Mean Meridional Circulations in the Middle Atmosphere

by

Kirill Semeniuk

A thesis submitted in conformity with the requirements for the degree of Doctor of Philosophy
Graduate Department of Physics
University of Toronto

© by Kirill Semeniuk 2000
The author has granted a non-exclusive licence allowing the National Library of Canada to reproduce, loan, distribute or sell copies of this thesis in microform, paper or electronic formats.

The author retains ownership of the copyright in this thesis. Neither the thesis nor substantial extracts from it may be printed or otherwise reproduced without the author’s permission.

L’auteur a accordé une licence non exclusive permettant à la Bibliothèque nationale du Canada de reproduire, prêter, distribuer ou vendre des copies de cette thèse sous la forme de microfiche/film, de reproduction sur papier ou sur format électronique.

L’auteur conserve la propriété du droit d’auteur qui protège cette thèse. Ni la thèse ni des extraits substantiels de celle-ci ne doivent être imprimés ou autrement reproduits sans son autorisation.

0-612-50060-8
Mean Meridional Circulations in the Middle Atmosphere
Doctor of Philosophy, 2000
Kirill Semeniuk, Department of Physics, University of Toronto

Abstract

In this thesis are presented results describing features of zonal-mean dynamics in the middle atmosphere which lie outside the framework described by the “downward control” principle. This principle applies to the nearly steady state wave-driven regime that is approached in the extratropics during solstice seasons. However, it does not apply in the tropics and transience in the extratropics can be significant.

One of the assumptions underlying the downward control principle is that local changes in wave drag or diabatic heating do not significantly modify wave drag elsewhere. Force feedbacks are investigated in the particular case of artificial “sponge layers” used in general circulation models. Such relaxational damping is found to couple with changes in forcing or diabatic heating in an unrealistic manner, leading to “upward control”.

In the tropics there exists the so-called nonlinear Hadley circulation, which is driven by a combination of gradient-wind balance and solar heating gradients rather than wave forcing. Previous work is extended by numerical experiments which support the validity of gradient-wind balance in the tropics. Inertial instability is found to have a potentially large impact on the structure of this circulation.

Another purpose of this work is to explain annual mean upwelling in the tropical stratosphere, an important element in the middle atmosphere climate. The inviscid linear view of the diabatic meridional circulation induced by extratropical wave forcing cannot explain the observed mean upwelling. The nonlinear Hadley circulation is found to make a large contribution to upwelling in the tropical upper stratosphere. In the tropical lower stratosphere, upwelling is produced by the wave-driven circulation. A nonlinear mechanism for this is demonstrated by numerical experiments.

The dynamical impact of spatio-temporal variation in the radiative damping rate is also investigated. In regions where the damping rate is weak there is increased lag between the evolution of wave forcing and the induced circulation. In such regions the temperature is also more sensitive to even a weak circulation. A time dependence in the radiative damping rate is shown to produce a transient-like signature in the annually averaged dynamical fields, which complicates the interpretation of annual mean anomalies.
Acknowledgements

My thanks go out to the following people who contributed to the development of this work: Prof. T. G. Shepherd; Dr. D. Sankey, Dr. J. Koshyk, Prof. R. Garcia, Prof. M. Geller and Dr. C. McLandress; D. Wirosoetisno, K. Ngan, P. J. Kushner and O. Bokhove; Prof. H.-R. Cho and Prof. G. W. K. Moore.

I am also grateful to the Department of Physics, University of Toronto, for providing a stimulating and open work environment.

Financial support for this work has come through the Ontario Graduate Scholarship, University of Toronto Open Fellowship, as well as research grants from the Natural Sciences and Engineering Research Council and from the Atmospheric Environment Service.
Contents

Abstract ................................................................. ii
Acknowledgements ......................................................... iii
Contents ................................................................. iv
List of figures ............................................................ vii

1 Introduction ............................................................. 1
  1.1 The Brewer-Dobson circulation ......................................... 1
    1.1.1 Zonal-mean dynamics ............................................. 3
  1.2 Issues ........................................................................... 6
    1.2.1 Relaxational mechanical damping ..................................... 6
    1.2.2 The stratospheric Hadley circulation ................................. 7
    1.2.3 Tropical upwelling ................................................ 9
    1.2.4 Aspects of extratropical transience .................................. 9

2 Sponge layer feedbacks in middle-atmosphere models .................. 11
  2.1 Introduction .................................................................. 11
  2.2 Idealized numerical solutions .......................................... 13
  2.3 Asymptotic solutions .................................................. 23
  2.4 The case of a uniform Rayleigh drag ................................ 27
  2.5 Conclusion ................................................................... 31

3 The nonlinear Hadley circulation and inertial instability ........... 37
  3.1 Introduction .................................................................. 37
  3.2 Model .......................................................................... 39
  3.3 Nonlinear Hadley circulation ........................................... 42
    3.3.1 Theory ..................................................................... 42
    3.3.2 Behaviour of the system in the absence of wave drag ............ 43
    3.3.3 Role of tropical wave drag ....................................... 50
    3.3.4 Validity of the balance model .................................... 51
  3.4 Influence of the wave-driven circulation on the nonlinear Hadley circulation 54
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.5</td>
<td>Nonlinear Hadley circulation in a middle atmosphere GCM</td>
<td>63</td>
</tr>
<tr>
<td>3.6</td>
<td>Conclusion</td>
<td>66</td>
</tr>
<tr>
<td>4</td>
<td>Tropical upwelling</td>
<td>71</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>71</td>
</tr>
<tr>
<td>4.2</td>
<td>Annual mean tropical upwelling</td>
<td>72</td>
</tr>
<tr>
<td>4.2.1</td>
<td>Role of viscosity in annual mean tropical upwelling</td>
<td>73</td>
</tr>
<tr>
<td>4.2.2</td>
<td>NLHC-driven tropical upwelling</td>
<td>74</td>
</tr>
<tr>
<td>4.2.3</td>
<td>Parcel trajectories</td>
<td>76</td>
</tr>
<tr>
<td>4.3</td>
<td>WDC-driven tropical upwelling</td>
<td>82</td>
</tr>
<tr>
<td>4.4</td>
<td>Effect of the NLHC on the WDC in the tropics</td>
<td>92</td>
</tr>
<tr>
<td>4.5</td>
<td>Contribution of the NLHC to upwelling in a middle atmosphere GCM</td>
<td>101</td>
</tr>
<tr>
<td>4.6</td>
<td>Conclusion</td>
<td>103</td>
</tr>
<tr>
<td>5</td>
<td>Aspects of transience in the extratropics</td>
<td>107</td>
</tr>
<tr>
<td>5.1</td>
<td>Introduction</td>
<td>107</td>
</tr>
<tr>
<td>5.2</td>
<td>Model</td>
<td>110</td>
</tr>
<tr>
<td>5.2.1</td>
<td>Unresolved dynamics</td>
<td>112</td>
</tr>
<tr>
<td>5.2.2</td>
<td>Approach</td>
<td>113</td>
</tr>
<tr>
<td>5.3</td>
<td>Response to switch-on forcing with spatially variable ( \alpha )</td>
<td>114</td>
</tr>
<tr>
<td>5.3.1</td>
<td>The ( \alpha(\phi) ) case</td>
<td>115</td>
</tr>
<tr>
<td>5.3.2</td>
<td>The ( \alpha(z) ) case</td>
<td>116</td>
</tr>
<tr>
<td>5.4</td>
<td>Response to periodic forcing with spatially variable ( \alpha )</td>
<td>118</td>
</tr>
<tr>
<td>5.4.1</td>
<td>Uniform ( \alpha ) case</td>
<td>119</td>
</tr>
<tr>
<td>5.4.2</td>
<td>The ( \alpha(\phi) ) case</td>
<td>123</td>
</tr>
<tr>
<td>5.4.3</td>
<td>The ( \alpha(z) ) case</td>
<td>123</td>
</tr>
<tr>
<td>5.4.4</td>
<td>The ( \alpha(t) ) case</td>
<td>125</td>
</tr>
<tr>
<td>5.5</td>
<td>Annual mean signature of ( \alpha(t) )</td>
<td>127</td>
</tr>
<tr>
<td>5.6</td>
<td>Long-lived transients</td>
<td>133</td>
</tr>
<tr>
<td>5.6.1</td>
<td>Persistence of transience and non-locality</td>
<td>134</td>
</tr>
<tr>
<td>5.7</td>
<td>Conclusion</td>
<td>136</td>
</tr>
<tr>
<td>6</td>
<td>Conclusions</td>
<td>139</td>
</tr>
<tr>
<td>6.1</td>
<td>Summary</td>
<td>139</td>
</tr>
<tr>
<td>6.2</td>
<td>Discussion</td>
<td>143</td>
</tr>
</tbody>
</table>
A  Hough mode decomposition ........................................ 145

B  Streamfunction equation .......................................... 147
    B.1  Inertial adjustment scheme .................................. 148

References. ............................................................ 151
List of figures

1.1 Schematic of meridional stratospheric structure and transport. 2
2.1 $F$ induced $\psi, \bar{T}$ and $\bar{u}$ with and without sponge layer. 16
2.2 $Q$ induced $\psi, \bar{T}$ and $\bar{u}$ with and without sponge layer. 20
2.3 As in Figure 2.1 but with $F$ inside the sponge layer. 22
2.4 Dependence of $\eta$ on $n$ and $\dot{r}/\alpha$. 25
2.5 $F$ induced $\psi, \bar{T}$ and $\bar{u}$ with a uniform background Rayleigh drag. 28
2.6 As in Figure 2.5 but for imposed $Q$. 29
2.7 $F$ induced $\psi, \bar{T}$ and $\bar{u}$ with a diffusive sponge layer. 33
3.1 Balance model January $\psi, \bar{T}, \bar{m}$ and $Q$ when $F \equiv 0$ with CIRA derived $\bar{T}_{rad}$. 44
3.2 Balance model latitudinal profile of $\bar{u}$ and $\bar{T}$ at 45 km. 45
3.3 CIRA January mean $\bar{m}$. 48
3.4 Balance model January fields when $\alpha = \alpha(z)$. 48
3.5 Balance model January fields using a comprehensive radiative transfer scheme. 49
3.6 Balance model vs. primitive equations model July with identical forcing. 52
3.7 Balance model January $\psi, \bar{T}, \bar{m}$ and $Q$ with imposed $F$ and $\bar{T}_{rad} = 240$ K. 57
3.8 Balance model January $\psi, \bar{T}, \bar{m}$ and $Q$ with imposed $F$ and CIRA $\bar{T}_{rad}$. 58
3.9 Numerical difference of $\psi$ and $Q$ from Figure 3.8 and Figure 3.1. 59
3.10 CMAM January mean gradient-wind imbalance. 64
3.11 CMAM $\bar{u}$ showing presence of inertial instability. 64
3.12 CMAM January mean $\bar{u}$ and $\bar{m}_{\phi}$ at 37 km vs. $\phi$. 65
3.13 CMAM and observed January mean $Q$ at 37 km vs. $\phi$. 65
3.14 CMAM January mean $\bar{m}$. 66
4.1 Balance model annual mean upwelling with $F \equiv 0$ and comprehensive $Q$ scheme. 76
4.2 Time series of vertical velocity at the equator. 77
4.3 NLHC parcel trajectories with $F \equiv 0$; parcels initially on the equator. 77
4.4 NLHC parcel trajectories with $F \equiv 0$; parcels initially at $12^\circ$N. 78
4.5 Parcel cloud distribution by the NLHC with $F \equiv 0$. 80
4.6 NLHC parcel trajectories with subtropical $F$; parcels initially on the equator. 81
4.7 NLHC parcel trajectories with subtropical \( \mathcal{F} \); parcels initially at 12°N.  
4.8 Parcel cloud distribution by the NLHC with subtropical \( \mathcal{F} \).  
4.9 Annual mean upwelling at 34 km with \( \mathcal{F} \) poleward of 15°N.  
4.10 Time series of the upwelling at 34 km on the equator.  
4.11 Annual mean upwelling on the 34 km level when \( \alpha = 0.025 \text{ day}^{-1} \).  
4.12 Annual mean upwelling at 34 km when \( \alpha = 0.025 \text{ day}^{-1} \) and \( \mathcal{F} \) is poleward of 22°N.  
4.13 Evolution of the diabatic heating during “off” phase.  
4.14 Time series of equatorial upwelling at 34 km in the inviscid model.  
4.15 Annual mean upwelling at 34 km in the inviscid model.  
4.16 Inviscid model \( \bar{m} \) at \( t = 105 \) days.  
4.17 Inviscid model annual mean upwelling at 34 km with \( \mathcal{F} \) reaching \( \phi = 0 \).  
4.18 Inviscid model equatorial upwelling time series at 34 km with \( \mathcal{F} \) reaching \( \phi = 0 \).  
4.19 Inviscid model upwelling time series at 11.25°S and 34 km.  
4.20 Background state \( \bar{m} \) vs. \( \phi \).  
4.21 \( \psi \) response to \( \mathcal{F} \) at 45°N for different background state \( \bar{m} \).  
4.22 Vertical velocity for the \( \psi \) in Figure 4.21.  
4.23 \( \psi \) response to \( \mathcal{F} \) at 30°N for different background state \( \bar{m} \).  
4.24 Vertical velocity for the \( \psi \) in Figure 4.23.  
4.25 Annual mean equatorial upwelling vs. height in CMAM and the balance model.  
4.26 CMAM mass streamfunction for January, April, July and October.  

5.1 Profiles of \( \alpha(\phi) \), \( \alpha(z) \) and \( \alpha(t) \).  
5.2 Switch-on forcing \( \psi \) for the \( \alpha(\phi) \) case.  
5.3 Switch-on forcing \( \bar{T} \) for the \( \alpha(\phi) \) case.  
5.4 Switch-on forcing \( \psi \) for the \( \alpha(z) \) case.  
5.5 Switch-on forcing \( \bar{T} \) for the \( \alpha(z) \) case.  
5.6 Time series of \( \bar{w}^* \) at 16 km and \( \phi = 45^\circ \) for uniform \( \alpha \) and \( \alpha(z) \).  
5.7 Time series of the start-up transient of \( \bar{w}^* \) for different \( \omega \).  
5.8 Periodic forcing \( \psi \) for the \( \alpha(\phi) \) case.  
5.9 Periodic forcing \( \bar{T} \) for the \( \alpha(\phi) \) case.  
5.10 Distribution of the month of maximum \( \bar{w}^* \) corresponding to Figure 5.8.  
5.11 Periodic forcing \( \psi \) for the \( \alpha(z) \) case.  
5.12 Periodic forcing \( \bar{T} \) for the \( \alpha(z) \) case.
5.13 Distribution of the month of maximum $\bar{w}^*$ corresponding to Figure 5.11. . . 129
5.14 Time series of $\bar{w}^*$ at 16 km and 42°N for uniform $\alpha$ and $\alpha(z)$. . . . . 130
5.15 Time series of $\bar{w}^*$ at 16 km and 42°N with $\alpha(t)$ for different $\omega$. . . . 131
5.16 Time series of $\bar{w}^*$ at 16 km and 42°N with uniform $\alpha$ and $\alpha(t)$ for different $\omega$. 132
5.17 Time series of $\delta T_{rad}$ induced $\Delta T$ for the $\alpha(t)$ case vs. the constant $\alpha$ case. . 133
5.18 Annual mean, $\delta T_{rad}$ induced $\Delta T$ for the $\alpha(t)$ case. . . . . . . . . . . . 134
5.19 Time series of $F$ induced $\Delta T$ for the $\alpha(t)$ case vs. the constant $\alpha$ case. . . . . 135
5.20 Annual mean, $F$ induced $\Delta T$ for the $\alpha(t)$ case. . . . . . . . . . . . . . . 136
5.21 Time series of $\bar{w}^*$ showing nonmonotonic behaviour. . . . . . . . . . . . . . 137
Chapter One

Introduction

1.1 The Brewer-Dobson circulation

Interest in the stratospheric climate and its response to anthropogenic and natural factors has increased since the discovery of near total ozone depletion over Antarctica during springtime. Understanding of the processes which govern the distribution of chemical constituents and temperatures is essential for the quantification of ozone loss. Such processes fall into two categories: those that belong to stratospheric chemistry, and those that belong to dynamics. Dynamics control the transport of chemical constituents and have a large impact on the temperatures which are important for chemistry.

There are two distinct forms of transport behaviour, which serve to define distinct regions of the stratosphere on account of the height-latitude profile of the tropopause. Below 100 hPa (approximately 18 km altitude) transport of tropospheric air and chemical constituents into the stratosphere can occur by quasi-horizontal isentropic mixing through the subtropical tropopause. In addition, there is a slow diabatic ascent at low latitudes and descent at middle to high latitudes. This so-called lowermost stratosphere is isolated from most of the tropics by the tropopause, which rises in altitude from about 300 hPa near the poles to 100 hPa in the subtropics (Fig. 1.1). The diabatic or Brewer-Dobson circulation becomes the primary agent of transport into the stratosphere above 100 hPa. Tropospheric air has to cross isentropes as it moves upward into the stratosphere through the tropical tropopause and continues in a large scale pattern of low latitude upwelling with poleward and downward motion. Isentropic mixing in the subtropics and midlatitudes acts to reduce the meridional gradients in constituent concentrations that tend to form through diabatic overturning.
The meridional circulation is much weaker than the zonal circulation and as a result there is a separation of transport timescales. Circumpolar mean air parcel drift takes on the order of a week but drift from the tropics to polar regions takes much longer than a year. This is reflected in the structure of the background state, especially above 100 hPa, which varies primarily in the meridional or height-latitude plane with relatively little longitudinal dependence. These properties enable a simplified description of the dynamics described below. In the lowermost stratosphere, however, the zonal homogeneity in constituent distributions begins to break down.

The existence of a global diabatic transport circulation was inferred by Brewer (1949) and Dobson (1956) on the basis of stratospheric water vapour and ozone distributions, respectively. However, theoretical understanding of this circulation took much longer to develop (see Dunkerton, 1978, for an overview of the history). The diabatic circulation inferred from observations (e.g. Eluszkiewicz et al., 1996, Fig. 7) takes the form of
two cells in the stratosphere. During solstice seasons there is a strong cell in the winter hemisphere and a weak cell in the summer hemisphere. The intensity and distribution of the two cells alternates between the two hemispheres, maximizing during solstices and minimizing during equinoxes. There is a single maximum of upwelling in the tropics which occurs in the summer hemisphere during solstice seasons.

Initially it was believed that the Brewer-Dobson circulation is driven by diabatic heating at low latitudes and diabatic cooling at high latitudes. However, subsequent work (see Andrews et al., 1987, or Dunkerton, 1978, for references) has shown that this is not a valid mechanism for the middle atmosphere extratropics. The nature of radiative transfer at these altitudes is relaxational, with damping rates much shorter than a season, which implies that imposed heating cannot maintain the persistent circulation required to explain observations in the extratropics. In fact, the extratropical part of the Brewer-Dobson circulation and a large component in the tropics is driven by the breaking and thermal dissipation of waves, primarily planetary waves, which originate in the troposphere. The seasonal evolution of the Brewer-Dobson circulation reflects the seasonal cycle of wave drag in the two hemispheres.

The direct effect of the planetary waves on the stratosphere is to decelerate the zonal flow so they act as a form of drag. The diabatic meridional circulation induced by this wave drag has a nonlocal character and acts to produce dynamical (or adiabatic) heating or cooling in regions remote from where it is forced. This circulation works against radiative damping, driving temperatures below radiative equilibrium in the tropics, where it leads to upwelling, and above radiative equilibrium at higher latitudes, where there is downwelling. For example, the seasonal cycle in lower tropical stratospheric temperatures is an indication of this dynamical control of temperatures and is believed to be a result of the asymmetry in the extratropical wave drag between the two hemispheres (Yulaeva et al., 1994). Such behaviour has led to the current view that extratropical wave drag acts as a “pump” for the upwelling in the tropics (e.g. Holton et al., 1995).

1.1.1 Zonal-mean dynamics

The response of an axially or zonally symmetric atmosphere in gradient-wind and hydrostatic balance to thermal (i.e. diabatic heating) and mechanical forcing (i.e. wave drag) was investigated by Eliassen (1951). He showed that a meridional circulation is induced
which keeps the system in thermal wind balance. This meridional circulation can be represented by a streamfunction which satisfies an elliptic partial differential equation. A point source of diabatic heating produces a circulation dipole aligned horizontally with vertical motion in the heating region, while a point source of mechanical forcing produces a dipole with vertical alignment and horizontal motion in the forcing region. The shape of the streamfunction is affected by the ratio of the Coriolis parameter to the buoyancy frequency.

The work of Eliassen was extended by Plumb (1982) who considered more realistic thermal forcing distributions. Plumb showed that the response depends on latitude. Forcing in the tropics produces a response with a high degree of tropical confinement. Extratropical forcing can induce a circulation with greater nonlocality extending into the opposite hemisphere.

A limitation of the work of Eliassen and Plumb is that it deals only with the instantaneous response to prescribed forcing, with no representation for the relaxational diabatic heating (infrared relaxation). The latter leads to an evolution of the streamfunction, the zonal wind and temperature. Garcia (1987) used a Newtonian cooling approximation for the diabatic heating in the thermodynamic equation and a Rayleigh drag in the zonal momentum equation to represent frictional dissipation. He found that for the typically small values of the ratio of the Rayleigh damping rate to radiative damping rate in the stratosphere, thermal forcing leads primarily to a local change in the temperature and zonal wind while mechanical forcing produces a strong meridional circulation.

The zonal-mean state has a strong stratification of isentropes and absolute angular momentum isopleths in the extratropics. The former is due to gravity through density stratification, while the latter is due to the Coriolis parameter. Air parcels conserve potential temperature and, in the zonal mean, absolute angular momentum, so there cannot be any meridional circulation in the absence of mechanical forcing and diabatic heating. The former is required for air parcels to cross absolute angular momentum isopleths while the latter is required for them to cross isentropes. In the tropics the stratification of absolute angular momentum breaks down due partly to the weakening of the Coriolis parameter which makes advective isopleth deformation easier.

A consequence of the relaxational diabatic heating and the horizontal stratification of absolute angular momentum in the middle atmosphere is that the circulation produced
by steady mechanical forcing will evolve into a state where streamlines are concentrated below the forcing region. In steady state the circulation cannot exist at latitudes with no mechanical forcing. This "downward control" behaviour, where the streamfunction at a given latitude and height is due to the forcing above it, was described by Haynes et al. (1991).

The theoretical foundation for simple zonally symmetric models with prescribed mechanical and thermal forcing was laid out by Andrews and McIntyre (1978). They developed the transformed Eulerian mean (TEM) formulation of the dynamics. In the Eulerian mean formalism the zonally asymmetric component of the atmospheric state gives rise to so-called eddy momentum and heat flux terms which act as a forcing on the zonal mean fields. It turns out that wave dissipation contributes to both the Eulerian mean meridional circulation and the flux terms and that the mean meridional circulation does not resemble the transport circulation inferred from observations. An ideal formulation should have the wave forcing only appear in the flux terms since the intent is to solve for the zonal mean fields instead of prescribing them. In the TEM framework this problem is largely resolved and wave dissipation appears as a forcing term in the mean zonal momentum equation (with a climatologically negligible heat flux term in the thermodynamic equation). This separation between mechanical and thermal forcing simplifies the interpretation of the dynamics. The mean meridional or residual circulation in this transformed formulation is qualitatively similar to the true transport circulation, which is a Lagrangian mean circulation that is difficult to formulate (Andrews et al., 1987). In addition, the forcing term has a clearer physical meaning in terms of wave-mean flow interaction theory, as it is the divergence of the Eliassen-Palm wave activity flux.

In the tropics there exists another type of circulation which is not driven by wave dissipation, as was first pointed out by Dunkerton (1989). Thermal wind balance prevents the temperature from relaxing to radiative equilibrium near the equator during solstice seasons. The resulting persistent diabatic heating drives a type of nonlinear Hadley circulation, which conserves absolute angular momentum and is an important component of the Brewer-Dobson circulation at low latitudes.
1.2 Issues

The basics of the response of the zonal mean system to imposed forcing are well understood; however, understanding of zonal mean dynamics is not complete. The zonal mean system has a closure problem in that the myriad of zonally asymmetric processes which contribute to its dynamical driving have to be parameterized in terms of the zonal mean fields or simply prescribed. The problem with the former is that such parameterizations are not justified a priori and existing approaches are rather ad hoc requiring tuning to give a reasonable model climatology. Prescribing the forcing has its own drawbacks. Such an approach would ideally be based on observations, but these are poor; in particular there are virtually no direct global observations of winds. Using assimilated data (Swinbank and O'Neill, 1994) allows a forcing to be calculated but it requires two levels of differentiation of coarse resolution fields which introduces significant errors. The same problem occurs when using general circulation model (GCM) data. The uncertainty in the wave drag distribution, specifically in the tropics, is an important problem in middle atmosphere research including the work presented in this thesis.

In spite of its limitations the zonal-mean view is useful for understanding the diabatic circulation. The aim of this thesis is to address some of the unanswered questions about the dynamics of the diabatic circulation in this framework. We examine how the system departs from downward control in the tropics and on account of transience at higher latitudes. In particular, we deal with the problems of the extratropical pump picture of tropical upwelling. The location of the solstitial tropical upwelling maximum, which occurs in the summer hemisphere, is inconsistent with the fact that the wave driving is concentrated in the winter hemisphere. Net annual mean upwelling in the tropics also cannot be viewed as a simple consequence of the one way "pumping" by extratropical wave drag. We consider the role of the nonlinear Hadley circulation in tropical upwelling, a subject which has not received much attention in the literature. Further details of the issues mentioned here are given in the following subsections.

1.2.1 Relaxational mechanical damping

The assumption underlying the downward control characterization of the diabatic circulation response to wave drag is that the wave drag can be separated from the effect it
produces. In other words, that any feedbacks on the wave drag itself and redistribution of wave drag elsewhere (through a nonlocal change in the zonal wind) are secondary, allowing a causal link between the circulation and the force to be established. It is not obvious \textit{a priori} that this view is correct in the context of middle atmosphere models, which form the basis for much of the research into the physics of the middle atmosphere. Specifically, there are certain features of such models that may violate downward control.

One of the simplest ways to represent mechanical forcing in the middle atmosphere is by Rayleigh drag, i.e. a linear “spring-like” damping on the zonal wind. This is a crude approximation to the dependence of the wave drag distribution on the zonal wind (e.g. gravity waves, Leovy, 1964) or to some kind of “eddy viscosity” (e.g. turbulent friction, Garcia, 1987). It is also typical for many middle atmosphere GCMs to include a Rayleigh drag “sponge layer” near the upper lid for numerical purposes to control spurious downward reflection of waves. However, there is a problem with using relaxational mechanical forcing in that it has an unlimited capacity for damping. As a result, it can couple with realistic mechanical forcing in an unrealistic manner. The strength of true wave drag is determined by its sources in the troposphere. Local changes in the zonal wind can redistribute this drag but cannot change its total amount. Therefore, wave breaking cannot act as a linear damping.

The concept of ubiquitous viscosity in the middle atmosphere also has little justification. Diffusive-like behaviour at these levels is associated with wave breaking and does not form a persistent background field. Molecular diffusion is vanishingly small in this layer of the atmosphere (Holton, 1992).

The effect of relaxational mechanical forcing on the behaviour of the zonal mean system is the focus of Chapter 2. The impact of sponge layers on GCMs and uniform background Rayleigh friction on zonal mean models is considered. It is found that sponge layers can produce a significant feedback leading to an “upward control” behaviour. The effect of a uniform viscosity is to keep the system in a transient-like state and away from the downward control limit.

1.2.2 The stratospheric Hadley circulation
The theory of the stratospheric nonlinear Hadley circulation put forth in Dunkerton (1989, 1991) involves a number of assumptions. We revisit the work of Dunkerton to
explore the effect of these assumptions in Chapter 3 and also consider the impact of the nonlinear Hadley circulation (NLHC) in a middle atmosphere GCM.

The key assumption underlying the existence of the thermally driven circulation in the tropical stratosphere is gradient wind balance. Only in the extratropics is there a sound basis for this type of balance, which is related to spherical geometry and the strong rotation of the Earth manifested through the Coriolis parameter. In the tropics the Coriolis term is weak and a dominant balance is not apparent. We test the validity of the gradient-wind balance assumption by using a three-dimensional primitive equations model and find that resolved adjustment processes in the tropics are sufficiently rapid to keep the system balanced in the zonal mean.

Another idealization in Dunkerton's models was the use of the so-called parabolic approximation, where the action of inertial instability was represented only as a modification of the streamfunction operator. It was argued that parameterizations such as inertial adjustment were also ad hoc and not necessary to investigate the NLHC in the steady-state limit (which is relevant for solstice seasons when the NLHC forms). Yet the impact of inertial adjustment on the zonal wind and temperature cannot be ignored since it modifies the streamfunction through the forcing terms as well, specifically the diabatic heating. An explicit inertial adjustment scheme is used in our numerical experiments. In the absence of wave drag it is found to produce a significant effect. However, wave drag reduces the region of inertial adjustment and its influence in the real atmosphere.

Dunkerton's models employed either a rigid lid near the stratopause or a mesospheric sponge layer to close off the NLHC. This was reasoned to be necessary to produce the saddle point in the absolute angular momentum distribution inferred from observations. In Chapter 3 it is shown that the NLHC begins to close off in the stratopause region and above due to the radiative transfer characteristics of the middle atmosphere rather than wave drag. Without wave drag the model absolute angular momentum distribution has many unrealistic features, however.

The theory of Dunkerton (1989) and to a lesser extent Dunkerton (1991) was based on the theory of Held and Hou (1980) for the tropospheric nonlinear Hadley circulation. The effect of extratropical wave drag was interpreted by Dunkerton as a modification of the "equal area rule" which defines the steady state latitudinal profile of the vertically averaged temperature, and therefore the meridional extent of the NLHC. A problem
with this interpretation is that even though it is consistent with the governing equation, it obscures the fact that the wave driven circulation and the NLHC can only couple in a limited sort of way. We pursue this further in Chapter 3. The effect of wave drag on the NLHC is considered in Chapter 3, while the impact of the NLHC on the wave-driven circulation is examined in Chapter 4.

1.2.3 Tropical upwelling

At first sight, annual mean tropical upwelling seems like an unambiguous feature of the Brewer-Dobson circulation given that wave drag in the stratosphere is dominated by the breaking of planetary waves, which act consistently to produce a negative forcing. However, there is a high degree of confinement of the tropics from the bulk of the wave drag which is concentrated in the "surf zone" distributed poleward of the subtropics in the lower stratosphere (Andrews et al., 1987). Constituent observations also indicate the presence of sharp gradients in mixing ratios in the subtropics, which imply a high degree of isolation of the tropics from extratropical isentropic mixing (Plumb, 1996). Consequently, the downward control behaviour of the meridional circulation limits its ability to produce the observed annual mean upwelling in the tropics (Mote et al., 1996).

