The general circulation of the atmosphere

Martin S. Singh

School of Earth, Atmosphere & Environment,
Monash University, Victoria, Australia

March 22, 2022

Acknowledgements
These notes draw heavily from teaching materials from a variety of sources on topics related to the general circulation. In particular, the development of these notes was strongly influenced by my direct experience as a student of Kerry Emanuel, Paul O'Gorman, and Richard Wardle and indirectly through my exposure to teaching materials authored by Isaac Held and Peter Stone. I thank Kate Bongiovanni, Nathan Eizenberg, and Corey Robinson for help proofreading these notes.

Contents

1 Introduction **1**
- 1.1 What is the general circulation? . . . 2
- 1.2 Observational preliminaries . . . 3
  - 1.2.1 Solar insolation forcing . . . 4
  - 1.2.2 Thermal structure of the atmosphere . . . 6
  - 1.2.3 The mean circulation . . . 9
  - 1.2.4 Eddies . . . 13
  - 1.2.5 The hydrological cycle . . . 14
  - 1.2.6 Summary . . . 15
- 1.3 Further reading . . . 16

2 Analysing the general circulation **19**
- 2.1 Governing equations . . . 19
  - 2.1.1 A statement of the problem . . . 19
  - 2.1.2 Conservation of mass . . . 20
  - 2.1.3 Conservation of momentum . . . 22
  - 2.1.4 Equation of state . . . 29
  - 2.1.5 The thermodynamic equation . . . 31
  - 2.1.6 Useful approximations and transformations . . . 32
- 2.2 Decomposing the circulation . . . 35
  - 2.2.1 Spatial and temporal averages . . . 35
  - 2.2.2 Combining the spatial and temporal averaging operators . . . 38
- 2.3 State estimation techniques . . . 39
  - 2.3.1 The Peixoto-Oort method . . . 39
  - 2.3.2 Modern analysis . . . 41

3 Radiative-convective equilibrium & Hide's theorem **43**
- 3.1 Why is there an atmospheric circulation? . . . 43
  - 3.1.1 Radiative equilibrium . . . 43

3.1.2 Radiative-convective equilibrium . . . 44
3.1.3 The tropical thermal structure . . . 47
3.2 Hide's theorem . . . 49
3.2.1 Column-by-column RCE with balanced zonal flow . . . 50
3.2.2 Constraints on the RCE state . . . 51
3.2.3 Angular momentum of the atmosphere . . . 51
3.2.4 Angular-momentum extrema in axisymmetric steady flow . . . 54
3.2.5 The violation of Hide's theorem . . . 56
3.3 Summary . . . 58

4 Axisymmetric Hadley Cells **59**
4.1 The Held & Hou model . . . 59
4.1.1 Conservation of angular momentum along streamlines . . . 60
4.1.2 Thermodynamic constraints on the Hadley Cell extent . . . 62
4.1.3 Tropopause zonal wind distribution . . . 64
4.1.4 Energy fluxes and the strength of the Hadley cell . . . 64
4.1.5 Summary of the Held & Hou model . . . 67
4.2 Limitations of the Held & Hou model . . . 67
4.2.1 Numerical simulations . . . 67
4.2.2 Angular-momentum sources and sinks . . . 68

5 The angular-momentum budget of the atmosphere **71**
5.1 Motivation . . . 71
5.2 Angular momentum of Earth's atmosphere . . . 72
5.2.1 The angular-momentum budget . . . 72
5.2.2 Angular-momentum budget of the upper troposphere . . . 73
5.3 Drag and the angular-momentum budget . . . 77
5.3.1 Surface friction . . . 79
5.3.2 Form drag . . . 79
5.3.3 Angular momentum budget and the midlatitude westerlies . . . 80
5.4 The global angular momentum cycle . . . 82

6 The maintenance of a barotropic jet **85**
6.1 Vorticity and circulation . . . 85
6.1.1 The vorticity equation . . . 85
6.1.2 Kelvin's circulation theorem . . . 87
6.2 Irreversible mixing and vorticity fluxes . . . 88
6.3 Momentum transport by Rossby waves . . . 90
6.3.1 Dynamics of Rossby wave momentum transport . . . 90
6.3.2 Kinematics of momentum transport . . . 92
6.4 Rossby wave propagation . . . 94

6.4.1 Jet maintenance: a conceptual picture 95

7 Eliassen-Palm fluxes and the transformed Eulerian mean 99
7.1 Quasigeostrophic preliminaries 100
7.2 Forcing of the zonal mean 102
7.2.1 Zonal-mean QG equations 102
7.2.2 Perturbation from a balanced state 104
7.2.3 Non-acceleration theorem 105
7.3 The Eliassen-Palm flux 107
7.3.1 the Transformed Eulerian Mean (TEM) 107
7.3.2 Quasigeostrophic potential vorticity (QGPV) 110
7.3.3 EP fluxes and Rossby waves 114
7.3.4 Summary and application to the atmosphere 117
7.4 Jets in a baroclinic atmosphere 119

8 The Hadley cell redux: eddy influences and the seasonal cycle 123
8.1 A non-axisymmetric Hadley Cell 123
8.1.1 Angular-momentum balance of the subtropics 123
8.1.2 Linear and angular-momentum conserving limits 125
8.1.3 What is the bulk Rossby number for the Hadley Cell? 127
8.2 Monsoons and the solsticial Hadley Cell 128
8.2.1 Geometrical constraints on the Hadley Cell 128
8.2.2 Extending Held & Hou to the solstice: the Lindzen & Hou model . . 130
8.2.3 Eddy-mean flow interactions and monsoons 132

Chapter 1

Introduction

This set of notes was created as a companion to the postgraduate-level unit *General circulation of the atmosphere* that I teach at Monash University. As such, it is not meant to be a comprehensive survey of all topics that one might consider under the umbrella of the general circulation of the atmosphere, but rather, it is a somewhat idiosyncratic tour through various concepts and models that could serve as an entry point to the vast general circulation literature that has built up over the last few decades.

While the notes include a brief introduction to the governing equations (see below), this should not be the student's first experience studying the dynamics of the atmosphere; ideally students should have taken undergraduate units in dynamical and physical meteorology and a basic introductory unit in fluid mechanics. For those that lack some of this background, a number of extra references are provided in section 1.3.

