Isentropic Formation of the Tropopause

MAARTEN AMBAUM

Royal Netherlands Meteorological Institute, De Bilt, the Netherlands

(Manuscript received 12 March 1996, in final form 7 August 1996)

ABSTRACT

The existence of the extratropical tropopause, which can be defined as the sharp transition between the high potential vorticity in the stratosphere and the low potential vorticity in the troposphere, is attributed to a dynamical equilibrium of diabatic effects and vortex stripping. The process is essentially isentropic and two-dimensional. This picture is examined and confirmed in a simple one-layer quasigeostrophic model of the atmosphere. It is possibly the simplest model context where the formation of a tropopause can be found.

1. Introduction

The World Meteorological Organization defines the tropopause as the boundary between the high-temperature lapse rates of the stratosphere and the low-temperature lapse rates of the troposphere.[1] The depth of this transition is relatively small, and this is the rationale for defining the tropopause as a seperate entity. In the Tropics the tropopause roughly coincides with an isentropic surface and in the extratropics with an isopotential vorticity surface. This is consistent with the fact that the tropopause has a quasi-material nature: for adiabatic motion, potential temperature as well as potential vorticity are materially conserved. In light of problems relating to stratosphere–troposphere exchange, it is in fact more convenient to define the extratropical tropopause as the 2-PVU isopotential vorticity surface[2] (Danielsen 1968; Holton et al. 1995). With this definition the exchange processes have a more clear-cut meaning: only diabatic and small-scale mixing effects are able to transport tracers across the tropopause. For adiabatic motions the tropopause thus becomes a material surface and is impermeable for passive tracers.

The practical equivalence of the two definitions of the tropopause follows from the definition of potential vorticity and from the invertibility principle (Hoskins et al. 1985). First of all, with the use of the hydrostatic approximation, potential vorticity is proportional to the potential temperature lapse rate with pressure. Therefore a jump in lapse rates should induce a jump in potential vorticity. At the same time, the invertibility principle states that under the balanced conditions of the extratropics a jump in potential vorticity must also induce a jump in lapse rates. In the rest of this article this practical equivalence of the two definitions of the extratropical tropopause will be assumed; the boundary between the high stratospheric potential vorticity values and the low tropospheric values will be referred to as the tropopause.

Isosurfaces of potential temperature and potential vorticity are generally not parallel. Therefore the isopotential vorticity surfaces, and more specifically the tropopause, will be cut by isentropic surfaces [see for example Thorncroft et al. (1993) for pictures of potential temperature on the 2-PVU surface.] The isentropes that intersect the tropopause [also called the 'Middleworld' (Hoskins 1991)] probe both the stratosphere and the troposphere. The potential vorticity on such an isentrope has high values around the poles and low values elsewhere. In Fig. 1 an example is plotted of a potential vorticity field on the 320-K Middleworld isentrope. The transition from troposphere to stratosphere is marked by a rather strong gradient in potential vorticity. This band of strong potential vorticity gradients will be referred to as the ''isentropic tropopause.'' In Fig. 2 the potential vorticity profile for this situation is plotted as a function of equivalent latitude (see figure caption for details). The enhanced gradients of potential vorticity at the tropopause can be clearly recognized in this profile. The stratosphere, and especially the troposphere, can be recognized as regions of diminished potential vorticity gradients. The picture is generic in the sense that the approximately piecewise uniform nature of the potential vorticity field is structurally stable: the tropopause can

[1] Strictly speaking, the tropopause is defined as the lowest level at which the lapse rate decreases to 2°C km^{-1} or less, provided that the average lapse rate between this level and all higher levels within 2 km does not exceed 2°C km^{-1}.

[2] A potential vorticity unit (PVU) is defined as 10^{-6} m^2 s^{-1} Kkg^{-1}.

Corresponding author address: Dr. Maarten Ambaum, Royal Netherlands Meteorological Institute, P.O. Box 201, 3730 AE De Bilt, the Netherlands.
E-mail: ambaum@knmi.nl

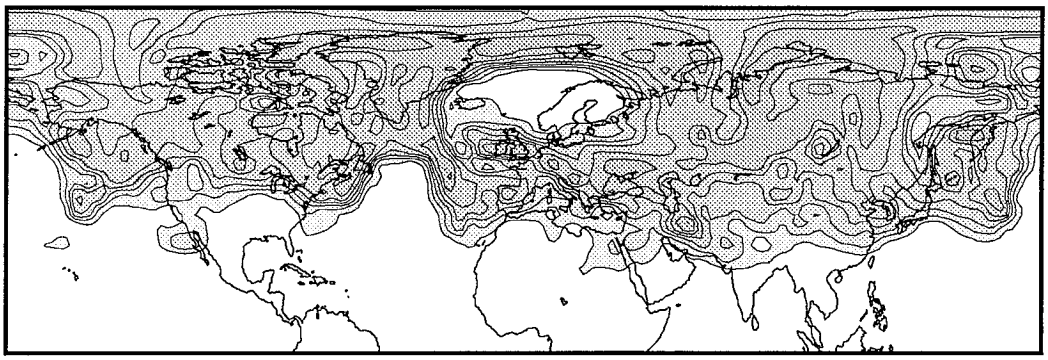

FIG. 1. Potential vorticity on the 320-K isentropic surface at 0000 GMT 15 February 1994, in lat–long projection for the Northern Hemisphere. Isolines are plotted every 1 PVU. Field produced by Dr. P. van Velthoven, using analyzed data from the European Centre for Medium-Range Weather Forecasts.

obtain a rather convoluted structure without affecting the piecewise uniform structure.

The problem of the origin of the tropopause can now be rephrased as a problem of the origin of the almost piecewise uniform nature of the potential vorticity field on certain isentropic surfaces that are therefore identified with the Middleworld. Observational studies of the stratospheric polar vortex (McIntyre and Palmer 1983) suggest that the phenomenon of vortex erosion, or stripping, can lead to sharp edges in the potential vorticity structure of the polar vortex. Vortex stripping is the phenomenon of weaker vorticity gradients being stripped off the boundary of a vortex, thus leaving a vortex with a steeper vorticity gradient. This process has been observed in subsequent numerical simulations (Juckes and McIntyre 1987; Haynes 1990) of the stratosphere and also in idealized numerical experiments (Legras and Dritschel 1993). McIntyre and Palmer (1984) suggested that the mechanism of vortex stripping might also play a central role in the formation of the tropopause, so that the tropopause may be viewed as the eroded edge of the transition between stratosphere and troposphere on Middleworld isentropes. Other theories like those of Held (1982) or Lindzen (1993) emphasize the role of baroclinic eddies in determining the height of the tropopause. These theories use baroclinic neutrality to determine this height. It may be wondered whether this is a correct criterion, because the large meridional variability of the tropopause does not suggest a baroclinically neutral state. In the vortex stripping argument eddies also play a crucial role, but now as a source of vortex erosion at the tropopause.

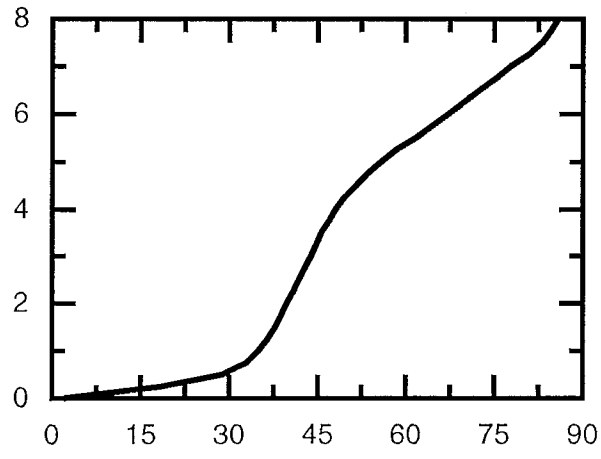

FIG. 2. Potential vorticity (in PVU) for the case of Fig. 1 plotted as a function of the area that is covered by potential vorticity values that are larger than the value at hand. Area is expressed as equivalent latitude, i.e., the latitude of a zonally symmetric contour enclosing the same area.

In this paper the argument of the tropopause as a vortex erosion surface is further examined in the context of a quasigeostrophic one-layer model of the atmosphere. In the next section some background will be given on the phenomena of vortex stripping and potential vorticity homogenization. It is argued that these processes are in fact the result of the same advective processes but that the advective processes have a systematic bias in the case of vortex stripping while they have no systematic bias in the case of homogenization. The advective processes involved are essentially two-dimensional and can therefore be observed in a one-layer model of the atmosphere. This model and its results are discussed in section 3. Section 4 is concerned with the relevance of the results for the three-dimensional atmosphere. It is argued how we may understand why the lowest isentrope to cut the tropopause generally grazes the earth's surface near the equator. A discussion and some concluding remarks are presented in section 5.