To a good approximation, dynamics in the extratropics can be captured by the quasi-geostrophic approximation of the governing equations. In the zonal mean the geostrophic equations are linear. This feature was exploited in previous studies (e.g. Eliassen, Plumb, Garcia). Although the assumptions underlying quasi-geostrophy fail in the tropics, the simplest model of the tropics is to assume that this approximation is valid near the equator as well. Since the time mean of the linear system has the form of the steady state (assuming no annual mean tendencies), the annual mean circulation in the tropics will be zero if the annual mean forcing is confined to the extratropics.

To get around this constraint requires either tropical mechanical forcing (wave drag or viscosity) or nonlinearity. The linear viscous mechanism for annual mean tropical upwelling was investigated by Plumb and Eluszkiewicz (1999). In Chapter 4 we examine a nonlinear mechanism for tropical upwelling as pertaining to the nonlinear Hadley circulation identified by Dunkerton (1989) which appears to be important in the upper stratosphere, and the wave driven circulation, which is responsible for the upwelling in the lower stratosphere. In addition, the effect of tropical dynamical characteristics, including the NLHC, on the wave driven circulation is investigated.
1.2.4 Aspects of extratropical transience

Much of the existing work on the dynamics of the mean meridional circulation in an idealized framework (e.g. Garcia, 1987; Haynes et al., 1991) employs simple Newtonian cooling with a constant radiative damping rate. Although some work on the dependence of solutions on the radiative damping rate has been performed (e.g. Holton et al., 1995), it involves simple variation of the amplitude. In fact, the radiative damping rate has a nonlocal and nonlinear dependence on temperature. However, using a comprehensive radiative transfer scheme for the diabatic heating complicates the analysis. In particular, it is difficult to study the separate effects of thermal and mechanical forcing since radiative nonlinearity requires a realistic distribution of wave drag to produce realistic radiative damping rates. The compromise followed here is to let the radiative damping rate be a function of space and time in the Newtonian cooling scheme.

Part of the motivation for this investigation is the sensitivity of the polar vortex interior to forcing by wave drag located at higher altitudes and lower latitudes. The maintenance of temperatures inside the Antarctic polar vortex is not well understood but is possibly due to high altitude wave drag (Garcia and Boville, 1994). The weak damping rate inside the vortex leads to a departure from downward control which reduces the penetration of the circulation induced by wave drag in the mesosphere and upper stratosphere to the lower stratosphere. At the same time penetration of the circulation from lower latitudes increases. Even if the externally controlled circulation inside the polar vortex is weak, the weak radiative damping rate translates into a proportionally stronger temperature response compared to midlatitudes.

Deviation from downward control also occurs in the lower stratosphere, where the damping rate is weak, and is an issue for stratosphere-troposphere exchange. The maximum downward mass flux out of the lowermost stratosphere occurs during late spring while the peak downward flux into it (through the 100 hPa surface) occurs during winter (Appenzeller et al., 1996). Some of this lag is the result of the departure from steady state that occurs in regions with long damping timescales.

In Chapter 5 we extend existing work to include the effect of spatial and temporal variation in the radiative damping rate on the response of the system to idealized steady and periodic forcing.
Chapter Two

Sponge layer feedbacks in middle-atmosphere models

2.1 Introduction

Haynes et al. (1991) elucidated the nature of the zonally symmetric response of the middle atmosphere to a quasi-steady force $\mathcal{F}$ imposed in the extratropics. The mechanical and thermal dissipation of Rossby and gravity waves propagating upward from the troposphere produces such a force on the zonal mean state. Under the assumption that the diabatic heating† is relaxational and that the planetary boundary layer acts like relaxational friction near the ground, Haynes et al. demonstrated that the meridional circulation develops a confinement to regions beneath the imposed $\mathcal{F}$ in the steady state limit. This "downward control" limit can be reached through the formation of an equal and opposite force in the frictional boundary layer which reacts to the surface zonal wind produced by the meridional circulation. Streamlines of the steady mean meridional circulation extend between the surface and the level of the imposed force. At levels where no force is present temperatures are pulled away from radiative equilibrium by adiabatic heating.

A problem not explicitly considered by Haynes et al. (1991) is the zonally symmetric response to an imposed extratropical quasi-steady diabatic heating $Q$. Such a thermal driving can be produced by a change in the radiative equilibrium temperature due to changes in concentration of radiatively active constituents, for example. Arguments presented by Haynes et al. in the $\mathcal{F}$ case imply that for an imposed $Q$, in the steady state limit, there is only a modification of the temperature in the vicinity of the imposed heating and no mean meridional circulation. (This thought experiment was discussed by McIntyre (1992, §6). See also Holton et al. (1995, §3).)

† In this thesis, "heating" is understood to be positive or negative depending on the context.
In general, the forcing and diabatic heating that exist in realistic models of the middle atmosphere and in the real atmosphere have a complex dependence on the winds and temperature. Wave-induced forcing leads to changes in the wind structure which modify the propagation of the waves and subsequently the distribution of the forcing itself. Diabatic heating can interact with the forcing as well. For example, ozone depletion at high latitudes has an impact on the wind distribution which affects the propagation of waves and formation of the forcing. The relevance of the thought experiments described above is that they capture the expected response to local changes in forcing or diabatic heating. Such local changes have a clearly defined meaning only as long as associated changes in the rest of the forcing or diabatic heating in the system are second-order effects. From another viewpoint, it is assumed that some hypothetical imposed $F$ or $Q$ does not result in a leading-order change in forcing or diabatic heating, but instead the dominant response involves changes in the boundary-layer frictional drag and radiative heating. The results of Haynes et al. (1991) are based on this assumption.

The relaxational responses envisioned by Haynes et al. are highly plausible on physical grounds. However, many middle atmosphere models have other relaxational mechanical drag processes besides boundary-layer friction which do not have a physical justification. It is common for two-dimensional models to have Rayleigh drag throughout the model domain, and for three-dimensional models to have either a Rayleigh-drag or diffusive "sponge" layer in the upper portion of the model domain. This additional frictional drag is intended to simulate missing gravity-wave drag, in order to produce a reasonable climate, or to control artificial model behaviour associated with finite resolution (such as wave reflection downward from the upper layer or numerical stability). Being a body force, frictional drag will induce a meridional circulation that conforms to the downward control principle. However, gravity-wave drag (and Rossby-wave drag) is not relaxational and cannot react with whatever strength is required to oppose changes in the wind driven by an applied force or diabatic heating. In fact, tropospheric gravity-wave sources and the wind distribution determine the strength of gravity-wave drag. So, local changes in wind conditions will only redistribute the finite amount of drag available. (The effect on tropospheric flow is weak from forcing at the level of the sponge layer, e.g., so the feedback on the strength of the gravity-wave sources should be a second-order effect.) Hence, the "downward control" exerted by frictional drag can be regarded as
spurious with the potential to contaminate model results below the sponge layer. The possibility that middle atmosphere models with a relaxational drag could respond to an imposed local force or diabatic heating in an unphysical manner was noted by Haynes et al. (1991).

In this chapter we investigate the long-term response of a zonally symmetric middle atmosphere model with a Rayleigh-drag sponge layer (in addition to a frictional boundary layer) to an imposed extratropical local force \( \mathcal{F} \) or diabatic heating \( Q \). (The effect of a diffusive sponge layer is briefly considered in section 2.5.) Idealized numerical simulations are presented in section 2.2. A departure from the physical downward control response is apparent, including significant temperature changes above the level of the imposed force or diabatic heating associated with the downward control exerted by the sponge layer. In section 2.3, the Haynes et al. (1991) asymptotic analysis of the long-term response is extended to include a Rayleigh-drag layer near the top boundary of the domain. This analysis gives some insight into the parameter dependences of the problem. Section 2.4 deals with the case of a uniform Rayleigh drag which is relevant to many two dimensional models. It is found that even a weak drag coefficient can produce a noticeable departure from the downward controlled response. Section 2.5 summarizes the results and includes conclusions.

### 2.2 Idealized numerical solutions

The problem of the zonally symmetric, balanced response to an imposed (switch-on) extratropical local force or diabatic heating is considered here. The model includes a lower frictional boundary layer and an upper sponge layer, and the diabatic heating is approximated by Newtonian cooling towards a reference state. A linearizing assumption is made in that the perturbations in the forcing and diabatic heating, \( \mathcal{F} \) and \( Q \), are sufficiently small that the response they induce is decoupled from the evolution of the reference state. Such a simplification is justified since the system of equations that governs the extratropical zonally symmetric dynamics has a high degree of linearity. Haynes et al. (1991) have shown that the essential nature of the dynamical response in the extratropical middle atmosphere can be captured in such a framework. (Such an idealization cannot be made in the tropics which are characterized by nonlinearity: see Chapters 3 and 4.)
The model is based on the zonally averaged primitive equations linearized about a state of rest. The zonal flow is assumed to be in geostrophic and hydrostatic balance (see e.g. Andrews et al., 1987):

\[
\frac{\partial u}{\partial t} - 2\Omega \mu v = F - r(z)u, \tag{2.1}
\]

\[
\frac{\partial T}{\partial t} + Sw = Q - \alpha T, \tag{2.2}
\]

\[
2\Omega \mu \frac{\partial u}{\partial z} = -\frac{R}{aH} (1 - \mu^2)^{1/2} \frac{\partial T}{\partial \mu}, \tag{2.3}
\]

\[
\frac{\partial}{\partial \mu} \left( (1 - \mu^2)^{1/2} v \right) + \frac{a}{\rho_0} \frac{\partial}{\partial z} (\rho_0 w) = 0. \tag{2.4}
\]

In the above, \(u, v\) and \(w\) are the zonal, meridional, and vertical velocities; \(\rho_0(z)\) and \(T_0(z)\) are reference profiles of density and temperature; \(T\) is the departure of the temperature from the reference profile; \(S = dT_0/dz + \kappa T_0/H\) is a measure of the static stability, where \(\kappa = R/c_p\); \(R\) is the gas constant and \(c_p\) the specific heat at constant pressure; \(H\) is the (constant) density scale height; \(\Omega\) and \(\alpha\) are the rotation rate and radius of the Earth; \(z\) is a log-pressure vertical coordinate; and \(\mu = \sin \phi\) where \(\phi\) is latitude. The non-divergence condition (2.4) allows the introduction of a meridional mass streamfunction \(\psi\) defined by

\[
v = -\frac{1}{\rho_0(1 - \mu^2)^{1/2}} \frac{\partial \psi}{\partial z}, \quad w = \frac{1}{\rho_0 a} \frac{\partial \psi}{\partial \mu}. \tag{2.5}
\]

The zonal momentum equation (2.1) includes an imposed (switch-on) force per unit mass \(F\) together with a Rayleigh drag with a coefficient \(r(z)\) that is zero except in the lower frictional boundary layer and the upper sponge layer. Similarly, the temperature or thermodynamic equation (2.2) has an imposed (switch-on) heating \(Q\) together with Newtonian cooling on a timescale \(1/\alpha\). It is evident that making \(Q \neq 0\) in a certain region is equivalent to changing the radiative equilibrium temperature there. The meridional circulation \((v, w)\) in (2.1)–(2.5) is closely related to the transformed Eulerian mean circulation \((\bar{v}^*, \bar{w}^*)\) in the zonally averaged equations of motion (Andrews et al., 1987). In the steady-state limit, \((v, w)\) can also be identified with the diabatic circulation. The boundary condition on the mass streamfunction \(\psi\) is zero vertical velocity at the top and bottom of the domain and no horizontal velocity at \(\phi = 0^\circ\) and \(\phi = 90^\circ\). This is equivalent to setting \(\psi = 0\) at the boundaries. (Haynes and Shepherd (1989) demonstrated that
zero vertical velocity at the lower boundary in pressure (and log-pressure) coordinates amounts to an approximation; but its effects are evanescent in the middle atmosphere.

We solve (2.1)-(2.5) numerically over the Northern Hemisphere using finite differences on a height-latitude grid. Parameter values are $T_0 = 240 \text{ K}$, $\alpha = 1/(5 \text{ days})$, $H = 7 \text{ km}$, $N^2 = 10^{-4} \text{ s}^{-2}$, $\rho_0(0) = 1.225 \text{ kg/m}^3$, and $\Omega = 2\pi/\text{day}$. The model top is placed at $z = 80 \text{ km}$, and the imposed force and diabatic heating are Gaussian distributions in $z$ and $\phi$ centred at 40 km and 45°N, with half-widths of $\Delta z = 5 \text{ km}$ and $\Delta\phi = 10^\circ$. Their maximum strengths are respectively $-10 \text{ m/s/day}$ and $1 \text{ K/day}$. The vertical grid spacing is 0.4 km, while the horizontal grid spacing is 0.9°. The planetary boundary layer and sponge layer (when applied) are represented by a Rayleigh drag with $r(z)$ given by

$$r(z) = \begin{cases} 
1 & \text{for } 65 \text{ km} \leq z \leq 80 \text{ km} \\
\frac{2}{4 \text{ days}} & \text{for } 60 \text{ km} \leq z \leq 65 \text{ km} \\
0 & \text{for } 5 \text{ km} \leq z \leq 60 \text{ km} \\
\frac{1 + \cos(\pi z/5)}{6 \text{ days}} & \text{for } 0 \text{ km} \leq z \leq 5 \text{ km} 
\end{cases}$$

(2.6)

The discretized system is numerically integrated for one model year. Steady state is not reached by this stage but the system is close to it and the results are indicative of the long-term response.

We first treat the case of a localized switch-on (westward) force $\mathcal{F} < 0$, with $Q = 0$ (Fig. 2.1). The initial (transient) response is shown in column $a$; the sponge-free response after one year is shown in column $b$; and the response after one year in the presence of the sponge is shown in column $c$. The “downward controlled” response $b$ can be considered the true response. Thus, differences between $b$ and $c$, are due to an artificial sponge-layer feedback. It should be repeated that this distinction is based on the assumption that the response of true wave drag at the levels of the sponge layer is second-order. Rayleigh drag, in contrast, produces a first-order response to perturbations.

The initial response (Fig. 2.1a) is in agreement with the theory of Eliassen (1951). The force locally decelerates the zonal flow and also drives a poleward mean meridional circulation which returns equatorward, partly above but mainly below the forcing region (Fig. 2.1a(i)). This meridional circulation induces adiabatic heating which, for the simple
Fig. 2.1 Numerical solutions of (2.1)-(2.5) for an imposed negative Gaussian force $F < 0$ centred at 40 km and 45°N with a maximum strength of $-10$ m/s/day. The model lid is at 80 km and the sponge begins at 60 km, ramping up to a constant value of $r = 1/(2 \text{days})$ between 65 km and 80 km; $H = 7$ km and $\alpha = 1/(5 \text{days})$. Column a shows the instantaneous response; b the model response after one year with no sponge layer; and c the model response after one year with the sponge layer. In each case the meridional mass streamfunction $\psi$ is shown in row (i); the temperature perturbation $T$ in row (ii); and the zonal-wind perturbation $u$ in row (iii): for the instantaneous response a, (ii) and (iii) show the tendencies $\partial T/\partial t$ and $\partial u/\partial t$. Contour intervals are as follows: a(i) 0.5 kg/(ms); b,c(i) 2.0 kg/(ms); a(ii) 0.1 K/day; b,c(ii) 2.0 K; a(iii) 0.4 m/s/day; b,c(iii) 20 m/s.
force distribution used here, appears as a quadrupole pattern of temperature tendencies \( \partial T / \partial t \) (Fig. 2.1a(ii)). Simultaneously the Coriolis torque produces zonal-wind tendencies \( \partial u / \partial t \) outside the force region (Fig. 2.1a(iii)). The combined wind and temperature tendencies are in thermal-wind balance (2.3). The non-local nature of the response is pronounced: the streamfunction satisfies an elliptic equation and is consequently non-zero everywhere, but mostly localized in the vicinity of the forcing.

The transient evolution is suppressed by radiative damping. The growth in the \( \alpha T \) term in response to increasing temperature leads to a decrease in the \( \partial T / \partial t \) term. This can be seen by the fact that both have the same sign and cannot balance each other. Instead, the sum of the two terms balances \(-Sw\) (the mean meridional circulation can be regarded as prescribed, on these timescales). In the vicinity of the force the tendency term essentially shuts down on a timescale of order \( 1/\alpha \) resulting in a quasi-steady local balance between adiabatic and diabatic heating. After this stage the meridional circulation cells continue to "burrow" both upwards and downwards away from the forcing region. The mass flux in the upper cell attenuates while in the lower cell it increases, locally asymptoting to the value in the forcing region. The cells eventually meet the upper and lower boundaries resulting in the formation of primarily positive drag forces in the relaxational friction layers.

When there is only frictional drag at the lower boundary, which is the scenario considered by Haynes et al. (1991), the upper circulation cell dies away leaving only the lower cell (Fig. 2.1b(i)). The associated temperature change is then confined between the ground and the forcing region (Fig. 2.1b(ii)). The temperature has a dipole pattern which results from the ascending and descending branches of the meridional circulation that maintains it. The absence of vertical variation in the steady-state mass streamfunction corresponds to a vertical velocity that varies as the inverse of the density below the forcing level. The temperature mirrors the exponential decrease of the vertical velocity. The dipole temperature structure translates into a vertically elongated tripole pattern in the zonal wind through (2.3). The positive meridional temperature gradient below the forcing region translates into a negative zonal wind which has vertical shear in the circulation region and no shear above. The outer flanks of the zonal wind are positive

\[\dagger\] The saturation timescale may be considerably longer if the forcing region is "tall"; see Haynes et al. (1991).
and are associated with the negative meridional temperature at the outer edges of the circulation region (Fig. 2.1b(iii)). Another explanation for the regions of positive zonal wind, consistent with the one above, is acceleration by the meridional circulation through the Coriolis torque term in (2.1) which has a sign opposite to that of $F$. Since there is no circulation above the forcing region in steady state, the zonal wind there represents the cumulative impact, or history, of the circulation during the transient regime. This is also the case below the forcing region and above the friction layer, where there is no steady state meridional velocity.

However, in the presence of an additional relaxational friction layer, or sponge layer, there are meridional circulation cells both above and below the region of imposed force in the long-term response. In the model the upper cell is very weak since the sponge is at an altitude over three density scale heights higher than the imposed force (Fig. 2.1c(i)). The temperature, on the other hand, bears a greater impact and has a vertically elongated quadrupole pattern (Fig. 2.1c(ii)). The zonal wind is structured similarly to the previous case below the region of imposed forcing, but decreases with height above the forcing region (Fig. 2.1c(iii)).

It is evident that the temperature just below the sponge layer has a similar exponential decay with height as the temperature below the imposed force. The upper circulation cell is not maintained indefinitely by the force below, but rather by the primarily positive force induced in the sponge layer by the zonal wind associated with the imposed force. This sponge-drag force drives a predominantly negative circulation which exhibits downward control. The attenuation of the zonal wind near the base of the sponge layer is consistent with a superposition of a positive zonal wind, flanked by negative values, originating from the sponge-drag force.

The sponge-layer feedback produces a temperature change comparable in magnitude to that of the imposed force. In both cases there is a rapid attenuation of the induced temperature with height below the forcing region. In the model response there is essentially no effect of the sponge at and below the level of the imposed force (compare Figs. 2.1b(ii) and 2.1c(ii)). Nevertheless, all levels of the model atmosphere below the sponge are usually considered of physical interest (in fact, the position of the sponge is selected on this basis). Hence, the sponge-induced temperature and zonal-wind effects are quantitatively significant.
The next case under consideration is that of a localized switch-on (positive) diabatic heating $Q > 0$, with $F = 0$ (Fig. 2.2). As expected the initial response (Fig. 2.2a) agrees with the theory of Eliassen (1951): the heating locally warms the atmosphere (Fig. 2.2a(ii)) and drives a rising mean meridional circulation which closes through descending branches on either side of the heating region (Fig. 2.2a(i)). This meridional circulation produces a quadrupole pattern of zonal-wind tendencies $\partial u/\partial t$ (Fig. 2.2a(iii)) through Coriolis forces. The temperature tendencies $\partial T/\partial t$ (Fig. 2.2a(ii)) are modified by the circulation which keeps the zonal wind and temperature tendencies in thermal-wind balance.

Unlike the case with an imposed force, the circulation does not persist in the long term when the frictional drag is confined to the lower boundary (Fig. 2.2b). Whereas an imposed force is balanced by a meridional velocity in (2.1), an imposed diabatic heating can be balanced by the $\alpha T$ term in (2.2). In this case the system is effectively unforced and unable to sustain a circulation. It takes several radiative damping timescales $1/\alpha$ for the $\alpha T$ term to respond, during which a purely transient circulation can exist before eventually collapsing. Some circulation is still present in the model results even after one year (Fig. 2.2b(i)), which shows the asymptotic nature of the radiative adjustment process. However, after one year the temperature response is basically confined to the vicinity of the imposed diabatic heating (Fig. 2.2b(ii)) with a vertically elongated dipole zonal wind distribution above (Fig. 2.2b(iii)).

This long-term barotropic zonal wind response above the diabatic heating leads to a very different system behaviour with the inclusion of an upper sponge layer (Fig. 2.2c). A sponge-drag dipole forms which drives two meridional circulation cells that persist indefinitely (Fig. 2.2c(i)). The sponge-drag circulation is not confined to the latitudes of the imposed diabatic heating and does not tend towards such a confinement. This can be seen by the vertical alignment of the streamlines outside regions of drag, which is especially apparent if the transient circulation of Figure 2.2b(i) is removed from Figure 2.2c(i), and which indicates close proximity to steady state. The breadth of the sponge-drag circulation merely reflects the absence of latitudinal confinement of the sponge-drag forcing.

The ascending branch of the sponge-induced meridional circulation is primarily in the latitude band of the imposed diabatic heating. As dictated by the linearity of the governing equations, the temperature response (Fig. 2.2c(ii)) is that of the sponge-free
Fig. 2.2 As in Figure 2.1, except for an imposed positive diabatic heating with maximum strength of 1 K/day. In row (i) the meridional cells are ascending at 45N. Contour intervals are as follows: a(i) 0.1 kg/(ms); b,c(i) 0.02 kg/(ms); a(ii) 0.04 K/day; b,c(ii) 0.4 K; a(iii) 0.05 m/s/day; b,c(iii) 2.0 m/s.

case (Fig. 2.2b(ii)) superimposed with a sponge-induced cooling that decreases exponentially below the sponge layer. Most of the sponge-induced cooling is confined to the ascending branch of the sponge-drag circulation and has a maximum value comparable
to the maximum value of the warming below. However the maximum value is not as large as in the imposed force case where the sponge-induced circulation has greater latitudinal confinement and correspondingly stronger values of vertical velocity.

As seen from the experiments described above, a sponge layer can have a significant impact on the zonally symmetric response to an imposed extratropical local force or diabatic heating. The sponge layer reacts to the barotropic zonal wind induced by the imposed force or diabatic heating, resulting in the formation of a sponge drag which is not confined to the latitudes of the imposed force or diabatic heating (this effect is especially evident for the imposed diabatic heating case as it is not obscured by a sponge-free circulation). In accordance with downward control, the sponge drag drives a circulation extending to the surface. The temperature and zonal wind changes associated with the sponge-layer feedback are attenuated by density stratification and essentially confined to a layer of a few density scale heights at the base of the sponge. However, the magnitude of the sponge-induced temperature and zonal wind does not depend on the height separation of the sponge from the imposed force or diabatic heating in the long term. This model-based observation is confirmed in an analytical context in the next section.

The parameter dependences of this sponge-layer feedback will be investigated by means of an asymptotic analysis in the next section. Before proceeding, an additional idealized numerical solution is considered. It is motivated by model results and observational inferences (e.g. Holton, 1982; Garcia and Solomon, 1985; Shine, 1989) indicating that most of the middle atmosphere gravity-wave drag occurs in the upper mesosphere and mesopause regions. The temperature structure in the mesosphere depends on this drag (e.g. Andrews et al., 1987). Gravity-wave drag may even significantly influence temperatures down to the middle stratosphere in polar regions (Haynes et al., 1991; Garcia and Boville, 1994). In most middle atmosphere general circulation models, however, the upper boundary is no higher than the mesopause level. Therefore any sponge layer acts in the upper mesosphere and overlaps the bulk of the force produced by gravity-wave drag parameterization schemes in such models. To investigate this scenario we examine the response to a force imposed within the sponge layer itself.

The results are shown in Figure 2.3, which corresponds to Figure 2.1 but with the imposed force centred at 73 km, inside the sponge layer. As before, the instantaneous
response is shown in panel a, the sponge-free long-term response in panel b, and the long-term response in the presence of the sponge in panel c. Comparing $b$ and $c$, a significant distortion of the long-term response is apparent. The mass circulation (Figs. 2.3b, c(i)) is reduced by about a factor of five in strength and is not confined to the latitudes of the
imposed force. The temperature response (Figs. 2.3b, c(ii)) is also weaker, by a factor of five to ten in its maximum values, and meridionally broadened. In addition, it is much shallower and basically confined inside the sponge layer. These differences imply that the response to the imposed force is almost completely nullified by the sponge, except for a spurious mass circulation which induces weak temperature changes over the whole hemisphere.

2.3 Asymptotic solutions

In this section we use the asymptotic approach of Haynes et al. (1991) to examine the long-time evolution of (2.1)-(2.5). For consistency with Haynes et al., we set $z = 0$ at the level of the imposed force or diabatic heating. The upper boundary is then at $z_t > 0$ and the lower boundary is at $z_0 < 0$. The sponge and the planetary boundary layer are represented with a Rayleigh-drag coefficient given by $r(z) = \tilde{r} \mathcal{H}(z - z_s) + \gamma \delta((z - z_0)/H)$, where $\delta(\cdot)$ is the Dirac delta-function and $\mathcal{H}(\cdot)$ is the Heaviside step function. The height of the sponge base, $z_s$, satisfies $0 < z_s < z_t$. The sponge is distributed in a deep uniform layer with a damping rate of $\tilde{r}$ and the boundary layer is very shallow with a damping rate of $\gamma$. The fields are decomposed into a Hough-mode representation (see Appendix A), in which case (A.1)-(A.4) may be combined into the single equation

$$
\left( \frac{\partial}{\partial t} + \alpha \right) \frac{1}{\rho_0} \frac{\partial}{\partial z} \left( \rho_0 \frac{\partial u_n}{\partial z} \right) + \frac{\epsilon_n N^2}{4 \Omega^2 a^2} \left( \frac{\partial}{\partial t} + \tilde{r} \right) u_n = \frac{\epsilon_n N^2 F_n}{4 \Omega^2 a^2} - \frac{\epsilon_n R}{2 \Omega \alpha \rho_0 H} \frac{\partial}{\partial z} (\rho_0 Q_n),
$$

where $N^2 \equiv RS/H$ is taken to be constant.

The response to an imposed force $F_n(z, t) = f_n \mathcal{H}(t) \delta(z/H)$ with $Q_n = 0$ is considered first. As described by Haynes et al., the intermediate-time solutions are characterized by an upward- and a downward- propagating component (see their (3.21)). The upward-propagating component reaches the sponge layer after a time $t \sim b_n z_s / \alpha H$ and induces a "sponge force". The sponge force drives a downward-propagating adjustment. The sponge-induced modification of the zonal wind, $u_n^s$, takes the form

$$
u_n^s \sim \eta \frac{f_n b_n}{2 \alpha} e^{(z-z_s)/H} \text{erfc}\left\{ -\left( t + \frac{b_n z'}{\alpha H} \right) \left[ \frac{H \alpha^2}{4 b_n (b_n + 1)|z'|} \right]^{1/2} \right\},
$$

(2.8)
where $z' \equiv z - 2z_s$, $b_n \equiv -\varepsilon_n N^2 H^2 / 4\Omega^2 a^2$ is an inverse rotational Froude number, and

$$\eta \equiv \frac{1 - \lambda}{1 + \lambda} \left( \frac{1 - \exp[\lambda(z_s - z_t)/H]}{1 - (\frac{1 - \lambda}{1 + \lambda})^2 \exp[\lambda(z_s - z_t)/H]} \right),$$

with $\lambda \equiv (1 + (4b_n \tau / \alpha))^{1/2}$. For individual modes the form of the sponge-force induced adjustment differs from that of the imposed force $f_n$ only by the factor $\eta$. However, $\eta$ depends on $n$ so that the latitudinal distribution of the sponge force and its adjustment will be different from the imposed force when there is more than one mode in the spectrum.

The dependence of $\eta$ on $n$ reflects a dispersion of the total fields inside the sponge layer. For example, the sponge force is weaker and distributed over a broader latitude span compared to the imposed force. The absolute value of the factor $\eta$ can be regarded as the efficiency with which individual modes of the Hough spectrum of $F$ induce a zonal wind response from the sponge layer. As indicated by Figure 2.4, modes with larger $n$ are more efficient. The relaxational sponge force cannot be larger than the imposed force so $|\eta| \leq 1$ for all $n$. The increase in efficiency with $n$ is linked to the change in structure of the induced sponge force. For Hough modes with higher $n$, the sponge force amplitude increases relative to that at the level of the forcing. In addition, the vertical profile of the sponge force, $\tilde{r}u_n$, becomes confined closer and closer to the base of the sponge layer. This shallower and stronger sponge force induces a zonal wind response, $u'_n$, which resembles more and more $u_n$ below the forcing level. Increasing the sponge damping rate, $\tilde{r}$, cannot change the dependence of $\eta$ on $n$ but increases $\lambda$, which brings $|\eta|$ closer to unity for a given mode by increasing the amplitude of the sponge force.