The notes may be roughly divided into three parts:

Part I (Chapters 1 & 2) includes a brief introduction to the basic characteristics of the general circulation and an outline of the main analytic tools used in the following parts. In particular, the *governing equations* of the atmosphere are introduced and the concept of *Reynolds averaging* (as it applies to temporal and zonal means) is described. Furthermore, we discuss some of the issues and techniques used in producing estimates of the general circulation from observations.

Part II (Chapters 2 & 3) considers the general circulation of the atmosphere from an *axisymmetric* perspective, where variations in the longitudinal direction may be neglected. While Earth's atmosphere is far from this limit, axisymmetric theory provides a reasonable zeroth order picture for (some aspects) of the zonally-averaged tropical circulation. Additionally, it has been argued (Held and Hou, 1980) that the importance of zonal asymmetries (i.e., *eddies*) in shaping the large-scale characteristics of the circulation may be

best appreciated by first considering the circulation in their absence.

Part III (Chapters 5-8) analyses the effects of eddies and their importance for determining the basic structure of the circulation from the perspective of the *angular-momentum budget*. In particular, the observed transports of angular-momentum within the atmosphere are introduced, and the relationship between latitudinal angular-momentum transport, zonal jets, and the meridional overturning is elucidated. The combined effects of eddy momentum and eddy heat transports on the zonal-mean circulation are then analysed through the powerful framework of the *Transformed Eulerian Mean*. Finally, the effect of eddies on the tropical circulation is considered, and the limitations of the axisymmetric theory presented in Part II are discussed.

In a future edition, these notes will include an additional Part IV in which two other important atmospheric budgets are discussed: those of *energy* and *water vapour*. For the present time, however, these budgets are not discussed in detail.

The notes are designed to be read in sequence, as certain concepts used later are introduced in preceding chapters, and they are designed to be covered in roughly one to one-and-a-half semesters. Parts I-III are taught as a one-semester postgraduate-level unit at Monash University. Because of this relative brevity, I have picked particular aspects of general circulation theory to focus on and skipped others entirely. More complete treatments of various topics can be found in a range of other texts; some of these are described in section 1.3.

1.1 What is the general circulation?

The subject of these notes is *the general circulation of the atmosphere*. But what do we actually mean by "general" circulation, and how does this differ from the study of atmospheric dynamics?

The American Meteorological Society's online glossary includes the following definition of the general circulation of the atmosphere:

> *"In its broadest sense, the complete statistical description of large scale atmospheric motions."*

This definition has two aspects. The first is straightforward–the general circulation has to do with "large-scale" atmospheric motions. Typically, this refers to motions that are at synoptic scales and larger. Secondly, the study of the general circulation is concerned with a statistical description of the flow. Thus, we are here concerned with aspects of the atmospheric flow that do not depend on the detailed initial conditions of a given realisation. We do not wish to explain or predict the particular position of the midlatitude jet next week, for example. Instead, we are interested in understanding the dynamical processes

that give rise to a jet in the first place, as well as the physical parameters that determine its mean position, the extent of its fluctuations, or indeed whether there should be multiple jets in each hemisphere.

Another key aspect of general circulation studies is a focus on how the atmosphere fulfils its role in the cycling of various quantities–energy, momentum, water, and other chemicals–through the climate system. In these notes we will place a particular focus on transports of angular momentum (Chapter 5), and future editions will include chapters on the energy and water budgets of the atmosphere. Ultimately, atmospheric transports of these various quantities are effected by the flow systems that make up the day-to-day weather that might be the focus of a course on dynamical meteorology. But in the study of the atmospheric general circulation, we are interested in weather systems primarily for their contribution to the overall atmospheric transports.

While a full description of the general circulation includes statistics of the three dimensional wind and thermodynamic fields and their spatial and seasonal variation, a particular focus historically has been the time- and zonal-mean meridional overturning circulation. Readers should be familiar with the components of this overturning–the Hadley, Ferrel, and Polar cells; a primary focus of these notes will be to explain how these circulations come about. A clear picture of this three-celled overturning circulation only emerged about a century ago. Until that time, considerable argument remained as to the meridional and vertical extent of the cells, and indeed the number of cells within each hemisphere. This historical development is beautifully summarised in the following concise and very readable essay:

Lorenz, E., 1983: A history of prevailing ideas about the general circulation of the atmosphere. *Bulletin of the American Meteorological Society*, **64**, 730–769

which students are encouraged to read.

1.2 Observational preliminaries

In this section we present an overview of the observed general circulation along with some questions that will motivate the presentations in later chapters. A first question one might ask is how best to estimate the general circulation. Most observations of the climate system (e.g., surface air temperature measurements) are taken at irregular time and/or space intervals, but to quantify time averages and other statistics, a regular grid is most useful. The plots shown in this chapter are based on datasets that use a variety of techniques to produce estimates of the circulation on a latitude-longitude grid. In particular, a number of plots in this chapter use the NCEP-DOE reanalysis dataset (Kanamitsu et al., 2002), which combines observations and the output of a numerical model to obtain an estimate of the three-dimensional state of the atmosphere through time. In Chapter 2, we will

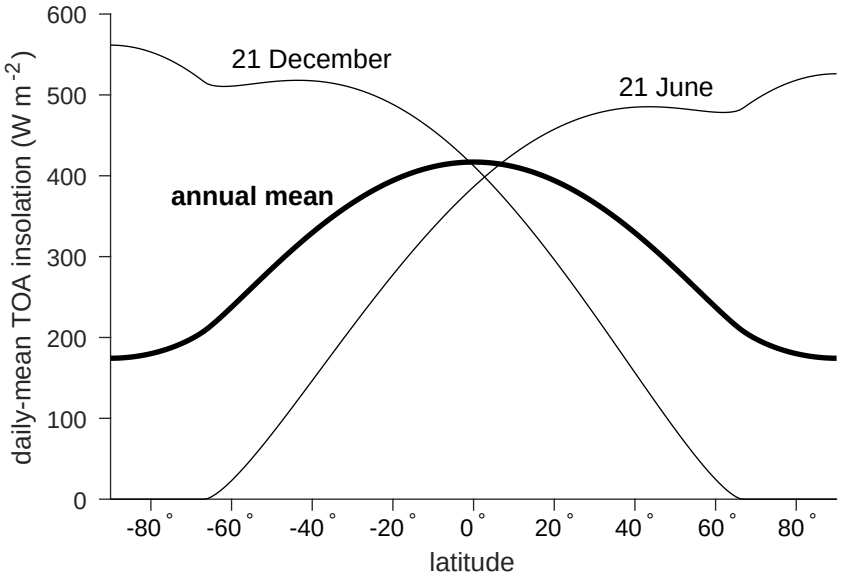

Figure 1.1: Daily-mean top-of-atmosphere insolation as a function of latitude for the Austral summer solstice (21 December), the Austral winter solstice (21 June), and the annual mean.

further discuss some of the techniques used for such *state estimation*, including those used in modern reanalyses.