2. Stripping and homogenization

The nonlinear evolution and breaking of vorticity waves may be viewed as an isentropic process, as long as the spatial scales are large enough to prevent irreversible potential temperature mixing and the temporal scales are small enough to be essentially uninfluenced by diabatic processes. These conditions are largely satisfied in synoptic-scale wave-breaking events, associ-

ated with baroclinic or barotropic eddies. Under these circumstances the wave breaking events will advect potential vorticity on isentropic surfaces. Due to the complicated behavior of the eddies, this advection will have a rather erratic character. This type of advection is called chaotic advection.

In general, chaotic advection of a tracer like potential vorticity leads to a fast homogenization of the tracer. This is called chaotic mixing. Homogenization of tracers was already studied theoretically by Batchelor (1955) and Rhines and Young (1982). The latter argue that the potential vorticity field within closed isolines of potential vorticity gets homogenized by the action of small-scale inertial processes. Essential to their argument is that the diabatic processes occur on a longer timescale than that of the inertial processes. In their paper the homogenization is not very explicitly attributed to chaotic mixing. They only assume that the turbulent motion inside and outside the vortex expands an ensemble of tracers, rather than contracts it. Pierrehumbert and Yang (1993) used isentropic velocity fields from operational analyses to advect a passive tracer. They found that larger-scale tracer anomalies get homogenized over the stratospheric part of the isentrope in about 10–15 days. Smaller-scale anomalies get homogenized on shorter timescales. The authors attribute the process to a combination of shear dispersion (mainly in the longitudinal direction) and turbulent diffusion (mainly in the meridional direction).

But isentropic advective processes do not always homogenize a tracer. In fact, there are situations where the same processes lead to strong intensification of tracer gradients, which is quite the opposite effect. This is the case in vortex stripping processes (Melander et al. 1987; Legras and Dritschel 1993). Here, wave-breaking events at the boundary of a vortex *systematically* transport area between two potential vorticity contours away from this boundary, largely without affecting the form of the boundary. Because the area between the two isolines of the vortex boundary is transported elsewhere, the isolines come closer to each other. This is equivalent to an increase of the potential vorticity gradient at the vortex boundary. The essential difference between homogenization and stripping is that in the former case the area transport has a more or less random direction, while in the latter case the transport has a definite direction. But why is advection at a boundary of a vortex so different in character than advection elsewhere in the fluid?

Suppose that there is a vortex, a coherent structure, in the fluid. The vortex may be the result of self-organization in the fluid (McWilliams 1984) or the result of diabatic effects. At the boundary of the vortex there is a region where the gradient of potential vorticity is largest. This region is also associated with a local maximum in velocity. At the inner and outer boundaries of the velocity maximum, bands of maximum radial velocity shear are located. Any wave breaking event at the position of the velocity maximum will, on average, lengthen potential vorticity isolines. The mechanisms that lead to these wave-breaking events are immaterial. Because the area between two isolines is conserved, the lateral distance between the two isolines must decrease, which is equivalent to an increase of the lateral potential vorticity gradient. The broken waves that transport interisoline area away from the maximum velocity zone could return again to this zone and there replenish the area again. This would be the case if there is no directional preference for potential vorticity flux. But after the breaking event, the waves arrive in the maximum shear region at the outskirts of the maximum velocity zone and are there torn into filamentary debris. In this manner, they will eventually reach the smallest scales and will be dissipated. The lack of area replenishment will lead to a net flux of interisoline area away from the maximum velocity zone. At the same time, these wave breaking events do not destroy the main structure of the vortex itself. This is because the broken waves, by their smallness, do not significantly contribute to the local velocity field of the large-scale vortex.

We can now imagine a situation where both homogenization and stripping occur. Suppose the potential vorticity front that is formed by the stripping processes exhibits large-scale aperiodic meridional motion. The source of the aperiodic motion might be orography, inducing a meridional background field, but again the source is immaterial. Any potential vorticity structure outside the potential vorticity front will get deformed and lengthened by the shear dispersion induced by the velocity maximum associated with the potential vorticity front. In this process it becomes effectively a passive tracer. Now the large-scale aperiodic velocity field will lead to chaotic mixing of the small-scale potential vorticity structures. This will eventually lead to a homogenized field, at least in a coarse-grained sense, on the timescale of the motion of the potential vorticity front. The global picture now is that of a more or less piecewise uniform potential vorticity field.

Summarizing, one can say that the gradient intensification as well as the homogenization of potential vorticity are the results of advective processes, which exhibit a systematic bias of area flux at the position of a vortex boundary but no systematic bias elsewhere. Both processes may occur at the same time in one vortex. In a way, the processes of vortex stripping and homogenization may be seen as complementary: the homogenization is associated with the expulsion of potential vorticity isolines from the interior of the vortex; these isolines get concentrated around the boundary as a result of vortex stripping.

These processes, which lead to a piecewise uniform potential vorticity field, have mostly been described in an adiabatic context. The influence of diabatic processes on stripping has not been studied very much. Mariotti et al. (1994) study the effect of viscosity and hyperviscosity on the structure of a vortex that decayed under the influence of an external strain field. They observe

that viscosity leads to the constant production of weak vorticity skirts at the outer side of the vortex. These skirts get stripped away, leaving a smaller bare vortex. Eventually this process destroys the vortex. Furthermore, the introduction of viscosity limits the maximum vorticity gradient. The stripping processes cannot increase the vorticity gradients indefinitely, because diffusive processes, which become stronger with higher gradients, tend to smooth gradients. A simple argument that dissipative processes generally tend to decrease the potential vorticity gradient can be found in the appendix. So the average vorticity gradient at the boundary of the vortex is determined by a dynamical equilibrium between stripping and dissipation. Note that in general this dynamical equilibrium may not be as clear-cut as in the case studied by Mariotti et al. (1994), or by Haynes and Ward (1993). The latter observed how gradient enhancing processes, which they model as a simple strain, can come to an equilibrium with realistic dissipative processes. In the case of the polar vortex or the tropopause, the advective processes are generally chaotic, leading to a stripping that is intermittent, spatially as well as temporally.

In the atmosphere, both forcing and dissipation play an important part in the maintenance of the general circulation. These processes, along with the inertial processes of homogenization and stripping, play the essential role in the picture that we now may form of the origin of the tropopause. First, there is the forcing due to the differential heating of the earth. Through geostrophic adjustment, the overall effect of this differential heating will be a buildup of potential vorticity near the poles. Dissipative processes, like longwave radiation, will counterbalance this buildup. In the hypothetical case in which both forcing and dissipation are very strong—that is, when their timescales are very small—these will determine the atmospheric flow completely. The atmosphere is then essentially a slave to the diabatic processes. The spatial scales of the equilibrium flow are of the same order as the scales of the forcing. A vortex characterized by weak potential vorticity gradients results. If the diabatic processes are weaker, nonlinear effects become important: inertial motion of the atmosphere will lead to instability of the large-scale forced flow pattern. As explained before, this motion will generally lead to homogenization except at the boundary of the vortex, where it leads to stripping. This will lead to the enhancement of the potential vorticity gradients at the boundary of the vortex, where at the same time dissipative processes will again prevent this transition from becoming infinitely sharp.

In the next section this picture is studied, and confirmed, in a spinup experiment with a one-layer quasigeostrophic model of the atmosphere. We will see how the described processes take place to form a sharp transition, the tropopause, from high potential vorticity values around the North Pole to low potential vorticity values around the equator. This is possibly the simplest model context in which the formation of a tropopause can be observed.