When the original downward-propagating component reaches the planetary boundary layer, an upward-propagating adjustment $u^b_n$ is induced which is given by (5.4) of Haynes et al. In the presence of both upper and lower boundaries, the long-term response then consists of an infinite series of rapidly diminishing adjustments propagating away from one boundary and toward the other. In the steady-state limit

$$u_n \sim \frac{f_n b_n}{\alpha(1 - \eta \xi e^{-(z_s - z_0)/H})} \begin{cases} 1 + \eta \xi e^{(z - z_t)/H} + \xi e^{z_0/H} + \eta \xi e^{(z - z_s + z_0)/H}, & 0 < z < z_s \\ e^{z/H} + \eta \xi e^{(z - z_t)/H} + \xi e^{z_0/H} + \eta \xi e^{-(z_s - z_0)/H}, & z_0 < z < 0 \end{cases}$$

(2.10)

where $\xi \equiv (\alpha / b_n \gamma) - 1$. The part of $u_n$ in (2.10) associated directly with the imposed force does not contain $\eta$ or $\xi$. Terms that involve $\eta$ but not $\xi$ represent the sponge response to
2.3 Asymptotic solutions

The dependence of \( \eta \) on \( n \) and \( \tilde{r}/\alpha \) is given in Figure 2.4. For reasonable values of the model parameters, \( \eta \approx -0.5 \). In fact it can be seen from (2.9) that \( \eta \sim -1 \) in the limit \( b_n \tilde{r}/\alpha \gg 1 \), a limit is relevant to middle atmosphere models. In short, \( \eta \) can be expected to be an order unity, negative number.
As a result of the sponge-layer feedback, the zonal wind induced by the imposed force is attenuated nearly to zero at the base of the sponge, with the effect exponentially decreasing as one moves down. This confirms the results seen in Figures 2.1b,c(iii). The corresponding effect on the temperature can be inferred from (A.3). In the steady-state limit, neglecting the terms involving ξ,

\[
T_n \sim \frac{f_n N^2 H^2}{2\Omega a R_0} \begin{cases} 
\eta e^{(z-z_s)/H}, & 0 < z < z_s \\
\eta e^{z/z_s} + \eta e^{(z-z_s)/H}, & z_0 < z < 0
\end{cases}
\]  

(2.12)

It is evident that the magnitude of the sponge-induced temperature change just below the sponge is comparable to the temperature change just below the imposed force and both decrease exponentially in \( z \) as one moves down. This supports the results seen in Figures 2.1b,c(iii).

The long-term effect on the meridional mass circulation can be inferred from the effect on \( T_n \) using the fact that in the steady-state limit (A.2) and (A.5) imply \( \psi_n = -(\rho_0 a \alpha / S) T_n \). The \( e^{z/H} \) factors in \( T_n \) seen in (2.12) cancel from the expression for \( \psi_n \), and there are meridional cells of opposite orientation above and below the forcing. The ratio in strength of the two cells is given by (ignoring the small corrections involving ξ)

\[
\frac{\text{mass flux in upper cell}}{\text{mass flux in lower cell}} \sim \frac{\eta e^{-z_s/H}}{1 + \eta e^{-z_s/H}}.
\]

(2.13)

It is clear that the effect of the sponge on the mass circulation is small when the forcing is separated by several scale heights. This agrees with the results seen in Figures 2.1b,c(i). Nevertheless, the weak upper cell has a significant effect on the temperature structure within one or two scale heights of the sponge.

The case of a thin frictional layer, namely the "surface" drag, imposed above the forcing level was considered briefly by Haynes et al. (1991, p. 668a). They found a reduction in the strength of the lower cell of order \( e^{-z_s/H} \) for \( b_n \gamma / \alpha \gg 1 \), which is consistent with (2.13).

We now consider the case of an imposed diabatic heating \( Q_n(z,t) = q_n \mathcal{H}(t) \delta(z/H) \) with \( F_n = 0 \), following the same analysis as above. The sponge responds similarly to the switch-on forcing case. For intermediate times, before feedback from the planetary boundary layer, the sponge-induced zonal wind change takes the form

\[
u_n^* \sim \eta \frac{q_n b_n \Omega a R}{N^2 H^2 \alpha} e^{(z-z_s)/H} \text{erfc}\left\{-\left(t + \frac{b_n z'}{\alpha H}\right) \left[\frac{H \alpha^2}{4 b_n (b_n + 1) |z'|}\right]^{1/2}\right\}
\]

(2.14)
2.4 The case of a uniform Rayleigh drag

and in the steady-state limit

\[
\begin{align*}
  u_n &\sim \frac{q_n b_n 2 \Omega R}{N^2 H^2 \alpha} \left\{ \begin{array}{ll}
    1 + \eta e^{(z-z_s)/H}, & 0 < z < z_s \\
    \eta e^{(z-z_s)/H} + \eta \zeta e^{-(z_z-\zeta_0)/H}, & z_0 < z < 0
  \end{array} \right. \\
\end{align*}
\]  

(2.15)

As above, the terms involving \( \zeta \), which reflect the feedback from the planetary boundary layer, are negligible for a diabatic heating imposed many scale heights away from the lower boundary, and the long-term response is given to good approximation by

\[
\begin{align*}
  u_n &\sim \frac{q_n b_n 2 \Omega R}{N^2 H^2 \alpha} \left\{ \begin{array}{ll}
    1 + \eta e^{(z-z_s)/H}, & 0 < z < z_s \\
    \eta e^{(z-z_s)/H}, & z_0 < z < 0
  \end{array} \right. \\
\end{align*}
\]  

(2.16)

This is essentially the same result seen in Figures 2.2b,c(iii). The corresponding steady-state temperature change is

\[
T_n \sim \frac{q_n \eta e^{(z-z_s)/H}}{\alpha} + \frac{q_n}{\alpha} \delta(z/H).
\]  

(2.17)

As in the case of the imposed force, the sponge-induced temperature change is of opposite sign to the temperature change at the level of the imposed heating, and achieves a comparable magnitude just below the sponge (with allowance for integration over the \( \delta \)-function). This confirms the large sponge-induced cooling above the level of imposed heating seen in Figure 2.2c(ii). The meridional mass circulation is given by \( \psi_n = -(\rho_0 \alpha \zeta / S) T_n \) and is entirely due to the sponge. The circulation extends through the depth of the atmosphere with a vertically uniform mass flux (Fig. 2.2c(i)). The dependence of \( \eta \) on \( n \) implies that the sponge-induced response has a different latitudinal structure than the imposed heating.

2.4 The case of a uniform Rayleigh drag

In Sections 2.2–2.3 we have considered the case of a vertically confined upper sponge layer, which is the situation of most relevance to three-dimensional middle atmosphere models. However, the case of a background Rayleigh drag throughout the domain, as is commonly used in two-dimensional middle atmosphere models (e.g. Garcia and Solomon, 1985; Garcia et al., 1992), is also of interest. Two-dimensional models have a smaller computational cost than three-dimensional models and consequently are still widely used for
chemical climate change and impact assessment studies. Garcia (1987) has considered the impact of Rayleigh drag on the long-term response to an imposed force or diabatic heating. Haynes et al. (1991) have suggested that downward control could be underestimated in the presence of a Rayleigh drag.

Figure 2.5 shows the result of an idealized numerical calculation corresponding to Figure 2.1c (namely an imposed local force centred at 40 km). In this case the sponge-layer drag is replaced by a weak uniform background Rayleigh drag imposed throughout the model domain, with \( r = 1/(100 \text{ days}) \). (This is the value used by Garcia et al. (1992).) The instantaneous response, and the long-term response in the absence of Rayleigh drag (with only the planetary boundary layer), are given as before by Figures 2.1a, b respectively. The sponge-free, "downward controlled" response (Fig. 2.1b) is treated here as the true response, as in Section 2.2. Since ubiquitous relaxational momentum damping is not a characteristic of the middle atmosphere, differences between Fig. 2.1b and Fig. 2.5 can be interpreted as a spurious Rayleigh-drag feedback. In the real atmosphere, drag processes will not act relaxationally. Gravity-wave drag is a non-local momentum transfer in the vertical, while small-scale eddy diffusion will also primarily act in the vertical. Neither introduces a net drag force in the vertical, as does Rayleigh drag or horizontal diffusion. The idealized experiment considered here is based on the same assumption that underlies the analysis of Haynes et al. (1991), namely that a localized imposed forcing of
2.4 The case of a uniform Rayleigh drag

The case of a uniform Rayleigh drag

Fig. 2.6 As in Figure 2.2c (case of imposed diabatic heating), except with the sponge layer replaced by a uniform Rayleigh drag throughout the depth of the atmosphere with \( r = 1/(100\ \text{days}) \). Contour intervals as in Figure 2.2c.

...the middle atmosphere does not induce leading-order net changes in the background wave drag field.

Comparing Figure 2.5 with Figure 2.1b, it is evident that the meridional mass circulation departs from a downward controlled regime. It is spread out meridionally, particularly towards lower latitudes. The total mass transport is smaller on account of the Rayleigh drag. There is also a distortion in the temperature response which is non-zero above the level of the imposed force. The zonal-wind response is reduced by about a factor of two.

Figure 2.6 shows the same fields as in Figure 2.5, but for the case of an imposed local diabatic heating as in Figure 2.2. Comparing Figure 2.6 with Figure 2.2b, a departure from downward controlled regime is evident. A meridional mass circulation extends over most of the model atmosphere. The temperature response, while not very different in terms of the maximum value, is spread out latitudinally leading to a zonal-wind response weaker in magnitude by about a factor of two.

Some quantitative understanding of these effects can be obtained by means of an asymptotic analysis, as in Section 2.3. We assume \( r < \alpha \), which is appropriate for two-dimensional middle atmosphere models. In this case it is convenient to ignore the boundaries altogether. The approach of Haynes et al. can be used but it fails as \( r \rightarrow \alpha \). (The propagating character of the solutions disappears when \( r = \alpha \), in which case the time dependence is the same for all \( z \), namely \( 1 - e^{-\alpha t} \).)
In the case of the switch-on force \( F_n(z, t) = f_n \mathcal{H}(t) \delta(z/H) \) with \( Q_n = 0 \), after a sufficiently long time the response is given by

\[
\begin{align*}
  u_n \sim & \frac{f_n b_n}{2\alpha} \begin{cases} 
    \frac{1}{\lambda} e^{(1-\lambda)z/2H} \text{erfc}\left(-\left(t - \frac{\tilde{b}_n z}{H\alpha\lambda}\right) \left[\frac{H\alpha^2\lambda^3}{4b_n(b_n + \lambda^2)z}\right]^{1/2}\right) & z > 0 \\
    \frac{1}{\lambda} e^{(1+\lambda)z/2H} \text{erfc}\left(-\left(t + \frac{\tilde{b}_n z}{H\alpha\lambda}\right) \left[\frac{H\alpha^2\lambda^3}{4b_n(b_n + \lambda^2)z}\right]^{1/2}\right) & z < 0
  \end{cases},
\end{align*}
\]

with \( \lambda \equiv (1 + (4\tilde{b}_n r/\alpha))^{1/2} > 1 \) as before and where \( \tilde{b}_n \equiv b_n(\alpha - r)/\alpha \). This has the steady-state limit

\[
\begin{align*}
  u_n \sim & \frac{f_n b_n}{\alpha} \begin{cases} 
    \frac{1}{\lambda} e^{(1-\lambda)z/2H}, & z > 0 \\
    \frac{1}{\lambda} e^{(1+\lambda)z/2H}, & z < 0
  \end{cases},
\end{align*}
\]

with associated temperature change

\[
\begin{align*}
  T_n \sim & \frac{f_n N^n H^2}{2\Omega a R} \begin{cases} 
    \frac{1 - \lambda}{2\lambda} e^{(1-\lambda)z/2H}, & z > 0 \\
    \frac{1 + \lambda}{2\lambda} e^{(1+\lambda)z/2H}, & z < 0
  \end{cases}.
\end{align*}
\]

Without Rayleigh drag, \( \lambda = 1 \); with Rayleigh drag, \( \lambda > 1 \). Without Rayleigh drag, (2.20) implies that the temperature change in steady state is proportional to \( e^{z/H} \) below the force and is zero above. With Rayleigh drag, there is a more rapid decay below the force, together with a temperature change above the force of opposite sign, decaying more slowly with \( z \) than the temperature change below (cf Garcia (1987), Figure 2.1). This explains the results seen in Figures 2.1b(ii), 5(ii). As before, the latitudinal profile of the Rayleigh-drag effect will be different from that of the imposed force, because \( \lambda \) depends on \( n \). (When \( \lambda = 1 \), in contrast, the steady-state response is latitudinally localized: this is part of downward control.)

In the case of the switch-on diabatic heating \( Q_n(z, t) = q_n \mathcal{H}(t) \delta(z/H) \) with \( F_n = 0 \), after a sufficiently long time the response is given by

\[
\begin{align*}
  u_n \sim & \frac{q_n b_n \Omega z}{2N^n H^2 \alpha} \begin{cases} 
    \frac{1 + \lambda}{\lambda} e^{(1-\lambda)z/2H} \text{erfc}\left(-\left(t - \frac{\tilde{b}_n z}{H\alpha\lambda}\right) \left[\frac{H\alpha^2\lambda^3}{4b_n(b_n + \lambda^2)z}\right]^{1/2}\right) & z > 0 \\
    \frac{1 - \lambda}{\lambda} e^{(1+\lambda)z/2H} \text{erfc}\left(-\left(t + \frac{\tilde{b}_n z}{H\alpha\lambda}\right) \left[\frac{H\alpha^2\lambda^3}{4b_n(b_n + \lambda^2)z}\right]^{1/2}\right) & z < 0
  \end{cases}.
\end{align*}
\]
This has the steady-state limit

\[ u_n \sim \frac{q_n b_n 2\Omega a R}{N^2 H \alpha} \begin{cases} \frac{1 + \lambda}{2\lambda} e^{(1-\lambda)z/2H}, & z > 0 \\ \frac{1 - \lambda}{2\lambda} e^{(1+\lambda)z/2H}, & z < 0 \end{cases}, \tag{2.22} \]

with associated temperature change

\[ T_n \sim \frac{q_n}{\alpha} \begin{cases} \frac{1 - \lambda^2}{4\lambda} e^{(1-\lambda)z/2H}, & z > 0 \\ \frac{1 - \lambda^2}{4\lambda} \delta(z/H), & z = 0 \\ \frac{1 - \lambda^2}{4\lambda} e^{(1+\lambda)z/2H}, & z < 0 \end{cases}. \tag{2.23} \]

Without Rayleigh drag, (2.23) implies that the temperature change in steady state is confined to the level of the imposed diabatic heating. With Rayleigh drag, the \( \delta \)-function temperature change at \( z = 0 \) is accompanied by an oppositely signed temperature change extending above and below. The drag-induced temperature change decays more slowly with \( z \) above than it does below. This explains the results seen in Figures 2.2b(ii),6(ii).

2.5 Conclusion

A feature of middle atmosphere models is the use of a sponge layer in the upper portion of the model domain. A major function of the sponge is to absorb vertically propagating waves in order to prevent spurious reflections off the model lid. In this respect there is no spurious feedback on the dynamics below. However, the sponge is also employed as a crude parameterization for missing gravity-wave drag. A reasonable climate is obtained by tuning the sponge drag coefficient so that the total sponge drag is roughly comparable to the missing gravity-wave drag. However, the sensitivity of this drag to model perturbations cannot be expected to be realistic. True gravity-wave drag is determined by the strength of the tropospheric gravity-wave sources, with refraction by the middle atmosphere wind profiles only determining where the drag is deposited and not its total strength. In contrast, the relaxational nature of the sponge allows it to couple to the dynamics at lower levels in an artificial manner. The sponge is an unlimited supply of damping momentum which reacts to the zonal wind changes associated with the perturbations in forcing or diabatic heating.
The initial observation by Haynes et al. (1991) on the possible effect of a sponge layer has been investigated in detail in this chapter by considering the problem of the long-term zonally symmetric response to an imposed extratropical quasi-steady local force or diabatic heating. The analysis of Haynes et al. has been extended by including an upper sponge layer in addition to the lower frictional boundary layer. In the case of an imposed force it was found that the "downward controlled" response is modified by the formation of a mean meridional circulation cell extending below the sponge layer which acts to divert a fraction of the meridional mass flux upwards. For realistic parameter values, this fraction is approximately equal to \( \exp(-\Delta z/H) \) where \( \Delta z \) is the distance between the forcing region and the sponge layer, and \( H \) is the density scale height. The effect of the sponge on the temperature field is much larger than on the mass circulation. The magnitude of the temperature changes near the base of the sponge layer are comparable to those found below the forcing region. In the case of an imposed diabatic heating the sponge induces a meridional circulation that is hemispheric in extent and persists in the long term. The temperature changes produced by this circulation below the sponge layer are of opposite sign, and comparable in magnitude, to those at the heating region.

The idealized numerical simulations presented in Figures 2.1 and 2.2 demonstrate the effects described above. These effects have also been explained by an asymptotic analysis of the near steady-state behavior (Section 2.3). A departure from the downward controlled response, which is particularly evident in Figure 2.2, is the fact that the sponge response is not as latitudinally confined as is the imposed force or diabatic heating. This feature is also accounted for by the asymptotics which show that the sponge-layer feedback depends on the spatial scale of the imposed force or diabatic heating.

The sponge-layer feedback is limited in its effect on temperatures and winds sufficiently far below it due to exponential attenuation from density stratification. Consequently, the impact of the sponge can be controlled, at a given altitude and assuming the perturbations in forcing and diabatic heating are below the sponge, by moving the sponge (and the upper boundary of the model) higher. For example, the effect on temperatures two scale heights below the sponge is roughly at the 10% level (i.e. \( e^{-2} \approx 0.1 \)).

If an imposed force is located within the sponge layer, as can be the case for parameterized mesospheric gravity-wave drag, then the response is essentially absorbed locally (Fig. 2.3). There is very little zonal wind or temperature change, mostly within
the sponge layer, and a very weak mass circulation that spans the model hemisphere. So, a parameterized force acting within a model sponge does not exert the appropriate influence on the state of the system below.

A case not considered in the previous sections is the effect of a diffusive sponge layer. A diffusive rather than a Rayleigh-drag sponge layer is implemented in some middle atmosphere models. The primary difference between the two is that the diffusive sponge layer is constrained to produce no net drag on a given level over the sphere. The effect of this constraint on the response is investigated by repeating the experiment shown in Figure 2.1 (the imposed force case) but with the Rayleigh sponge drag $-ru$ replaced by a horizontal diffusion $\nu \nabla^2 u$. The diffusion coefficient $\nu$ was ramped up between $z = 60$ km and $z = 65$ km to a value of $5 \times 10^5$ m$^2$/s. (This corresponds approximately to the diffusion used by Hack et al. (1994).) The precise vertical dependence of $\nu$ is the same as in (2.6). Comparing the response shown in Figure 2.7 to that shown in Figure 2.1c, it is evident that the sponge-layer feedback is qualitatively the same in the two cases.

The presence of sponge-layer feedbacks in comprehensive middle atmosphere models has not been addressed directly in this chapter. Experiments with a GCM were performed by Shepherd et al. (1996), who demonstrated that the zonally symmetric model captures the basic sponge layer effect on the dynamics.
The sponge-layer feedback which has been the focus of this chapter is a consequence of the highly nonlocal character of the zonally symmetric balance dynamics. A perturbation in wave drag (or diabatic heating; see Section 2.3) induces a zonal wind response that propagates upward, as shown by Haynes et al. (1991). The zonal-wind change eventually encounters the sponge layer leading to the formation of a sponge drag that drives a downward-controlled circulation through the depth of the atmosphere. The sponge-drag circulation affects the temperature over a density scale height below the sponge layer to the same extent as the force-driven circulation over a scale height below the force maximum. A similar situation develops in response to a diabatic heating perturbation. This is also indicated by a simple scaling analysis based on the assumption (see Section 7 of Haynes et al., 1991) that in the transient stage \( w \) is roughly independent of height above the level of the imposed force or diabatic heating. However, different latitudinal profiles of the upward propagating zonal wind produce different sponge-drag profiles and subsequently variations in the temperature response.

A secondary focus of this chapter is on the effect of uniform background Rayleigh drag which is commonly employed in two-dimensional models. The drag coefficient used in the uniform case is typically much smaller than that of the sponge drag. Nevertheless, uniform drag acts in the regions of interest and there is the possibility of a spurious feedback in such models, as noted by Haynes et al. (1991). This problem has been studied in Section 2.4, where it is shown that even a weak Rayleigh drag \( (r = 1/(100 \text{ days}) \) compared with a Newtonian cooling coefficient of \( \alpha = 1/(5 \text{ days}) \) leads to a distortion of the response to an imposed extratropical local force or diabatic heating. The effects can be clearly seen by comparing Figures 2.5 and 2.6 with Figures 2.1b and 2.2b, respectively. These results are consistent with the earlier work of Garcia (1987).

The results in this chapter are predicated on the assumption of extratropical quasi-steady dynamics of the middle atmosphere. In the tropics, the downward control arguments fail to apply (Haynes et al., 1991; Holton et al., 1995). An imposed force centered on the equator is not effective at driving a meridional circulation and instead produces local zonal wind acceleration. There is a weaker zonal wind response above an imposed diabatic heating as well. As a result, the coupling with the sponge-layer is diminished. In addition, there exists in the tropics a nonlinear, angular-momentum-conserving meridional overturning in response to radiative driving (Dunkerton, 1989). The way in which
this nonlinear tropical regime connects to the extratropical downward controlled regime is considered in Chapters 3 and 4.
Chapter Three

The nonlinear Hadley circulation and inertial instability

3.1 Introduction

During solstice seasons there are strong latitudinal gradients of solar heating in the middle atmosphere tropics, particularly in the upper stratosphere and lower mesosphere. If gradient wind balance is assumed to hold in the tropics then the zonal wind corresponding to the radiative equilibrium temperature is singular at the equator where the Coriolis parameter vanishes. In the absence of mechanical forcing by waves, conservation of absolute angular momentum implies that the temperature cannot attain radiative equilibrium in the tropics even in steady state, since there is not enough momentum available in the system. Wave momentum deposition tends to be easterly in the tropics, with some exceptions such as the westerly phases of the SAO and QBO. However, there is not enough westerly momentum to overcome the formation of a thermally driven circulation, as pointed out by Dunkerton (1989). Dunkerton referred to this meridional circulation as the nonlinear Hadley circulation (NLHC). It is remarkable in that it persists in steady state even with relaxational radiative damping.

There is no a priori reason for gradient-wind balance to hold in the tropics and the singularity problem may not arise at all. This issue has not been explored, but results from primitive equations models (e.g. GCMs), which are not based on the gradient-wind balance assumption, indicate that the temperature has no latitudinal gradient at the equator in the course of solstice seasons and that gradient-wind balance is a good approximation in the stratosphere (see section 3.3.4). The impact of approximate gradient-wind balance on the strength of the NLHC is considered in this chapter.

Evidence for the existence of a thermally driven circulation also comes from the tropical upwelling profile. The observed upwelling maximum tends to occur on the summer
side of the equator. This behaviour is not consistent with the wave drag distribution inferred from observations, which minimizes in the equatorial stratosphere except for the upper layers (e.g. Rosenlof, 1995, Fig. 8). In the absence of other processes, the circulation produced by this wave drag should have an upwelling pattern with a strong maximum in the winter hemisphere and a weak maximum in the summer hemisphere, as is evident from the results of Plumb and Eluszkiewicz (1999). As noted by them, an additional tropical thermal forcing is necessary to explain the upwelling profile (compare their Fig. 14(b) and Fig. 14(c)). The process most likely responsible for the upwelling profile is the tropical diabatic circulation described by Dunkerton (1989).

Dunkerton (1989, 1991) also dealt with the influence of extratropical wave drag on the NLHC. It was concluded that remote mechanical forcing can intensify and increase the latitudinal extent of the NLHC. However, it was not made clear whether it is the NLHC component of the total diabatic circulation that is enhanced or the tropical branch of the wave-driven circulation. The latter was alluded to in Dunkerton (1991) where a "sideways control" behaviour was noted in streamlines produced by subtropical wave drag adjacent to a region of nearly horizontal absolute angular momentum isopleths such as forms in the stratopause region. It was also concluded in Dunkerton (1991) that extratropical wave drag contributes significantly to the formation of this region of severely deformed $\overline{m}$ isopleths. Here we look at the effect of the wave-driven circulation on the NLHC and leave consideration of the modification of the former in the tropics for Chapter 4.

Another gradient-wind-balance-related effect of tropical solar heating gradients is the movement of the absolute angular momentum maximum off the equator into the winter hemisphere. This leads to inertial instability which erodes absolute angular momentum equatorward, in the process homogenizing it as well as temperature in a band on the winter side of the equator. Inertial instability acts on the absolute angular momentum distribution much like the advection produced by the NLHC, except on a much shorter timescale. This process was omitted in Dunkerton (1989, 1991) on account of the parabolic approximation which was argued to become more accurate in the steady state limit. The parabolic approximation, however, only accounts for the effect of inertial adjustment though the coefficients of the streamfunction equation instead of the $\overline{u}$ and $\overline{T}$ distributions which affect the forcing of the streamfunction. It will be shown below that for the upper stratosphere, inertial adjustment has a significant direct impact on the
temperature and zonal wind structure and consequently on the diabatic heating driving the NLHC.

A description of the numerical model used is given in section 3.2. This model differs from the one used by Dunkerton in that no attempt is made to close off the NLHC in the mesosphere (with either a rigid lid at the stratopause or a sponge layer in the mesosphere) and the effect of inertial instability is explicitly parameterized. Section 3.3 reviews the NLHC in light of the relaxed model assumptions. The effect of the vertical structure of the radiative equilibrium temperature is considered. The validity of gradient-wind balance in the tropics is tested using a spectral primitive equations model. The influence of the wave-driven circulation on the NLHC is revisited in section 3.4. Evidence for the NLHC in middle atmosphere GCMs is the focus of section 3.5.

3.2 Model

The results of this chapter are for the most part based on numerical solutions of the transformed Eulerian mean (TEM) form of the governing equations (e.g. Andrews et al., 1987)

\[ \bar{u}_t + \bar{v}^* \left( \frac{\bar{u}_\phi}{a} - f - \bar{u} \tan \phi \right) + \bar{w}^* \bar{u}_z = \mathcal{F}, \]  

\( (f + 2 \bar{u} \frac{\tan \phi}{a}) \bar{u}_z = - \frac{R}{aH} \bar{T}_\phi, \)  

\[ \frac{1}{a \cos \phi} (\bar{v}^* \cos \phi)_\phi + \frac{1}{\rho} (\rho \bar{w}^*)_z = 0, \]

where \( a \) is the planetary radius, \( f = 2 \Omega \sin \phi \) is the Coriolis parameter, \( \rho = \rho_o \exp(-z/H) \) is the reference density profile with \( H \) the constant density scale height, \((\bar{v}^*, \bar{w}^*)\) is the residual circulation, and \( S \equiv \bar{T}_z + \frac{1}{H} \bar{T} \) is the static stability parameter. The Eliassen-Palm flux divergence representing forcing by wave drag is denoted by \( \mathcal{F} \), while \( Q \) denotes the diabatic heating. An eddy heat flux term has been neglected on the rhs of (3.2); this term is conventionally assumed to be small, an assumption that will be checked \textit{a posteriori} in experiments using a three-dimensional primitive equations model. With this
assumption \((\vec{v}', \vec{w}')\) is equivalent to the transient diabatic circulation. The diabatic heating can be represented by a Newtonian cooling to some time-dependent radiative equilibrium temperature, or by a more comprehensive radiative transfer scheme. Equation (3.3) represents thermal-wind balance (a combination of gradient-wind and hydrostatic balance), while (3.4) represents mass continuity. The surface friction scheme is based on the stress condition of Held and Hou (1980) with a drag coefficient of \(C = 0.005 \, s^{-1}\).

Mass continuity allows a streamfunction representation for the residual circulation. Defining \(\psi_z = -\rho \cos \phi \vec{v}'\) and \(\psi_\phi = a \rho \cos \phi \vec{w}'\) and using thermal wind balance, (3.1)–(3.4) can be combined into a single instantaneous equation (see Appendix B for details):

\[
\mathcal{L}[\psi] \equiv C_{\phi\phi} \psi_\phi + C_{\phi z} \psi_{\phi z} + C_{zz} \psi_{zz} + C_{\phi} \psi_\phi + C_z \psi_z = D, \tag{3.5}
\]

where the coefficients of the \(\psi\) terms are functionals of \(\vec{u}\) and \(\overline{T}\), and \(D\) is a functional of \(\vec{u}\), \(\mathcal{F}\) and \(Q\).

The operator on the lhs of (3.5), \(\mathcal{L}\), can be mixed elliptic, parabolic or hyperbolic depending on the distributions of \(\vec{u}\) and \(\overline{T}\). The behaviour of the coefficients is largely determined by the Coriolis parameter: the latter dominates the contribution of \(\vec{u}\) and \(\overline{T}\) to \(\mathcal{L}\) in the extratropics leading to ellipticity, but weakens in the tropics allowing parabolic and hyperbolic regions to form.

Hyperbolic regions are characterized by \(\vec{u}_\phi - a f - \vec{u} \tan \phi\) positive for \(\phi > 0\) or negative for \(\phi < 0\), in which case the coefficient of \(\vec{v}'\) in (3.1) has an opposite sign to \(-f\). These regions can form if the absolute angular momentum maximum is moved off the equator. As will be shown in section 3.3, the maximum is shifted into the winter hemisphere in the upper stratosphere and above by cross-equatorial advection, or by increasing subtropical zonal wind amplitude. The latter effect occurs near the stratopause and in the mesosphere, resulting in a hyperbolic region on the winter side of the equator.

The formation of hyperbolic regions represents a breakdown of the assumptions underlying the balance model (3.1)–(3.4), since such regions are symmetrically unstable (Eliassen, 1951); the basic state, although in gradient-wind and hydrostatic balance, cannot be realized. Symmetric instability and the associated adjustment are not captured by the balance model (3.1)–(3.4) and have to be parameterized. The adjustment prevents \(\mathcal{L}\) from becoming hyperbolic, but \(\mathcal{L}\) can be parabolic in regions of marginal stability.
The parameterization implemented here is based on the one used by Holton and Wehrbein (1980) and presumes that symmetric instability is primarily inertial instability; the details are given in Appendix B. It involves an angular-momentum-conserving redistribution of the zonal wind at every time step in regions where

\[ f(f - \bar{u}\phi/a) < 0. \]

The \( \bar{u}\tan\phi/a \) term is dropped since it was found not to have a significant effect on the solutions but contributed to numerical instability of the adjustment scheme. The temperature is also adjusted to maintain thermal-wind balance with the modified zonal wind distribution. Vertical and horizontal diffusion are included in (3.1) and (3.2) with constant coefficients of 0.25 m\(^2\) s\(^{-1}\) and 5000 m\(^2\) s\(^{-1}\), respectively, to control numerical instability. The solutions were not sensitive to this choice for the diffusion coefficients and were very close to those for much smaller values.

In the model of Dunkerton (1989), the action of inertial instability is represented by making the streamfunction operator parabolic where it becomes hyperbolic but leaving the zonal wind and temperature untouched. The formulation of the streamfunction equation used here (see below) was found to be sensitive to sharp transitions in the coefficients, so this method was not used. Dunkerton's approach also has a limitation in that the effect of inertial adjustment on the temperature and thereby the diabatic heating is not taken into account.

The numerical scheme used to solve the system given by equations (3.1), (3.2) and (3.5) is inspired by that of Kinnersley (1996). The system is discretized and then the streamfunction equation is obtained by imposing the condition that thermal-wind balance hold at the end of each timestep. The aim is to preserve thermal-wind balance while using both prognostic equations [(3.1) and (3.2)]. This avoids the need to obtain \( \bar{u} \) from \( \bar{T} \) or vice versa, which is uncertain up to some additive function of height or latitude.