1.2.1 Solar insolation forcing

We begin by considering the ultimate driver of the general circulation, the input of energy from the sun. As we shall see in later chapters, the fundamental reason for the existence of any large-scale circulation in Earth's atmosphere is the uneven distribution of solar radiation impinging on the Earth at different latitudes. Figure 1.1 shows this latitudinal variation for different times of year. Plotted is what is known as the insolation–the total energy impinging on a given area of the Earth's surface per unit time–at each latitude, measured at the top of the atmosphere (TOA) and averaged over a day. In the annual mean, the TOA insolation peaks at the equator, decreasing monotonically to the pole. This is because at high latitudes, the angle of the sun in the sky tends to be low, spreading out the incoming rays over a larger area of Earth's surface.

At the Solstices, however, the maximum daily-mean insolation occurs at the summer pole. At this time of year, latitudes poleward of the Antarctic/Arctic circle are either in perpetual daylight or perpetual night, depending on the season. The effect of having 24 hours of continuous sunlight overwhelms the effect of solar zenith angle, allowing the summer pole to receive more radiation than any other point on the Earth's surface. Given this distribution

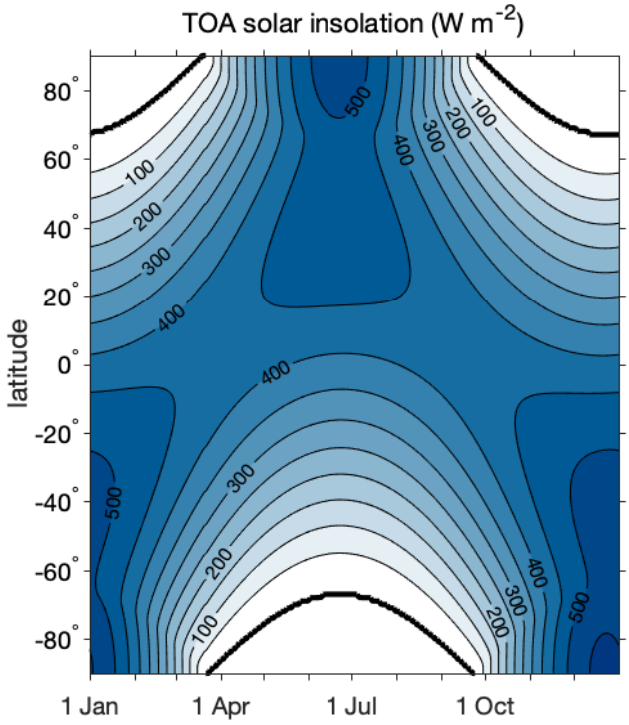

Figure 1.2: Daily-mean top-of-atmosphere insolation as a function of latitude and time of year in units of Watts per metre squared. Thick line shows the zero contour; regions poleward of this contour are experiencing the polar night.

of insolation, one might ask why the warmest place on Earth on the 21 of December is not the South Pole. More generally, we might ask the question: what sets the position and seasonal migration of the maximum surface temperature on the Earth?

The seasonal cycle of solar insolation is shown in more detail in Fig. 1.2. It is clear from this figure that the strongest meridional gradient in the TOA insolation occurs in the winter hemisphere. Does this also imply that the winter hemisphere has stronger meridional temperature gradients? How does this affect the general circulation and its seasonal cycle?

It may also be seen in both Fig. 1.1 and Fig. 1.2 that the maximum daily-mean insolation is higher at the Austral summer solstice than at the Boreal summer solstice. This is a result of the Earth's elliptical orbit around the sun; at present, the Earth reaches its closest approach to the sun (*perihelion*) around the 4th of January, close to the summer solstice in the southern hemisphere. This means that, from the point of view of the incoming insolation, the Southern hemisphere has more intense summers, but also shorter summers,

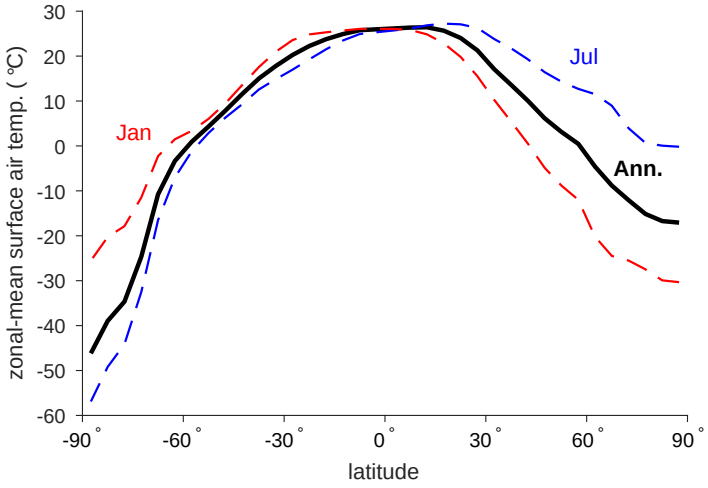

Figure 1.3: Zonal- and time-mean surface air temperature for January (red), July (blue), and the annual mean (black) calculated based on the CRUTEM4 gridded dataset for the years 1961-1990.

since a planet moves fastest around its orbit when it is closest to its star according to Kepler's Laws.