3. The tropopause of a one-layer atmosphere

The simple model context, which will be used to study the formation of the tropopause, consists of a single isentropic surface. The vertical structure is fixed by assuming that the potential temperature does not change with height, so that the model effectively represents an isentropic layer. It can be shown that an atmospheric model consisting of a single isentropic layer in hydrostatic equilibrium is governed by the shallow-water equations (Verkley 1996, manuscript submitted to *J. Atmos. Sci.*). Because we will only be interested in extratropical phenomena, we can make the additional assumption of a small Rossby number. This leads to a further simplification of the model by using the quasigeostrophic approximation to the shallow-water equations, commonly referred to as the equivalent barotropic vorticity equation (Pedlosky 1979). This is the equation that will be used in the following subsection.

a. Model setup

The isentropic layer is assumed spherical with radius $a = 6.371 \times 10^6$ m rotating with angular velocity $\Omega = 7.292 \times 10^{-5}$ s^{-1}. The scale height H of the layer is taken to be 10 km. Lengths are expressed in units of a and time in units of Ω^{-1}. The equivalent barotropic vorticity equation reads

$$\frac{\partial q}{\partial t} + \mathbf{v}\cdot\boldsymbol{\nabla} q = \frac{S_0}{\tau_F} + \frac{F}{\tau_R}(\psi - \psi_0), q \tag{1a}$$

$$= f + \zeta - F\psi + f\frac{h}{H}. \tag{1b}$$

The second equation is the definition of the quasigeostrophic approximation to Ertel's potential vorticity. The first contribution to the potential vorticity is the planetary vorticity. In units of Ω it equals $f = 2 \sin \phi$, where ϕ is the latitude. The second contribution is the relative vorticity ζ, that is, the vertical component of the curl of the horizontal velocity field $\mathbf{v}$. The nondivergent velocity field $\mathbf{v}$ is given in terms of the streamfunction ψ by $\mathbf{v} = (-\partial\psi/\partial\phi, (1/\cos\phi)\,\partial\psi/\partial\lambda)$. From this it follows that the relative vorticity can be written as $\zeta = \nabla^2\psi$. The third contribution to the potential vorticity is the "stretching term," $-F\psi$; it is the quasigeostrophic representation of the vortex stretching effect due to the variations in the depth of the isentropic layer. The Froude number F in this term is defined as $F = L_R^{-2}$, where L_R is the Rossby radius in units of a. Vortex stretching by orography is represented by the last term in the expression for q. Here h is the height of the orography and it is scaled with the scale height H. Eq. (1a) states that the potential vorticity q, apart from being advected, changes due to a fixed source term, S_0/τ_F,

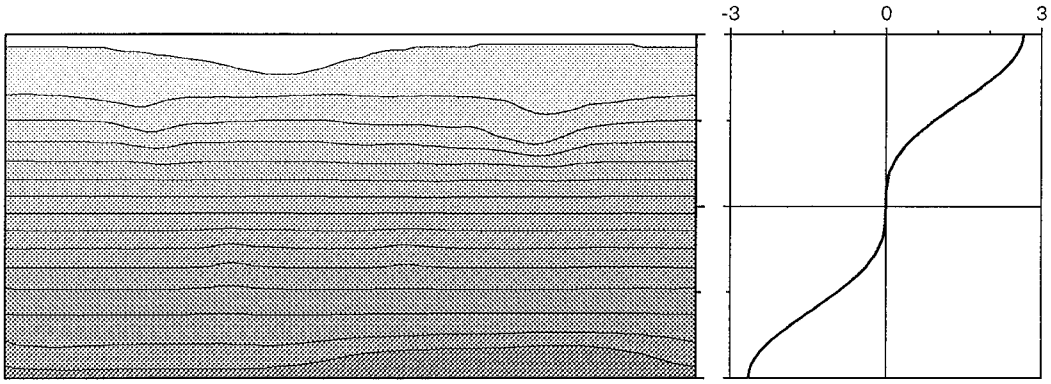

FIG. 3. The left panel shows the ψ_0 field to which the streamfunction thermally relaxes, in lat–long projection for the whole globe. This field corresponds to $q = 0$ everywhere. It is dominated by easterlies, due to the Coriolis effect. In the Northern Hemisphere, mountains show up as high streamfunction areas. The right panel shows the meridional profile of the source field S_0. The vertical axis represents the latitude from the South Pole to the North Pole. The source field S_0 is the combination of the Y_{01} and Y_{03} spherical harmonics that makes the meridional derivative vanish at the equator. It is normalized such that the global average of $S_0^2 = 1$.

acting on a timescale τ_F, and relaxes thermally to some prescribed streamfunction ψ_0, on a timescale τ_R.

We will take $F = 100$, which amounts to a Rossby radius of 637.1 km. This Rossby radius is smaller than what should be expected from the physical parameters of the system, but it is chosen to give realistic velocities in the case of a piecewise uniform potential vorticity field (see also Verkley 1994 and Ambaum and Verkley 1995). In the early practice of numerical weather prediction, this stretching term was referred to as the Cressman term. It was introduced to correct the excessive retrogression of the largest planetary waves in barotropic models. Cressman himself (1958) suggested an optimal value of $F = 60$. In a more recent survey (Rinne and Järvinnen 1993) of the Cressman term, it is suggested that any reasonable choice of the Froude number works well but that the most optimal choice is a Froude number that is not a constant but a function of position.

The forcing field is written as S_0/τ_F, where τ_F is a forcing timescale and S_0 is a fixed potential vorticity source field. In this experiment S_0 is a combination of the Y_{01} and Y_{03} spherical harmonics. The meridional derivative of S_0 vanishes at the equator. It is normalized such that the global average of S_0^2 equals 1. The forcing varies smoothly from positive values around the North Pole to negative values around the South Pole (see Fig. 3). There is no zonal structure in this source field, which is of course not very realistic. It is meant to mimic the differential heating of the atmosphere, leading to a buildup of potential vorticity around the poles. The most important feature that we want to emphasize is that it varies smoothly from equator to pole, so that a strong gradient in potential vorticity is not forced a priori. Apart from that our choice has been made as aspecific as possible.

For the thermal relaxation toward ψ_0, we have chosen the streamfunction that corresponds to $q = 0$ everywhere (see Fig. 3). The field ψ_0 can be determined by making the right-hand side of Eq. (1b) equal to zero. This streamfunction has no jetlike structure. It varies smoothly from positive values around the North Pole toward negative values around the South Pole. The corresponding velocity field is everywhere westward (easterly). The relaxation toward this ψ_0 corresponds to relaxation towards the most well-mixed situation, that is, the situation where q vanishes everywhere. Note from Eq. (1a) that a given relaxation field ψ_0 is equivalent to an additional forcing field $-F(\tau_F/\tau_R)\psi_0$.

This model was implemented numerically as a spectral model on the sphere. In the experiments shown, a $T85$ spectral truncation was used. The high truncation was chosen in order to resolve small-scale eddies explicitly. Time stepping was done with a leapfrog scheme, using a time step of 15 min. A small fourth-order hyperviscosity was included, in order to improve the inertial range of the energy spectrum. The hyperviscosity damps the wavenumber 85 spherical harmonics on a timescale of 40 h. Numerical stability for this term was obtained by implementing it semi-implicitly. Mariotti et al. (1994) warned that hyperviscosity could lead to spurious effects in the process of vortex stripping. This warning does not apply to our case; in the present model, the effects of hyperviscosity on the stripping processes are swamped by the effects of the much stronger thermal damping.

The parameters that we will consider in more detail are τ_F and τ_R, the forcing and dissipation timescales. If these two timescales are very small, that is, if the forcing and dissipation are very large, the atmosphere is essentially a slave to these processes. A time-independent atmospheric state will be established that, according to Eq. (1a), satisfies

$$\tau_R \mathbf{v}\cdot\boldsymbol{\nabla} q = \frac{\tau_R}{\tau_F} S_0 + F(\psi - \psi_0). \tag{2}$$

It can be solved iteratively for small τ_R and a fixed ratio

of τ_R and τ_F. The solution can be written as a series expansion in τ_R, which can be regarded as a Reynolds number. We will not consider the stability of these states nor the convergence of the series. In the limit for $\tau_R \to 0$ it is clear, though, that the series converges. In the absence of orography the resulting streamfunction is zonally symmetric and satisfies the Rayleigh condition for stability. In this case the advective term in Eq. (2) vanishes, and it can be easily shown that the potential vorticity field would relax to the field q_s defined as

$$q_s = -\frac{\tau_R}{\tau_F}\frac{(\nabla^2 - F)S_0}{F}. \tag{3}$$

Time integrations of the model with very small Reynolds numbers immediately lead to a time-independent, stable state.