The particular approach used here is to replace (3.1) with a prognostic equation for the variable \( (\bar{m}/a)^2 = \cos^2\phi (\bar{u} + a\Omega \cos \phi)^2 \). This involves no loss of sign information for \( \bar{u} \) since \( \bar{m} \), the absolute angular momentum, is positive under climatological conditions so \( \bar{u} \) can be uniquely determined from it. This formulation does restrict the amplitude of forcing that can be used in the model. The advantage is that the thermal-wind relation
between \((\overline{m}/a)^2\) and \(\overline{T}\) is linear. Therefore the streamfunction equation is also linear and the necessity for a nonlinear iterative solution at every time step is avoided. The streamfunction is obtained at every timestep using the MUDPACK multigrid solver (numerical library created by John C. Adams at NCAR). This elliptic solver can handle parabolicity in \(L\) over portions of the domain.

### 3.3 Nonlinear Hadley circulation

The basics of the tropical angular-momentum-conserving circulation described by Dunkerton (1989) are presented in this section. Some aspects which were not addressed previously in this context are considered. These pertain to the distinct roles of symmetric instability and the diabatic circulation in affecting the distribution of \(\overline{m}\) in the tropics. In addition, the accuracy of the balance model solutions is tested by comparing them to those of a three-dimensional primitive equations model.

#### 3.3.1 Theory

During solstice seasons there are strong latitudinal gradients of solar heating in the tropical upper stratosphere and lower mesosphere. However, the observed temperature does not follow the solar heating pattern and basically has no latitudinal gradient at and near the equator. As noted by Dunkerton (1989), the temperature distribution is constrained by gradient-wind balance. In the balance system there can be no meridional gradient in the geopotential, and hence temperature, at the equator where the Coriolis parameter vanishes. This constraint leads to an increase of the temperature on the winter side of the equator and a decrease on the summer side, as compared to the distribution that would be produced by radiative relaxation (and wave-induced heating) alone. This heating is balanced through the formation of a single-cell angular-momentum-conserving meridional circulation. In effect, balance acts like a mechanical forcing which maintains the diabatic heating against radiative relaxation but does not locally accelerate the zonal flow or induce a diabatic circulation directly.

In the steady-state limit, in the absence of wave drag, the zonal wind is deformed by the NLHC towards a parabolic profile in latitude (Dunkerton, 1989, Fig. 8), and \(\overline{m}\) is homogenized. If the static stability \(\frac{R}{H}S\) is independent of latitude then there is no
net heating on pressure surfaces since there cannot be any net flow across them. For the
Newtonian cooling approximation this results in an equal area rule for the temperature
profile (Dunkerton, 1989, Fig. 9). The timescale needed to reach steady state in the trop-
ics is longer than a year, although the state is approached closely during solstice seasons
(Dunkerton, 1989). This reflects the slow homogenization of $\overline{m}$ through advection by the
NLHC. In addition to this slow process there is rapid homogenization of $\overline{m}$ in inertially
unstable regions through inertial adjustment. The implications of this adjustment were
not addressed by Dunkerton's model.

The work in Dunkerton (1989) is a stratospheric extension of the theory of Held and
Hou (1980) for the inviscid tropospheric nonlinear Hadley circulation. One of the elements
in this theory is a rigid lid to close off the Hadley circulation and produce realistic zonal
winds. In Dunkerton's model, extending the domain to include the mesosphere results
in unrealistically strong equatorial easterlies in the mesosphere (Dunkerton, 1989, Fig.
19d). In Dunkerton (1991) the model has a deeper domain but uses a mesospheric Rayleigh damping layer, meant to represent gravity wave drag, to produce a realistic $\overline{m}$
distribution.

3.3.2 Behaviour of the system in the absence of wave drag

Here we present results of simulations similar to Dunkerton (1991) for the case with
no mechanical forcing, using the balance model described in section 3.2. The model
uses a domain extending from the surface to 100 km and has no Rayleigh damping in
the mesosphere. This latter choice is made to avoid spurious feedbacks arising from
the unphysical nature of Rayleigh drag (see Chapter 2), and is a significant difference
from the Dunkerton (1991) model. Another difference is the presence of an explicit
inertial adjustment scheme in our model. In Dunkerton's model only the coefficients
of the streamfunction equation are adjusted to keep it from becoming hyperbolic, an
approximation which becomes more accurate near steady state. We perform several
experiments to test these model assumptions as well as the impact of the choice of diabatic
heating representation.

In the first experiment the diabatic heating term, $Q$, is approximated by Newtonian
cooling with a constant damping rate of 0.1 day$^{-1}$ to a realistic time-dependent radi-
tive equilibrium temperature. The radiative equilibrium temperature was dynamically
Fig. 3.1 Results from the annual cycle of the balance model simulation for January with a constant Newtonian cooling timescale of 10 days: (a) streamfunction \(2 \text{ kg m}^{-1}\text{s}^{-1}\), (b) temperature \(8 \text{ K}\), (c) absolute angular momentum \(10^8 \text{ m}^2\text{s}^{-1}\), (d) diabatic heating \(3 \times 10^{-6} \text{ Ks}^{-1}\). The flow is northward across the equator.

determined (e.g. Fels, 1987) with a comprehensive middle atmosphere radiative transfer scheme by integrating the thermodynamic equation without the advection terms for three years. The radiative transfer scheme is based on the middle atmosphere part of the scheme described by Fomichev and Blanchet (1995). This part of the scheme is valid only above 16 km and as implemented here uses zonally averaged climatological distributions of ozone and other radiatively active gases. The derived \(\overline{T}_{\text{rad}}\) was extended below 16 km by interpolation with a simple height-dependent temperature profile. Above 80 km there are large equatorial gradients of \(\overline{T}_{\text{rad}}\) and hence thermal driving of a distinct NLHC originating in the thermosphere. As a result, strong cross-equatorial flow develops near the model lid which is enhanced by the confinement effect of the lid. This can lead to
Fig. 3.2 (a) Latitudinal profile of balance model zonal wind (solid) at \( z = 45 \) km compared with the angular-momentum-conserving parabolae corresponding to \( 10^\circ \) (long dash) and \( 30^\circ \) (dash). (b) Latitudinal profile of \( \bar{T} \) (solid) and \( \bar{T}_{\text{rad}} \) (dash) at the same level as (a).

Numerical instability; since the focus here is primarily on the stratosphere, the thermal driving above 80 km is removed by interpolating \( \bar{T}_{\text{rad}} \) to a uniform value of 240 K.

The model state at the NH winter solstice of a seasonal cycle simulation without wave drag is shown in Figure 3.1a–d (the SH winter solstice fields are similar but flipped about the equator). The lowermost (tropospheric) and uppermost (thermospheric) parts
of the model domain are not shown since they are regions where $\overline{T}_{rad}$ is artificially constructed. The model basically has no troposphere and, in particular, it lacks the Hadley circulation. During equinoxes, the NLHC dies away since the latitudinal gradient of the radiative equilibrium temperature vanishes at the equator. The state represented in Figure 3.1 is not steady, but reflects transience associated with the seasonal cycle. Complete homogenization of $\overline{m}$ in the region of overturning through advection by the NLHC is not attained, as can be seen in the latitudinal profile of $\overline{u}$ (Fig. 3.2a) which approaches a parabola only in the inertial adjustment region.

In the tropical stratosphere, isopleths of $\overline{m}$ (Fig. 3.1c) are increasingly deformed with altitude. This behaviour has two causes. As the altitude increases, the solar heating gradient increases, which leads to a stronger NLHC and stronger advection of $\overline{m}$ towards the winter pole. At the same time, the radiative equilibrium zonal wind field has increasingly strong westerlies in the winter hemisphere, which moves the associated $\overline{m}$ maximum towards the winter pole. When the $\overline{m}$ maximum moves off the equator it is eroded through inertial adjustment, producing a strip of homogenized $\overline{m}$ on the winter side of the equator. Above the stratopause, $\overline{m}$ is seen to be homogenized over a broad region extending to the winter hemisphere midlatitudes. This feature is due to the second mechanism, i.e. inertial adjustment triggered by the large subtropical meridional shear in the radiative equilibrium zonal wind, rather than advection by the NLHC. Advection of $\overline{m}$ into the winter hemisphere by the NLHC is limited in the adjustment region since the meridional gradient of $\overline{m}$ is negligible. The annual mean $\overline{w^*}$ distribution (see Chapter 4) indicates that there is a transition from an NLHC induced adjustment regime in the stratosphere to a shear-driven inertial adjustment regime in the mesosphere. The winter hemisphere edge of the homogenized $\overline{m}$ region moves further from the equator with height and reaches a terminal latitude in the mesosphere. This pattern reflects the vertical profile of the westerly zonal wind and hence the distribution of its latitudinal shear.

The diabatic heating associated with the NLHC (Fig. 3.1d) peaks near the tropical stratopause. Its width broadens considerably and its magnitude increases between the lower stratosphere and the stratopause, reflecting the strong increase of solar heating with altitude. The diabatic heating correspondingly diminishes above the stratopause as the solar heating decreases. The distribution of the diabatic heating is much broader
than can be explained by the balance constraint on the equatorial temperature profile. This is due, in part, to transience in the NLHC. But inertial adjustment also plays a role, its effect becoming apparent in the upper stratosphere and mesosphere. The cooling maximum shifts into midlatitudes with height following the edge of the homogenized $\bar{m}$ region (Fig. 3.2b). The diabatic heating is modified on account of thermal wind balance which necessitates a temperature redistribution in the process of $\bar{m}$ homogenization on the winter side of the equator. This is evident in the temperature field (Fig. 3.1b) at higher levels where contours are flattened out deep into the winter hemisphere. The inertial adjustment induced temperature change results in a heating dipole (cooling in the winter hemisphere and warming in the summer hemisphere and tropics) which superimposes itself on the “pure” NLHC heating dipole, shifting it towards the winter pole. The zonal-wind-shear induced adjustment persists well above the lower mesosphere, but it is not able to produce significant heating at altitudes with little equatorial solar heating gradient.

The NLHC (Fig. 3.1a) progressively shifts towards the winter pole with altitude. This deformation is tied to the increasing deflection of $\bar{m}$ isopleths towards the winter pole and the corresponding shift in the diabatic heating dipole off the equator. In the lowest part of the tropical stratosphere, however, $\bar{m}$ remains close to its planetary value and the diabatic heating is centered on the equator, so the NLHC is nearly symmetric about the equator. The tendency for streamline alignment with $\bar{m}$ contours arises from the same constraint on evolution to steady state as is exhibited by “downward control” (Haynes et al., 1991). The difference is that in the extratropics, $\bar{m}$ isopleths are vertically aligned and have limited mobility so that they organize the streamfunction distribution. In the tropics, by contrast, the increasing mobility of $\bar{m}$ becomes an important factor in the evolution to steady state since the streamfunction significantly affects the $\bar{m}$ distribution forming a nonlinear feedback. Indeed, it is this feature that allows a steady circulation to develop in the absence of wave drag or viscosity (Held and Hou, 1980).

In the second experiment, the limitations of simple Newtonian cooling are explored. In particular, we look at the influence of the choice of the vertical profile of the radiative relaxation rate on the structure of the circulation and $\bar{m}$. The effect of using a comprehensive radiative transfer scheme on the solutions is also considered.

The flattening of $\bar{m}$ contours in the tropical stratopause region of the balance model is due to the distribution of the diabatic heating which drives the NLHC, and which
Fig. 3.3 CIRA January mean $\bar{m} \; (10^8 \text{ m}^2\text{s}^{-1})$.

Fig. 3.4 Same as Figure 3.1 but for the balance model simulation with a monotonically decreasing Newtonian cooling timescale with height.

happens to maximize in this region. This feature is suggestive of the pattern seen in
Fig. 3.5 Same as Figure 3.1 but for the balance model simulation using a comprehensive radiative transfer scheme.

observations (Fig. 3.3) and GCM simulations if the mesosphere is ignored. However, the assumption of a uniform radiative damping rate used in the simulation above is unrealistic given the generally decreasing optical thickness of the atmosphere with altitude. Yet if the constant radiative damping rate is naively replaced with a monotonically increasing function of height, the diabatic heating and the region of horizontal $\overline{w}$ contours in the model shift upwards in an unrealistic manner (Fig. 3.4).

This pathological behaviour can be explained by the fact that the balance constraint on the latitudinal temperature profile is insensitive to the damping rate (except possibly in the upper mesosphere where the damping timescale is on the order of a day): the deviation of $\overline{T}$ from $\overline{T}_{rad}$ depends only on $(\overline{T}_{rad})_\phi$. It follows that the amplitude of the diabatic heating near the equator is affected primarily by the radiative damping rate. Thus, the radiative damping rate acts as a weighting function which, in the monoton-
ically increasing case, shifts the heating to higher altitudes even though the radiative equilibrium temperature gradient is weaker there.

In fact, the radiative damping rate is sensitive to temperature. The damping rate increases in the upper stratosphere, maximizing near the stratopause (Andrews et al., 1987), reflecting the vertical profile of the radiative equilibrium temperature. Consequently the region of largest solar heating gradients also corresponds to a local maximum in the damping rates. The balance model simulation was therefore repeated using the comprehensive radiative transfer scheme of Fomichev and Blanchet (1995) instead of Newtonian cooling (Fig. 3.5). The Q dipole (Fig. 3.5d) is stronger compared to the one in the first experiment (Fig. 3.1d), reflecting the shorter radiative damping time scale and its dependence on the local temperature. The diabatic heating is also confined closer to the stratopause. This feature manifests itself in the streamfunction (Fig. 3.5a) as more concentrated and horizontal streamlines just above the stratopause. In the lower tropical stratosphere the streamfunction is weaker since the radiative damping time is much longer than 10 days.

To summarize the results of this section, the inclusion of a more realistic inertial adjustment scheme and extension of the domain into the mesosphere (without Rayleigh drag) has yielded additional insight into the NLHC problem. The zonal wind and temperature distribution can be significantly affected by inertial adjustment in the upper stratosphere and mesosphere. This is reflected in the diabatic heating associated with the NLHC, which extends to higher latitudes of the winter hemisphere. It has also been demonstrated that flattening of $\overline{m}$ contours in the stratopause region is a feature of the NLHC diabatic heating distribution, which is determined by the vertical distribution of radiative heating and radiative damping. It thus appears that the excessive equatorial mesospheric easterlies in the model of Dunkerton (1989) are an artefact of the radiative driving used. The absence of an explicit inertial adjustment of the zonal winds likely played a role too, since transport of angular momentum to the equator by inertial adjustment diminishes the magnitude of the equatorial easterlies.

### 3.3.3 Role of tropical wave drag

Wave drag has so far been neglected, but is an important element in the formation of the middle atmosphere zonal wind and temperature distributions. The model $\overline{m}$ field shown
in Figure 3.1 departs significantly from the observed one (Fig. 3.3) above the stratopause; in the real atmosphere the region of horizontal $\bar{u}$ contours is confined to the stratopause and lower mesosphere region. The degree of contour deflection into the winter hemisphere in the stratosphere is also smaller.

Dunkerton (1991) used a relaxational damping layer in the mesosphere to represent the action of wave drag and produce a qualitatively correct distribution of $\bar{u}$. Plumb and Eluszkiewicz (1999) have also noted the stabilizing effect of weak relaxational damping on the $\bar{u}$ distribution of the tropical stratosphere. However, relaxational forcing is artificial and does not represent the response of wave drag to zonal wind changes (see Chapter 2). In particular, the momentum flux available for wave drag is fixed (aside from variations in tropospheric sources), and zonal wind changes lead only to a redistribution of this fixed amount of wave drag. In contrast, a relaxational damping layer is an infinite reservoir for local momentum dissipation.

Nevertheless, relaxational damping imitates some forms of wave drag in the tropics which act to pull the zonal wind to zero. This is suggested by the restoration of vertical $\bar{u}$ isopleths in the middle and upper tropical mesosphere (Fig. 3.3) even in the presence of strong cross-equatorial flow. The reduced mobility of $\bar{u}$ in the upper tropical stratosphere is most likely a consequence of the same sort of tropical wave drag.

Deceleration of the midlatitude wintertime zonal wind by stationary planetary waves in the stratosphere and by gravity waves in the mesosphere also reduces the latitudinal shear of the zonal flow in the subtropics. As a result, zonal-wind-shear-driven inertial instability and non-advective $\bar{u}$ homogenization are expected to be confined closer to the equator.

### 3.3.4 Validity of the balance model

A significant physical assumption underlying the results presented above is gradient-wind balance. However, the latter is an empirically motivated approximation and not an instantaneous constraint that is satisfied by the real atmosphere, or indeed by more comprehensive models which resolve wave and instability processes. Gradient-wind balance is predicated on the separation of timescales between inertial adjustment and the slowly evolving basic state (Eliassen, 1951). This allows the separation of the dynamics
Fig. 3.6 (a) Balance model \( \overline{m}, \psi \) and \( \overline{T} \) fields in July. Contour intervals as in Figure 3.1. (b) Same fields as (a) for a 3D primitive equations model Saravanan (1992), with a resolution set to T42 and 60 levels.

into a slow and fast part, with the effects of the fast part being parameterized rather than resolved.

We use a mechanistic model based on the three-dimensional primitive equations (Saravanan, 1992) to test the validity of the balance model simulations. The same meridional domain and Newtonian cooling scheme are used in both models. The models have the
3.4 Influence of the WDC on the NLHC

same vertical diffusion but different horizontal diffusion, but this is not considered to be important for the basic structure of the solutions. Both models were run for 91 days to the model summer solstice from a state of rest with a gradual ramping in the first month to a seasonally evolving radiative equilibrium temperature. This integration period is much longer than the inertial adjustment timescale in the three-dimensional model (a few days) and sufficient to establish the pattern of evolution of the slow component. The primitive equations model captures zonally asymmetric and unbalanced dynamics. Comparing the zonal mean states at the end of the simulation period (Fig. 3.6) it is evident that they have a high degree of similarity on large scales except for the presence of large-amplitude small-scale disturbances in the primitive equations model state. These disturbances occur in the same inertially unstable region as found in the two-dimensional balance model, and are evidence of explicitly resolved inertial instability.

Inertial instability is more prominent in the primitive equations model when it is run in axially symmetric mode, as found by Saravanan (1992). The relative weakness of the disturbances in the zonal mean fields in the three-dimensional simulation is apparently due to the development of a large zonally asymmetric component which is obscured by zonal averaging. This is consistent with the stability analysis of Dunkerton (1983) which predicts zonally asymmetric modes of inertial instability as being the fastest growing. Nevertheless, zonally symmetric modes are not negligible and are active in the winter hemisphere adjustment as can be seen from the zonal mean fields (see also Clark and Haynes, 1996).

Of note is the apparent insensitivity of the NLHC to the details of the balance process. The overall form and amplitude of the NLHC in the primitive equations model matches that of the balance model, except for the additional direct contribution from inertial instability which is largest in the winter hemisphere mesosphere. On the basis of this simulation, it appears that gradient-wind imbalance has only a local effect. It also appears that the balance model inertial adjustment scheme is responding correctly to the distribution of the zonal wind (consider the $\vec{m}$ homogenization in the primitive equations model).
3.4 Influence of the wave-driven circulation on the nonlinear Hadley circulation

A feature of the nonlinear balance system is the lack of self-advection in the meridional circulation. On account of gradient-wind and hydrostatic balance, the inertial terms, namely $D\bar{u}^* /Dt$ and $D\bar{w}^* /Dt$, are negligible in the meridional and vertical momentum equations. The inertial terms contain the nonlinear (self-advection) and tendency parts of these equations. Neglecting the tendency part makes the meridional circulation an instantaneous, nonlocal field which requires the presence of mechanical or diabatic forcing to exist. Dropping the nonlinear part of the meridional and vertical momentum equations makes the streamfunction equation linear. A streamfunction equation can be derived even if balance is not assumed, but it will be nonlinear and have time derivatives as well. Adveective nonlinearity in the balance system manifests itself only as redistribution of the $\bar{m}$ and $\bar{\theta}$ fields, which can modify $\psi$ by changing the coefficients of the operator $\mathcal{L}$ and $\mathcal{Q}$.

In general, wave drag and diabatic heating are nonlinear, nonlocal functions of the zonal wind and temperature, respectively. Here we assume that wave drag is a prescribed field which is not determined as part of the solution but affects its evolution. For the diabatic heating the linear Newtonian cooling approximation is used. These assumptions simplify the problem but leave enough to investigate the basic zonal-mean interaction of the NLHC and subtropical wave drag. The tropical region is defined here as extending between $20^\circ$S and $20^\circ$N, based roughly on the subtropical edge of the surf zone in the lower stratosphere (above 16 km). The influence of the NLHC on the subtropical wave drag distribution is best dealt with in a more comprehensive model which can resolve waves and wave breaking, so will not be addressed here. The effect of the NLHC on the wave-driven circulation (WDC) is considered in Chapter 4.

To a good approximation, extratropical zonal-mean dynamics are linear due to the stable stratification of $\bar{m}$ in latitude and $\bar{\theta}$ in altitude. Specifically, $\bar{m}_\phi /a$ is close to $(\Omega a^2 \cos^2 \phi) /a = -a \cos \phi f$ and much larger than $\bar{m}_z = a \cos \phi \bar{u}_z$, while $\bar{\theta}_z$ is close to $e^{az/H} S_0$ (where $S_0$ is the static stability for an isothermal atmosphere with $\bar{T} = 240$ K, although $S_0$ based on a standard reference $\bar{T}$ profile is more accurate) and is much larger than $\bar{\theta}_\phi /a$. As a result, to leading order the absolute angular momentum equation is

$$\bar{m}_t - a \cos \phi f \bar{u}^* = a \cos \phi \mathcal{F} \quad \text{i.e.} \quad \bar{u}_t - f \bar{u}^* = \mathcal{F}$$
and the potential temperature equation is

$$\bar{\theta}_t + e^{\kappa z/H} S_0 \bar{w}^* = e^{\kappa z/H} Q \quad \text{i.e.} \quad \bar{T}_t + S_0 \bar{w}^* = Q.$$  

The linear form of $\mathcal{F}$ and $Q$ considered then result in solutions for the WDC and the NLHC being independent (see Chapter 5). In the tropics, however, the Coriolis term, $f$, is small so that the advective (nonlinear) terms in the zonal-wind equation become relatively more important and cannot be neglected (i.e. $\bar{m}_\phi/\alpha$ can deviate from $-\alpha \cos \phi f$ significantly and can be comparable to $\bar{m}_z$ in magnitude). This allows the WDC and NLHC to couple in the tropics through the dependence of $\mathcal{L}$ on $\bar{m}$ and $\bar{\theta}$ and indirectly through redistribution of $Q$.

Since $\bar{\theta}$ is stably stratified its variation has a limited effect on $\mathcal{L}$, as in the extratropics. The advective influence on $\bar{m}$ appears to be strongest in the stratopause region (e.g. Fig. 3.3) where $\bar{w}^*$ is typically much larger than in the lower tropical stratosphere. In addition to advection, the $\bar{m}$ distribution can be modified through thermal-wind balance and the associated inertial adjustment as described in section 3.3.

The tropics are characterized by the breakdown of ellipticity in the streamfunction operator $\mathcal{L}$. In the linear system, non-ellipticity is confined to the equator itself where $\mathcal{L}$ becomes parabolic. In the nonlinear system, the non-elliptic behaviour can occur over a broad span of tropical latitudes during solstice seasons, as shown in section 3.3.

Even the minimal equatorial parabolicity which occurs in the linear system has a significant effect on the WDC. The tropical $\bar{w}^*$ distribution produced by forcing in the subtropics can have no meridional gradient at the equator. This constraint is the same as for the temperature profile and originates from the degeneracy of thermal-wind balance at the equator: it is not possible for subtropical wave drag to produce a meridional gradient in the temperature across the equator. This prevents the WDC from either enhancing or suppressing the NLHC through modification of the diabatic heating in the immediate vicinity of the equator. However, this is not necessarily the case at higher latitudes. In the nonlinear case this constraint also applies only at the equator since generally $\bar{m}_z \neq 0$ in the finite parabolic regions adjacent to the equator.

The purpose here is to investigate how the WDC-induced redistribution of $\bar{m}$ (and $\bar{\theta}$) in the tropics modifies the NLHC. We conduct a numerical test involving a prescribed
region of mechanical forcing with annual periodicity in the model extratropics. Another approach would be to use a more comprehensive wave drag parameterization (e.g. Garcia, 1991) but it is believed that the essential aspects can be captured with an idealized wave drag. The imposed wave drag has the form

\[ F(\phi, z, t) = F_0 Y(\phi)^{1/2} Z(z)^{1/2} H(t) \]

with \( F_0 = -7 \text{ m s}^{-1} \text{ day}^{-1} \). The latitudinal dependence is defined by

\[
Y(\phi) = \begin{cases} 
  \frac{1}{2} - \frac{1}{2} \cos \left( \frac{\pi}{2} \frac{\phi - \phi_1}{\phi_2 - \phi_1} \right) & \text{for } \phi_1 \leq \phi \leq \phi_2 \\
  \frac{1}{2} + \frac{1}{2} \cos \left( \frac{\pi}{2} \frac{\phi - \phi_2}{\phi_3 - \phi_2} \right) & \text{for } \phi_2 \leq \phi \leq \phi_3 \\
  0 & \text{otherwise}
\end{cases}
\]

with \( \phi_1 = 15^\circ \) being the subtropical limit of the forcing, \( \phi_2 = 38^\circ \) the latitude of its maximum, and \( \phi_3 = 80^\circ \) its northern limit. The vertical dependence is defined by

\[
Z(z) = \begin{cases} 
  \frac{1}{2} + \frac{1}{2} \cos \left( \frac{2\pi}{2} \frac{z - \frac{1}{2}(z_1 + z_2)}{z_2 - z_1} \right) & \text{for } z_1 < z < z_2 \\
  0 & \text{otherwise}
\end{cases}
\]

with \( z_1 = 16 \text{ km} \) the lower limit and \( z_2 = 84 \text{ km} \) the upper limit of the forcing region. The time dependence is defined by

\[
H(t) = \begin{cases} 
  \sin^2 \left( \frac{\pi}{91} \frac{t + 30}{365} \right) & \text{for } 0 \leq t \leq 61, 335 \leq t \leq 365 \\
  0 & \text{for } 61 \leq t \leq 335
\end{cases}
\]

(with \( t \) in days) so that it acts only during the winter season. In order to qualitatively separate the effect of the WDC on \( \bar{m} \) alone or in combination with the NLHC, the radiative equilibrium temperature was respectively prescribed to be either isothermal or climatologically derived (as in section 3.3).

In the isothermal \( T_{rad} \) case (Fig. 3.7), where there is no NLHC, the effect of the imposed \( F \) on the \( \bar{m} \) distribution in the tropics is limited. There is no significant isopleth deformation but a strong gradient in \( \bar{m} \) forms in the subtropics at the edge of the forcing region. The response in the climatological case (Fig. 3.8) is quite different, with greater \( \bar{m} \) isopleth deformation in the upper stratosphere and mesosphere tropics than is evident.
Fig. 3.7 January distribution of (a) $\psi$ (8 kg m$^{-1}$s$^{-1}$), (b) $\bar{T}$ (5 K), (c) $\bar{m}$ ($10^6$ m$^2$s$^{-1}$) and (d) $\mathcal{Q}$ ($3 \times 10^{-6}$ Ks$^{-1}$) from the nonlinear balance model with $T_{\text{rad}} = 240$ K and forcing specified in the text.

in Figure 3.1. There is also a noticeable alteration of the diabatic heating which occurs on both sides of the equator and at the stratopause. The diabatic heating shifts downward and intensifies on account of the forcing (Fig. 3.8d) which superimposes a quadrupole heating pattern of its own (Fig. 3.7d).

The $\bar{m}$ deformation in the upper tropical stratosphere and lower tropical mesosphere in the climatological case persists much longer than the duration of the forcing and lasts into the following year. This slow relaxation towards radiative equilibrium in the tropics is due to the "flywheel effect" (Scott and Haynes, 1998) resulting from the decoupling of the temperature and the zonal wind near the equator. The range of altitudes where the severe $\bar{m}$ isolopleth distortion occurs is partly determined by the height of the imposed wave drag maximum (at 50 km), which corresponds to the peak cross-equatorial flow and is reflected in the intensification of the diabatic heating. However, it can be seen from
Figure 3.8 Same as Figure 3.7 but with $T_{rad}$ based on CIRA temperatures.

Figure 3.7c that without the NLHC and inertial adjustment the WDC cannot produce the sort of isopleth distortion seen in Figure 3.8c. In the climatological case the tropical state is conditioned in the stratopause region in a way that enhances the cross equatorial flow associated with the WDC (see Chapter 4). This tropical stratopause region $\overline{m}$ deformation (Fig. 3.3) is consistent with observational inferences, which point to an increase in both the equatorward penetration and amplitude of wave drag with height. In the model subtropics, the imposed wave drag acts to partially suppress the advective and inertial adjustment deformation of $\overline{m}$ (compare Fig. 3.8c with Fig. 3.1c), which brings the horizontal $\overline{m}$ stratification closer to that observed.

The enhancement of the WDC in the tropics in the presence of the NLHC is evident in Figure 3.9, which shows the net contribution of the forcing to the fields of the model with realistic $\overline{T}_{rad}$. Such a comparison is not strictly valid on account of nonlinearity, but has some justification if the intensity and distribution of the NLHC is assumed not to
change significantly. This assumption is discussed below. When compared with Figure 3.7 the streamfunction is seen to be of slightly larger amplitude and greater latitudinal extent in the hemisphere without the forcing. There is some enhancement of the tropical and subtropical diabatic heating that also extends deeper into the summer hemisphere. The diabatic heating, however, exhibits a much stronger cooling region. This difference can partly be attributed to advection due to large meridional temperature gradients in this region (a feature which does not occur in the model with uniform radiative equilibrium temperature).

The effect of the WDC on the NLHC is difficult to discern without an exact quantification of both in the presence of each other. It would seem that a precise quantitative separation of the WDC from the NLHC is possible if $Q$ is linear in $\bar{T}$. Separating the component of $Q$ due to wave drag from that due to the NLHC would allow this since limited nonlinearity in the balance system makes $\mathcal{L}$ linear in $\psi$. However, this rests on being able to separate the temperature in a similar manner. Even with a linear $Q$, the nonlinear coupling of the NLHC and WDC through the $\overline{m}$ (and to a lesser degree $\overline{T}$) dependence of $\mathcal{L}$ will be reflected in the distribution of $\overline{T}$, making the problem non-separable. The NLHC that forms when the WDC is present has a different structure than the one that forms without wave drag, and vice versa. As will be discussed below, the model experiments conducted here indicate that this interaction has a significant impact on the $\overline{m}$ distribution but is rather limited in terms of the bulk features of the diabatic heating.

