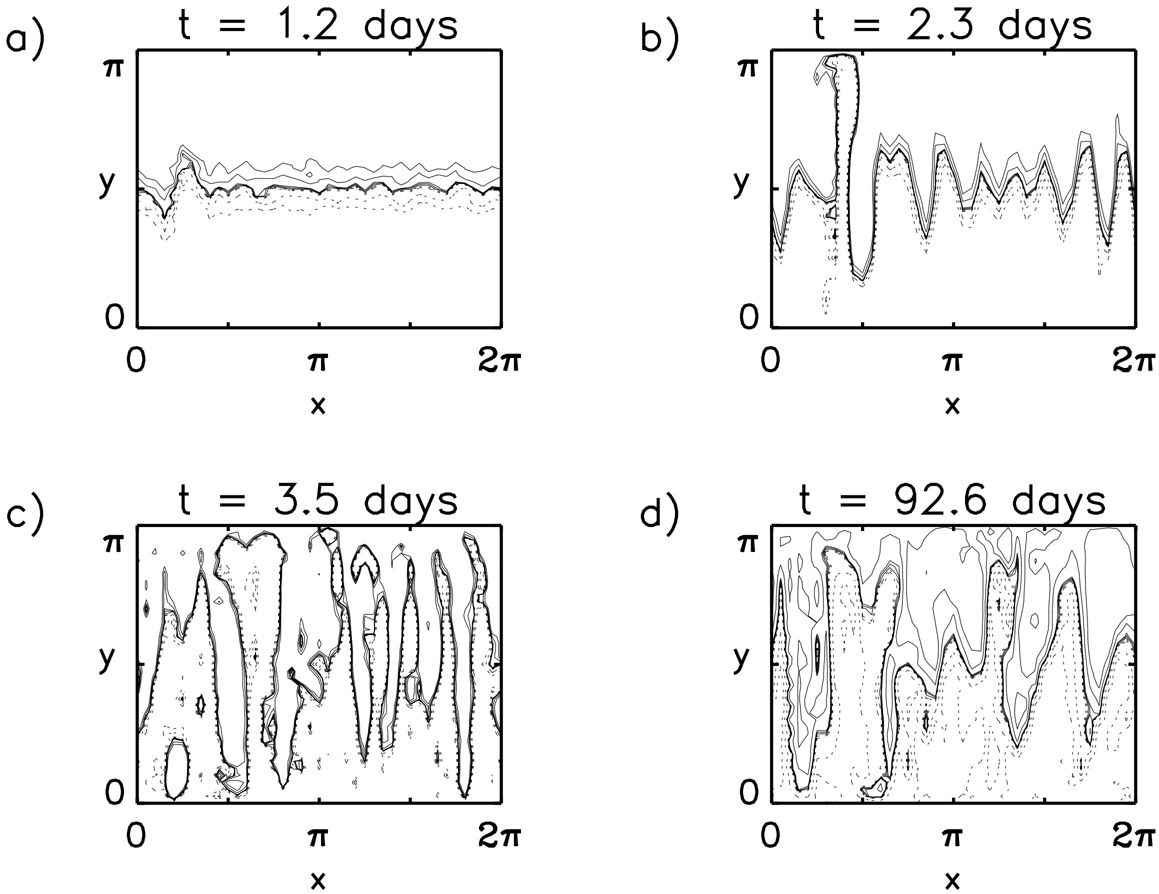
60
50
40
30
20
10
0
ΔT_cr (K)
1 2 3 4 5 6 7 8 9 10 11 12 13 14 15
m
ΔT_eq,90



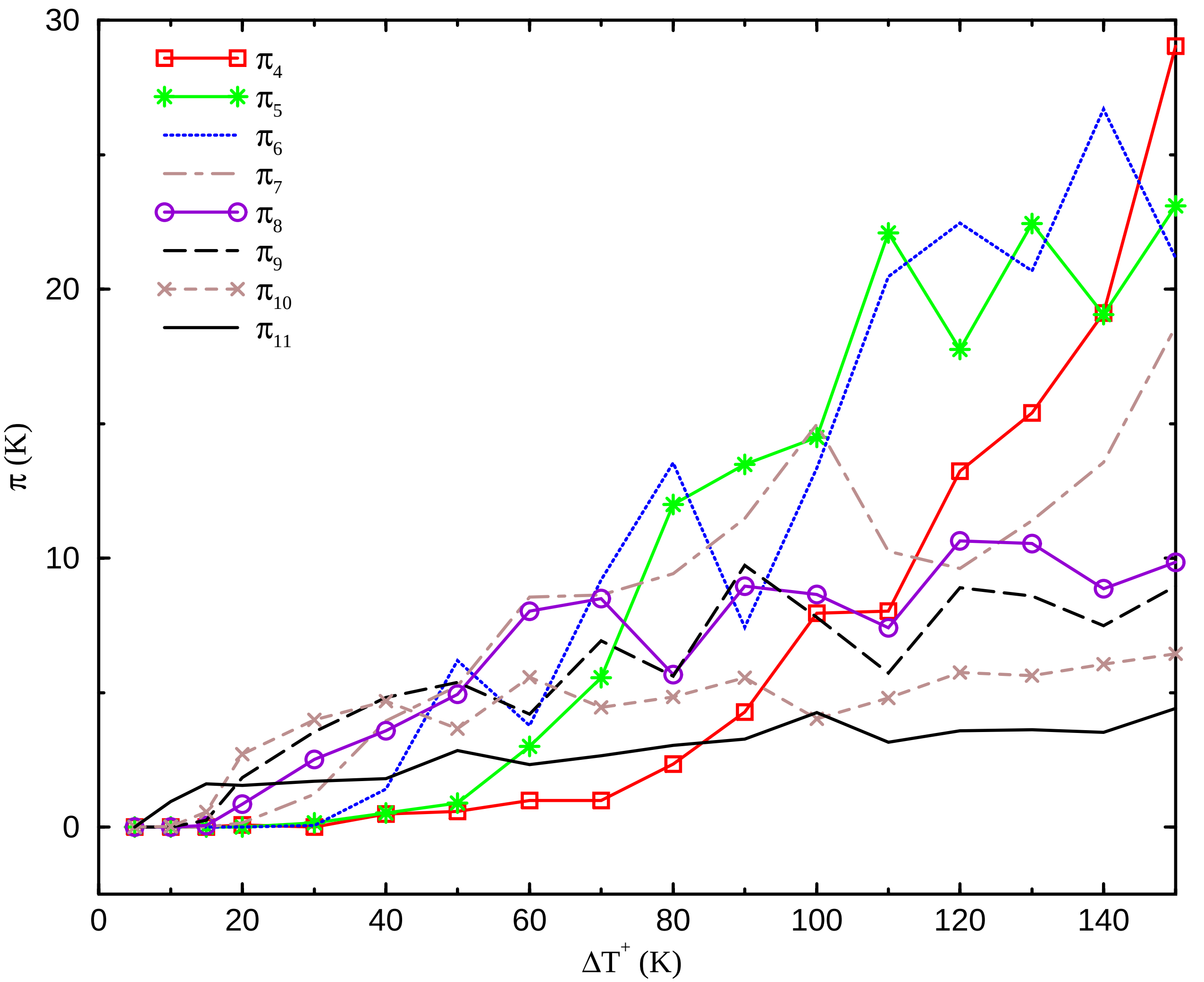
10^{-2}