The timing of perihelion relative to the solstices is not fixed, but rather it varies as part of Earth's precessional cycle with a period of roughly 25 thousand years. This precession of Earth's orbit is known to exert an influence on the climate. In the mid-Holocene period (roughly 6000 years ago), perihelion occurred during the Northern Hemisphere summer, resulting in stronger Northern Hemisphere monsoons, and contributing to the "greening of the Sahara" in which the region that is now the Sahara desert was able to support vegetation such as is found in tropical savannahs.

The precessional cycle, in combination with similar cycles in the eccentricity and obliquity (tilt of Earth's rotation axis relative to the orbital plane) of Earth's orbit are collectively known as Milankovitch Cycles after the Serbian mathematician who first studied them. These cycles are known to be important in determining the timing of the periodic glacial episodes that have characterised Earth's climate variability for the past two million years.

1.2.2 Thermal structure of the atmosphere

Near-surface temperature

We next consider the observed thermal structure of the atmosphere. Figure 1.3 shows the time-mean surface air temperature, averaged around latitude circles (what is known as a zonal mean) according to the CRUTEM4 dataset developed by the Climate Research Unit

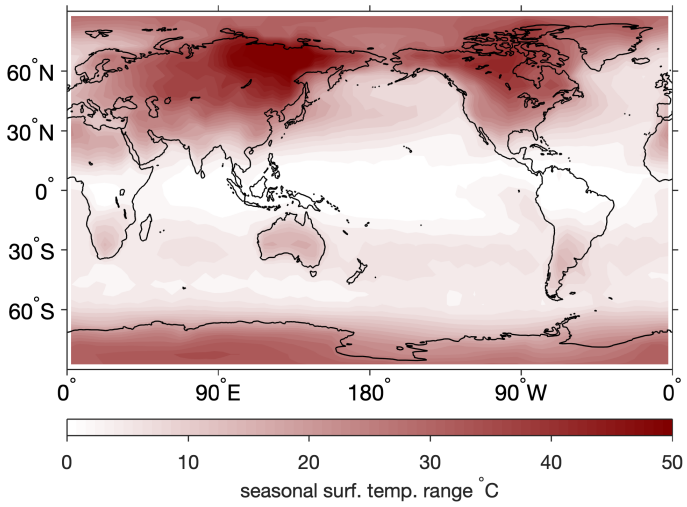

Figure 1.4: Seasonal range in surface air temperature, taken as the difference between the maximum and minimum monthly mean calculated based on the CRUTEM4 gridded dataset for the years 1961-1990.

at the University of East Anglia (Osborn and Jones, 2014). Unlike the solar insolation, the peak in the zonal-mean temperature remains relatively close to the equator throughout the year. Also unlike the solar insolation, northern hemisphere summer temperatures are higher than those at equivalent southern latitudes in the Austral summer. Nevertheless, the seasonal variation in temperature is clearly largest in polar regions, where the seasonal variation in insolation is also largest, suggesting that the insolation does play a major role in determining the observed temperature distribution.

A clue as to the reason why temperatures do not simply follow the pattern of insolation may be found in Fig. 1.4. The seasonal range of surface air temperature varies strongly across the Earth's surface. In addition to a general increase in seasonal range with latitude, consistent with the distribution of solar insolation, there is a strong land-ocean contrast in this quantity. The further one is away from the ocean, the larger the seasonal range in temperature. The reason for this is primarily the difference in the thermal heat capacity of the land compared to the ocean. In oceanic regions, a change in surface temperature on seasonal timescales must be accompanied by a change in the temperature of a large portion of the ocean mixed layer, which has a depth of tens of metres. In continental regions, seasonal temperature changes occur in the upper few centimetres of the soil layer. The amount of heat required to warm the air near the surface is therefore much greater in oceanic regions than over a continent. This allows continental regions to respond much more quickly to the seasonal variation in solar insolation and to undergo much wider swings

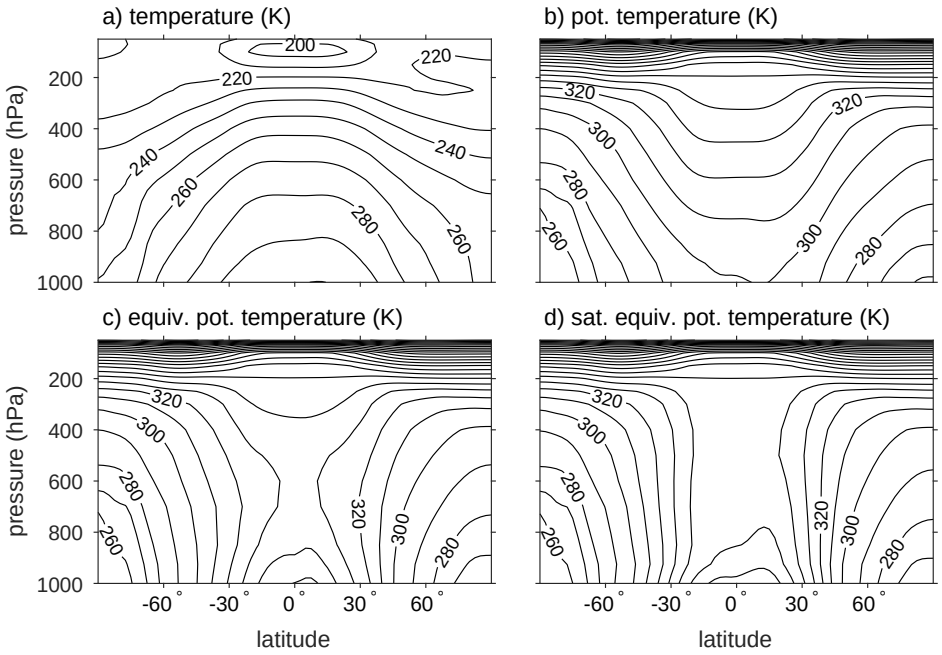

Figure 1.5: Annual- and zonal-mean (a) temperature, (b) potential temperature, (c) equivalent potential temperature, and (d) saturation equivalent potential temperature as a function of latitude and pressure according to the NCEP-DOE reanalysis for the years 1981-2010. Potential temperatures calculated based on the time and zonal-mean pressure, temperature, and (for equivalent potential temperature) relative humidity and neglecting ice processes.

in temperature on seasonal (and diurnal) timescales.