Fig. 3.9 Contour plot of $\psi$ (left panel) and $Q$ (right panel) when the fields corresponding to Figure 3.1 are subtracted from those of Figure 3.8. To be compared with (a) and (d) of Figure 3.7.
and circulation fields. In fact, the limited type of nonlinearity in the system makes the problem pseudo-linear in the atmospheric case. This is consistent with the climatological distributions of $\bar{\eta}$ and $\bar{T}$, which exhibit a simple and fairly predictable time-dependent structure in the course of the annual cycle. Consequently, the NLHC must also have a regular climatological structure and to a large extent can be considered superposable on the WDC and vice versa.

With the Newtonian cooling approximation, the action of the WDC (and $F$) can be treated as a modification of $\bar{T}_{\text{rad}}$ from the perspective of the NLHC. As noted previously, the WDC is prevented from producing equatorially asymmetric heating near the equator. Consequently, the dynamically modified radiative equilibrium temperature $\bar{T}'_{\text{rad}}$ retains the meridional gradient of the true $\bar{T}_{\text{rad}}$ near the equator and the NLHC is neither suppressed nor enhanced. Away from the equator the WDC can produce a gradient in $\bar{T}'_{\text{rad}}$ different from that in $\bar{T}_{\text{rad}}$. The stratospheric wave drag distribution is such that the WDC pulls the temperature below radiative equilibrium mostly on the winter side of the equator and to a lesser extent on the summer side. Overall, $\bar{T}'_{\text{rad}}$ is then colder than $\bar{T}_{\text{rad}}$ in the tropics and has a steeper decline from the summer hemisphere to the winter hemisphere. The NLHC dipole heating pattern of the $F \equiv 0$ case considered in section 3.3 forms relative to $\bar{T}_{\text{rad}}$. In the presence of wave drag this dipole forms relative to $\bar{T}'_{\text{rad}}$ and will have a different structure due to changes in inertial adjustment. The reasoning above applies in steady state as well, but $\bar{T}'_{\text{rad}}$ may not differ from $\bar{T}_{\text{rad}}$ over most of the tropics if the wave drag is confined to the extratropics where the downward control principle applies.

The impact of the WDC on the NLHC in a hypothetical steady state is limited by the location of the wave drag. The steady state NLHC region is characterized by homogenized $\bar{\eta}$, which is the result of advection by the NLHC and inertial adjustment. If there is a separation between the region of homogenized $\bar{\eta}$ and the wave drag then there is no effect since all the WDC streamlines are confined beneath the region of wave drag in accordance with downward control. If the wave drag overlaps the steady state NLHC region then there is an interaction. One effect is the reduction of $\bar{\eta}$ isopleth deformation in the overlap region seen in the transient regime in the climatological model results. This leads to a confinement of the homogenized $\bar{\eta}$ region closer to the equator, which reduces the latitudinal extent of the diabatic heating produced by inertial adjustment. Another
effect is that streamlines of the WDC are not constrained in latitude in the homogenized \(\overline{w}\) region and extend into the tropics (where they are subject to the constraint of no latitudinal gradient in \(\overline{w}'\), as in the transient case). This steady state penetration of the WDC into the tropics could potentially be associated with a modified distribution of the homogenized \(\overline{m}\) region and the NLHC. The steady state case has not been investigated in the course of this study, although more analysis of this state would be useful.

The overlapping wave drag case corresponds to that considered in Dunkerton (1991) (see his Fig. 1b) where it was reasoned on the basis of potential temperature conservation, in particular the increased tropical departure from radiative equilibrium, that extratropical wave drag can enhance the NLHC. However, this constraint applies to the total circulation that includes the WDC, and the latter makes a significant contribution to the increased steady state \(\overline{T}\) deviation from \(\overline{T}_{rad}\) in the tropics. So it is not clear how the NLHC itself is affected.

As indicated by the observed \(\overline{m}\) field, true wave drag limits the formation of inertially unstable regions in the subtropics induced by zonal wind shear, particularly in the upper stratosphere and mesosphere. Consequently, the reduced diabatic heating will diminish the latitudinal span of the NLHC at these altitudes compared to the case without wave drag. Judging by the absence of a strong warming pattern in the adjustment region in Figure 3.9, indicative of reduced cooling, the wave drag in the forced model is too weak to suppress the adjustment process in the extratropics. This points to the limitation of using an idealized prescribed forcing field; what is needed is a wave drag parameterization that can respond in a physically plausible manner to changes in the zonal wind.

Some qualitative sense of the impact of the WDC on the NLHC distribution can be inferred from the numerical experiments. The poleward tilt of the NLHC (as defined by the \(\psi_0 = 0\) line roughly bisecting the streamfunction in latitude, i.e. the \(\overline{w}' = 0\) contour separating the upwelling and downwelling regions) follows the zero heating contour which starts on the equator near the tropical tropopause and moves to 15°N at the stratopause level (Fig. 3.1d). The equatorial asymmetry in the diabatic heating is due to deformation of \(\overline{m}\), which leads to horizontal transfer of heat towards the winter pole in order to restore thermal wind balance. If the \(\overline{m}\) isopleths were vertical at all latitudes and \(\overline{m}\) was equatorially symmetric, then the \(Q\) distribution and the "NLHC" would be centered on the equator as in the zonal mean quasi-geostrophic (linear) system. The addition of the
body force to the model reduces poleward deflection of $\bar{m}$ isopleths in the subtropics but increases their deflection in the tropics above 40 km. As a result, the NLHC component of the diabatic heating field should retain most of its equatorial asymmetry. This is indicated by the location of the zero heating contour separating the warming and cooling regions (Fig. 3.8d) at about 15°N over the levels spanning the upper stratosphere and lower mesosphere. The component of the diabatic heating associated with the body force and its circulation is not dominating the tropical heating pattern, otherwise the zero heating contour would be closer to 30°N. Consequently, the NLHC diabatic heating and streamfunction must exhibit a poleward tilt similar to the unforced case. The midlatitude reduction of NLHC diabatic heating, which is only partial in the model, should shift the streamline pattern seen in Figure 3.1a equatorward.

In general, the nonlocal effect of strong extratropical wave drag is to counteract the local stabilization of $\bar{m}$ isopleths by relatively weak tropical wave drag via the WDC, and contribute to the poleward deformation of the streamlines of the total diabatic circulation (through the $\bar{m}$ dependence of $L$ and $Q$).

The latitudinal extent of the NLHC can be defined by the region where there is not enough absolute angular momentum available for the temperature to reach radiative equilibrium and where there is persistent inertial adjustment. The latter can increase the breadth of the NLHC significantly. It was pointed out in Dunkerton (1989) that the NLHC can be expanded in latitude by increasing the $T_{rad}$ gradient, which amounts to increasing the latitudinal span of the region of large $\bar{u}$ amplitude about the equatorial singularity. This can be regarded as the true measure of the intensity of the NLHC. The WDC contributes to the circulation and heating in the tropics and subtropics but apparently only redistributes the NLHC and is unable to affect its intensity. Given the significant effect that the WDC has on the tropical $\bar{m}$ distribution it is surprising that the effect on the diabatic heating is weak (Fig. 3.9). Wave drag, on the other hand, can suppress the formation of subtropical heating by inertial adjustment. Inasmuch as this can be regarded as part of the NLHC, some suppression of the latter can occur.

Although Figure 3.7d suggests that the wave-drag-induced diabatic heating in the tropics and subtropics cannot dominate the structure of $Q$, the amplification of the WDC that occurs in the presence of the NLHC has to be taken into account. It increases the diabatic heating driven by the WDC. Therefore, subtracting the field in Figure 3.7d from the field in Figure 3.8d will not yield the modified NLHC diabatic heating distribution.
3.5 Nonlinear Hadley circulation in a middle atmosphere GCM

While the balance model is a reasonable starting point for the study of the NLHC, the goal is to identify this process in the real atmosphere. Current efforts to understand the atmosphere include GCM studies. Given the limited nature of observational data, particularly the tropical winds, these comprehensive models provide a proxy data set. At the same time, such analysis is a useful test of the physical realism of a GCM.

The accuracy of zonal-mean tropical gradient-wind balance was considered in section 3.3.4 in a mechanistic model. Here we consider the Canadian Middle Atmosphere Model (CMAM) (documented in Beagley et al., 1997) in a similar light. The time averaged rms difference between the two dominant terms in the meridional momentum equation, namely the meridional gradient of the geopotential and the Coriolis term, is shown in Figure 3.10 for the monthly averaged January zonal-mean state. It appears that gradient-wind balance is a good approximation over most of the stratosphere and begins to break down only in the upper stratosphere and lower mesosphere where the cross-equatorial flow is the largest. This corresponds to regions where inertial instability is most active, as is discussed below. Nevertheless, the breakdown of balance is relatively slight in the monthly mean.

Inertial instability is active in the upper stratosphere of CMAM. There are $2\Delta z$ cells (2-grid structures) (Fig. 3.11) with a strong zonally symmetric component mainly on the winter side of the equator, similar to those found by Hunt (1981). The fact that the cells are on the grid scale (in the vertical) is not a reflection of computational instability; rather, it reflects the fact that the physical instability has no scale selection, and occurs on the smallest allowable vertical scale (i.e. it is limited by numerical diffusion). The latitudinal zonal wind profile at altitudes in the middle and upper stratosphere is consistent with the presence of inertial adjustment in CMAM. In particular, a kink in the zonal wind forms on the winter side of the equator (Fig. 3.12a) which allows the zonal wind to satisfy the inertial stability criterion and which cannot be explained by radiative forcing or wave drag. However, there are no parabolic regions in the stratosphere and above near the equator. Unlike the balance model parameterization, the explicitly resolved inertial adjustment in CMAM does not homogenize $\bar{m}$ in regions of inertial instability and only weakens $\bar{m}_\phi$ in small bands on both sides of the equator (Fig. 3.12b). There is
Fig. 3.10 Latitude-height plot of the January average of the rms gradient-wind imbalance from CMAM (left panel) and January mean $\Phi/a$ ($\Phi$ is the geopotential) (right panel); contour interval is $2 \times 10^{-4}$ m s$^{-2}$ in both panels. There is little breakdown of gradient wind balance in the CMAM tropical stratosphere and lower mesosphere (except very close to and at the equator).

Fig. 3.11 CMAM January mean zonal-mean meridional residual velocity field; the contour interval is 0.2 m s$^{-1}$). The presence of inertial instability is indicated by grid-trapped circulation cells near the equator (stacked layers of alternatingly poleward and equatorward flow). In the course of the annual cycle these cells alternate from one side of the equator to the other, minimizing during equinoxes.
certainly no indication of \( \overline{m} \) homogenization occurring at higher latitudes such as in the balance model runs of section 3.3. The absence of \( \overline{m} \) homogenization suggests either that angular momentum is not conserved or that there is significant damping by wave drag in the tropical upper stratosphere and mesosphere. The presence of a strong semi-annual oscillation (SAO) in CMAM (Beagley et al., 1997) is suggestive of the latter.

The CMAM temperature field gives an indication of the presence of the NLHC. The
Fig. 3.14 January mean CMAM absolute angular momentum distribution (10^8 m^2 s^-1).

Latitudinal profile of the diabatic heating (Fig. 3.13 (solid curve)) reveals that the tropical cooling maximum is in the summer hemisphere as it is in the diabatic heating inferred from observed temperatures (Fig. 3.13 (dashed curve)). Such a temperature profile is consistent with a superposition of the NLHC heating dipole and the tropical heating distribution produced by wave drag. Summer hemisphere dynamical heating is too weak (Rosenlof, 1995) to pull the temperature away from radiative equilibrium to this extent, while winter hemisphere drag would produce a heating maximum on the winter side of the equator (or on the equator if it is sufficiently close to it (related to the PE effect)).

A major feature of the tropical $\overline{m}$ distribution in both CMAM simulations and CIRA observations is the region of flattened contours in the stratopause region (Figs. 3.3 and 3.14). This is another possible indication of the presence of the NLHC in CMAM. As shown above, horizontal deformation of $\overline{m}$ contours in the tropical stratopause region can be produced by the NLHC. But the degree of $\overline{m}$ contour deformation suggests that this feature is the combined result of the NLHC and the WDC (cf Fig. 3.8). The necessary WDC can be plausibly produced by the climatologically inferred wave drag in the upper stratosphere, which has close proximity to the equator (Rosenlof, 1995).

3.6 Conclusion

The NLHC is a significant factor in the dynamics of the tropical middle atmosphere. This is borne out by the robustness of gradient-wind balance in the tropics, which validates the
balance model, and by indirect evidence such as the diabatic heating distribution in the tropics. The former gives some confidence in the strength of the NLHC in more realistic models such as CMAM which are more accurate representations of the real atmosphere. The latter underscores the significance of the NLHC in the structure of the tropical state. The total diabatic circulation pattern is heavily modified in the tropics, with the upwelling maximum shifting into the summer hemisphere.

The results presented here based on a realistic $T_{rad}$ distribution lead to different conclusions from those of Dunkerton (1989) regarding the mesosphere. Dunkerton found that extending the domain in his model to 84 km from 42 km along with the thermal driving gave unrealistically strong easterlies with a peak magnitude of $-280$ ms$^{-1}$ near the equator (see his Fig. 19d). It is not clear what form $T_{rad}$ took in his extended model but clearly the excessive easterlies were its artefact. In the model used here the easterlies did not exceed a value of $-135$ ms$^{-1}$ and formed a subtropical jet in the summer hemisphere. This amplitude is greater than the observed winds (by about a factor of two), as can be expected from the absence of mesospheric wave drag in the model. In the tropics the easterlies were weaker (less negative) than $-30$ ms$^{-1}$ over most altitudes. Dunkerton (1991) used a $T_{rad}$ with more reasonable characteristics but its impact was obscured by the use of a sponge layer in the mesosphere to represent gravity wave drag.

Dunkerton (1991) argued that the formation of the saddle point in $\overline{m}$ near the stratopause was the result of mesospheric friction and extratropical wave drag. It is clear that the flattening of $\overline{m}$ contours which occurs in the stratopause region most prominently during solstices is due to significant cross-equatorial flow into the winter hemisphere. However, the location of this flow is partly the result of the concentration of the thermal driving producing the NLHC near the stratopause. The largest meridional flow component of the NLHC is in the mesosphere where the NLHC begins to close off as its thermal driving diminishes. The meridional velocity produced by the NLHC is too weak over most of the stratosphere, especially in the lower part, to bend over $\overline{m}$ isopleths.

Extratropical wave drag affects the tropics through the WDC in an opposing manner to that of tropical zero-phase-speed wave drag. Streamlines of the WDC have a significant horizontal component in the tropics in the course of the annual cycle, and the NLHC makes them even flatter in the upper stratosphere. This translates to more cross-equatorial flow and more deformation of tropical $\overline{m}$ isopleths towards the winter
hemisphere. In addition, planetary-wave drag reaches deeper into the tropics in the upper stratosphere and mesosphere, which should contribute to the observed flattening of $\bar{m}$ isopleths.

The reason why severe $\bar{m}$ isopleth deformation does not extend through the mesosphere in reality (e.g. CMAM, CIRA) is presumably the increasing impact of near-zero-phase-speed equatorial and gravity-wave drag which acts to stabilize $\bar{m}$. Similar stabilization by wave drag occurs in the stratosphere as well, so that the significant $\bar{m}$ homogenization seen in the unforced balance model in the subtropics does not occur in nature.

The direct impact of the WDC on the NLHC is limited by the nature of the system, which prevents the WDC from developing a meridional gradient near the equator. As a result, the WDC cannot affect the temperature near the equator in a way which counteracts the formation of the diabatic heating that drives the NLHC. Another way to look at this is to start from the quasi-geostrophic model where the WDC and the NLHC are completely independent (assuming a linear dependence of the diabatic heating of the temperature) whence it becomes evident that any coupling is due to nonlinearity. Except for regions such as the tropical stratopause, nonlinearity is relatively weak so that the $\bar{m}$ distribution retains a coherent structure, which implies that the features of the WDC and NLHC remain intact, with some deformation, when they coexist. The deformation in the WDC and the NLHC follows the deformation in $\bar{m}$.

A potentially critical limitation of the balance model used here is the inertial adjustment scheme. Its appropriateness has been tested in section 3.3.4 by direct comparison with a primitive equations model. The balance model is remarkably accurate, although it fails to capture the zonally asymmetric component of inertial instability which appears to mix absolute angular momentum across the equator and in the summer hemisphere. With regard to the real atmosphere, however, there are clearly missing processes. Compared to the CIRA and CMAM $\bar{m}$ distribution the angular-momentum-conserving balance model adjustment scheme is too strong. Part of this is the result of the absence of important wave drag in the balance model. However, it also could be related to the physics of inertial adjustment which may not be angular momentum conserving and could be more diffusive in nature. Certainly in CMAM, inertial instability could be contributing significantly to vertical diffusion due to the fact that it occurs at the grid scale. This highlights
the limitations of GCMs which have low resolution, producing inertial adjustment that is unrealistically coarse in both space and time.

In CMAM, gradient-wind balance starts to break down in the stratopause region and above due to the strong cross-equatorial flow (Fig. 3.10). The processes that maintain balance, like inertial instability, appear to lag the advective redistribution of $\overline{\mathbf{u}}$. If the balance state is only partly achieved, one would expect the NLHC to be weaker than predicted by the balance model. Based on a simulation in a spectral mechanistic model (Saravanan, 1992) this does not appear to be the case over most of the stratosphere. However, the breakdown of balance in the mesosphere is significant as indicated by the formation of large adjustment-related circulation cells. Planetary wave breaking in the stratopause region and above also acts to produce inertial instability (O'Sullivan and Hitchman, 1992). This effect is not captured in the mechanistic model experiment comparison but may be important in driving the system away from balance in the mesospheric tropics.
Chapter Four

Tropical upwelling

4.1 Introduction

The dominant feature of the tropical branch of the Brewer-Dobson circulation is upwelling. Diabatic ascent in the tropics accounts for most of the transport of tropospheric air parcels and their load of chemical constituents into the stratosphere at altitudes above the tropical tropopause. Observations of water vapour and methane transport (Mote et al., 1996) point to the persistence of tropical upwelling over the course of the seasonal cycle, leading to a positive annual mean upwelling. Even though great strides have been made in the understanding of the dynamics of the diabatic circulation, the origin of annual mean tropical upwelling is unresolved.

The current paradigm for tropical stratospheric transport is that planetary wave drag distributed in midlatitudes and subtropics acts like an extratropical "vacuum pump" pulling air parcels upward and poleward in the tropics through the induced nonlocal diabatic circulation (Holton et al., 1995). This circulation is also argued to be the agent by which extratropical wave forcing exerts nonlocal control on tropical tropopause temperatures (Yulaeva et al., 1994).

There is a problem with this picture, however, since it cannot explain either the observed annual mean upwelling or the location of the upwelling maximum. Wave drag confined to the extratropics of one hemisphere, as is the case for the bulk of stratospheric wave drag, is unable to produce a circulation with upwelling maximizing in the opposite hemisphere. In addition, the diabatic circulation driven by such a wave drag distribution, under certain simplifying assumptions (see section 4.2), is constrained by annual mean downward control to be confined to the subtropics and higher latitudes.

71
Annual mean upwelling can be obtained with either tropical wave drag or nonlinearity acting to circumvent the annual mean downward control restriction. Plumb and Eluszkiewicz (1999) (hereafter PE) proposed weak tropical wave drag as a mechanism for annual mean upwelling in the lower stratosphere. They used a QG model with viscous damping to represent tropical wave drag. As discussed in Chapter 2, there are negative aspects to using relaxational drag. Their model used a uniform radiative equilibrium temperature so it did not form the NLHC. Consequently, the potential impact of the NLHC on the WDC, through the $\overline{m}$ distribution, was not addressed. Sankey (1998) included the effect of nonlinearity in similar numerical experiments to PE and found it to have a suppressing effect on annual mean upwelling. However, this conclusion was based on the comparison of a linear and nonlinear model with viscosity.

The need for a tropical thermal forcing to explain the latitudinal profile of the seasonal tropical upwelling was pointed out by PE. The results of Chapter 3 have demonstrated that the NLHC indeed plays an important role in the structure of the Brewer-Dobson circulation and temperature in the tropics.

In this chapter we investigate the previously unaddressed role of the nonlinear Hadley circulation in annual mean tropical upwelling, and examine the impact of the NLHC on the WDC. We also consider the (nearly) inviscid nonlinear mechanism of annual mean tropical upwelling due to the WDC. The numerical results presented here are based on the model described in section 3.2. Section 4.2 deals with upwelling driven by the NLHC. The nonlinear mechanism for wave-driven upwelling is considered in section 4.3. The effect of the tropical $\overline{m}$ distribution on the wave-driven circulation is treated in section 4.4. In section 4.5 the idealized model results are compared to the CMAM GCM.

### 4.2 Annual mean tropical upwelling

A property of the inviscid linear TEM system is that its annual average has the same form as the steady state in the absence of interannual variability in wave drag and radiative forcing. To a first approximation, then, isolation of the tropics from the bulk of the wave drag in the stratosphere, except possibly in the upper layers where the surf zone penetrates closer to the equator, precludes annual mean upwelling in the tropics. Rather, the circulation exhibits downward control and the streamlines are confined to higher
latitudes. During the year, however, the system is not required to reach steady state and the circulation can penetrate deep into the tropics and across the equator.

This annual mean downward control behaviour may be demonstrated as follows. Consider the annual mean $\langle \cdot \rangle$ of the linear TEM system driven by annually periodic $\mathcal{F}$ and $\overline{T}_{\text{rad}}$. The zonal wind equation (2.1) then takes the form

$$f \langle \bar{v}^* \rangle = \langle \mathcal{F} \rangle,$$  

(4.1)
since the tendency term averages to zero and there are no correlation terms that would arise from nonlinear advection. The Coriolis factor $f$ is zero only at the equator, so if $\langle \mathcal{F} \rangle$ is zero in some latitude range then so must be $\langle \bar{v}^* \rangle$. Without viscosity, mechanical forcing is guaranteed to be zero outside the latitudes of wave drag. Then the time-averaged mass continuity equation (3.4) implies

$$\frac{1}{\rho} (\rho \langle \bar{w}^* \rangle)_z = 0$$

in this region since $(\langle \bar{v}^* \rangle \cos \phi)_\phi = 0$. This equation can be solved by integrating towards the surface from $z = \infty$ where $\rho \langle \bar{w}^* \rangle = 0$ (otherwise there would be an unphysical increase of $\langle \bar{w}^* \rangle$ faster than $\rho^{-1}$ with height). Hence, $\langle \bar{w}^* \rangle$ is also zero throughout the region. Thus there cannot be a net annual mean upwelling at these latitudes.

The inviscid QG form of the balance TEM system supports a linear Hadley circulation (LHC). As in the nonlinear system it is a consequence of the thermal-wind relation and tropical solar heating gradients. However, this circulation is not angular momentum conserving since the $\bar{m}$ profile is fixed. In the absence of mechanical damping the LHC will produce unrealistically persistent acceleration of the zonal wind (through the Coriolis torque term) preventing the system from approaching steady state under fixed solstitial conditions.

As can be seen by setting $\langle \mathcal{F} \rangle$ to zero in (4.1), the linear analogue of the NLHC cannot produce annual mean upwelling without interannual variability. This follows from the behaviour of the diabatic heating, which alternates in sign from one solstice season to the next in response to the tropical solar heating gradient, and which averages to zero in the annual mean. So if the NLHC is to contribute to annual mean upwelling it will have to be through nonlinearity or viscosity.
4.2.1 Role of viscosity in annual mean tropical upwelling

The behaviour of the WDC and the NLHC in the linear TEM system is modified by viscosity. Viscosity is a relaxational mechanical forcing which responds to the zonal wind changes produced by the diabatic circulation or wave drag. In the case of the WDC the effect is to allow the circulation to cross $\bar{m}$ ($f$ in the linear case) contours in steady state. Streamlines can then extend outside the latitudes of the wave drag in both steady state and the annual mean. Viscosity also enables the linear analogue of the NLHC to have a steady state associated with it, although the approachability of this state during solstice seasons depends on the value of the viscosity coefficient.

As pointed out in Dunkerton (1989), for the NLHC to evolve to steady state requires either $\bar{m}$ advection or some form of relaxational mechanical forcing to permit sufficient flow across isopleths of $\bar{m}$. However, the steady state for the linear viscous system does not resemble that reached in the nonlinear system (Dunkerton, 1989, Fig. 19), especially in the upper stratosphere. So nonlinearity is an essential element in the dynamics in these regions.

It was shown by Plumb and Eluszkiewicz (1999) that a significant annual mean tropical upwelling can arise from the WDC even for weak viscosity, provided the wave breaking region extends close enough to the equator. If the forcing falls within an equatorial “boundary layer” then the spreading of streamlines of the force-induced circulation across the tropics is significantly enhanced. The width of this boundary layer is given by $L_R P^{1/4}$ where $L_R$ is the equatorial Rossby radius and $P$ is the ratio of radiative to viscous relaxation times. As will be shown in section 4.4, this effect is linked to parabolic behaviour of the streamfunction operator at the equator.

In contrast, the LHC makes no contribution to annual mean upwelling in the presence of viscosity. This follows from equation (9) of Plumb and Eluszkiewicz (1999) with $D$ and $\theta_p$ set to zero. The origin of this difference is that wave drag leaves an annual mean signature in the linear system while the NLHC diabatic heating does not. The mechanical forcing produced by viscosity reflects this. This leads to the conclusion that annual mean upwelling driven by the NLHC diabatic heating $Q_{NLH}$ is only possible with nonlinearity.
4.2.2 NLHC-driven tropical upwelling

The tendency for streamlines and $\bar{m}$ isopleths to deform together in the nonlinear system has a significant effect on the behaviour of tropical upwelling. In the case of the NLHC, as seen in Chapter 3, this leads to a deflection of the streamlines away from the equator into the winter hemisphere. As a result there is an asymmetry in $\psi$ which does not average out in the annual mean. The behaviour of the circulation induced by subtropical wave drag is more complicated since it has both a mechanical and a diabatic heating component. In the linear problem the diabatic heating component acts to cancel out upwelling produced during the active phase of the forcing in the course of the annual cycle, at latitudes outside of the forcing region. But it will be shown below that subtropical wave drag can produce an annual mean tropical upwelling through nonlinearity.

The annually averaged $\bar{w}$ distribution from the section 3.3 simulation without wave drag but with the comprehensive radiative transfer scheme used previously is given in Figure 4.1a. There is a region of upwelling maximizing near the stratopause which is flanked by regions of downwelling. Most of the upwelling occurs in the mid to upper stratosphere where the advection of $\bar{m}$ is greatest, which also explains the increasing latitudinal span of the region of net flow with height.

There appears to be a separation of the dynamics between the upper and lower tropical stratosphere. The region of annual mean upwelling almost disappears below approximately 27 km (see blow-up in Figure 4.1b) indicating a transition to a predominantly linear regime. There is a change in the latitudinal structure of the upwelling at this level which appears to argue against simple density attenuation as the explanation for the decrease. It should be noted that this pattern depends on the radiative transfer scheme used. If the radiative scheme is Newtonian cooling with uniform damping then the annual mean upwelling region penetrates deeper into the lower stratosphere (to around 21 km) and higher into the lower mesosphere, reflecting the deeper nature of the diabatic heating distribution (cf Figure 3.1d).

A difference is also evident in Figure 4.1a between the pattern of upwelling in the mesosphere and stratosphere. Above the stratopause, the single upwelling maximum at the equator splits in two and moves away from the equator. The model mesospheric pattern reflects the dominant contribution of the zonal-wind-shear-driven inertial instability.
Fig. 4.1 Annual mean vertical velocity for the balance model simulation (average over year 3). (a) Upper stratosphere and mesosphere \( (2 \times 10^{-4} \text{ m s}^{-1}) \). (b) Lower stratosphere \( (3 \times 10^{-5} \text{ m s}^{-1}) \).

to the diabatic heating, evident in the split in the heating dipole (Figure 3.5d) at these altitudes.

Figure 4.2 shows the annual cycle of upwelling at the equator for different altitudes for the same simulation. There is persistent but periodically modulated upwelling near the equator throughout the seasonal cycle in the upper stratosphere. Below about 30 km the upwelling begins to undergo periodic sign reversals. Peak upwelling occurs earlier in the year and is stronger at higher levels (e.g. 228 days into the year at 40 km compared with 257 days at 29 km). This can be explained partly by the shortening radiative damping timescale with height. But there is also a lagged penetration of the circulation produced near the heating maximum in the upper stratosphere and lower mesosphere down to lower levels. The mass streamfunction at lower altitudes is primarily driven by the regions of weak \( Q_{NLH} \) there, as can be seen from the increasing streamfunction amplitude in the lower stratosphere. But at these levels \( Q_{NLH} \) is not strong enough to produce significant \( \bar{\omega} \) isopleth mobility and persistent upwelling.
4.2 Annual mean upwelling

Fig. 4.2 Balance model vertical velocity at the equator versus time for three different altitudes: 29 km (solid), 35 km (dash), and 40 km (long dash).

Fig. 4.3 Parcel trajectories for a three year advection by the balance model residual circulation in the case with no wave drag. Parcels are initially placed on the equator in the middle of January at 23.3 km (solid), 36.7 km (dash) and 50 km (long dash).
4.2.3 Parcel trajectories

Some sense of the effect of the NLHC on tropical diabatic transport can be obtained from parcel trajectories in the meridional plane. We use the meridional circulation field from the simulation of Chapter 3 integrated over several years beginning in January. Parcels initially on the equator move off into the SH and circulate with a net descent (Fig. 4.3). The latitudinal span of the circulation and net movement of the parcels off the equator increases with initial parcel altitude. On the other hand, parcels initially located off the equator at 12 degrees in the NH oscillate about the equator in latitude with increasing amplitude as they drift upward (Fig. 4.4). For parcels initially at 12° in the SH the behaviour is much the same as for those initially on the equator.

The apparent meridional asymmetry in the NLHC indicated by the parcel trajectories in Figure 4.3 is a consequence of the initial state. If the meridional circulation is started from July then a similar trajectory pattern develops but with initial latitudes mirrored about the equator. The initial-state-dependent asymmetry is not particular to solstice seasons: parcels initialized during the spring and fall equinoxes also do not behave the same on opposite sides of the equator. This emphasizes the difference between the Lagrangian-mean and Eulerian-mean upwelling.