10^{-3}
10^{-4}
10^{-5}
10^{-6}
E'(m)
1
10
100
m



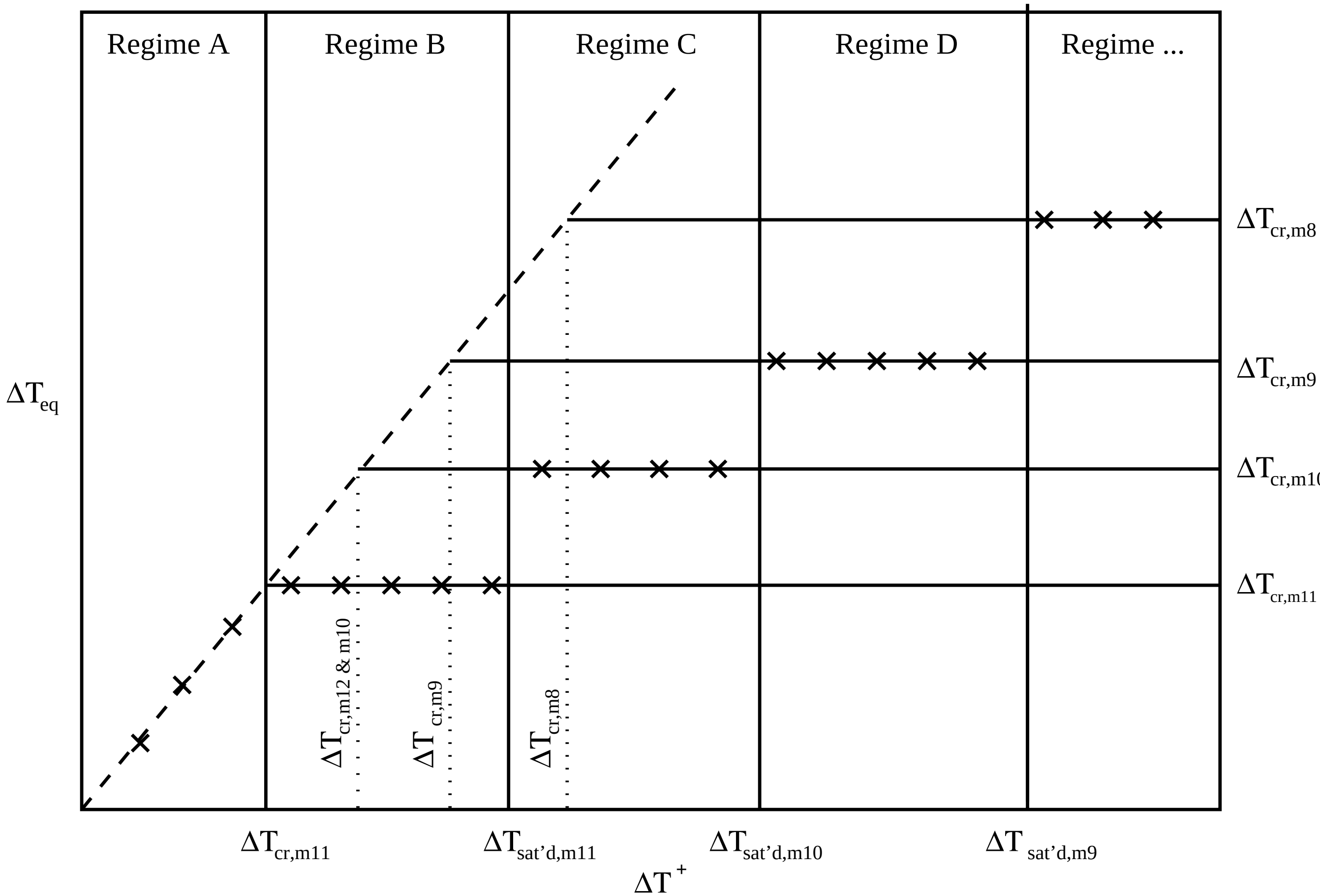
degrees Kelvin
0
20
40
60
80
100
120
140
ΔT⁺ (K)



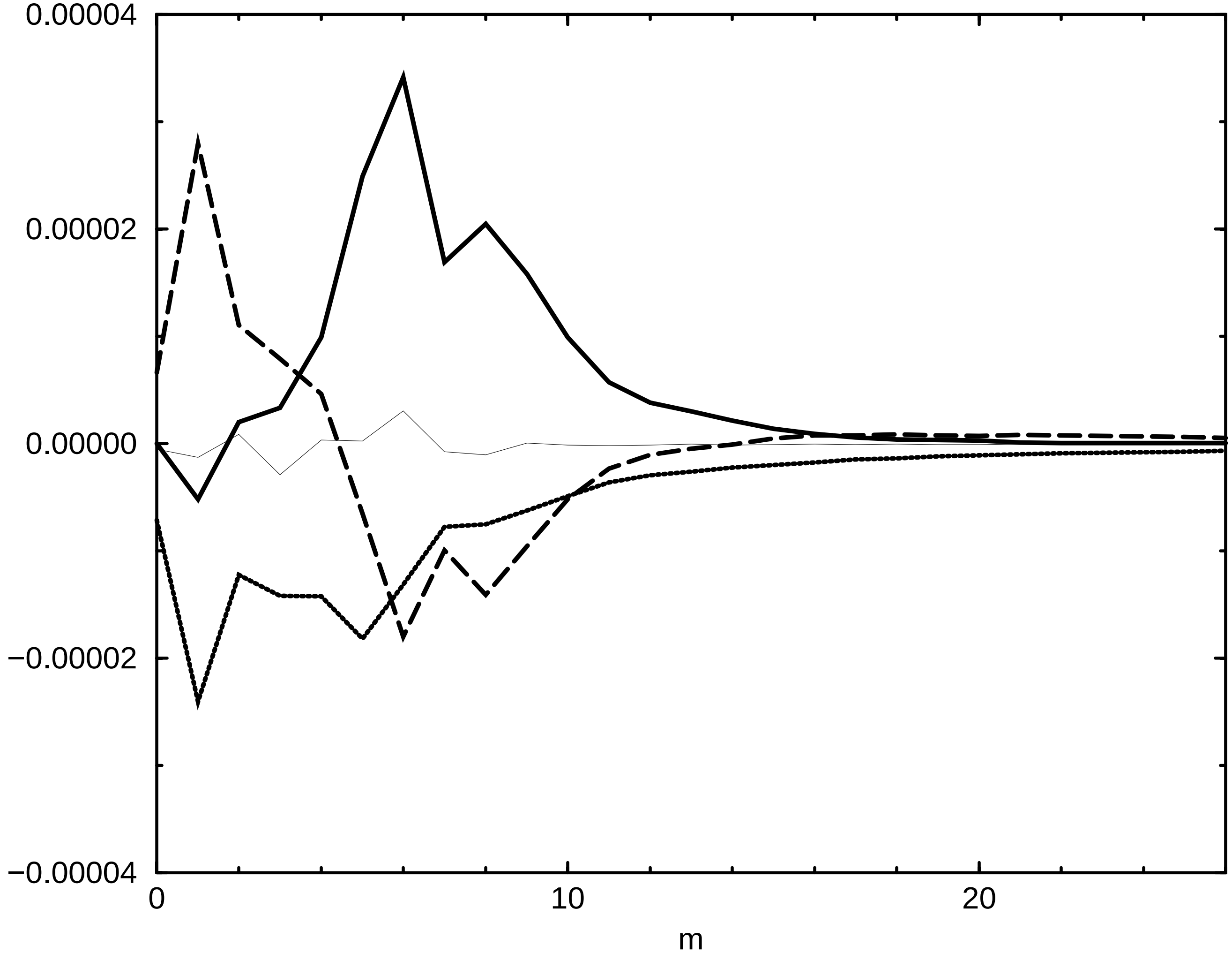