The finite timescale on which surface air temperature responds to changes in heating rates partially accounts for the observed differences between the distribution of solar insolation and the observed distribution of surface air temperature. Other effects include different albedos in different regions that affect the amount of solar insolation that is actually absorbed by the Earth and atmosphere, and the effect of the atmospheric circulation, which acts to spread the influence of oceanic regions to land regions and vice versa.

Vertical structure

It is well known that the temperature of the atmosphere typically decreases with height up to the tropopause and increases above. But what sets the lapse rate in the troposphere? And what sets the height of the tropopause?

Figure 1.5 shows the zonal- and time-mean thermal structure of the atmosphere estimated from the NCEP-DOE reanalysis using four different variables: the temperature, potential

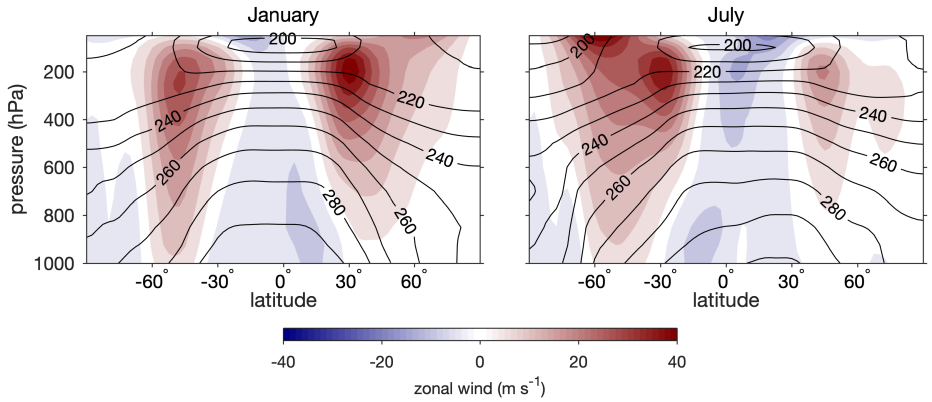

Figure 1.6: Time- and zonal-mean temperature (contours; K) and zonal-wind (colours) for January (left) and July (right) as a function of latitude and pressure according to the NCEP-DOE reanalysis for the years 1981-2010.

temperature, equivalent potential temperature, and saturation equivalent potential temperature[1]. As expected, temperature decreases with height within the troposphere and, as required by gravitational stability, potential temperature increases with height at all latitudes.

The distributions of the equivalent potential temperatures are more interesting: in the tropical regions, the equivalent potential temperature θ_e tends to have a local minimum in the mid troposphere, while the saturation equivalent potential temperature θ_e^* is almost constant with both height and latitude, at least within the free troposphere. Can we understand why the atmosphere maintains a state of roughly constant saturation equivalent potential temperature in tropical regions? Conversely, what is different about extratropical regions that causes θ_e^* to increase with height? More generally, what controls the thermal stratification of the atmosphere? Despite their apparent simplicity, these questions are still the subject of active research. We will discuss how the tropical atmospheric thermal structure is maintained in chapter 3.

1.2.3 The mean circulation

A large component of the study of the general circulation of the atmosphere is devoted to understanding how the zonal- and time-mean circulation is maintained. In this section we present an overview of the mean circulation in Earth's atmopshere.

Figure 1.6 shows the zonal winds as a function of latitude and pressure for January and July. The main features of the plot are the strong westerly jets in the subtropical to midlatitude

[1]Readers unfamiliar with these variables can find a lucid description in ch. 4 of Emanuel (1994) or a slightly simpler treatment in ch. 3 of Wallace and Hobbs (2006).

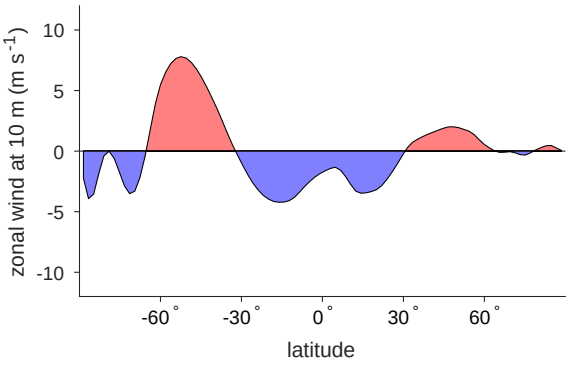

Figure 1.7: Time- and zonal-mean zonal wind at 10 m above the ground according to the NCEP-DOE reanalysis for the years 1981-2010. Red shading represents westerly winds and blue shading represents easterly winds.

regions of each hemisphere and the weak easterlies existing in the tropical region in between. There is typically one jet in each hemisphere, although evidence of a double jet structure exists in the Austral winter, and the jets are strongest in the winter season. What is the cause of these jets in the upper atmosphere? The temperature contours plotted in Fig. 1.6 demonstrate that the strongest upper tropospheric winds coincide with regions of strong horizontal temperature gradients. Indeed, readers familiar with the thermal wind relation (see chapter 3) will recognise that, for an atmosphere close to a state of geostrophic and hydrostatic balance, the vertical increase in the zonal wind is proportional to the rate of increase of temperature toward the equator. On Earth, where solar insolation is highest at the equator (in the annual mean), this implies westerly winds in the upper atmosphere. Furthermore, in winter, where the solar insolation gradient is highest, the temperature gradient is also largest, and the westerlies are strongest.

The argument above accounts for some aspects of the observed westerly jets based on the radiative forcing alone. But does this simple balance argument provide a complete explanation of the wind distribution? For example, Fig. 1.7 shows the time- and zonal-mean zonal wind distribution at 10 m above the Earth's surface (as estimated by the NCEP-DOE reanalysis). There is a clear pattern of westerly winds at midlatitudes and easterly winds in tropical regions. Since the thermal wind relation relates the vertical gradient of the wind to the temperature distribution, it cannot be used to reason about the winds at the atmosphere's lower boundary. What drives the pattern of westerly and easterly surface winds on the Earth?