The more interesting behavior occurs for larger Reynolds numbers τ_R. The aforementioned iterative procedure might fail or the time-independent states become unstable.[3] The main experiment we performed corresponds to $\tau_F = \tau_R = 31$ days. The initial condition, which is not important for the long-term behavior, corresponds to the zero-energy state. This is the state for which the streamfunction vanishes everywhere. In the following subsection, the main emphasis will be on the Northern Hemisphere. In section 5 the Southern Hemisphere will be discussed. With these parameters and this initial condition the model was set to work and integrated for 1000 days.

b. Model results

The spinup phase can be defined as the phase in which the energy relaxes toward its equilibrium value. This value has a meaning only in a time-averaged sense because the system relaxes to a time-dependent state. At the same time other quantities, for example the potential enstrophy, reach their equilibrium values. The spinup phase takes about 250 days (see Fig. 4). At the start of this phase the behavior of the atmosphere shows little inertial motion. It is characterized by strong direct orographic influences. This can be clearly seen in Fig. 5, where the potential vorticity distribution after 3 and 6 weeks is plotted. The orography induces relatively large potential vorticity gradients to the southeast of the main orographic features. This is the result of the familiar behavior of air parcels being advected over orography in the presence of the β effect. The conservation of potential vorticity implies that if an air parcel reaches a mountain, its absolute vorticity must decrease due to the vortex squeezing effect. This results in a southward deflection of the air parcel. This southward movement will only stop after the parcel has crossed the mountain.

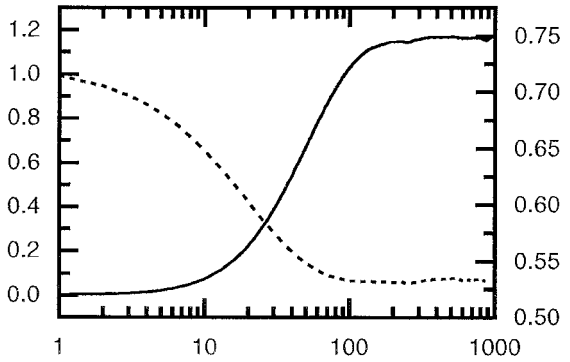

FIG. 4. Plot of the energy (solid line, left axis) and the potential enstrophy (dashed line, right axis) as a function of time (in days). Both quantities are in dimensionless units. The forcing and dissipation timescale is 31 days.

After this the parcel will be located south of the position where it first encountered the mountain. This means that due to the β effect its planetary vorticity is lower so that its relative vorticity must have become higher. The latter effect results in a northward deflection of the parcel. After this the air parcel will attain an oscillatory motion, periodically exchanging planetary and relative vorticity. This picture becomes more complicated if the mountain has a finite southward extent, as is the case for the orography of the earth. In this case there are also air parcels that move southward of the mountain. They will not experience the orographic vortex squeezing and therefore will not be deflected toward the south. The meeting of the latter air parcels and the deflected air parcels will result in a region of isoline convergence southeast of the mountain. In Fig. 6 a schematic drawing of the situation is depicted.

When the velocities become larger, the inertial motion of the atmosphere becomes more important. The regions of higher potential vorticity gradients, as induced by the orography, are advected eastward. The onset of this advection of gradient can be seen by comparing Figs. 5a and 5b. The general structure of the potential vorticity now becomes established: the forcing increases the values of the potential vorticity around the North Pole; the vorticity contrast between pole and equator increases. This buildup of vorticity is counteracted by the thermal damping, which becomes more important with increasing streamfunction values. A dynamical equilibrium sets in for the potential vorticity values around the North Pole. The wave breaking events, especially those induced by the orography, sharpen the transition between the high and low potential vorticity values, as explained in the previous section. In the appendix it is shown that the local rate of change in the potential vorticity gradient as a result of the radiative damping in our model can be estimated as

$$\frac{\partial|\nabla q|}{\partial t} = -\frac{F}{\tau_R}(v_{\|} - v_{0\|}). \tag{4}$$

[3] We have, in fact, found Reynolds numbers where there are time-independent stable states, while the series does not converge.

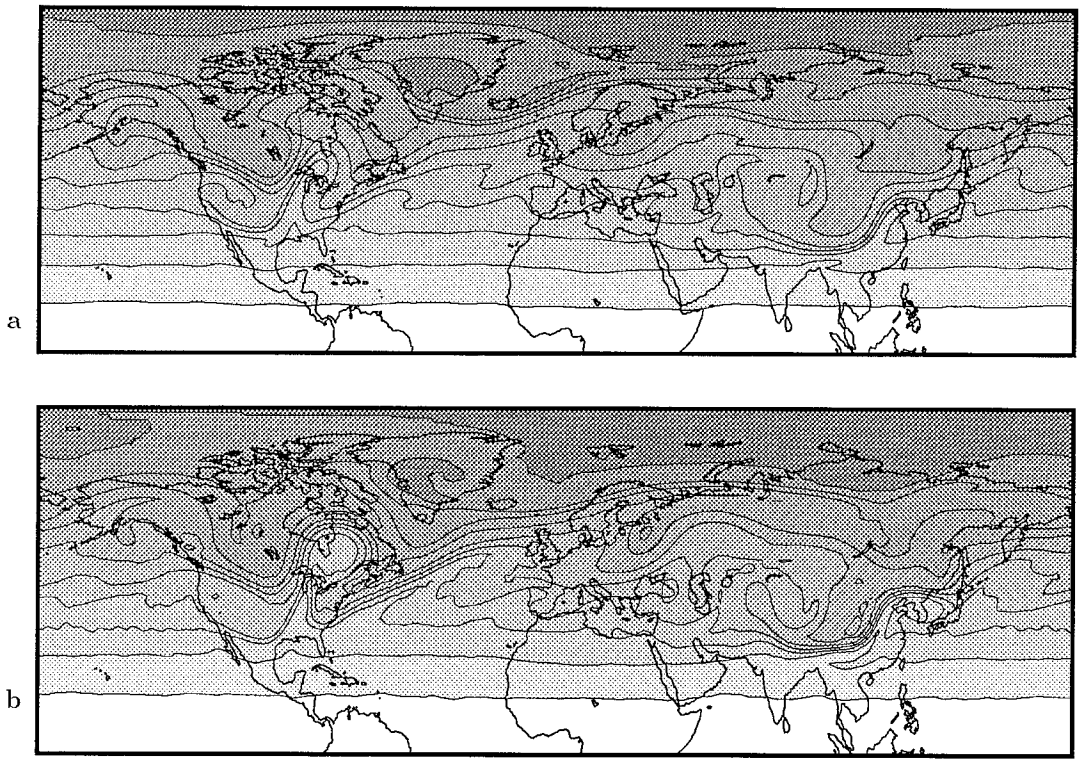

FIG. 5. The potential vorticity field after 3 weeks (a) and after 6 weeks (b). Note the strong convergences of the flow southeast of orographic features. Isolines are plotted every 0.3 dimensionless potential vorticity unit. The most southern isoline corresponds to $q = 0.3$.

Here $\mathbf{v}_{\|}$ and $\mathbf{v}_{0\|}$ represent the velocity perpendicular to the main gradient (and so parallel to the main jet) as a result of ψ and ψ_0, respectively. As the right-hand side of Eq. (4) is generally less than zero, radiative damping counteracts the gradient enhancing effect of stripping. A dynamical equilibrium is established between the stripping and damping processes, and this determines the width of the transition zone between high and low potential vorticity values. The model layer behaves as if it were part of the Middleworld: the polar regions of large potential vorticity values are associated with the stratosphere. The equatorial regions of low potential vorticity values are associated with the troposphere. The abrupt transition between the two is the tropopause.

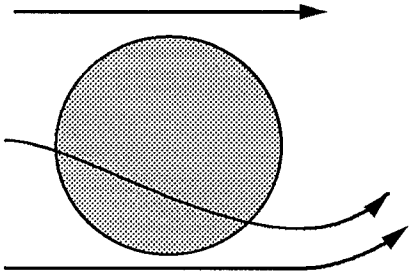

FIG. 6. A schematic drawing of the physical origin of the convergences of the flow pattern southeast of a mountain range. The shaded area represents a mountain. Particles flowing over the mountain will be deflected southward, due to the vortex squeezing effect of the mountain. At the southeast of the mountain these air parcels will meet the air parcels that went south of the mountain.

A single stripping event is shown in Fig. 7. It can be seen how a breaking wave transports potential vorticity away from the boundary of the vortex. The area associated with this broken wave is subtracted from the area that was originally between two isolines of the vortex boundary. Locally, this leads to the enhancement of the potential vorticity gradient. This enhanced gradient is advected eastward with the flow. The evolution of the vorticity gradient can be followed in the small graphs on the right in the figure. This development very much resembles the nonlinear decaying phase of unstable baroclinic waves, as for example, studied by Simmons and Hoskins (1978). They also observe that this phase, which evolves quasi-barotropically, leads to a strengthening of the jet stream that supported the wave.

The general structure of the potential vorticity is now dominated by a strong gradient along some isolines that encircle the earth. In Fig. 8 a nice example is shown of such a situation. Here the tropopause position is well defined everywhere. This clear-cut structure of the tropopause is not always evident. Due to large-scale breaking events and strain, the potential vorticity gradient can become diluted again, which is consistent with observations.