For parcels with initial positions at latitudes well away from the equator (towards the subtropics) in the middle stratosphere, the overall tendency is less net upward drift
4.2 Annual mean upwelling

and recirculation. This is due to the decreasing $m$ mobility at higher latitudes which translates into increasing linearity. At these latitudes, parcels feel a periodic circulation with little or no annual mean flow. The increase in the latitudinal excursion of parcels with altitude reflects the streamline behaviour. The meridional velocity, $\overline{v}$, is large in the upper stratosphere but still has no annual mean component. It is however possible for equatorial parcels to enter the midlatitudes during the winter season in the middle stratosphere, effectively extracting them from the "tropical pipe" (Plumb, 1996).

From the tendency of parcels to move downward apart from a very narrow latitude range on one side of the equator, it is evident that the WDC is required to account for the broad band of tracer upwelling in the tropics (Mote et al., 1998). Even though there is a broad region of annual mean upwelling due to the NLHC in the upper stratosphere, it is flanked by downwelling regions. As seen by the degree of latitudinal excursion of parcels at these altitudes in Figure 4.3, there will therefore be a tendency for parcels to move down away from the equator. (Certainly, the behaviour away from the equator can be expected to change in the presence of the wave-driven circulation.)

In Figure 4.5 is shown the evolution of many parcels initially situated in a band about the equator. An upward and poleward drift is evident in the southern hemisphere for parcels initially situated close to the equator. Due to the choice for the initial circulation state, the bulk of the particles get advected into the summer hemisphere where the group that is the furthest from the equator experiences net downwelling. If the snapshots of parcel positions six months out of phase are also considered (not shown here) a sloshing about the equator is apparent.

Individual parcel trajectories in the presence of subtropical forcing are presented in Figures 4.6 and 4.7. The initial positions are the same as before. The form of the wave drag is given in the next section. It has a magnitude of $-1.5 \, \text{ms}^{-1}\text{day}^{-1}$ with the peak value at 35 km and 30° latitude. The forcing is applied for half the year overlapping the winter season. The presence of the WDC alters the behaviour of parcels initially on the equator. In the lower stratosphere there is a significant enhancement of the upward drift. In the middle stratosphere parcels are pulled completely out of the tropics into the winter hemisphere. There is little effect in the upper stratosphere since the forcing distribution used does not produce a significant WDC there.
Fig. 4.5 Plot of the positions of $10^4$ particles at the beginning of the integration (upper left) at the end of year one (upper right), two (lower left) and three (lower right). Particles are initially placed on a uniform grid ($101 \times 100$) between $-20^\circ$ and $20^\circ$ latitude and between 25 km and 35 km altitude.

For parcels starting at $12^\circ$N the drift towards the pole is clearly evident except for the parcel in the upper stratosphere which is put on a trajectory downward in the summer hemisphere. Once again, this is due to the weakness of the WDC at these levels, but indicates that parcel trajectories are sensitive to relatively small path alterations. It also points to the requirement for a realistic wave drag distribution to produce the correct transport behaviour.

The pattern of transport of many parcels initially situated in a band about the equator (Fig. 4.8) is modified considerably. Instead of a sloshing about the equator with a slow drift upward close to the equator, as in the no-wave-drag case (Fig. 4.5), there
is now a stronger upwelling near the equator and extraction of parcels into the winter hemisphere midlatitudes. There is still a net downwelling in the summer hemisphere for parcels situated furthest from the equator. This indicates that the WDC, in this model case, is not producing a significant effect on the circulation pattern at these latitudes. Tracer distributions at annual intervals six months out of phase with those presented
Fig. 4.8 Same as Figure 4.5 but for the case with wave drag.

indicate a large amount of vertical distortion in the parcel "cloud". There is a certain degree of reversibility in the tracer transport in the system which can be seen in the course of the annual cycle.

4.3 WDC-driven tropical upwelling

The limited contribution of the NLHC to annual mean upwelling in the lower tropical stratosphere necessitates consideration of other mechanisms. One such mechanism could be nonlinearity in the WDC. This mechanism for annual mean tropical upwelling is not predicated on viscosity. For simplicity, as with previous studies (Plumb and Eluszkiewicz, 1999; Sankey, 1998) the diabatic heating is approximated by Newtonian cooling with
the radiative equilibrium temperature set to a uniform constant value to suppress the formation of the NLHC and inertially unstable regions. This is also partly justified by the nearly linear behaviour of the NLHC at lower altitudes.

The response of the linear and nonlinear systems to an idealized wave drag poleward of 5° and 25° from the equator was considered by Sankey (1998). That study was an extension of the work of PE to include nonlinear effects. The forcing had a simple localized distribution with a sin² time dependence with an annual period and a magnitude of $-10 \text{ ms}^{-1}\text{day}^{-1}$. For wave drag approaching the equator to within 5°, Sankey found strong upwelling extending deep into the opposite hemisphere, in both the linear and nonlinear versions of the model. The nonlinear system, however, produced a weaker mean upwelling in the tropics than did the linear system. The strong annual mean upwelling in the linear system is due to the wave drag overlapping the viscous boundary layer described by PE. With the forcing situated poleward of 25° the region of upwelling was found to diminish significantly and to be confined close to the forcing region both with and without nonlinearity. Sankey did not consider the role of viscosity in the upwelling produced by the nonlinear viscous model.

It happens that there is a qualitatively different response when the equatorial edge of the forcing region is situated at latitudes intermediate to those considered by Sankey (1998), e.g. 15°. To demonstrate this, the balance model is used with a uniform radiative equilibrium temperature of 240 K and an imposed wave drag distribution of the form $\mathcal{F}(\phi, z, t) = F_0 Y(\phi)Z(z)H(t)$. The forcing magnitude, $F_0$, is $-3 \text{ ms}^{-1}\text{day}^{-1}$. The latitudinal dependence is defined by

$$ Y(\phi) = \begin{cases} 
\frac{1}{2} - \frac{1}{2} \cos\left(\pi \frac{\phi - \phi_1}{\phi_2 - \phi_1}\right) & \text{for } \phi_1 \leq \phi \leq \phi_2 \\
\frac{1}{2} + \frac{1}{2} \cos\left(\pi \frac{\phi - \phi_2}{\phi_3 - \phi_2}\right) & \text{for } \phi_2 \leq \phi \leq \phi_3 \\
0 & \text{otherwise}
\end{cases} \quad (4.2) $$

with $\phi_1$ being the subtropical limit of the forcing, $\phi_2$ the latitude of its maximum, and $\phi_3$ its northern limit. The vertical dependence is defined by

$$ Z(z) = \begin{cases} 
\frac{1}{2} + \frac{1}{2} \cos\left(2\pi \frac{z - \frac{1}{2}(z_1 + z_2)}{z_2 - z_1}\right) & \text{for } z_1 < z < z_2 \\
0 & \text{otherwise}
\end{cases} \quad (4.3) $$
with $z_1 = 16$ km the lower limit and $z_2 = 50$ km the upper limit of the forcing region. The time dependence is defined by

$$H(t) = \begin{cases} 
\frac{1}{2} - \frac{1}{2} \cos \left( \frac{\pi}{t_2 - t_1} t_1 \right) & \text{for } t_1 \leq t \leq t_2 \\
1 & \text{for } t_2 \leq t \leq t_3 \\
\frac{1}{2} + \frac{1}{2} \cos \left( \frac{\pi}{t_4 - t_3} t_3 \right) & \text{for } t_3 \leq t \leq t_4 \\
0 & \text{otherwise}
\end{cases} \quad (4.4)$$

The forcing ramps up between day $t_1 = 0$ and day $t_2 = 46$, persists, and then diminishes between day $t_3 = 137$ and day $t_4 = 183$. The results should be qualitatively similar for all localized and single-signed subtropical forcing within the same latitudinal limits.

The model horizontal diffusion is reduced by a factor of ten from the values used in Chapter 3 (and can be removed altogether) since the simulations with uniform $T_{rad}$ are more numerically stable, primarily on account of the absence of inertial instability (a consequence of the inability of wave drag situated off the equator to induce a temperature gradient at and near the equator through the WDC, as discussed in Chapter 3).

Latitudinal profiles of the annual mean $\bar{w}$ for the linear and nonlinear cases at 34 km, the level of the forcing maximum, are given in Figure 4.9 for $\phi = 15^\circ$. At this latitude, the forcing region is not overlapping the viscous equatorial boundary layer described by Plumb and Eluszkiewicz (1999). This can be verified explicitly by making the approximation

$$\alpha_u \bar{u} \sim K_{zz} \frac{\partial^2 \bar{u}}{\partial z^2} + K_{yy} \frac{\partial^2 \bar{u}}{\partial y^2}$$

so that

$$\alpha_u \sim \frac{K_{zz}}{L_z^2} + \frac{K_{yy}}{L_y^2}$$

where $L_z$ and $L_y$ are the vertical and meridional scales of $\bar{u}$. For the simulations done here $K_{zz} = 0.3 \, \text{m}^2\text{s}^{-1}$, $K_{yy} = 500 \, \text{m}^2\text{s}^{-1}$, $L_z \approx 10000 \, \text{m}$ and $L_y \approx 3 \times 10^5 \, \text{m}$. So the vertical diffusion dominates and $\alpha_u \approx 0.00074 \, \text{day}^{-1}$ which gives a viscous boundary layer width of $L_c \sim \left( \alpha_u \right)^{1/4} L_R \approx 650 \, \text{km}$ or $5.8^\circ$ ($L_R \approx 2500 \, \text{km}$).

\[1\] These length scales are derived from consideration of the actual values of diffusion in the model.
Fig. 4.9 Meridional profile of the annual mean $\overline{w}^*$ at 34 km in the sixth year of the simulation for the linear (solid) and nonlinear (dash) balance model. $\phi_1 = 15^\circ$, $\phi_2 = 30^\circ$ and $\phi_3 = 80^\circ$.

The fact that the forcing lies outside the equatorial boundary layer is also seen by the fact that in the linear case the upwelling region is confined in latitude at the equatorial edge of the forcing region. In the nonlinear case, by contrast, there is a region of weak upwelling penetrating into the opposite hemisphere. The balance of terms in the zonal wind equation shows that viscosity is not responsible for the formation of this feature. This is in contrast to the situation with $\phi_1 = 5^\circ$ when near-equatorial upwelling in the linear case is stronger than in the nonlinear case (similar to the results of Sankey (1998)).

If we consider the time dependence of $\overline{w}^*$ at 34 km on the equator (Fig. 4.10), then it is apparent that in the nonlinear case the upwelling phases dominate the downwelling phases. It is also evident that the nonlinear system reaches equilibrium faster than the linear system from the time the periodic forcing is turned on; the linear evolution asymptotes over a much longer period to a situation where the upwelling and downwelling phases are balanced. In both systems there is still some annual mean transience in the tropics even after six model years of simulation, but the balance of terms in the tropics implies that this transience is not responsible for the upwelling pattern in the nonlinear case.
Fig. 4.10 Time dependence of $w^*$ at 34 km on the equator for the simulation of Figure 4.9 with the linear model (solid) and nonlinear model (dash).

An important parameter implicit in the above is the radiative damping rate. Using a much smaller value of $\alpha$ (0.025 day$^{-1}$ instead of 0.17 day$^{-1}$), more characteristic of the lower tropical stratosphere, results in a much longer transient phase. After the same simulation period of six years, the upwelling in the linear case is still almost equal to that of the nonlinear case (Fig. 4.11). However, the enhanced transience does not increase the amplitude of the upwelling for forcing located further away from the equator (Fig. 4.12). The viscous boundary layer half-width in the small $\alpha$ case is about 1000 km or 9.3°, so that no overlap occurs with the two forcing distributions used. For this reason the linear model has no annual mean upwelling on the equator after the switch-on transients have decayed.

It should be noted that the balance of terms in the linear and nonlinear models is different. In the linear model the balance in the annually averaged zonal momentum equation is between the acceleration and the Coriolis term in the tropics. By the continuity equation, then, it is the annual mean tendency that drives the upwelling. However, in the nonlinear model the balance of terms indicates that transience is not responsible for the mean upwelling, which in fact becomes slightly higher after the start-up transients have died away (not shown).
4.3 WDC-driven upwelling

![Graph](image)

**Fig. 4.11** Meridional profile of the annual mean $\bar{\omega}$ at 34 km in the sixth year of the simulation for the linear (solid) and nonlinear (dash) balance model with $\alpha = 0.025 \text{ day}^{-1}$. The forcing is the same as for Figure 4.9.

![Graph](image)

**Fig. 4.12** Same as Figure 4.11 except $\phi_1 = 22^\circ$, $\phi_2 = 35^\circ$ and $\phi_3 = 80^\circ$.

The net annual mean upwelling in the nonlinear system seen in Figure 4.9 points to major differences in the pattern of evolution compared to the linear system. The diabatic heating induced during the active forcing phase produces a downwelling during the inactive phase (the time dependence used here is for clarity; a more slowly varying
forcing will still be subject to the same constraint), which in the inviscid linear system cancels the upwelling in the annual mean. In the nonlinear system this heating experiences a deformation as it decays (Fig. 4.13), which increases the rate of attenuation of the associated \( \bar{w} \) in the tropics. In contrast, the diabatic heating region in the linear system
is much more spatially fixed and takes a longer time to decay. At a given latitude of the tropics outside the forcing region, the upwelling phase will not be cancelled by the downwelling phase in the nonlinear system so that annual mean downward control will not apply.

To test whether diffusion was important for the nonlinear system the diffusion coefficients were reduced to 0.02 m$^2$s$^{-1}$ for the vertical and 5 m$^2$s$^{-1}$ for the horizontal. This had very little effect on the annual mean upwelling profile. So diffusion cannot explain
Fig. 4.15 Annual mean of $\bar{w}$ on the 34 km level after 12 years of simulation for the inviscid linear (solid) and nonlinear (dash) models.

The above results demonstrate that it is possible for subtropical wave drag no closer than 15° to the equator to produce annual mean upwelling through a nonlinear mechanism. However, the upwelling produced is too weak to explain the observed values in the lower tropical stratosphere even though the forcing amplitude is larger than typical values for subtropical wave drag (Rosenlof, 1995, Fig. 8). It must be noted that a more realistic force distribution would produce a stronger upwelling due to contributions from other levels which are not forced in the model. However, for upwelling at a given level, density stratification limits the contribution from higher levels (assuming that any force magnitude increase with height is much slower than exponential — as appears to be the case in the atmosphere). One of the assumptions underlying the work here is that the wave drag is confined to the surf zone, which at these altitudes is located further from the equator than in the upper stratosphere. While there is strong evidence of the isolation of the lower tropical stratosphere based on tracer observations (Grant et al., 1996), this does not necessarily imply that there is isolation in terms of wave drag. In fact, GCMs such as CMAM do not have the wave drag confined to the surf zone and there is weak EP-flux divergence extending to the equator.
Fig. 4.16 Distribution of $\overline{m}$ for the inviscid nonlinear model at $t = 105$ days in year 50 of the simulation.

Therefore, the above experiments were repeated to see what effect a forcing distribution extending closer to the equator and having a more realistic magnitude ($-1.5 \text{ ms}^{-1}\text{day}^{-1}$ at $30^\circ$, decaying to zero at $6^\circ$) would have on the upwelling in the absence of explicit model diffusion (grid diffusion is unavoidable). The results are given in Figures 4.14 and 4.15 for the 12 model years of the simulation. The linear model equatorial upwelling shows a monotonic trend to a state with little or no upwelling, whereas the nonlinear model upwelling appears to equilibrate about a positive value reached after half the integration time. It appears that grid diffusion is playing a minor role possibly giving a very small annual mean upwelling in the linear model in the long-time limit. There is some interannual variation in the nonlinear case (Fig. 4.14). The shorter radiative damping rate ($1/6 \text{ day}^{-1}$) had little effect on the duration of this feature.

A constraint on the possible distributions of the imposed wave drag is the degree of $\overline{m}$ deformation that results. PE argued that the pattern of $\overline{m}$ isopleths in the lower tropical stratosphere is an indication of a linear regime. Considering the $\overline{m}$ field at the peak deformation by the forcing in the inviscid experiment (Fig. 4.16), the deflection of isopleths is limited except near the equator. The pattern would at first appear to differ from that of CMAM but the comparison is misleading since the contour with the largest value (at this contour interval) closes off near the tropopause in the latter, indicating an
erosion of absolute angular momentum near the equator or weak easterlies. Consequently, nonlinearity cannot be excluded from consideration based on analysis of the \( \overline{m} \) field alone.

However, the spatial distribution of the forcing used in the inviscid experiment above still does not produce the right magnitude of annual mean upwelling on the equator. The CMAM EP-flux divergence of resolved waves points to wave drag extending up to the equator as one possibility, which we will consider below. Another possible explanation is the \( \overline{m} \) distribution in the idealized model with uniform \( \overline{T_{ra}} \), which lacks the equatorial asymmetry of the CMAM \( \overline{m} \) field in the lower stratosphere. As will be shown in the next section, the \( \overline{m} \) distribution affects the intensity of tropical upwelling.

The results of Chapter 3 indicate that the equatorial asymmetry of \( \overline{m} \) in the lower stratosphere is not a feature of the radiative equilibrium state. This asymmetry is likely the result of the pattern of zonal wind acceleration in the subtropics by higher latitude wave drag through the diabatic circulation. This is an angular-momentum-conserving process since it is formed through advection of the \( \overline{m} \) maximum off the equator by a relatively weak cross-equatorial flow (indicated by the horizontal component in CMAM diabatic circulation streamlines over the equator). The apparent uniformity of the \( \overline{m} \) asymmetry with height is consistent with the distribution of extratropical wave drag which extends over a deep layer. Since the resolved wave drag in CMAM is negative in the winter hemisphere subtropics, it cannot produce such an asymmetry in the \( \overline{m} \) profile through an acceleration of the zonal wind directly.

In Figures 4.17, 4.18 and 4.19 are presented the results of an inviscid model run with the forcing extending to the equator. The upwelling magnitude is almost double that for the previous experiment in the nonlinear case, while the linear model is now producing similar levels of upwelling on the equator. So even without viscosity the behaviour of the linear model is very similar to the viscous case in the region of weak forcing near the equator. From Figure 4.18 it is clear that this is not a transient effect. However at latitudes where the forcing does not reach there is an indication of a slow decay in the annual mean upwelling with time (Figure 4.19).

4.4 Effect of the NLHC on the WDC in the tropics

As noted in Chapter 3, the streamfunction operator is non-elliptic at the equator and possibly adjacent regions. This has implications for the behaviour of the WDC in the tropics.
Just as the WDC can redistribute $\bar{u}$ and $\bar{T}$ in the tropics and modify the evolution of the NLHC, so can the NLHC modify the evolution of the WDC through redistribution of these fields. On the basis of the results of Chapter 3 this effect is expected to be strongest in the upper stratosphere. The NLHC modification of $\bar{m}$ and $\bar{T}$ reflects the influence of $\bar{T}_{rad}$ and thermal wind balance. There are three related but distinct effects acting on the $\bar{m}$ and $\bar{T}$ distributions. One is advection by the NLHC. Another is simply the link between the zonal wind wind distribution and $\bar{T}_{rad}$ through thermal wind balance. Even though the balanced thermal wind is not defined near the equator during solstice seasons when $\bar{T}_{rad}$ has a meridional gradient at the equator, there is an impact in the form of the $\bar{m}$ maximum moving into the winter hemisphere. This appears to be the case even in the lower tropical stratosphere considering the January CMAM $\bar{m}$ field (Figure 3.14). In addition to these two effects, inertial adjustment can also significantly modify $\bar{m}$ and $\bar{T}$. The basic effect of the advective and radiative processes in the tropics is to extend parabolicity in $\mathcal{L}$ to higher latitudes. The impact of such $\bar{m}$ deformations on the WDC is the focus of this section.

We consider the effect of the variable properties of $\mathcal{L}$ on the WDC in a linearized version of the model described in section 3.2 forced with a switch-on $\mathcal{F}$. This approach captures the essence of the instantaneous response in the nonlinear system since the
absence of self-advection means that the WDC only sees the distribution of $\bar{u}$ and $\bar{T}$ which can be regarded as given. However, the linearized model fails to respond to the feedback of changes in $\bar{u}$ and $\bar{T}$ on the evolution of the system so it will differ from the nonlinear one.

Parabolicity in $\mathcal{L}$ spanning a range of latitudes can produce the equatorial behaviour of the WDC at higher latitudes. We consider three highly idealized basic states: one with a broad parabolic region, one with a weak $\mathcal{M}_\phi$ region, and one which corresponds to no zonal flow (see Figure 4.20). The streamfunction and vertical velocity in each of the cases
is presented in Figures 4.21 and 4.22 for switch-on forcing of the form

\[ F(\phi, z) = F_0 \exp \left[ \left( -\frac{\phi - \phi_0}{10^\circ} \right) - \left( \frac{z - 40 \text{ km}}{10 \text{ km}} \right) \right], \]

with \( F_0 = -2 \times 10^{-5} \text{ m s}^{-2} \), centered at \( \phi_0 = 45^\circ \text{N} \), and in Figures 4.23 and 4.24 for the same forcing centered at \( \phi_0 = 30^\circ \text{N} \). As the region of near or exact parabolicity approaches the location of the forcing, both the instantaneous and long-term streamfunction penetrate further away from the forcing region and intensify (compare the panels of

[Fig. 4.19] Same as Figure 4.18 but for a point 11.25° in the hemisphere opposite the forcing.
Fig. 4.20 Profiles in latitude of the reference state $\overline{m}$ for the three cases used to test the behaviour of the streamfunction to different levels of parabolicity: no zonal wind (solid), weak $\overline{m}_\phi$ (dash) and broad parabolic (long dash).

Fig. 4.21). A similar effect occurs when $F$ is moved to lower latitudes (compare Fig. 4.23 and Fig. 4.21).

The expression for $L$ linearized around these height-independent basic states simplifies to

$$\frac{C_{\phi\phi}}{\cos \phi} \frac{\partial}{\partial \phi} \left( \frac{1}{\cos \phi} \frac{\partial \psi}{\partial \phi} \right) + C_{zz} \rho \frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial \psi}{\partial z} \right) = D \quad (4.5)$$

($C_{\phi\phi}$, $C_{zz}$ and $D$ are the same as given in Appendix B but with terms in $M_z$ set to zero.)

If we take $\Delta z$ to be the vertical scale of $\psi$ and $a \Delta \phi$ to be its latitudinal scale, then in regions with $D \sim 0$ scaling analysis as in Holton et al. (1995) gives the relation

$$\Delta z \sim \max \left[ \frac{\cos \phi}{a} \sqrt{\frac{C_{zz}}{C_{\phi\phi}}} a \Delta \phi, \frac{\cos^2 \phi C_{zz} (a \Delta \phi)^2}{a^2 C_{\phi\phi} H} \right] \quad (4.6)$$

where the first relation follows when $\Delta z < H$ and the second when $\Delta z > H$. So in regions of weak $C_{zz}$, such as in the regions with homogenized $\overline{m}$ (where $M_\phi$ is small or zero), the latitudinal scale of the streamfunction has to be broader for the vertical derivative term to be balanced by the latitudinal derivative term. It follows that the streamlines will extend further from $F$. In the case of overlap between $D$ and a parabolic region the two dominant terms are $C_{\phi\phi} \psi_{\phi\phi}$ and $D$ in the region of overlap. This also leads to a broadening of the streamfunction. The effect appears more dramatic than
in the nonoverlapping case discussed above (Figs. 4.21 and 4.22) since the number of streamlines affected is much greater.
Fig. 4.22 Same as Figure 4.21 but for the vertical velocity. The contour interval for panels on the left is $3 \times 10^{-5}$ m s$^{-1}$ and $1 \times 10^{-4}$ m s$^{-1}$ for those on the right.

The change in the number of streamlines or, equivalently, the amplitude of $\psi$ in Figure 4.23 (early time limit) comes from the $\sqrt{M}$ factor multiplying $F$ in the expression for $D$. The basic state $\overline{m}$ profiles used here all produce a different value of $M$ at the latitudes of the forcing. The distribution of the zonal wind in the forcing region affects the streamfunction response through the rhs of the streamfunction equation and not just through the coefficients of $L$. 
The formation of parabolic or weak $\bar{m}_\phi$ regions which extend away from the equator affects the diabatic circulation pattern induced by subtropical wave drag. Not only is there a deeper penetration of streamlines into the tropics, there is an enhancement of the circulation and upwelling as well. If significant homogenization of $\bar{m}$ in the real atmosphere were to occur, then the WDC component of tropical upwelling would be boosted. There would also be a broadening of the tropical “boundary layer” of PE. Weak
$\bar{m}_p$ does indeed occur near the tropical stratopause possibly making this region more sensitive to subtropical wave drag (and leading, perhaps, to an enhanced easterly phase of the SAO).

This effect is a generalization of the fact that the equatorward penetration of $\bar{w}^+$ away from a low-latitude forcing region is greater compared to the response when the forcing region is located at higher latitudes. This fact was noted by Plumb and Eluszkiewicz.
(1999); the tropical "boundary layer" identified by them is associated with the non-ellipticity of the linear system at the equator.

Acceleration of the zonal wind by the diabatic circulation is not possible in regions of homogenized $\overline{m}$. In such regions, it is only through adiabatic heating and the thermal wind relation that the meridional circulation can influence the zonal wind. The system can still evolve to steady state, but streamlines of the circulation are no longer forced to collapse onto $\overline{m}$ surfaces confined to the latitudes of the forcing region.

Non-ellipticity near the equator partly explains the rapidly decreasing rate of evolution to steady state and the breakdown of "downward control". The cause is a different pattern of diabatic heating response to low latitude mechanical forcing compared to the extratropics. Since streamlines of the force-driven circulation are less concentrated below the forcing level, the diabatic heating it induces is more spread out. The circulation is also less effective at inducing a diabatic heating due to the decoupling of $\overline{T}$ from $\overline{u}$ on account of the diminishing Coriolis factor in the thermal-wind relation. In addition, no significant meridional gradients in diabatic heating can be supported near the equator in the balanced meridional response.

The central element of "downward control" is a feedback between the component of the circulation driven by the relaxational diabatic heating, $\psi_Q$, and the distribution of the heating itself. In regions of horizontally stratified $\overline{m}$, the heating quadrupole induced by a localized region of wave forcing induces a circulation which adds to the force-driven circulation component, $\psi_F$, below $\overline{F}$ and diminishes it at other latitudes. In the tropics, $\psi_Q$ is much less effective in this role due to its distribution and intensity. Consequently, the essential feedback in the "downward control" effect breaks down.

4.5 Contribution of the NLHC to upwelling in a middle atmosphere GCM

Since the balance model appears to reasonably capture the intensity of the NLHC (section 3.3), a quantitative inference about the importance of the NLHC to CMAM can be made. The annual mean vertical profiles of upwelling on the equator in CMAM and in the balance model simulation of Chapter 3 using the comprehensive radiative transfer scheme are given in Figure 4.25. The quantitative agreement in the upper stratosphere and stratopause region suggests that a large part of the CMAM upwelling originates from
Fig. 4.25 Annual mean $\overline{w}$ for the nonlinear balance model without mechanical forcing (solid) and CMAM (dash) on the equator.

Streamline and $\overline{m}$ mobility similar to that of the NLHC near the equator. Upwelling in the mid and lower stratosphere, in contrast, is apparently dominated by the wave-driven component since the NLHC component is much too weak.

The increasing deformation of $\overline{m}$ isopleths with altitude (as is apparent from the CMAM and CIRA $\overline{m}$ field) implies that the QG approximation ceases to be valid for the system sufficiently far above the tropopause. QG is based on the assumption that $\overline{m}$ is close to $\Omega a^2 \cos^2 \phi$, which is characterized by vertical isopleths. Below 30 km the $\overline{m}$ distribution has a higher degree of equatorial symmetry and nearly vertical isopleths over most of the tropics. However, there is some deformation of $\overline{m}$ even in the lower tropical stratosphere. Given the unforced model results (see Figure 3.5), this deformation is likely due to subtropical wave drag. Viscosity does not play the same role in the nonlinear system that it does in the linear one, as can be inferred from the balance of terms in the zonal momentum equation. This obviates the need for the linear viscous mechanism of PE.

The nature of the mobility of $\overline{m}$ is such that the NLHC acts in concert with the WDC to produce annual mean upwelling near the equator. The streamfunction of the
4.6 Conclusion

From the work in this chapter it can be concluded that the NLHC makes a direct contribution to upwelling only in the upper stratosphere. At these levels it can explain a
large part of the upwelling seen in the CMAM GCM near the equator. However, the role of wave drag in the tropics remains important at all levels and particularly in the lower stratosphere.

Mechanical forcing is required to explain annual mean upwelling in the lower stratosphere. As seen in section 4.3, the mechanism for this upwelling can be nonlinearity, and appeal to viscosity is not necessary when wave drag is confined away from the equator. However, the observed amplitude of the upwelling suggests that wave drag must penetrate quite close to the equator. Also, in the lower tropical stratosphere there is weak inertial adjustment consistent with the climatological $T_{\text{rad}}$, so it is possible that inertial instability contributes to the formation of a viscous equatorial boundary layer in the real atmosphere. The linear viscous model becomes more relevant if mechanical forcing extends across the tropics.

The inviscid nonlinear upwelling mechanism involves an asymmetric evolution in the force-induced diabatic heating compared to the linear case. The induced diabatic heating
decays more rapidly on account of advection and fails to produce enough downwelling to cancel the force-driven upwelling during the declining phase of the forcing and afterwards. Increased upwelling amplitude in the course of the seasonal cycle can translate into increased annual mean upwelling through greater nonlinearity.

The effect of a reduced meridional \( \overline{m} \) gradient on the WDC considered in section 4.4 is likely important for low-latitude wave drag in the CMAM stratopause region and possibly near the equator in the vicinity of inertial instability. The parabolicity of the streamfunction operator near the region of highly deformed \( \overline{m} \) contours close to the stratopause increases the cross-equatorial flow and upwelling. Some of the resulting enhanced upwelling could reach lower levels but decaying exponentially, e-folding at most over one density scale height. In the lower stratosphere the CMAM \( \overline{m} \) distribution also has a region of weak meridional gradient on the winter side of the equator. This can be expected to enhance the magnitude of the seasonal tropical upwelling, although not as significantly as in the upper stratosphere.

Another mechanism for the formation of weak \( \overline{m}_\phi \) regions in the tropics, not considered here, is the quasi-biennial oscillation (QBO). During the easterly phase of the QBO there may be significant formation of near-parabolic regions at mid to lower levels of the tropical stratosphere. This has the potential to locally enhance tropical upwelling.
Chapter Five

Aspects of transience in the extratropics

5.1 Introduction

One of the characteristic properties of the middle atmosphere is the relaxational nature of the diabatic heating. A consequence of this is the zonal-mean downward control behaviour of the diabatic circulation induced by extratropical wave drag (mechanical and thermal dissipation of waves) (Haynes et al., 1991). The spatio-temporal structure of the response to wave drag is sensitive to the radiative damping rate. If the timescale of the wave drag is much longer than the radiative damping time then the response has a steady-state character; if it is much shorter, then the response has a transient character.