Another gap in our argument is that the thermal wind relation relates upper-tropospheric westerly jets to tropospheric temperature gradients, but it doesn't explain why those temperature gradients should be localised in the extratropics. This point can be made particularly clearly by considering the winds not in their time- and zonal-mean, but in a daily

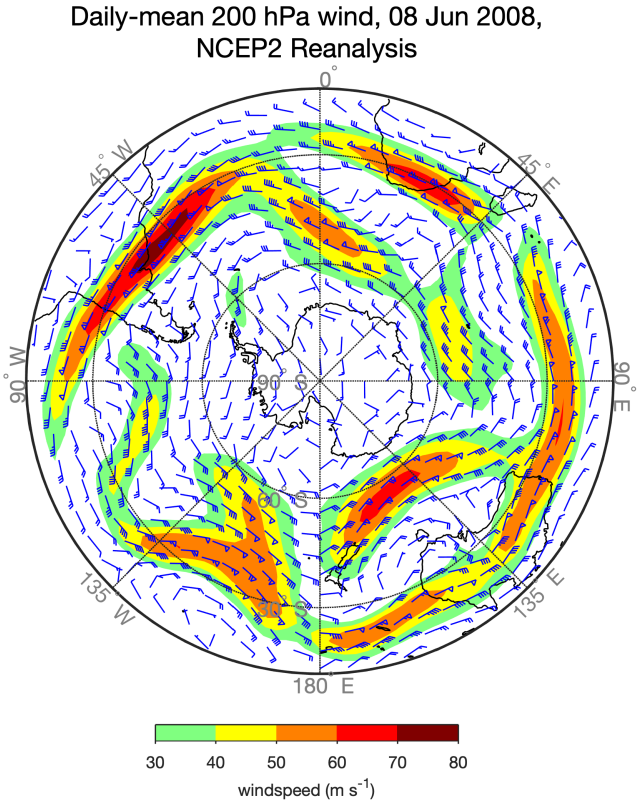

Figure 1.8: Daily-mean winds at the 200 hPa level on the 8th of June, 2008 according to the NCEP-DOE reanalysis. The direction and speed of the winds are shown by wind barbs in units of m s^{-1}, with regions of strong winds highlighted by colours.

snapshot. Fig. 1.8 shows the daily-mean horizontal winds at the 200 hPa level over the southern hemisphere for a single day in the Austral winter of 2008. The figure shows that, on daily timescales, jets in the atmosphere are not broad features occupying tens of degrees of latitude, but relatively narrow and sharply defined regions of intense winds that vary with both latitude and longitude. In the particular snapshot shown, evidence can also be seen of multiple jet, with a set of wind maxima at roughly 30°S (the subtropical jet) as well as further poleward around 50°S (the subpolar jet).

By thermal wind balance, these sharp localised jets must be related to localised temperature gradients in the troposphere. But why should such temperature gradients form given the relatively smooth distribution of solar insolation? And what determines the number of jets/regions of strong temperature gradients? Clearly there are dynamical process that favour the formation of jets and their associated sharp temperature gradients. In chapter 6,

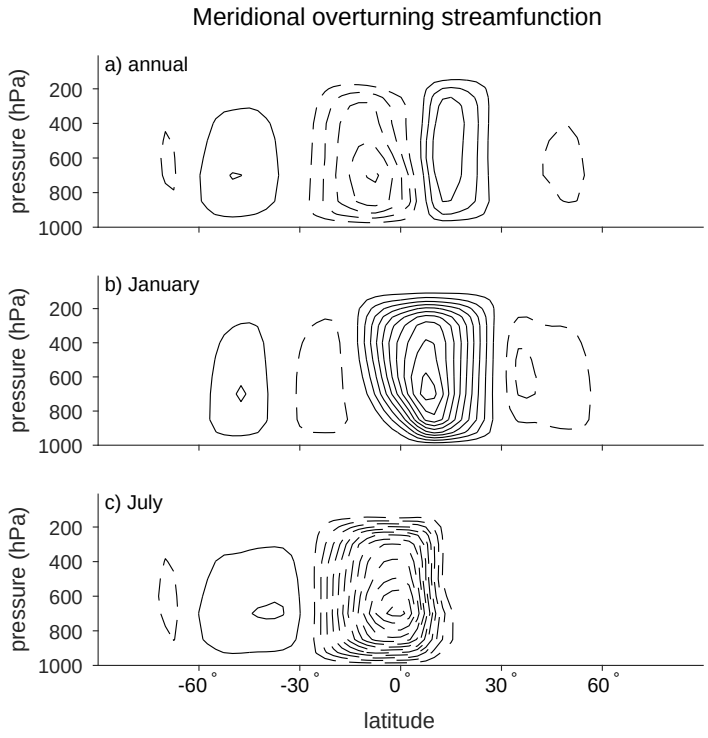

Figure 1.9: The meridional overturning streamfunction (see section 2.1.6) calculated based on NCEP-DOE reanalysis data for the years 1981-2010 and for (a) the annual mean, (b) January, and (c) July. Solid lines represent clockwise motion and dashed lines represent anticlockwise motion and the contour interval is 2×10^9 kg s^{-1}.

we will discuss some of the mechanisms that lead to jet formation in the atmosphere.

As mentioned above, the meridional overturning circulation in Earth's atmosphere consists of three cells in each hemisphere–the Hadley, Ferrel, and Polar cells. Figure 1.9 visualises these cells using the meridional overturning streamfunction Ψ. Contours of Ψ are parallel to the direction of the zonal-mean flow in the latitude-height plane, and the density of Ψ-contours represents the size of the mass flux within the circulation (see also section 2.1.6). Of note is that the Hadley cells dominate the meridional overturning, particularly in the solsticial seasons, while the polar cell is barely visible at the contour interval used. In the solsticial seasons, the winter Hadley Cell expands and increases dramatically in strength, while the summer Hadley cell weakens to be even weaker than the winter Ferrel cell. This seasonal rearrangement of the tropical circulation is the zonal-mean expression of Earth's monsoons.