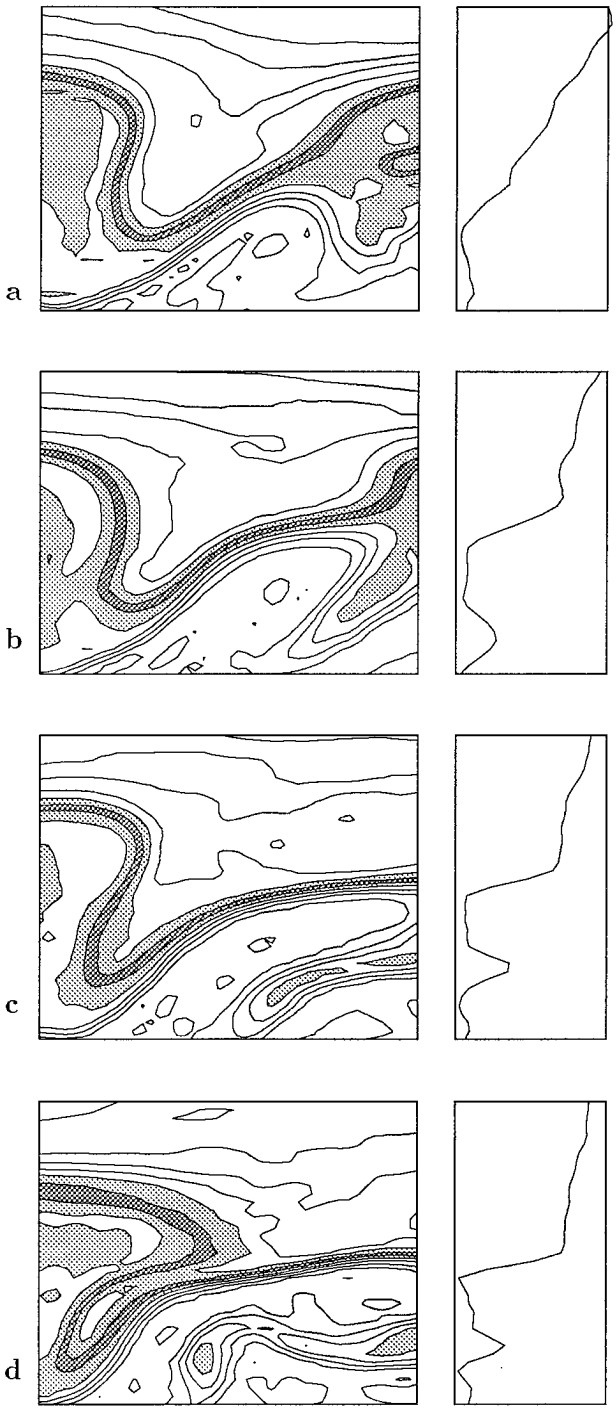

FIG. 7. An example of a single stripping event. The panels are taken with a 2-day interval between them. The fields are on a lat–long grid, ranging from 100° longitude to 160° longitude and from 20° latitude to 65° latitude. Isolines are plotted every 0.25 dimensionless potential vorticity unit. Shaded are the potential vorticity values that can be associated with the tropopause (1.00–1.75). Note the breaking waves at the lower part of the trough in the potential vorticity field. The evolution of the corresponding meridional potential vorticity profile (sampled at the right-hand side of the picture) is plotted on the right. During the wave-breaking event, the gradient gets sharpened.

As mentioned in the previous section, closed potential vorticity isolines can give rise to homogenization inside the isolines, as long as the timescales of forcing and dissipation are larger than the inertial timescale. In these experiments we have chosen 31 days for the former. Indeed, the inertial timescales are smaller than this. The dimensionless shear can be estimated as of order 1, which gives a shearing timescale of order 1 day. The latter is the timescale on which weak vorticity structures get deformed, mainly in the longitudinal direction. The timescales of forcing and dissipation are also larger than the timescale that Pierrehumbert and Yang (1993) found for total homogenization in the real atmosphere (10–15 days). As a result of this, the homogenization is rather effective. This can be seen from the meridional profile of the potential vorticity field in Fig. 9.

The profile in Fig. 9 shows a strong agreement with the profiles found by Polvani et al. (1995), who carried out shallow-water simulations of the stratospheric polar vortex. This suggests that the quasigeostrophic model is as capable of producing a realistic vortex as the shallow-water model. The model profile shows two regions of diminished potential vorticity gradients: one at the stratospheric side and one at the tropospheric side of the tropopause. In the context of simulations of the stratospheric polar vortex, the latter region is usually identified with a so-called ''surf zone'' (McIntyre and Palmer 1984). In the profile of the observed field, the tropospheric region of diminished potential vorticity gradients extends all the way down to the equator. Furthermore, it can be seen that the degree of homogenization on the stratospheric side of the tropopause is lower in the observation than in the model. Probably this is due to a larger influence of diabatic effects in the real stratosphere, which makes the homogenization less effective.

The occurence of both vortex stripping and homogenization makes the resulting field more or less piecewise uniform. As such, it fits observations (Hoskins et al. 1985) and it is ideally suited to be implemented as a contour dynamics model, as presented in Verkley (1994) and Ambaum and Verkley (1995). These articles give a view of planetary waves as interfacial Rossby waves [for an early survey of this idea, see Queney (1952)] in an exactly piecewise uniform potential vorticity field. The discontinuous jump represents the tropopause. By potential vorticity conservation, this jump is advected by the local fluid velocity. Quasigeostrophic balance leads to a straightforward invertibility principle: the velocity field can be obtained by a contour integral over the discontinuity. Though the longtime behavior of the spectral model is strongly influenced by diabatic processes, it may be expected that the shorttime behavior is essentially adiabatic, and as such can be simulated with the contour dynamics approach. The implementation of diabatic processes in contour dynamics is nontrivial and some discussion on the subject is deferred to the last section. For now we just show how strongly

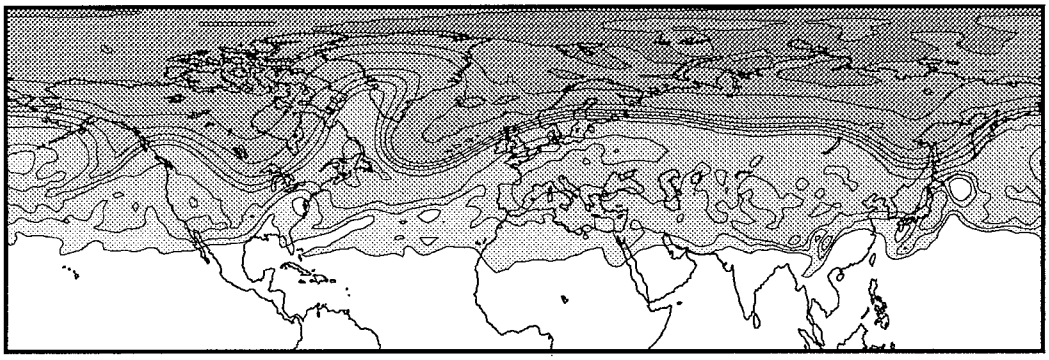

FIG. 8. An example of a situation (at 914 days in the integration) where the tropopause has a very distinct structure for all longitudes. Isolines as in Fig. 5.

this piecewise uniform structure of the potential vorticity is conserved. In Fig. 10 a picture of day 849 in the simulation is given. It shows an example of a flow situation that can be identified as an Atlantic dipole blocking. The piecewise uniform structure of the potential vorticity remains very clear (see Fig. 11), despite the rather convoluted form of the isentropic tropopause.