a)
t = 1.2 days
b)
t = 2.3 days
c)
t = 3.5 days
d)
t = 92.6 days
π
y
0
0
π
2π
x



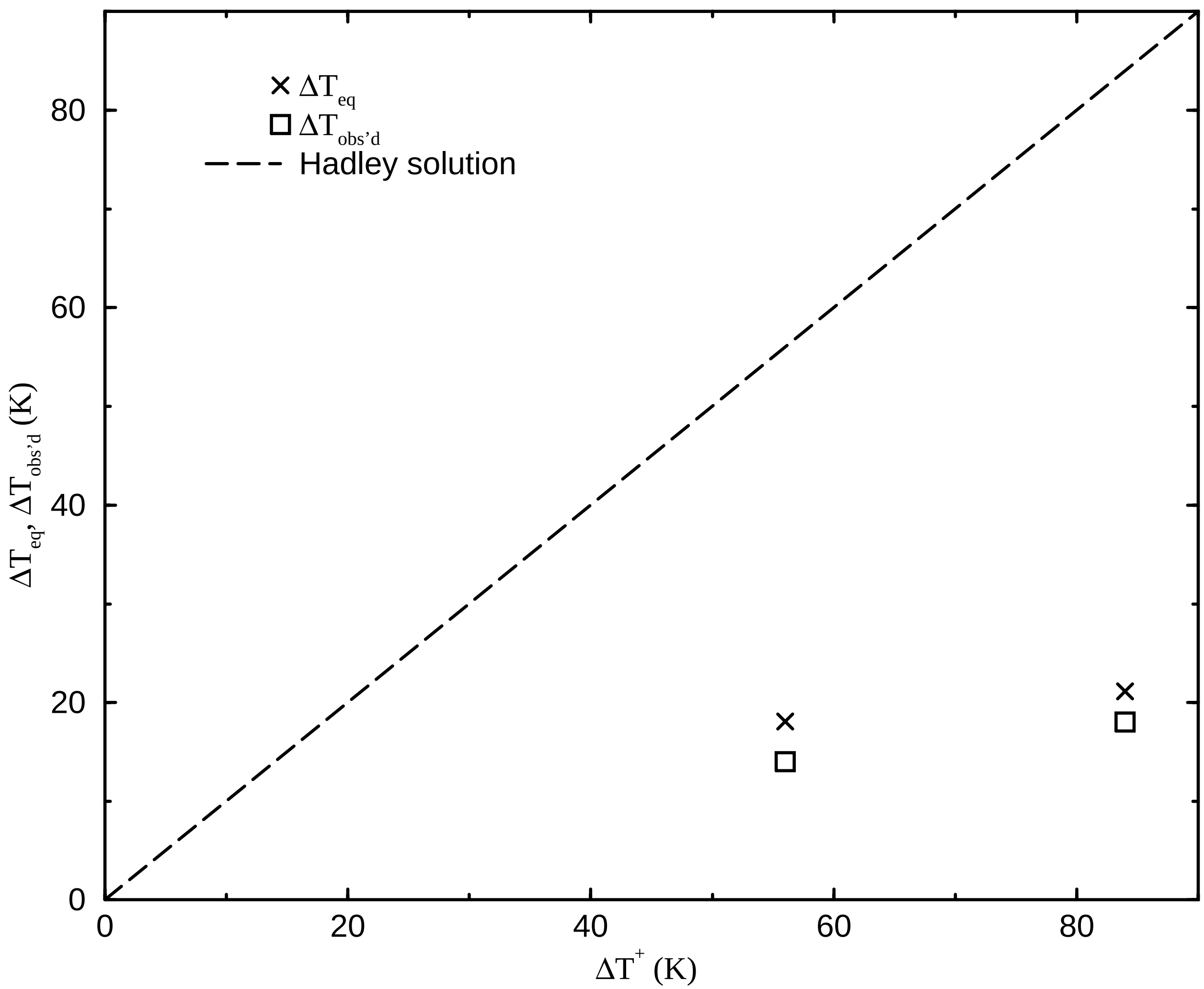
π4
π5
π6
π7
π8
π9
π10
π11
π (K)
ΔT+ (K)
0
10
20
30
0
20
40
60
80
100
120
140



Regime A
Regime B
Regime C
Regime D
Regime ...
ΔT_eq
ΔT_cr,m8
ΔT_cr,m9
ΔT_cr,m10
ΔT_cr,m11
ΔT_cr,m12 & m10
ΔT_cr,m9
ΔT_cr,m8
ΔT_cr,m11
ΔT_sat'd,m11
ΔT_sat'd,m10
ΔT_sat'd,m9
ΔT^+



0.00004
0.00002
0.00000
−0.00002
−0.00004
0
10
20
m



ΔT_eq
ΔT_obs'd
Hadley solution
0
20
40
60
80
ΔT^+ (K)
ΔT_eq, ΔT_obs'd (K)