In all seasons, the Hadley cells in Earth's atmosphere are confined to the tropical regions,

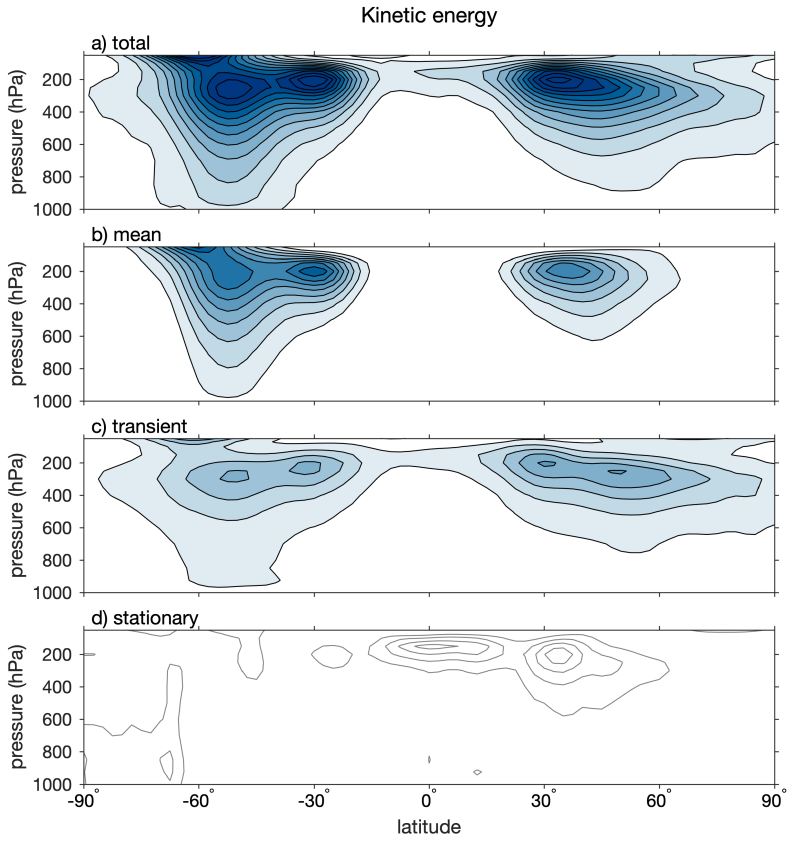

Figure 1.10: Kinetic energy as a function of latitude and pressure according to the NCEP-DOE reanalysis. (a) Total kinetic energy, and kinetic energy broken into (b) zonal- and time-mean component, (c) transient eddy component, and (d) stationary eddy component. Contour interval is 50 $m^2\ s^{-2}$ except in panel (d) in which it is 10 $m^2\ s^{-2}$.

generally not extending beyond 30° north or south of the equator. In contrast, the Hadley cells on Mars, as well as the Saturnian moon Titan, extend almost from pole to pole. What sets the latitude at which the Hadley Cell terminates, and how might this change under a warmer climate? Moreover, the Ferrel cell that exists at midlatitudes is thermally indirect; it lifts cool subpolar air upwards and pushes warm subtropical air downwards. How is such a circulation maintained, and why is it necessary? These questions are considered in chapters 5 & 6.

1.2.4 Eddies

So far we have considered the atmosphere primarily from a time-mean and zonal-mean perspective. But variations from the mean are obviously important for day-to-day weather,

and we shall see that they are important for the general circulation through their transport of energy, water and momentum. In these notes, we will refer collectively to variations from the time and zonal mean as eddies. This is irrespective of wether these variations take the traditional form of eddies or if they are more wave-like. Furthermore, we will distinguish between two types of eddies; transient eddies that vary with time, and stationary eddies, that vary zonally, but not in time. This definition will be made more mathematically precise in chapter 2.

Figure 1.10 shows an example of the above decomposition for the kinetic energy $\frac{1}{2}|\mathbf{u}|^2$, where $\mathbf{u}$ is the vector velocity. Panel (a) shows the total kinetic energy calculated by averaging the square of the velocity, while panels (b-d) show the contribution of the zonal- and time-mean wind, the transient eddies, and the stationary eddies, respectively. While the mean flow accounts for a substantial fraction of the kinetic energy in the vicinity of the jet, a considerable fraction of the kinetic energy in the atmosphere is associated with transient eddies, particularly at the flanks of the jet. We will see in later chapters that these eddies effect important transports of energy and momentum that cannot be ignored in any complete account of the general circulation.

Stationary eddies are relatively weak in Fig. 1.10, suggesting that their importance for the general circulation is limited. However, it should be noted that the decomposition of the eddy field into stationary and transient components depends somewhat on the definitions used. For example, in Fig 1.10, we have taken the time mean as an annual mean, so all deviations from this annual mean count as transient eddies. If, instead, the time mean was taken for a given month of the year, then zonally asymmetric seasonal variations (such as the monsoons) would be included in the stationary eddy component, and the stationary eddy contribution to the kinetic energy would be considerably higher.

The division of circulation statistics into a mean and eddy component is a powerful tool for analysing the general circulation. It amounts to an application of Reynold's decomposition, as may be familiar to readers who have studied turbulence. We discuss this technique in more detail in chapter 2.

1.2.5 The hydrological cycle

Finally, we turn to the observed hydrological cycle. Water enters the atmosphere through evaporation from the surface, and leaves it through precipitation. Both these processes are strongly influenced by the atmospheric circulation, with evaporation depending on the low-level windspeed and thermodynamic properties of the air, and precipitation being strongly associated with upward motion within the atmosphere.