4. The three-dimensional structure of the tropopause

In the previous section we have seen how a one-layer isentropic model with a simple forcing and damping leads to the typical potential vorticity structure: high potential vorticity values around the North Pole and low potential vorticity values around the equator. The transition between the two is almost discontinuous and is identified as the tropopause. In the traditional picture we now have an isentropic surface of the Middleworld, cutting through the tropopause. The polar regions of the isentropic surface lie in the stratosphere and the equatorial regions lie in the troposphere. Our study, on the other hand, leads to the viewpoint that the existence of large isentropic potential vorticity gradients on Middleworld isentropes is not the *result* of the fact that these isentropes cut through the tropopause. Rather, a tropopause *is formed* on certain isentropes that we therefore associate with the Middleworld. In this respect McIntyre and Palmer (1984) illuminatingly state that ''for some purposes the tropopause may, as the isentropic viewpoint itself suggests, be more like a 'wall' than a 'ceiling.' ''

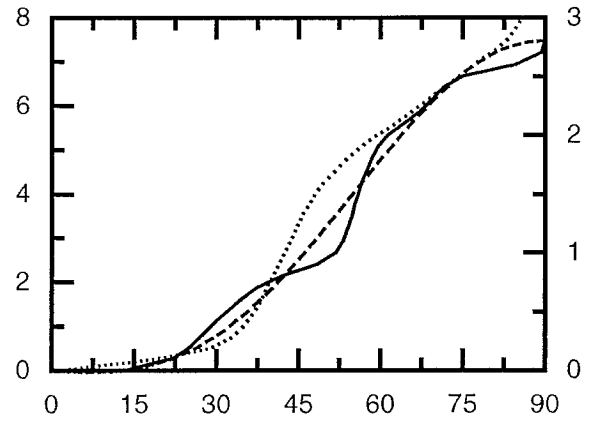

FIG. 9. Potential vorticity as a function of equivalent latitude (definitions as in Fig. 2) for the situation depicted in Fig. 8 (solid line, right axis), for the field q_s that corresponds to a zonally symmetric steady state [see Eq. (3) for the definition] (dashed, right axis) and for the observed field of Fig. 1 (dotted, left axis).

But the three-dimensional picture is more complicated, and some remarks on the applicability of our two-dimensional picture to the real atmosphere are now appropriate. It seems as if a tropopause will be formed on any isentrope as long as there is some gradient in the forcing and as long as there is enough wave activity. This would mean that every isentrope would have to cut through the tropopause. This is not the case in the real atmosphere: we know that there are ''Overworld'' isentropes that lie everywhere in the stratosphere and ''Underworld'' isentropes that lie everywhere in the troposphere. This contradiction is not so severe as it might seem at first sight.

As for the Overworld, we can say the following. The tropopause was defined throughout this paper as a given isosurface of the potential vorticity. In this definition of the tropopause there is no Overworld. The potential vorticity varies from positive values around the North Pole to negative values around the South Pole, so every isentrope that encompasses the globe must cut through the tropopause. On the higher isentropes, which are usually associated with the Overworld, the tropopause thus defined is very close to the equator. But this means that for these isentropes, due to the lack of geostrophic balance near the equator, the potential vorticity definition of the tropopause is not very useful. As mentioned in the introduction, the equatorial tropopause can generally be identified with an isentropic surface. Therefore the Overworld corresponds to the isentropes of higher potential temperature than that of the equatorial tropo-

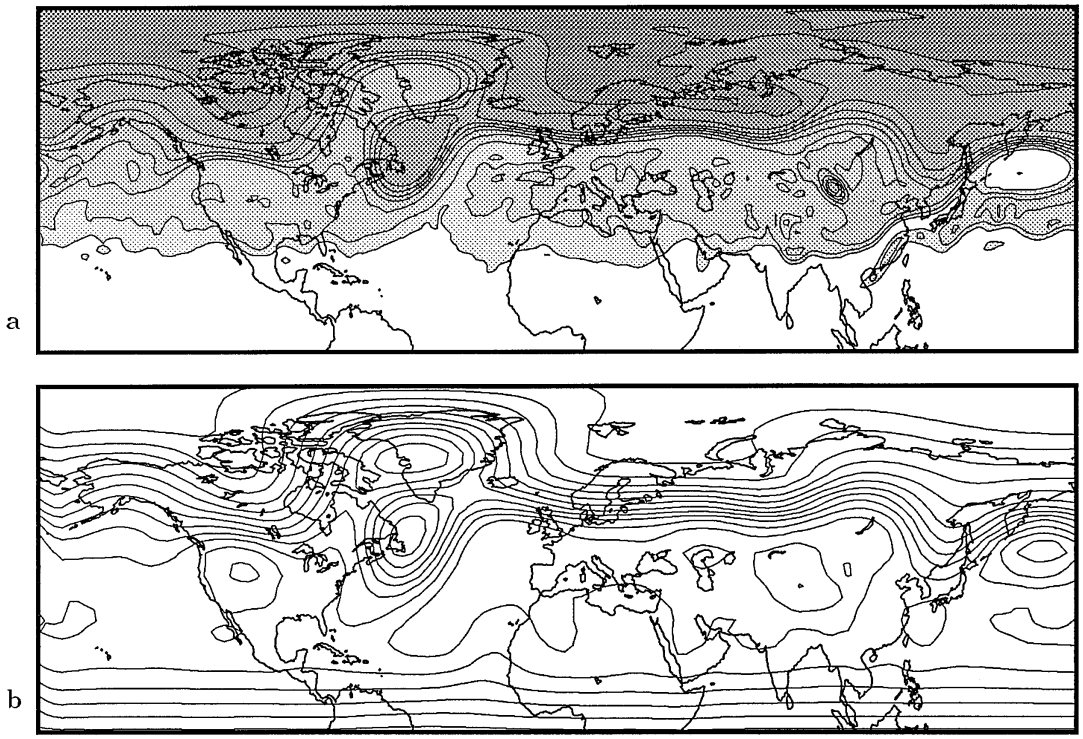

FIG. 10. An example of an Atlantic blocking event. (a) Potential vorticity field is depicted. Isolines as in Fig. 5. From the corresponding streamfunction in (b) the dipole structure becomes very clear.

pause. Obviously a more complete theory of the tropopause should be able to predict which isentropic surface defines the tropical tropopause.

For the Underworld we note the following. The Underworld isentropes in the real atmosphere do not cut the tropopause either. As noted in the introduction, these isentropes generally cut the earth's surface somewhere.

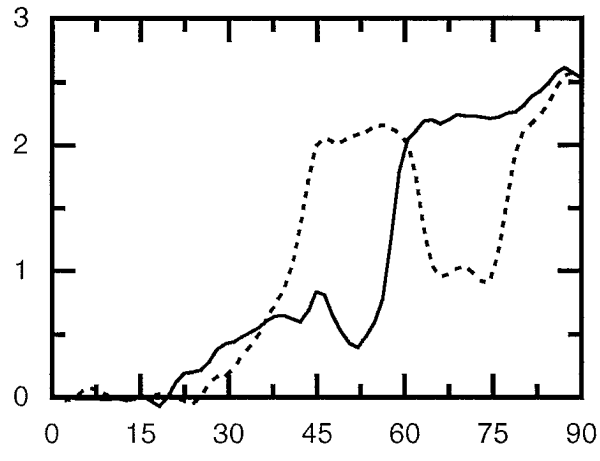

FIG. 11. Meridional crosssections through the potential vorticity field as depicted in Fig. 10 taken at $-11°$ long (solid) and at $-48°$ long (dashed). Despite the rather intricate structure of the tropopause in Fig. 10, the piecewise uniform nature of the potential vorticity field remains intact.

This means that these isentropes have lateral boundaries. Some discussion on the effect of these boundaries on the potential vorticity budget can be found in Hoskins (1991). It is clear that our simple system cannot describe the dynamics of these isentropes. One may speculate, though, that the lateral boundaries act as a leak of potential vorticity and that this leak may prevent the formation of high values of potential vorticity. This could explain why isentropes that cut the earth's surface generally do not cut the tropopause. An informal argument for the leaking effect of lateral boundaries may be found in the fact that isentropic surfaces cut the atmospheric boundary layer rather than the earth's surface. Because of convective instability in the boundary layer, the stratification, and consequently the potential vorticity, is very low. This means that the turbulent exchange between an isentropic layer and the turbulent boundary layer must lead to an effective local decrease in the absolute value of the potential vorticity in the isentropic layer, under the assumption that on the average there is no net mass flux in the process. By mixing processes across the isentropic layer, this low potential vorticity air is transported northward. In this way, the potential vorticity on an isentrope, cutting the atmospheric boundary layer, relaxes toward a zero potential vorticity state. This mechanism is analogous to that of the ventilated thermocline in the ocean, where an isopycnic surface cuts the turbulent boundary layer (e.g., Woods 1985),

though in this process the local net mass flux is generally not equal to zero.

Hoskins (1991) remarks in his article on the PV-θ view of the general circulation that the lowest isentrope of the Middleworld usually grazes the earth's surface in the Tropics and that this presumably happens by chance. This study indicates that this might not happen by chance. Indeed, our study suggests that *if* an isentrope does not intersect the earth, a tropopause will be formed and therefore the isentrope must cut the tropopause. Or equivalently, *if* an isentrope does not cut through the tropopause, then it must somewhere cut the earth's surface. In terms of the ''worlds,'' an isentrope lying everywhere above the earth's surface is part of the Middleworld and an Underworld isentrope must always cut the earth's surface. Note that this still leaves room for isentropes that cut both the tropopause as well as the earth's surface. But if the argument of the ''ventilated isentropes'' is valid, there may be a more strict separation between isentropes that cut the tropopause and isentropes that cut the earth's surface.