Previous studies of the basic zonal-mean dynamics (Garcia, 1987; Haynes et al., 1991; Holton et al., 1995; Sankey, 1998) have used quasi-geostrophic equations and the Newtonian cooling approximation with constant radiative damping rate for the diabatic heating. These studies considered the effect of varying the magnitude of the damping rate but did not address the influence of spatio-temporal variation in the damping rate. Although varying the value of a uniform radiative damping rate gives a sense of the dependence of solutions on this important parameter, there are features that can only be captured with a non-uniform spatial or temporal distribution.

In fact, the radiative transfer characteristics of the middle atmosphere vary significantly in space and time (Andrews et al., 1987). This stems from the pattern of solar heating and from increasing optical thickness with decreasing altitude. The solar heating maximum and hence the radiative equilibrium temperature maximum migrate from pole to pole during the annual cycle, and very cold radiative equilibrium temperatures form in polar night regions. While dynamical heating tends to mitigate some of the
more extreme annual temperature variations, especially in the polar night, the underlying radiative equilibrium has a major impact on the state of the middle atmosphere. The radiative damping rate is sensitive to the local background temperature: lower background temperatures imply a slower damping rate. As a result, the radiative damping rate has significant latitudinal dependence and has the smallest values at a given height in the winter hemisphere polar regions. There is also a significant variation in the radiative damping rate with height due to density stratification. Larger densities at lower levels translate into greater optical thickness and smaller damping rates. For a Newtonian cooling approximation to capture these effects in the simplest possible form, the cooling coefficient needs to vary with latitude, height and time. A parameterization of the cooling coefficient in terms of temperature would make the system nonlinear and complicate the analysis. However, the quasi-geostrophic formulation of the dynamics as used in previous studies suffices if the problem is treated as some perturbation of a climatological basic state.

The impact of spatial variation of the radiative damping rate on the dynamics will be considered in the context of several problems. One such problem is the influence of wave drag on the polar night temperature, which is important for ozone chemistry, while another is extratropical transport out of the "overworld" into the lowermost stratosphere. The central element in these problems is local deviation from downward control due to weak damping rates. Beyond spatial variation, temporal variation in the radiative damping rate can affect the interpretation of annual mean temperature anomalies and the behaviour of transients that are present with interannual variability.

With regard to the first problem, numerical experiments indicate that gravity-wave drag can have a significant remote influence on the stratospheric southern hemisphere high latitudes during the winter season (Garcia and Boville, 1994). An important element in this result is the very weak damping rates that prevail at these latitudes. Due to the weaker planetary-wave drag in the southern hemisphere, gravity-wave drag plays a relatively more important role compared to the northern hemisphere in producing dynamical heating. This problem is related to the "cold pole" problem of GCMs which have difficulty resolving gravity-wave drag and its effect on the evolution of the model atmosphere. It is not clear which regions of gravity-wave drag have an important effect on the mid-stratosphere polar night temperature. The downward control principle would
suggest that high-latitude wave drag in the mesosphere makes the largest contribution (Haynes et al., 1991). However, with a weak radiative damping rate the concentration of streamlines induced by high-latitude mesospheric wave drag below the forcing level would take longer than a season to occur. So the strength of the downwelling is weaker than downward control would suggest, although the weak damping rate means that the temperature is more sensitive to the circulation.

Part of the effect described by Garcia and Boville (1994) could be due to lower latitude stratospheric wave drag. Most wave drag in the middle and lower stratosphere is believed to be confined outside the polar vortex. While the damping rate may be large enough in midlatitudes for downward control to be achieved on a seasonal timescale, the weak damping rate inside the vortex could prevent the concentration of diabatic circulation streamlines to the equatorial side of the vortex edge. The manner in which the spatial distribution of the radiative damping in the polar night affects the impact of adjacent regions of wave drag therefore requires investigation.

With regard to the second problem, calculation of the seasonal cycle of mass in the lowermost stratosphere (Appenzeller et al., 1996) shows a lag of the mass outflow into the troposphere compared to the peak wave drag above 100 hPa (about 16 km). The mass flux through the 100 hPa surface is an important part of the analysis. The weak radiative damping at these altitudes means that downward control of the diabatic circulation by wave drag is not attained and the pattern of mass flux will reflect this.

In addition, the radiative damping rate is a key factor affecting transients associated with interannual variations in forcing (e.g. wave breaking or ozone heating) in the middle atmosphere. At high latitudes and in the lower stratosphere, transients will last longer due to the smaller radiative damping rate. As is evident from the results presented here, transients can persist longer than a year and therefore can obscure interpretation of annual mean diagnostics. For example, the annual mean temperature increase observed above the region of seasonal ozone loss in the southern hemisphere polar regions (Ramaswamy et al., 1996, Fig. 1) could be partly explained by transience. This is in addition to warming attributable to wave drag redistribution or additional infrared heating from the troposphere (through the cold, ozone-depleted region).

The annual behaviour of transients associated with the formation of the ozone hole will also be affected by the seasonal variation of the radiative damping rate. The time
dependence of the radiative damping rate is greatest in the polar regions owing to the extreme swing in temperatures between the winter solstice (polar night) and summer solstice. In contrast, the variations in tropical temperature and radiative damping rate are smaller as the sun passes over the equator twice a year and the region is always sunlit.

In this chapter we look at the effect of the spatio-temporal dependence of the radiative damping rate on the zonal mean dynamics pertaining to the issues described above. In section 5.2 the model is presented. The effects of spatial and temporal variation in the radiative damping rate on switch-on and steady periodic forcing cases is considered in sections 5.3 and 5.4. In section 5.5 the annual mean signature of temporal variation in the radiative damping rate is examined. Long-lived transients are considered in section 5.6.

5.2 Model

The zonal-mean dynamics of the middle atmosphere can be reasonably modeled using the TEM equations. The Coriolis term typically dominates the nonlinear terms in the zonal momentum equation in the extratropics and the system is close to being in thermal wind balance. Thus for this study, the quasi-geostrophic (linear) version of the TEM equations is employed:

\begin{equation}
\frac{\partial \tilde{u}}{\partial t} - f \tilde{v} = \mathcal{F}
\end{equation}

\begin{equation}
\frac{\partial \tilde{T}}{\partial t} + S \tilde{w} = Q
\end{equation}

\begin{equation}
\frac{1}{a \cos \phi} (\tilde{v} \cos \phi) + \frac{1}{\rho} (\rho \tilde{w})_z = 0
\end{equation}

\begin{equation}
f \tilde{u}_z = - \frac{R}{a H} \tilde{T}_\phi
\end{equation}

where \( f = 2\Omega \sin \phi \) is the Coriolis parameter.

The diabatic heating consists of two parts: a solar heating \( Q_s \) and a temperature dependent long-wave cooling \( Q_l(\tilde{T}) \). A reasonable approximation is Newtonian cooling \( Q \equiv Q_l(\tilde{T}) + Q_s = -\alpha (\tilde{T} - \tilde{T}_{rad}) \) where \( \alpha \) depends on \( \phi, z \) and \( t \). Idealized profiles of the radiative damping rate (Fig. 5.1) are inferred from a comprehensive radiative transfer scheme (Fomichev and Blanchet, 1995) applied to the CIRA climatology. The largest variations in space and time occur at high latitudes where the system moves in and out of polar night with the march of the seasons. The radiative damping rate also
Fig. 5.1 Profiles of $\alpha(\phi)$ (upper panel), $\alpha(z)$ (middle panel) and $\alpha(t)$ (lower panel) in units of 1/day. Profiles are based on sections of January $\alpha$ from a realistic radiative transfer scheme in latitude at $z = 39$ km, height at $\phi = 45^\circ$ and time at $\phi = 60^\circ$ and $z = 30$ km, respectively.

varies significantly with height due to the exponential increase in density with decreasing altitude.

The dependence of $\alpha$ on $\bar{T}$ is not allowed for in this formulation, which makes the system separable into a $\mathcal{F}$-driven and $Q_*$-driven part. The latter can be further
subdivided into $\overline{T}_{rad}$ and a hypothetical perturbation $\delta \overline{T}_{rad}$ (e.g. ozone loss). Imposed diabatic forcing is equivalent to a change in $\overline{T}_{rad}$ also. Defining the wave drag component $\overline{u}_F$, $\overline{T}_F$, $\overline{v}_F$ and $\overline{w}_F$ by

$$\frac{\partial \overline{u}_F}{\partial t} - f \overline{v}_F = \mathcal{F}$$

$$\frac{\partial \overline{T}_F}{\partial t} + S \overline{w}_F = -\alpha (\overline{T}_F)$$

the diabatic forcing perturbation component $\overline{u}_Q$, $\overline{T}_Q$, $\overline{v}_Q$ and $\overline{w}_Q$ by

$$\frac{\partial \overline{u}_Q}{\partial t} - f \overline{v}_Q = 0$$

$$\frac{\partial \overline{T}_Q}{\partial t} + S \overline{w}_Q = -\alpha (\overline{T}_Q - \delta \overline{T}_{rad})$$

and the radiative equilibrium component $\overline{u}_R$, $\overline{T}_R$, $\overline{v}_R$ and $\overline{w}_R$ by

$$\frac{\partial \overline{u}_R}{\partial t} - f \overline{v}_R = 0$$

$$\frac{\partial \overline{T}_R}{\partial t} + S \overline{w}_R = -\alpha (\overline{T}_R - \overline{T}_{rad}),$$

the combined fields satisfy the full equations

$$\frac{\partial (\overline{u}_F + \overline{u}_Q + \overline{u}_R)}{\partial t} - f (\overline{v}_F + \overline{v}_Q + \overline{v}_R) = \mathcal{F}$$

$$\frac{\partial (\overline{T}_F + \overline{T}_Q + \overline{T}_R)}{\partial t} + S (\overline{w}_F + \overline{w}_Q + \overline{w}_R) = -\alpha (\overline{T}_F + \overline{T}_Q + \overline{T}_R - (\overline{T}_{rad} + \delta \overline{T}_{rad}))$$

We are interested in the first two components of the solution. The forcing, $\mathcal{F}$, can be some idealized distribution or based on GCM or observed climatology.

5.2.1 Unresolved dynamics

This model does not include a number of important elements. One is advective nonlinearity, which prevents the separation of solutions into non-interacting mechanically and radiatively forced components. This is justified by the relative weakness of advection terms in the extratropics.

Another form of nonlinearity is in the diabatic heating and wave drag. This also makes the problem non-separable. With a comprehensive radiative transfer scheme for
the diabatic heating, if the system is to be kept in the climatological regime with the right temperature range and $\alpha$ distribution then a realistic $F$ field is required. The wave drag would also respond to changes in the zonal wind and cannot really be regarded as fixed. Since we are interested in the basic dynamics, we treat the wave drag as fixed and use the Newtonian cooling approximation for the diabatic heating.

5.2.2 Approach

Following previous studies (e.g. Haynes et al., 1991) we consider the response of the system with variable $\alpha$ to switch-on and steady periodic forcing. This allows a straightforward comparison with existing work and captures the essential dynamical aspects. Using the system's separability property we set $\bar{T}_{rad}$ to zero in all the experiments. For simplicity, the three different cases of $\alpha = \alpha(\phi)$, $\alpha = \alpha(z)$ and $\alpha = \alpha(t)$ are treated separately. Variation in $\alpha$ introduces new complexities to the asymptotic analysis of Haynes et al. (1991) so a numerical analysis using ad hoc localized forcing is performed instead. (A similar asymptotic analysis may be possible in the $\alpha = \alpha(z)$ case.)

The $\alpha(\phi)$ case is motivated by the large latitudinal variation of $\alpha$ at high latitudes in the winter hemisphere at the transition to polar night. This case is intended to analyze the influence of wave drag outside the polar vortex on the interior of the vortex. The effect of variation of $\alpha$ with altitude is treated in the $\alpha(z)$ case and is relevant for the problem of the influence of mesospheric gravity wave drag on the stratosphere and the problem of the time-dependence of mass flux through the 16 km (100 hPa) surface. The effect of seasonal variation in $\alpha$ on the annual mean temperature signature and transients associated with forcing on the seasonal timescale is addressed in the $\alpha(t)$ case. Other cases of $\alpha$ depending on more than one dimension or a combination of time and spatial dependence are not investigated here.

Experiments use an idealized forcing of the form $F = F_0 Y_F(\phi) Z_F(z) H(t)$ where the time-dependent part, $H(t)$, is a step function or $\cos(\omega t)$ and where

$$Y_F(\phi) = \begin{cases} 
\frac{1}{2} + \frac{1}{2} \cos\left(\frac{2\pi}{\phi_2 - \phi_1}\left[\phi - \frac{1}{2}(\phi_1 + \phi_2)\right]\right) & \text{for } \phi_1 \leq \phi \leq \phi_2 \\
0 & \text{otherwise}
\end{cases}$$

(5.5)
and
\[
Z_F(z) = \begin{cases} 
\frac{1}{2} + \frac{1}{2} \cos \left( \frac{2\pi}{z_2 - z_1} \left[ z - \frac{1}{2} (z_1 + z_2) \right] \right) & \text{for } z_1 \leq z \leq z_2 \\
0 & \text{otherwise}
\end{cases}
\] (5.6)
with \( F_0 = -5 \text{ m s}^{-1} \text{day}^{-1} \). The position and latitudinal scale of the forcing is varied depending on the profile of \( \alpha \) being investigated.

The idealized diabatic heating perturbation used in experiments dealing with the ozone hole is given by \( \Delta Q = Q_0 Y(\phi) Z(z) H(t) \) with \( Q_0 = -0.35 \text{ K day}^{-1} \) and where
\[
Y(\phi) = \begin{cases} 
\sqrt{\frac{1}{2} + \frac{1}{2} \cos \left( \frac{\phi - \phi_1}{\phi_2 - \phi_1} \right)} & \text{for } \phi_1 \leq \phi \leq \phi_2 \\
0 & \text{otherwise}
\end{cases}
\] (5.7)
with \( \phi_1 = 65^\circ \) and \( \phi_2 = 90^\circ \), and
\[
Z(z) = \begin{cases} 
\frac{1}{2} + \frac{1}{2} \cos \left( \frac{2\pi}{z_2 - z_1} \left[ z - \frac{1}{2} (z_1 + z_2) \right] \right) & \text{for } z_1 \leq z \leq z_2 \\
0 & \text{otherwise}
\end{cases}
\] (5.8)
with \( z_1 = 10 \text{ km} \) and \( z_2 = 26 \text{ km} \), and
\[
H(t) = \begin{cases} 
\frac{t_2 - t}{t_2 - t_1} & \text{for } t_1 \leq t \leq t_2 \\
\frac{1}{2} + \frac{1}{2} \cos \left( \frac{2\pi}{t_3 - t_2} \left[ t - \frac{1}{2} (t_2 + t_3) \right] \right) & \text{for } t_2 \leq t \leq t_3 \\
0 & \text{otherwise}
\end{cases}
\] (5.9)
with \( t_1 = 1.5 \), \( t_2 = 4 \) and \( t_3 = 5.5 \) months.

5.3 Response to switch-on forcing with spatially variable \( \alpha \)

The switch-on forcing regime corresponds to infra-seasonal timescales when, for example, the winter season surf zone forms and persists until the spring equinox season. As a point of reference we take the uniform damping case which has been studied by Haynes et al. (1991). The response of the system to a forcing with a step-function time dependence has been investigated in some detail by them with asymptotic and numerical solutions (their section 3(c) and Fig. 6, respectively). It evolves towards a steady state which has a "downward control" character. The steady-state limit is not reached by the atmosphere but is approached closely during the winter solstice seasons (Rosenlof and Holton, 1993).
5.3  Response to switch-on forcing

Some aspects can be inferred \textit{a priori} about the steady state with spatially variable $\alpha$. In steady state, $\overline{w}$ is independent of $\alpha$ and is determined by the latitudinal distribution of streamlines which are all vertically aligned below the region of $\mathcal{F}$ (this follows from the QG downward control formula). So from the steady version of (5.2) with $T_{rad} \equiv 0$, which is $\overline{w} S = -\alpha \overline{T}$, it can be seen that the distribution of $\overline{T}$ will vary depending on $\alpha$. This is in addition to the variation imposed by $\overline{w}$ which is zero above the forcing region and decays inversely to the density below the forcing region. For $\alpha$ varying only in latitude (Fig. 5.1a) this requires that $\mathcal{F}$ overlap the high latitude regions for there to be a steady-state effect at high latitudes. If this is the case then there will be an amplified $\overline{T}$ response in the region overlapping the smaller values of $\alpha$. The effect can be quite large at the altitudes of the forcing. With an $\alpha$ profile given by Figure 5.1b there will be an increased $\overline{T}$ response in steady state at lower altitudes. The increase in $\overline{T}$ will be small if the wave drag is over a scale height above the weak $\alpha$ region. If $\alpha$ increases proportionally to the density below the forcing region, however, then the $\overline{T}$ response can be unifrom. However, the rate of evolution to steady state is roughly proportional to $\alpha$ and this situation will not develop in the course of a season.

The forcing used here has the distribution described in section 5.2. The system is marched in time for 120 days after the forcing is turned on.

5.3.1  The $\alpha(\phi)$ case

The streamfunction and temperature are shown in Figures 5.2 and 5.3 at a couple of times during the progression to steady state of the system using the $\alpha$ profile given in Figure 5.1a. If the forcing sits completely outside the region of weak radiative damping, then it evolves to steady state at a rate as if this region did not exist. But fields penetrating into this region exhibit some differences in the transient regime. Streamlines tend to reach to higher latitudes and the induced temperature is stronger (not shown).

These effects are more persistent if there is some overlap of the forcing with the region of weaker $\alpha$ as is the case for Figures 5.2 and 5.3. Compared to the uniform damping case the differences in the mass streamfunction are small. The temperature, however, exhibits a much larger amplitude in the region of small $\alpha$. The temperature maximum of the warming region below the forcing and nearest the pole moves to higher latitudes. The same $\overline{w}$ is producing a larger $\overline{T}$ change in regions of smaller $\alpha$. The
larger temperature response will also apply to the steady state as can be seen from the thermodynamic equation ($\bar{w}^* S = -\alpha \bar{T}$).

5.3.2 The $\alpha(z)$ case

In Figures 5.4 and 5.5 are presented snapshots of the evolution of the system to steady state with the model $\alpha$ set to the profile in Figure 5.1b. It is apparent that the transience persists for longer periods at lower levels. In the lowermost panels of Figure 5.4 it can be seen that at the forcing levels, where the damping rate is the same, the streamfunction has a nearly identical distribution in the two cases. Where the damping rates are longer the streamfunction is still in an earlier stage in its evolution. In particular, it is less confined in latitude and streamlines take a longer time to concentrate onto absolute
angular momentum surfaces below the forcing. Even though the circulation is weaker at these altitudes the temperature response is greater than in the uniform $\alpha$ case. As noted above this is due to the weaker radiative damping at these levels. The effect on the streamfunction is more noticeable than in the $\alpha(\phi)$ case because of the asymmetric nature of the circulation response with streamlines penetrating downward instead of sideways.

Reducing the latitudinal scale of the forcing reduces the degree of temperature penetration downward at a given time since the transience is increased. The maximal penetration below the forcing level is in steady state, when the temperature decays exponentially below the forcing level with the same scale height as the density and does not depend on the latitudinal scale of the forcing.

The forcing distribution in the real atmosphere (or GCMs) extends through all levels of the midlatitude stratosphere. To see the behaviour of forcing at lower levels we repeat the experiment of above forcing a lower level. On the 16 km surface, where $\alpha(z) =$
Fig. 5.4 Contour plots of the streamfunction at 10 days (top row) and 95 days (bottom row) for forcing at 50 km. Left panels are for the $\alpha(z)$ case. Right panels correspond to a uniform $\alpha = 1/(6 \text{ days})$. The force distribution parameters are $\phi_1 = 15^\circ$, $\phi_2 = 75^\circ$, $z_1 = 40 \text{ km}$ and $z_2 = 60 \text{ km}$. Contour interval is 0.5 kg m$^{-1}$s$^{-1}$ for all panels.

1/(60 days), the system takes longer than a season to approach steady state (Fig. 5.6). The damping rate is slow enough that using a forcing of larger horizontal scale did not result in a significant change in the rate of evolution to steady state.

5.4 Response to periodic forcing with spatially variable $\alpha$

The switch-on forcing case can only give insight into part of the response to time-dependent forcing. Here we consider the time-harmonic behaviour of the system. On time scales longer than a season middle atmosphere wave drag has a cyclical time dependence following the annual cycle. Mesospheric gravity wave drag has opposite sign in the two solstice seasons and a roughly zero annual mean. Planetary wave drag in the stratosphere
5.4 Response to periodic forcing

occurs primarily during the winter season and is single-signed, so it has a non-zero annual mean component.

5.4.1 Uniform $\alpha$ case

The periodic forcing case for the uniform $\alpha$ problem has been treated by several studies (Garcia, 1987; Haynes et al., 1991; Holton et al., 1995; Sankey, 1998). The spatio-temporal structure of the solutions is sensitive to the value of the ratio $\omega/\alpha$, where $\omega$ is the frequency of the forcing. The adiabatic or transient limit corresponds to $\omega/\alpha \to \infty$ and the steady state limit to $\omega/\alpha \to 0$. This is demonstrated by a scaling analysis in Holton et al. (1995), reproduced below, which also applies for the intermediate time regime.

The system (5.1)-(5.4) with the Newtonian cooling approximation can be rearranged into a single equation for $\overline{w}^*$. Assuming that the time dependence is harmonic, with
constant frequency $\omega$, the substitutions $F = \text{Re}(Fe^{i\omega t})$, $Q_z = \text{Re}(Qe^{i\omega t})$ and $\bar{w}^* = \text{Re}(w e^{i\omega t})$ give this equation the form

$$
\frac{\partial}{\partial z} \left( \frac{1}{\rho} \frac{\partial (\rho w)}{\partial z} \right) + \left( \frac{i \omega}{i \omega + \alpha} \right) \frac{N^2}{4 \Omega^2 a^2} \cos \phi \frac{\partial}{\partial \phi} \left( \frac{\cos \phi}{\sin^2 \phi} \frac{\partial w}{\partial \phi} \right) =
$$

$$
\left( \frac{i \omega}{i \omega + \alpha} \right) \frac{R/H}{4 \Omega^2 a^2} \cos \phi \frac{\partial}{\partial \phi} \left( \frac{\cos \phi}{\sin^2 \phi} \frac{\partial Q}{\partial \phi} \right) + \frac{1}{2 \Omega a \cos \phi} \frac{\partial}{\partial \phi} \left( \frac{\cos \phi}{\sin \phi} \frac{\partial F}{\partial z} \right) \tag{5.10}
$$

The solution to this equation consists of a linear superposition of a component due to solar heating, the first term on the rhs, and a component due to the imposed force, the second term on the rhs. We are interested in the solution pertaining to the imposed force, so the first term on the right hand side can be excluded. The adiabatic limit corresponds to the limit $i \omega/(i \omega + \alpha) \rightarrow 1$. In this case, scaling analysis of (5.10), taking $a \Delta \phi$ as the horizontal scale and $\Delta z$ as the vertical scale of $\rho w$, yields

$$
\Delta z \sim \max \left[ \frac{2 \Omega \sin \phi}{N} a \Delta \phi, \frac{4 \Omega^2 \sin^2 \phi (a \Delta \phi)^2}{N^2 H} \right] \tag{5.11}
$$

outside the forcing region, where the terms on the lhs of (5.10) balance. The first value in the rhs of (5.11) corresponds to $\Delta z < H$ and the second value corresponds to $\Delta z > H$. So for shallow forcing the ratio of the vertical scale to the meridional scale is the ratio of
the Coriolis frequency to the buoyancy frequency. For deep forcing, which is the relevant atmospheric case, the vertical scale is much larger, varying as the square of the meridional scale.

In the intermediate case when \( \alpha > \omega \), scaling analysis gives

\[
\Delta z \sim \max \left[ \left( \frac{\alpha}{\omega} \right)^{\frac{1}{2}} \frac{2\Omega \sin \phi}{N} a \Delta \phi, \left( \frac{\alpha}{\omega} \right) \frac{4\Omega^2 \sin^2 \phi (a \Delta \phi)^2}{N^2} \frac{1}{H} \right]
\]

(5.12)

outside the forcing region. The additional factors involving \( \alpha/\omega \) imply that the vertical scale of the solution is greater for a given meridional scale than in the adiabatic regime. Specifically, the coefficient of the second term on the lhs of (5.10) gets smaller as \( \alpha/\omega \) increases, which requires the first term on the lhs to decrease also in order to maintain balance. As a result the vertical variation of the solution decreases, or equivalently its vertical scale increases. In the steady state limit, \( \alpha/\omega \to \infty \), the second term on the lhs of (5.10) disappears; then \( \rho w \) has no vertical variation outside the forcing region and is zero outside the latitude band containing the force (i.e. downward control).

The solutions have a component in phase with the forcing and a component \( \pi/2 \) out of phase (Sankey, 1998), which arise from the induced diabatic heating. The out of phase component reflects the finite radiative damping timescale in the system. If \( \alpha = \infty \) then the solutions are completely in phase with the forcing. This simple time dependence follows from the fact that in this case the system has no homogeneous modes and can have no evolution without forcing.

The temporal behaviour was considered in some detail by Sankey (1998). The relative lag of \( \bar{w}^* \) and \( \bar{T} \) is given by \( \tan^{-1}(\omega/\alpha) \) at all points in the meridional plane. The lag between the upwelling and the temperature therefore increases for decreasing \( \alpha \) and decreases for increasing \( \alpha \). Since this is a pointwise result, the relation will still apply if \( \alpha \) is spatially variable. However, the result is based on the time-harmonic decomposition of the governing equations so it will not apply if \( \alpha \) is time dependent, in which case there is mixing of harmonics.

The lag relationship between \( \bar{w}^* \) and the forcing is more complicated and depends on position relative to the forcing region. A useful diagnostic of this was used by Sankey (1998) who considered distributions of the month of maximum \( \bar{w}^* \) in the height-latitude
plane to get a sense of this lag (cf his Fig. 2.17). Below and equatorward of the forcing maximum the upwelling lags the forcing, while immediately poleward of the forcing maximum the downwelling leads the forcing. At very high latitudes $\bar{w}^*$ lags the forcing.

Although start-up transients produced by switch-on forcing have been considered by Haynes et al. (1991), those associated with steady oscillatory forcing have not received much attention. The behaviour of switch-on transients for different values of $\omega$ is given in Figure 5.7. Shown is a time series lasting two periods at a fixed point in the height-latitude plane of the difference between $\bar{w}^*$ immediately after start up and $\bar{w}^*$ near the end of the model run (approximate long time limit). The value of $\alpha$ in these runs is 1/6 day$^{-1}$. The decay rate evidently has no dependence on $\omega$, although the presence of the undulation in some of the curves indicates that it takes longer for the response to forcing with larger $\omega$ to equilibrate.

The effect of changing $\alpha$ to 1/40 day$^{-1}$ on the switch-on transient corresponding to the annual cycle $\omega$ is also shown in Figure 5.7. It is apparent that reducing the radiative
damping rate has a significant impact on the duration of the transient. This result has implications for annually periodic forcing such as the ozone hole, which is located in the lower polar winter stratosphere where the damping rate is weak \( \alpha \approx 1/40 \text{ day}^{-1} \).

### 5.4.2 The \( \alpha(\phi) \) case

This case is somewhat artificial for periods longer than a season since \( \alpha \) changes significantly during the annual cycle at high latitudes. We thus consider \( \alpha \) to be an annual mean value (for consideration of the time dependence in \( \alpha \) see below).

Figures 5.8 and 5.9 show the \( \cos(\omega t) \) component of the streamfunction and temperature for forcing at two different latitudes. Noticeable differences in the streamfunction occur only if streamlines overlap the region of small \( \alpha \) at high latitudes, which is the case in the lower panels. The equilibrated oscillatory state at high latitudes has a weaker \( \bar{\omega} \) but a stronger temperature response. There is also more spreading of the circulation into high latitudes which contributes to the increase in temperature inside the weak \( \alpha \) region.

The lag between \( \bar{\omega} \) and the forcing is represented in Figure 5.10 (based on Sankey, 1998). Below the forcing level, \( \bar{\omega} \) leads the forcing equatorward of the forcing maximum, but lags it poleward of the forcing maximum. This pattern results from the fact that upwelling on the equatorward side involves downwelling on the poleward side. Upwelling on the poleward side occurs half a period out of phase behind that on the equatorward side. Above the forcing this pattern is reversed since the circulation is in the opposite sense. The spatial variation of the lag or lead time also reflects the effect of radiative damping, which introduces an increasing delay between the time of peak forcing and peak upwelling as one moves away from the forcing in the vertical and horizontal. In the case here, introducing horizontal variation in the radiative damping rate leads to an increase in the lag between \( \bar{\omega} \) and the forcing near the pole (compare the left panel to the uniform damping case in the right panel: the region of dashed contours near the pole is diminished and solid contours extend to higher latitudes).

### 5.4.3 The \( \alpha(z) \) case

The results with \( \alpha \) varying with height below the forcing region are presented in Figures 5.11–5.13. As in the steady forcing case, the effect on mass streamline behaviour is significant since streamlines penetrate downward into the region of weak \( \alpha \) regardless of
the latitude or altitude of the forcing. In this region the system appears to be trapped in a transient state (lower panels, Fig. 5.11). Such a combination of transience and weak damping rate leads to a stronger temperature response even near the pole (Fig. 5.12, lower panel). Once again, the diminished concentration of streamlines in latitude translates into a weaker $\overline{w}^*$ but one that extends further poleward (and equatorward). This is also evident in Figure 5.13 from the spreading of the region of $\overline{w}^*$ lagging $\mathcal{F}$ towards the pole (and the leading region towards the equator) at lower altitudes.

The behaviour of $\overline{w}^*$ on the 16 km (100 hPa) surface compared to the uniform $\alpha$ case is given in Figure 5.14. The dependence of the phase lag on $\alpha$ is most apparent for the large $\omega$ case (lower panel) where there is a noticeable shift in the peaks of the solid and long dash curves. Although the maximum vertical velocity leads the forcing by a longer
time interval for small \( \omega \) case (upper panel), reaching about 3.5 months in the annually periodic forcing case, the phase lag is decreasing in this limit. This is consistent with the system having a steady-state behaviour in the limit \( \omega \to 0 \).