5. Discussion and conclusions

The potential vorticity in the stratosphere is distinctly higher than the potential vorticity in the troposphere. The transition between the two, which is rather abrupt in terms of potential vorticity, defines the tropopause. In this paper it is argued that an isentropic view of this structure gives clear insights on the formation and maintenance of the tropopause. Diabatic effects explain why the polar parts of the isentrope have the high, stratospheric values and the equatorial parts have the low, tropospheric values. But the diabatic effects alone cannot explain why the transition between the two is so abrupt, that is, why there is a tropopause. It is argued that adiabatic inertial motion of the atmosphere contributes significantly to the ultimate potential vorticity structure. As these motions are largely adiabatic, they occur *on* the isentropic surface considered. The process of stripping plays a central role. This process enhances the isentropic potential vorticity gradient and thus leads to the formation of the tropopause. At the same time homogenization leads to diminished isentropic potential vorticity gradients throughout the rest of the stratosphere and the troposphere. The general structure of the isentropic potential vorticity is now determined by a dynamical equilibrium between forcing and dissipation, which explains the potential vorticity contrast between the stratosphere and troposphere, and an equilibrium between stripping processes and diabatic effects, which explains the isentropic width of the tropopause. This dynamical equilibrium becomes established essentially on an isentropic surface and, as such, can be observed in a one-layer isentropic model of the atmosphere.

The ubiquitous nature of a tropopause in an isentropic layer that encompasses the earth suggests that *if* an isentrope does not cut through the tropopause, this *implies* that it must somewhere cut the earth's surface. This may still leave room for isentropes cutting the tropopause as well as the earth's surface, but if the argument of the ventilated isentropes, as presented in the previous section, is valid, it follows that isentropes cutting the earth's surface generally do not cut through the tropopause. This may then explain why the lowest isentrope of the Middleworld grazes the earth's surface. This contrasts with the opinion of Hoskins in his 1991 paper on the PV-θ view of the general circulation, where he states that this presumably occurs by chance (Hoskins 1991).

The role of orography in the formation of the tropopause is important in our one-layer view of the atmosphere. The mountains act as a source of enhanced potential vorticity gradients (see Figs. 5 and 6), as well as a source of wave breaking events. As explained in section 2, these events lead to gradient enhancement, and therefore to the formation of the tropopause and homogenization elsewhere. The actual source of wave activity was not referred to in the argument. In a three-dimensional atmosphere, baroclinic instability is another candidate mechanism. For example, Simmons and Hoskins (1978) show that baroclinically unstable waves grow fastest in the upper troposphere. The decay of these disturbances, that is, the nonlinear stage at which the waves are breaking, show a quasi-barotropic behavior. They also show that the maximum southward flux of potential vorticity occurs during this decay stage and that this flux is mainly located southward of the jet core. In this way the jet becomes stronger in much the same way as is shown in this article. So their experiments show that baroclinic instability may induce the required wave activity as does the orography in our equivalent barotropic model. Experiments with multilayer primitive equation models (P. H. Haynes and J. F. Scinocca 1995, personal communication), show that baroclinically induced waves indeed can play the same role in the formation of the tropopause. So as long as there is a buildup of a polar vortex on an isentropic surface and as long as there is some mechanism that induces wave activity, we will find a tropopause.

That the orography's role in our one-layer view is rather large can also be observed in the behavior of the model in the Southern Hemisphere. In Fig. 12 we see the Southern Hemisphere of the same situation as depicted in Fig. 8. While the northern tropopause is very clear, the southern tropopause is almost completely absent. This is the case throughout the complete integration period. Besides that, the variability of the southern flow pattern is also very low. This observation is consistent with the mechanism for tropopause formation as presented in this study: the lack of variability, which is the source of stripping events, results in the absence of a clear tropopause. The complementary process of homogenization inside and outside the polar vortex will

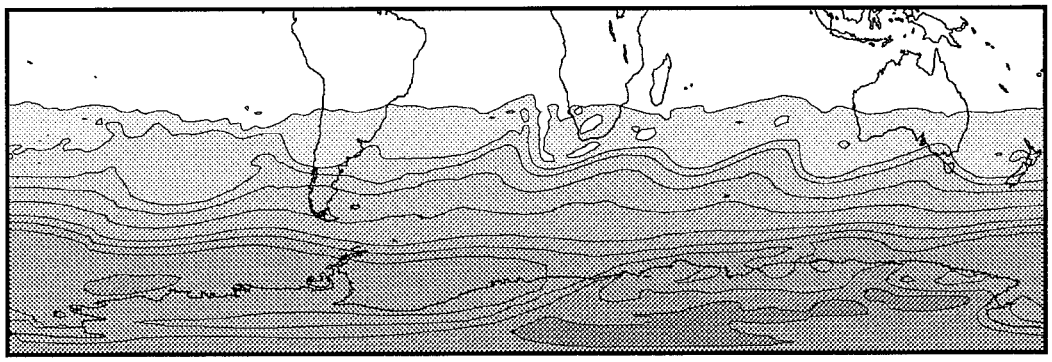

FIG. 12. The Southern Hemisphere in the situation depicted in Fig. 8. There is little wave activity, which inhibits the onset of a clear tropopause. Effective homogenization does not occur either. Isolines are plotted every 0.3 dimensionless potential vorticity units. The most northern isoline corresponds to $q = -0.3$.

also be inhibited by the lack of variability. The profile in Fig. 13 shows these effects clearly. Whereas in the Northern Hemisphere the flow is constantly stirred by the orography, the southern orographic features are too weak to induce enough wave activity. Baroclinic instability seems to play the key role in the formation of the Southern Hemisphere tropopause.

The tropopause shows a strong structural stability (Fig. 10), which is indicative for the low rates of stratosphere–troposphere exchange of air. The potential vorticity signature of stratospheric and tropospheric air can be blurred only by diabatic effects (including small-scale potential temperature mixing). Note that these diabatic effects (excluding small-scale potential temperature mixing) do not directly affect the mixing ratio of passive tracers. Measurements (P. H. Haynes 1995, personal communication; Haynes and Ward 1993) indeed show that the mixing ratio of some chemicals show steeper contrasts between stratosphere and troposphere than the potential vorticity. Though the local flux of tracers over the tropopause is very difficult to determine, the net flux can be estimated using global budgets (Holton et al. 1995).

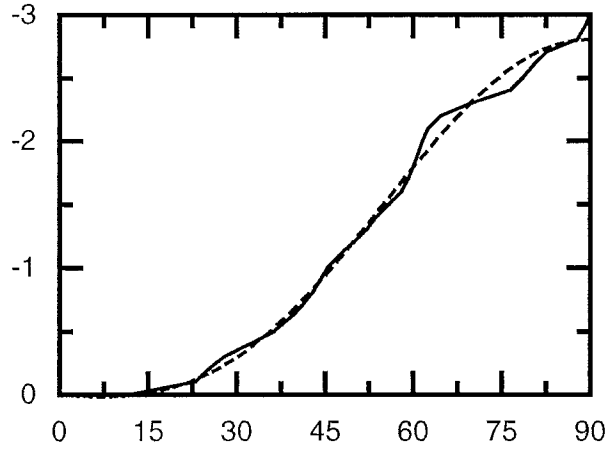

FIG. 13. Potential vorticity as a function of equivalent latitude (definitions as in Fig. 2) for the situation depicted in Fig. 12 (solid line) and for the field q_s that corresponds to a zonally symmetric steady state [see Eq. (3) for the definition] (dashed).

As shown in this article, the model intrinsically leads to an almost piecewise uniform isentropic potential vorticity, which is ideally suited to a contour dynamics approach. But we have obtained more than that. In the contour dynamics model, the area of the vortex and the potential vorticity contrast between stratosphere and troposphere are fixed parameters. However, in the present spectral model they are dynamically obtained variables. This suggests the possibility of changing the two parameters in the contour dynamics model into dynamical variables that in a global sense satisfy the diabatic budgets on the isentropic surface. This extension may enhance the applicability of contour dynamics to the atmosphere considerably. Presently we are working on an implementation of this view into a contour dynamics code.

The fact that the generic potential vorticity structure on Middleworld isentropes is approximately piecewise uniform reduces the number of degrees of freedom drastically. Among these degrees of freedom are the area of the stratospheric part of a Middleworld isentrope and the typical potential vorticity contrast between stratosphere and troposphere. These parameters may well be important parameters for the large-scale behavior of the atmosphere. In Ambaum and Verkley (1995), for example, it is shown that the structure of planetary waves is strongly dependent on the area.

Acknowledgments. Wim Verkley is gratefully thanked for the many discussions and valuable comments on the manuscript. He also provided the spectral model that was used in this study. The author was supported by the Netherlands Geosciences Foundation (GOA), with financial aid from the Netherlands Organization for Scientific Research (NWO). The anonymous reviewers are also acknowledged for their inspiring comments.

APPENDIX

The Influence of Diabatic Processes on Potential Vorticity Gradients

As made clear in this paper, isentropic advective processes can influence potential vorticity gradients in a systematic way. But diabatic effects, for example, the source term on the right-hand side of Eq. (1a), can also influence these gradients. Following Eq. (1a) we have, apart from advective processes,

$$\frac{\partial}{\partial t}\nabla q = \nabla S, \tag{A1}$$

where S stands for the right-hand side of Eq. (1a). The potential vorticity gradient has a direction, say **n**, that is perpendicular to the isolines of potential vorticity. One of these isolines can be identified with the tropopause. The formula above describes the change of the tropopause width as a result of diabatic effects. To demonstrate this, Eq. (A1) is written as

$$\frac{\partial}{\partial t}\left(\frac{\partial q}{\partial n}\right) = \frac{\partial S}{\partial n}, \tag{A2}$$

where $\partial/\partial n$ is the derivative in the direction **n**. If we assume that the width of the isentropic tropopause can be estimated as D, and the potential vorticity contrast over this width as δq, we may estimate the gradient $\partial q/\partial n$ as $\delta q/D$. If we now assume that δq remains constant as a function of time, we may estimate from Eq. (A2)

$$\frac{\partial D}{\partial t} = -\frac{D^2}{\delta q}\frac{\partial S}{\partial n}. \tag{A3}$$

It can now be easily calculated what effect different diabatic terms have on the potential vorticity gradient and the tropopause width. Results for radiative damping, Ekman friction and diffusion are given in Table A1.

TABLE A1. The expected behavior of the potential vorticity gradient and the isentropic tropopause width as a result of different diabatic effects.

	S	$\frac{\partial}{\partial t}\left(\frac{\partial q}{\partial n}\right)$	$\frac{\partial D}{\partial t}$
Radiative damping	$\frac{F}{\tau_R}(\psi - \psi_0)$	$-\frac{F}{\tau_R}(v_\parallel - v_{0\parallel})$*	$\frac{FD^2}{\delta q \tau_R}(v_\parallel - v_{0\parallel})$
Ekman	$-\kappa\zeta$	$\kappa\frac{\partial^2 v_\parallel}{\partial n^2}$	$-\frac{\kappa D^2}{\delta q}\frac{\partial^2 v_\parallel}{\partial n^2}$
Diffusion**	$\nu\nabla^2 q$	$\approx -\nu\frac{\delta q}{D^3}$	$\approx \frac{\nu}{D}$

* The term $v_\parallel$ represents the velocity perpendicular to the main potential vorticity gradient, that is, parallel to the tropopause.

** Here it is assumed that $\nabla^2 q$ scales as $\delta q/D^2$.

It can be seen from Table A1 that these diabatic processes generally lead to a decrease in the potential vorticity gradient, that is, to a widening of the isentropic tropopause. These effects may be counteracted by the effect of vortex stripping. It may also be observed that diffusion leads to a singular behavior of the width D as a function of time for vanishing D. This leads to a value of D proportional to the square root of time, as is the familiar behavior for diffusion. The other terms are nonsingular for vanishing D. In the case of thermal damping, the change of tropopause width is proportional to $v_\parallel$, the velocity parallel to the tropopause. One may note that generally $v_\parallel$ is strongly dependent on D, so it is not very useful to integrate the expressions in the table for a constant $v_\parallel$. On dimensional grounds, for example, one may expect $\partial^2 v_\parallel/\partial n^2$ to scale as $1/D^2$, which, in the Ekman case, would lead to a linear increase of D with time.

REFERENCES

Ambaum, M. H. P., and W. T. M. Verkley, 1995: Orography in a contour dynamics model of large-scale atmospheric flow. *J. Atmos. Sci.,* **52,** 2643–2662.

Batchelor, G. K., 1955: On steady laminar flow with closed streamlines at large Reynolds number. *J. Fluid Mech.,* **1,** 177–190.

Cressman, G. P., 1958: Barotropic divergence and very long atmospheric waves. *Mon. Wea. Rev.,* **86,** 293–297.

Danielsen, E. F., 1968: Stratospheric–tropospheric exchange based on radioactivity, ozone and potential vorticity. *J. Atmos. Sci.,* **25,** 502–518.

Haynes, P. H., 1990: High-resolution three-dimensional modelling of stratospheric flows: Quasi-two-dimensional turbulence dominated by a single vortex. *Topological Fluid Mechanics,* H. K. Moffatt and A. Tsinober, Eds., Cambridge University Press, 345–354.

——, and W. E. Ward, 1993: The effect of realistic radiative transfer on potential vorticity structures, including the influence of background shear and strain. *J. Atmos. Sci.,* **50,** 3431–3453.

Held, I. M., 1982: On the height of the tropopause and the static stability of the troposphere. *J. Atmos. Sci.,* **39,** 412–417.

Holton, J. R., P. H. Haynes, M. E. McIntyre, A. R. Douglas, R. B. Hood, and L. Pfister, 1995: Stratosphere–troposphere exchange. *Rev. Geophys.,* **33,** 403–439.

Hoskins, B. J., 1991: Towards a PV-θ view of the general circulation. *Tellus,* **43AB,** 27–35.

——, M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quart. J. Roy. Meteor. Soc.,* **111,** 877–946.

Juckes, M. N., and M. E. McIntyre, 1987: A high-resolution one-layer model of breaking planetary waves in the stratosphere. *Nature,* **328,** 590–596.

Legras, B., and D. G. Dritschel, 1993: Vortex stripping and the generation of high vorticity gradients. *Appl. Sci. Res.,* **51,** 445–455.

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

Mariotti, A., B. Legras, and D. G. Dritschel, 1994: Vortex stripping and the erosion of coherent structures in two-dimensional flows. *Phys. Fluids,* **6,** 3954–3962.

McIntyre, M. E., and T. N. Palmer, 1983: Breaking planetary waves in the stratosphere. *Nature,* **305,** 593–600.

——, and ——, 1984: The 'surf zone' in the stratosphere. *J. Atmos. Terr. Phys.,* **46,** 825–849.

McWilliams, J. C., 1984: The emergence of isolated coherent vortices in turbulent flow. *J. Fluid Mech.,* **146,** 21–43.

Melander, M. V., J. C. McWilliams, and N. J. Zabusky, 1987: Axisymmetrization and vorticity-gradient intesification of an isolated two-dimensional vortex through filamentation. *J. Fluid Mech.,* **178,** 137–159.

Pedlosky, J., 1979: *Geophysical Fluid Dynamics.* Springer-Verlag, 710 pp.

Pierrehumbert, R. T., and H. Yang, 1993: Global chaotic mixing on isentropic surfaces. *J. Atmos. Sci.,* **50,** 2462–2480.

Polvani, L. M., P. W. Waugh, and R. A. Plumb, 1995: On the subtropical edge of the stratospheric surf zone. *J. Atmos. Sci.,* **52,** 1288–1309.

Queney, P., 1952: Les ondes atmosphériques considérées comme associées aux discontinuités du tourbillon. *Tellus,* **4,** 88–111.

Rhines, P. B., and W. R. Young, 1982: Homogenization of potential vorticity in planetary gyres. *J. Fluid Mech.,* **122,** 347–367.

Rinne, J., and H. Järvinen, 1993: Estimation of the Cressman term for a barotropic model through optimization with use of the adjoint model. *Mon. Wea. Rev.,* **121,** 825–833.

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

Thorncroft, C. D., B. J. Hoskins, and M. E. McIntyre, 1993: Two paradigms of baroclinic-wave life-cycle behaviour. *Quart. J. Roy. Meteor. Soc.,* **119,** 17–55.

Verkley, W. T. M., 1994: Tropopause dynamics and planetary waves. *J. Atmos. Sci.,* **51,** 509–529.

Woods, J. D., 1985: Physics of thermocline ventilation. *16th International Colloquium on Ocean Hydrodynamics,* J. C. J. Nihoul, Ed., Elsevier Oceanography Series, Vol. 40, Elsevier, 543–590.