### 5.4.4 The \( \alpha(t) \) case

Here we consider the effect of an annually periodic but spatially uniform \( \alpha \) given by

\[
\alpha(t) = \frac{1}{6 \, \text{day}} \left[ \frac{2}{3} - \frac{1}{3} \cos\left(\frac{2 \pi t}{365}\right) \right].
\]

Periodic forcing with the same three different frequencies as before was applied to the system. Consideration of the \( \cos \omega t \) solution component (not shown) indicated that for the higher frequency forcing this harmonic is least affected by the annual periodicity of the radiative damping rate. From Figure 5.15 (lower panel) it is evident that \( \alpha(t) \) is imposing an annual modulation of the response but is not deforming the sinusoidal form.
Fig. 5.10 Contour plots of the month of maximum $\bar{w}'$ corresponding to Figure 5.8. Solid lines are positive, dotted lines represent the zero contour and dashed lines are negative. Positive values imply that $\bar{w}'$ is lagging the forcing while negative values imply that it is leading the forcing. Contour interval is about 30 days.

of the solution as much as with the smaller frequency forcing (upper panel). The reason for this is that higher frequency forcing will "see" the low and high values of $\alpha$, thereby generating lower frequency modulation of the solutions (Fig. 5.15, lower panel). On the other hand, lower frequency forcing that has periods longer than a year (not considered here) will "see" an average $\alpha$ value and hence are less sensitive to its time-dependence. For forcing with the same frequency as that of the radiative damping rate the relative phase is important for the time dependence of $\bar{w}'$. In the upper panel of Figure 5.15 the "climatological" in-phase solution still has significant modulation. Since $\alpha$ is time dependent there are higher harmonics, but they are evidently of secondary importance.

A comparison of $\bar{w}'$ for uniform and time-dependent $\alpha$, taken at a point in the height-latitude plane, is given in Figure 5.16. As before, the lead time between the peak
forcing and peak upwelling increases as $\omega$ decreases, but the phase difference becomes smaller (for the choice of frequencies used here this is a small effect). For forcing with sufficiently long periodicity only the time mean value $\alpha(t)$ will matter and the system will approach a steady state.

The annually varying radiative damping rate leads to a more transient distribution in the streamfunction compared to the constant damping case. The degree of transience increases with the frequency of the forcing.

5.5 Annual mean signature of $\alpha(t)$

A property of the QG equations is that under conditions of no interannual variability
and time-independent $\alpha$, their annual mean form is the same as their steady state form:

$$-f \langle \bar{v}^* \rangle = \langle \mathcal{F} \rangle$$

$$S \langle \bar{w}^* \rangle = -\alpha \langle (\bar{T} - \bar{T}_{rad}) \rangle \equiv -\alpha \langle \Delta \bar{T} \rangle$$

This gives rise to the constraint of annual mean downward control. Time-dependence in $\alpha$ introduces additional annual mean correlation terms in the QG annual mean thermodynamic equation

$$S \langle \bar{w}^* \rangle = -\langle \alpha \rangle \langle (\bar{T} - \bar{T}_{rad}) \rangle - \langle \alpha' (\bar{T}' - \bar{T}'_{rad}) \rangle$$

The annual mean zonal wind equation has no additional terms and annual mean downward control still holds when $\alpha$ is a function of time (as long as $\langle \bar{u}_t \rangle = 0$).

The significance of the correlation terms in the thermodynamic equation is that the temperature will have a different annual mean distribution than the steady-state one. In
the case of an imposed diabatic forcing alone (e.g. ozone hole case), which can be treated as a modification of the radiative equilibrium temperature, $\delta T_{\text{rad}}$, we have

$$< \alpha \Delta \overline{T} > = < \alpha \delta T_{\text{rad}} >$$

since $< \overline{w^*} > = 0$ everywhere. Hence,

$$< \alpha > < \Delta \overline{T} - \delta T_{\text{rad}} > = - < \alpha' (\Delta \overline{T} - \delta T_{\text{rad}})' > .$$

Thus, unlike in the case of time-independent $\alpha$, the temperature change $\Delta \overline{T}$ is not constrained to overlap the diabatic forcing region in the annual mean. This can be seen in Figure 5.18 which shows the annual mean temperature response of the model forced with an annually periodic heating $\delta T_{\text{rad}}$ when $\alpha$ varies with time, and the difference with the constant $\alpha$ case. The time dependence of $\alpha$ is given by

$$\alpha(t) = \frac{1}{10 \text{ day}} \left[ \frac{2}{3} - \frac{1}{3} \cos \left( \frac{2\pi t}{365} \right) \right] .$$
Fig. 5.14 Time series of $\bar{w}$ on the 16 km surface at 42°N for uniform $\alpha = 1/(6 \text{ days})$ (solid) and for $\alpha(z)$ (long dash). The time dependence of the forcing is represented (short dash) with an arbitrary amplitude. The upper panel is for $\omega$ corresponding to a period of 1 year, and the lower panel for a period of 1/3 year. The force distribution parameters are $\phi_1 = 20^\circ$, $\phi_2 = 50^\circ$, $z_1 = 40 \text{ km}$ and $z_2 = 40 \text{ km}$.

The time behaviour of the idealized $\alpha$ used here is chosen based on the annual cycle of the $\alpha$ inferred from a comprehensive radiative transfer scheme (Fomichev and Blanchet, 1995) for the region of interest (high latitudes at about 20 km). The value of $\alpha$ used in the time-independent case was $\alpha(t_2)$ where $t_2 = 4 \text{ months}$ is the time of peak diabatic forcing. This choice is arbitrary but makes the comparison more straightforward since the magnitude of the temperature (and zonal wind) response is determined by the typical values of $\alpha$ when the forcing is engaged.
Another effect of time-dependent $\alpha$ is to change the rate of evolution to steady state at different times of the year. As a result, the transient induced during the late-winter and early-spring heating considered above will decay faster during the summer season than if $\alpha$ were time independent. During the winter season the weaker radiative damping rate results in a larger temperature but weaker diabatic circulation. This is shown in Figure 5.17 in a time series of $\Delta\overline{T}$ at a point in the heating region. Thus, $\alpha = \alpha(t)$ will introduce a seasonal asymmetry in the evolution compared to the time-independent $\alpha$ case resulting in a transient signature in $<\Delta\overline{T}>$ (Fig. 5.18). For the diabatic heating...
considered here, this signature is weak. This can be partly attributed to the weakness of the transience during the seasonal cycle. The region of warming above the region of cooling has a peak amplitude less than 10% of the latter during most of the heating period.

A similar experiment to the one above was performed using a forcing with a time dependence roughly resembling stratospheric planetary wave drag. In this case the time dependence of $\alpha$ is given by

$$\alpha(t) = \frac{1}{8.1 \text{ day}} \left[ 1 - 0.35 \cos \left( \frac{2\pi t}{365} \right) \right].$$
Fig. 5.17 Time series of $\delta T_{rad}$ induced $\Delta T$ at 25 km and 76°N for the $\alpha(t)$ case (solid) and constant $\alpha$ case (dash), corresponding to Figure 5.18.

For the reference constant $\alpha$ case the minimum value of $\alpha(t)$, which is 1/12.5 day$^{-1}$, was used since it corresponded to the time of peak forcing. It is evident that the increased damping during the summer season diminishes the transient tail of the winter season forcing (Fig. 5.19). This translates into a transient pattern in $<\Delta T>$ (Fig. 5.20), with a response above the forcing, even as the system reaches a steady oscillatory state. The larger $<\Delta T>$ amplitude in the constant $\alpha$ case is due to the fact that the damping rate is weaker than in the $\alpha(t)$ case except at a single point in time. This highlights the fact that the $<\Delta T>$ pattern in the time-dependent $\alpha$ case is not an artefact of start-up transience related to the finite duration of the model simulation. If that were the case, start-up transience would appear in the uniform $\alpha$ case also, which has a weaker radiative damping rate.

5.6 Long-lived transients

There is a significant degree of interannual variability in the middle atmosphere. It is linked to external forcing (mechanical or diabatic) and internal modes. In the zonal mean both are folded into the EP flux divergence and radiative equilibrium temperature variations on a mean seasonal cycle. Phenomena such as the ozone hole have not yet
equlibrated and ozone loss has been increasing in magnitude during the last 20 years (WMO, 1999). Wave drag (Rossby wave and gravity wave breaking and thermal dissipation) also exhibits interannual variation in its distribution (e.g. sudden warmings, QBO link to polar vortex intensity (Holton and Tan, 1980)). As shown in section 5.4.1 there is a start-up transient which persists for the first period of a steady oscillatory forcing when starting the system from rest. So a forcing with annual periodicity and interannually varying amplitude can be expected to maintain a transient component (on top of the equlibrated mean).

5.6.1 Persistence of transience and non-locality
The persistence of transients is somewhat counter-intuitive since the radiative damping
timescale is much shorter than a year even at lower levels and in the polar night. The reason for their long life is the diabatic circulation, which acts to divert part of the energy input by the forcing (mechanical or diabatic) into a nonlocal thermal reservoir by driving the temperature away from radiative equilibrium (Snieder and Fels, 1988). When the forcing is switched off, some of the stored thermal energy gets converted into a meridional circulation which redistributes this energy. This circulation acts to produce adiabatic cooling or warming, respectively. The net result is that the temperature does not decay exponentially with an e-folding rate of $\alpha$, but at a much slower rate. In the absence of the diabatic circulation the problem would, in contrast, be completely local and exhibit exponential decay. The zonal wind is damped through the temperature to which it is coupled via thermal wind balance. Consequently, it evolves over similar time scales.

The nonlocality (ellipticity) which is associated with the diabatic circulation leads to a complicated pattern of evolution. Even though no homogeneous (unforced) dynamics are possible in the balance TEM system, there is an oscillatory type of behaviour that results from the coupling of nonlocality with relaxational radiative damping. Regions of forcing (mechanical or radiative) induce heating in adjacent regions. This leads to a decaying pattern of heating and circulation propagation which is particularly apparent
Fig. 5.20 Annual mean $\Delta T$ for $\alpha(t)$ (upper panel) and for the constant $\alpha$ case (lower panel), for the case of an imposed force corresponding to Figure 5.19. The contour interval is 0.2 K in both panels.

after a finite duration forcing is switched off. In Figure 5.21 is shown a time series of $\overline{w^r}$ which demonstrates the non-monotonic behaviour after the forcing is switched off: there is an oscillation with a rapidly decaying amplitude which persists several years.

5.7 Conclusion

We have investigated the response of the quasi-geostrophic TEM system to an imposed switch-on or periodic forcing, given prescribed radiative damping profiles $\alpha(\phi)$, $\alpha(z)$ and $\alpha(t)$. Spatial variation of radiative damping introduces spatial dependence of the rate of evolution to steady state; temporal variation of radiative damping leads to the rate of evolution towards steady state being time dependent. Compared to the uniform $\alpha$ case, the lag between $\overline{w^r}$ and $\overline{T}$ is no longer the same at every point in the meridional
Fig. 5.21 Time series of $\overline{w}^*$ at 39 km and $-14^\circ$ for 36 model months (left panel). The force is active only during the first 100 days with a uniform amplitude of $-2.5 \text{ ms}^{-1}\text{day}^{-1}$ and distribution specified by $\phi_1 = 35^\circ$, $\phi_2 = 65^\circ$, $z_1 = 30$ km and $z_2 = 50$ km. The radiative damping rate has a constant value of $1/6 \text{ day}^{-1}$. The right panel shows a blow-up of the same time series starting at 10 model months.

plane for steady-state oscillatory forcing. If $\alpha$ varies only in space the lag depends on its local value. Time-dependent $\alpha$ introduces a temporally local modulation of the solution amplitude. The only other factor which introduces a spatial variation in the rate of evolution to steady state is the Coriolis parameter (the time taken to reach steady state increases as the Coriolis parameter decreases at low latitudes.)

The spatial distribution of solutions is affected by the local value of the radiative damping rate. Regions with weaker $\alpha$ have reduced latitudinal confinement of streamlines and less streamline penetration below the forcing level. This means a weaker $\overline{w}^*$ response but generally greater induced $\overline{T}$. In the steady periodic regime the solutions are trapped in a transient pattern in such regions. The solutions of the system with steady periodic forcing reflect the behaviour of the radiative damping rate, provided that the frequency of the forcing is not too large — in which case the system is in the transient limit and not sensitive to $\alpha$. But we are generally interested in cases where the forcing period is longer than the damping timescale.

For the missing gravity-wave-drag problem of Garcia and Boville (1994), the important aspect is the sensitivity of the temperatures inside the polar vortex on account of the longer radiative damping timescale there. This applies to sub-polar wave drag in
the stratosphere and to high-latitude drag in the mesosphere. Even if the circulation is weaker than is the case with a faster damping rate, the temperature can be driven further from radiative equilibrium. Based on the idealized simulations performed here, it appears that wave drag outside the polar vortex in the lower stratosphere cannot be excluded as an explanation for the temperatures inside the southern hemisphere vortex. The penetration of the temperature response below the forcing level is limited by exponential density stratification as is evident in the $\alpha(z)$.

The results presented here pertaining to the mass flux through the 16 km (100 hPa) surface indicate that for the annually periodic case the maximum upwelling cannot be regarded as in phase with the forcing. It must be noted that the lag of approximately two months is for the specific force distribution used here. The force distribution inferred from observations is latitudinally broader, which according to the results of Haynes et al. (1991) should translate into a more steady-state character of the circulation and therefore less lag. Nevertheless, deviation from the steady-state streamline distribution increases at lower altitudes and cannot be expected to hold below the 16 km surface. The implication for the results of Appenzeller et al. (1996) is that the delay between the peak wave drag in the “overworld” (above 16 km) and the mass outflow into the troposphere from the lowermost stratosphere is not all due to seasonal evolution of tropopause height.
Chapter Six

Conclusions

6.1 Summary

The work in this thesis can be described as an exploration of the limitations of the "downward control" picture of the zonal mean meridional circulation in the middle atmosphere. This theme has been approached in three different contexts: mechanical forcing feedbacks, the nonlinear Hadley circulation, including its effect on tropical upwelling, and transience associated with spatio-temporal variability in radiative damping. The importance of processes that do not fit into the picture of the diabatic circulation as being near steady state and wave-driven is apparent in the tropics and in regions with small radiative damping rates, such as the lower stratosphere and the high-latitude polar night. The sensitivity of the zonal mean system to the details of wave drag is evident from the difficulty in explaining annual mean upwelling in the tropical lower stratosphere and the dynamical control of temperatures in the wintertime Antarctic lower stratosphere.

Chapter 2 considered the impact of relaxational damping on the response to an imposed force (local change in wave drag) or diabatic heating (local change in the radiative equilibrium temperature). Many GCMs include a frictional layer near their upper boundary which serves to control spurious downward wave reflection and also acts as a proxy for unresolved gravity-wave drag. It was found that such sponge layers, whether based on Rayleigh friction or enhanced diffusion, produce a significant and unrealistic effect on the temperature and zonal wind within a couple of density scale heights of their lowest level. In addition, the sponge-drug force spreads out beyond the latitudinal band of the imposed force or diabatic heating, producing a weak meridional circulation over the whole of the hemispheric model domain. True wave drag is also expected to respond
to an imposed force or diabatic heating, but in a much more limited way. In particular, it does not have an unbounded capacity to produce local zonal momentum damping like relaxational friction.

In Chapter 3 the nonlinear stratospheric Hadley circulation was investigated in an extension of the work of Dunkerton (1989, 1991). It was verified, using a three-dimensional primitive equations model and analysis of GCM data, that zonal mean gradient-wind balance, which is essential for the occurrence of the nonlinear Hadley circulation, is a reasonable assumption for the tropics. The structure of the nonlinear Hadley circulation is determined by the radiative transfer properties (radiative equilibrium temperature and radiative damping rate) in the tropical middle atmosphere. In particular, the closing off of the circulation around and above the stratopause is a radiative feature reflecting the distribution of the diabatic heating that drives the circulation. Although wave drag is required to get a realistic distribution of the absolute angular momentum in the mesosphere, and to a lesser extent in the stratosphere, the nonlinear Hadley circulation makes a contribution to the formation of the region of overturning absolute angular momentum isopleths in the tropical stratopause region. In the process, this thermally driven circulation contributes to the development of the easterly phase of the SAO.

Another important effect on the structure of the nonlinear Hadley circulation is inertial instability. Inertial adjustment acts on a much faster time scale than advection by the NLHC to latitudinally homogenize absolute angular momentum. The redistribution of absolute angular momentum modifies the properties of the streamfunction operator and, through thermal-wind balance, the diabatic heating as well. The change in diabatic heating is not accounted for in Dunkerton's parabolic approximation, and leads to an intensification of the nonlinear Hadley circulation further from the equator in the winter hemisphere. Inertial instability is triggered by advection of the absolute angular momentum maximum off the equator and large subtropical meridional shear in the radiative equilibrium zonal wind. However, both the unforced two-dimensional balance model and the three-dimensional primitive equations model used here produced an unrealistically broad region of inertial instability in the upper stratosphere and mesosphere. As indicated by GCM simulations and observations, wave drag acts to confine the region of inertial instability closer to the equator. So the enhancement of the nonlinear Hadley circulation by inertial adjustment is expected to be limited to lower latitudes as well.
The role of the nonlinear Hadley circulation in the specific problem of tropical upwelling was examined in Chapter 4. Persistent upwelling in the tropics throughout the annual cycle — as is found in the real atmosphere — is not a direct result of the wave-driven circulation or the extratropical “vacuum pump”. Instead, either a nonlinear mechanism or “diffusive*-like tropical wave drag (much like zero-phase-speed gravity-wave drag) is necessary to explain the observations.

Due to the limited nonlinearity in the zonal mean balance system, which appears to be a good approximation even in the tropics, the NLHC is able to superimpose, with some deformation, its single-cell pattern onto the wave-driven circulation. The wave-driven circulation is unable to substantially interfere with the formation of the NLHC, due to the same gradient-wind balance constraint near the equator that brings the NLHC into existence in the presence of latitudinal solar heating gradients. The NLHC accounts for the occurrence of the region of maximum upwelling in the summer hemisphere and enhances the seasonal deformation of the absolute angular momentum distribution in the upper stratosphere. This seasonal mobility of absolute angular momentum isopleths translates into a deflection of the meridional circulation streamlines towards the winter pole, which gives rise to annual mean upwelling in the tropical upper stratosphere.

In the lower tropical stratosphere the NLHC has little effect on annual mean upwelling. Nonlinearity in the wave-driven circulation was therefore examined as a possible explanation for annual mean upwelling. Based on numerical experiments it was concluded that realistic upwelling values require wave drag extending deep into the tropics. There is some indication that wave drag penetrates to the equator in the lower stratosphere from the CMAM GCM, even though most of the surf zone is confined away from the equator.

The key aspect of the nonlinear upwelling mechanism is the difference in the evolution of the wave-drag-induced diabatic heating. Nonlinearity leads to significant deformation of the induced diabatic heating which acts to diminish its ability to cancel the upwelling produced during the “on” phase (when the wave drag is active). In the linear or quasi-geostrophic case, the induced heating preserves its structure during the “off” phase (after the wave drag stops acting) and produces a more persistent downwelling than in the nonlinear case. This difference in the evolution is expected to occur for forcing cases which do not have a clear separation into an “on” and “off” phase; for example, when
the force varies gradually, increasing to a maximum from zero and then decreasing back to zero over a seasonal time scale.

The work in Chapter 5 indicates that the small radiative damping rate inside the wintertime polar vortex makes this region sensitive to dynamical heating from wave drag outside the vortex. Specifically, there is the possibility that sub-polar wave drag in the lower stratosphere (the region of largest ozone loss) may be exerting a significant effect horizontally on temperatures inside the vortex. This is somewhat at odds with the vertical downward control thinking. In fact, transience is a significant aspect of the dynamics at high latitudes and lower levels of the stratosphere where the radiative damping rate is small. In such regions, the system behaves differently from the downward control limit. Streamlines of the mean meridional circulation are not latitudinally confined, and the lag between the forcing and the circulation is greater. However, the downward effect of higher altitude (upper stratosphere or mesosphere) wave drag on the lower stratosphere cannot be excluded. It may be of sufficient magnitude to overcome the attenuation by density and horizontal streamline spreading due to transience that limits the amplitude of $w^*$ induced in the lower polar stratosphere. Even for weak $w^*$ the temperature response can be large given the weak radiative damping rate in this region.

Reflective of the transient regime at these altitudes, upwelling (and downwelling) through the "lid" of the model lowermost stratosphere at 16 km peaks over two months after the force when the period of the latter is a year. At still lower levels, this lag is expected to get progressively greater. This indicates that the delay in the mass flow out of the lowermost stratosphere into the troposphere compared to peak wave drag above 16 km has its origins in the behaviour of the diabatic circulation at these altitudes. The position of the tropopause acts as a modulation on the mass outflow (Appenzeller et al., 1996).

The seasonal variation of the radiative damping rate leaves an annual mean signature in the temperature distribution. For wave drag this effect is noticeable as a temperature response above the forcing level even in midlatitudes where the temporal variation of $\alpha$ is not as great as in the polar regions. Similarly, some of the annual mean warming seen above the Antarctic ozone hole may not be related to radiative transfer but may be such a transient effect. So care must be taken in interpreting annual averages.
The zonal mean balance system is characterized by nonlocality, which acts to increase the duration of transients. Instead of simple local exponential relaxation to radiative equilibrium, the system spins up a meridional circulation which redistributes zonal momentum and temperature. The energy needed to drive this circulation is stored in the form of the nonlocally induced diabatic heating which drives a circulation of its own that, in turn, also delays radiative relaxation (Snieder and Fels, 1988). The significance of this thermal “flywheel” is that start-up transients last much longer than a few radiative damping timescales. Interannual variation in the ozone hole, for example, can produce a weak temperature anomaly lasting from one year to the next, which will also appear as a transient pattern in the annually averaged fields.

An interesting feature of the system is a dissipative wave-like character of solutions during the decay phase of the diabatic heating induced by some forcing of finite duration. This non-monotonic behaviour is due to the redistribution of the diabatic heating through nonlocality.

6.2 Discussion

A logical extension of the zonal mean modelling performed in this thesis would be to include a comprehensive representation for wave drag, possibly based on the parameterization of Garcia (1991) for Rossby wave drag and Hines (1997) for gravity-wave drag. Such parameterizations have limitations, described below, but offer a path to increase the realism of the simple two-dimensional model without having to employ a more complicated and computationally expensive three-dimensional model. With a wave drag scheme that reacts in a plausible manner to changes in the zonal wind, the investigation of Chapter 2 could be generalized to include the response of realistic wave drag to changes in forcing and diabatic heating at lower levels. A quantification of the effect of wave drag on the nonlinear Hadley circulation would also be possible.

However, questions such as the control of temperatures in the Antarctic lower stratosphere, and possibly tropical upwelling, require a level of detail in the wave drag that cannot be captured with a simple parameterization in terms of the zonal wind. The particulars of the southern hemisphere wave sources, such as the weaker contribution of topography compared to the northern hemisphere, and unresolved sources, such as
convection, have to be accounted for. Wave drag schemes are “tuned” through several parameters to get a reasonable model climatology. Even with tuning, these schemes are a coarse approximation to reality and cannot be expected to reproduce the true wave drag “climatology” (probably the tuning problem is non-unique). Nevertheless such schemes are, in practice, better than simply prescribing wave drag inferred from observations.

One of the underlying assumptions of the work described in this thesis is that the radiatively determined state can be regarded as independent of wave drag. However, the radiative equilibrium temperature is determined by the distribution of ozone and other agents that get transported by the wave drag directly and through the diabatic circulation. So the problem is ultimately non-separable. In this light, the results of Chapters 3 and 4 have to be interpreted as pertaining to components of the meridional circulation which behave appropriately only when they coexist. For example, there would be little benefit from studying the behaviour of a transport model which determines $\overline{T}_{rad}$ through photochemistry if wave drag is absent. We are interested in the formation of the nonlinear Hadley circulation in the presence of wave drag; since the wave-driven circulation and the nonlinear Hadley circulation do not have a strong direct coupling, a climatological $\overline{T}_{rad}$ is appropriate.
Appendix A

Hough mode decomposition

The Hough-mode decomposition of the fields (Plumb, 1982) takes the form

\[ v = \sum_n v_n(z, t)B_n(\mu), \quad w = \sum_n w_n(z, t)\Theta_n(\mu), \quad \psi = \sum_n \psi_n(z, t)(1 - \mu^2)^{1/2}B_n(\mu), \]

\[ u = \sum_n u_n(z, t)\mu B_n(\mu), \quad T = \sum_n T_n(z, t)\Theta_n(\mu), \]

\[ F = \sum_n F_n(z, t)\mu B_n(\mu), \quad Q = \sum_n Q_n(z, t)\Theta_n(\mu), \]

where the functions \( \Theta_n(\mu) \) and \( B_n(\mu) \) satisfy

\[ \frac{d}{d\mu} \left( \frac{1 - \mu^2}{\mu^2} \frac{d}{d\mu} [\Theta_n(\mu)] \right) = \epsilon_n \Theta_n(\mu), \quad \frac{d}{d\mu} [(1 - \mu^2)^{1/2}B_n(\mu)] = \Theta_n(\mu), \]

and \( \epsilon_n \) are the Hough-mode eigenvalues (listed in Table 1 of Haynes et al., 1991). Note that \( \epsilon_n < 0 \) and \( |\epsilon_n| \gg 1 \).

With this decomposition, (2.1)–(2.5) take the spectral form

\[ \frac{\partial u_n}{\partial t} - 2\Omega v_n = F_n - r(z)u_n, \quad (A.1) \]

\[ \frac{\partial T_n}{\partial t} + Sw_n = Q_n - \alpha T_n, \quad (A.2) \]

\[ \frac{\partial u_n}{\partial z} = -\frac{\epsilon_n RT_n}{2\Omega a H}, \quad (A.3) \]

\[ v_n + \frac{a}{\rho_0} \frac{\partial}{\partial z} (\rho_0 w_n) = 0, \quad (A.4) \]

\[ v_n = -\frac{1}{\rho_0} \frac{\partial \psi_n}{\partial z}, \quad w_n = \frac{\psi_n}{\rho_0 a}. \quad (A.5) \]
Appendix B

Streamfunction equation

To be able to use both prognostic equations in the numerical model requires a discretization scheme that preserves thermal-wind balance. The scheme used here is motivated by that of Kinnersley (1996). The numerical model streamfunction is obtained from the discretized governing equations using the thermal-wind balance condition discretized at the end of the timestep. Discretizing the continuous streamfunction amounts to using the thermal-wind balance relation discretized at the beginning of the timestep. This method gives a nonlinear streamfunction when using $\bar{u}$ as one of the prognostic variable since the thermal-wind relation is nonlinear in $\bar{u}$. The nonlinearity can be avoided by using a variable that makes the thermal-wind relation linear. In particular, the zonal wind is replaced by $M = (\bar{m}/a)^2$, $\bar{m} = a \cos \phi (\bar{u} + a\Omega \cos \phi)$. Under climatological conditions $\bar{m} > 0$ so $\bar{u}$ can be easily obtained from $M$. With the zonal wind equation replaced by an equation for $M$, the streamfunction equation coefficients are:

$$C_{\phi} = \left( (1 - \kappa) M_z + R \cos^2 \phi S \right) \frac{\tan \phi}{a^2 H}$$

$$C_{\phi\phi} = \frac{R}{a^2 H} \cos^2 \phi S$$

$$C_{\phi z} = 2 M_z \frac{\tan \phi}{a}$$

$$C_{zz} = -M_{\phi} \frac{\tan \phi}{a^2}$$

$$C_z = \left( \frac{1}{a^2 \cos^2 \phi} + 3 \frac{\tan^2 \phi}{a^2} \right) M_z - M_{\phi} \frac{\tan \phi}{a^2 H}$$

$$D = \rho_0 \cos \phi \left( 2 \frac{\sin \phi}{a} (\sqrt{M} \mathcal{F})_z + \frac{R}{aH} \cos^2 \phi Q_\phi \right)$$

147
where $S = \bar{T}_z + (\kappa/H)\bar{T}$. The discretized version of $D$ has an additional correction term, which originates from the thermal-wind-balance-conserving discretization scheme, namely
\[
D \rightarrow D + \frac{\rho_0 \cos \phi}{\Delta t} \left( M_z \frac{\tan \phi}{a} + \frac{R}{aH} \cos^2 \phi \bar{T}_\phi \right)
\]
where $\Delta t$ is the timestep. The term in the brackets is the deviation from thermal-wind balance, and is equal to zero in the continuous system.

B.1 Inertial adjustment scheme

The inertial adjustment scheme is an iterative redistribution of $\bar{u}$ at every timestep at levels where the condition $f(\bar{u}_\phi - a f) > 0$ is satisfied. The version used here is a simplified form of the one used by Holton and Wehrbein (1980). In the original scheme, $\bar{u}_\phi$ is modified where the zonal wind satisfies the barotropic instability criterion. However, barotropic instability is a separate issue and is not an essential element of tropical dynamics, so we ignore this possibility.

It was found that using $f + \bar{u} \tan \phi/a$ instead of $f$ in the determination of the region of symmetric instability led to some numerical instability. Since there was not much difference in the regions requiring adjustment either way, the simpler form was used.

The discrete system domain is rectangular and defined by $\{(i, j) | 1 < i < N_\phi, 1 < j < N_z\}$. Here, $i$ is the latitudinal index and $j$ is the vertical index. The model grid was 65 by 65. For every $j$ on the domain grid, ranges of $i$ where the inertial stability condition is violated are identified at every timestep and take the form $\{N_{\phi 1}(j) < i < N_{\phi 2}(j)\}$. On $\{N_{\phi 1}, N_{\phi 2}\}$ the discretized zonal wind field $\bar{u}_{i,j}$ undergoes a conservative iterative adjustment of the form
\[
\bar{u}_{i,j} \rightarrow \bar{u}_{i,j} + (\bar{u}_{i+1,j} - \bar{u}_{i,j})/2
\]
\[
\bar{u}_{i+1,j} \rightarrow \bar{u}_{i+1,j} - (\bar{u}_{i+1,j} - \bar{u}_{i,j})/2
\]
(the value of $\bar{u}$ at a tropical or subtropical grid-point in the winter hemisphere is averaged with the value at the adjacent grid point on the polar side; no vertical adjustment is performed) in the NH until $\bar{u}_\phi = a f$ is satisfied. In the SH the scheme is similar but with the substitution $i + 1 \rightarrow i - 1$ and $i \rightarrow i$ in the expressions above. This scheme leads to the accumulation of angular momentum near the equator with a very sharp latitudinal
gradient at $\phi = 0$. This gradient is removed by an ad hoc extension of the adjustment region into the opposite hemisphere where $f$ is replaced by a function rapidly decaying to zero.

The discretized temperature field $\bar{T}_{ij}$ is also relaxationally adjusted at every time step to be in thermal wind-balance with the redistributed $\bar{u}_{ij}$. 
References


HUNT, B. G. 1981. The maintenance of the zonal mean state of the upper atmosphere as represented in a three-dimensional general circulation model extending to 100 km. J. Atmos. Sci. 38, 2172–2186.