The imprint of the general circulation on the hydrological cycle may be seen at a glance by examining estimates of the observed precipitation distribution in Fig. 1.11. High precipitation rates are concentrated in the intertropical convergence zone over the oceans

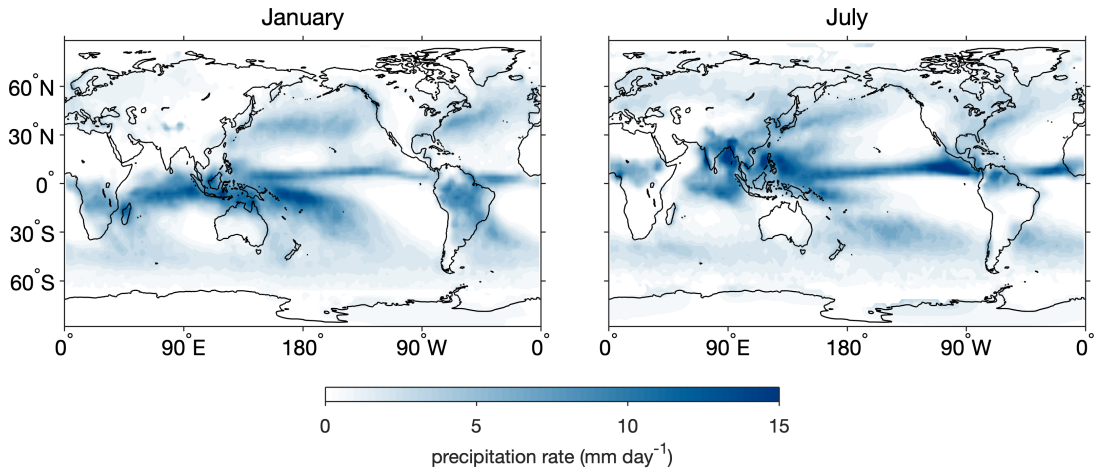

Figure 1.11: Time-mean precipitation rate for (left) January and (right) July calculated based on the CMAP gridded dataset (Xie and Arkin, 1997) for the years 1981-2010 .

and in monsoon regions over tropical continents. These regions correspond to the rising branches of the atmospheric circulation, which, in the zonal mean, corresponds to the Hadley circulation.

At midlatitudes, high precipitation rates are concentrated in the storm tracks, regions where midlatitude weather systems are most often found. These are prominent in the western side of the north Pacific and Atlantic oceans, as well as in the Southern Ocean. Unlike in the tropics, the zonal-mean overturning circulation provides a poor indicator of the regions of strong precipitation. This is because of the importance of eddies and the positive-definite nature of precipitation; transient eddies produce both upward and downward motion in the atmosphere, but while upward motion is associated with precipitation, downward motion does not correspond to negative precipitation. As a result, an eddy that produces no zonal-mean ascent can nevertheless produce substantial zonal-mean precipitation.

1.2.6 Summary

The preceding discussion presented an overview of some of the main features of the general circulation of Earth's atmosphere and motivated a number of questions about the general circulation and the processes that maintain it. In this set of notes we will only address a subset of these questions, and we will not necessarily provide full answers to those that are addressed. This is both to keep the scope of these notes manageable, but also because not all the answers are fully known. Questions as simple as "what sets the thermal stratification of the atmosphere?" are still actively researched to this day.

The purpose of these notes are therefore to provide the reader with the tools, and hopefully the inspiration, to engage with research on the large-scale atmospheric circulation.

1.3 Further reading

Like any set of teaching materials, these course notes draw heavily from other texts and the literature. In particular, the following resources have been particularly influential in the development of the general circulation of the atmosphere unit at Monash University.

1. Vallis, G. K., 2017: *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. 2nd ed., Cambridge University Press, 946 pp.

   This book provides an up-to-date and comprehensive treatise on large-scale atmospheric dynamics, from the derivation of the governing equations to the general circulation. Almost every topic covered in these notes are covered in more detail within this book, I cannot recommend it highly enough. Chapters 14, 15 and 18 are particularly relevant.

2. Stone, P., 2005: General circulation of the Earth's atmosphere. *MIT Open Courseware*, Woods Hole Oceanographic Institution, URL https://ocw.mit.edu/courses/earth-atmospheric-and-planetary-sciences/12-812-general-circulation-of-the-earths-atmosphere-fall-2005/.

   Lecture notes from a course on the general circulation of the atmosphere at the Massachusetts Institute of Technology, originally taught by Peter Stone. Provides an overview of the atmospheric budgets of angular momentum, energy, water vapour, and also includes a section on the entropy budget of the atmosphere.

3. Held, I. M., 2000: The general circulation of the atmosphere. *Proc. Geophysical Fluid Dynamics Program*, Woods Hole Oceanographic Institution, 1–54, URL https://www.whoi.edu/fileserver.do?id=21464&pt=10&p=17332.

   Notes from a beautiful set of lectures given by Isaac Held at the Woods Hole Oceaographic Institution's summer school in the year 2000. Perhaps a bit dated now, but well worth a read. Chapters 6 and 7 of the current notes are highly influenced by these lectures.

4. Peixoto, J. P. and A. H. Oort, 1992: *Physics of Climate*. AIP Press, 520 pp.

   One of the first attempts to produce a coherent estimate of the general circulation from observations. Discusses the main budgets of the atmosphere and shows estimates of the terms within them. While modern reanalysis techniques are considerably more sophisticated, the simplicity of the Piexoto-Oort method is useful for its transparency

and the insight it provides to the process of state estimation. This method is described in chapter 2 of these notes.

For those students that lack a solid background in dynamical meteorology, the following book provides a concise introduction to atmospheric dynamics that may be of use:

- Vallis, G. K., 2019: *Essentials of Atmospheric and Oceanic Dynamics*. Cambridge University Press, 366 pp.

For those students with a strong background in fluid dynamics, but a lack of specific meteorological knowledge, the following texts provide an introduction to many of the important concepts of atmospheric science and climate science

- Hartmann, D. L., 1994: *Global physical climatology*. Academic Press, 411 pp.
- Wallace, J. and P. Hobbs, 2006: *Atmospheric science: an introductory survey*. 2d ed., Academic Press.

Finally, these notes assume a reasonable familiarity with atmospheric thermodynamics, including the various conserved variables used within the atmospheric sciences. A basic treatment is provided by Wallace and Hobbs (2006) given above. A more advanced treatment is given in

- Emanuel, K. A., 1994: *Atmospheric convection*. Oxford University Press, 580 pp.

Chapter 2

Analysing the general circulation

This chapter presents a short introduction to the governing equations of the atmosphere that will be used to analyse the general circulation. I assume that students have been exposed to these equations before, and so the derivations contained here are not complete and not rigorous. For a more complete treatment of the equations for large-scale flow within the atmosphere, the reader is invited to consult chapters 1 and 2 of Vallis (2017).

2.1 Governing equations

2.1.1 A statement of the problem

The state of the atmosphere at a given time may be wholly described by the specification of six key variables:

- three components of the vector velocity $\mathbf{u} = (u, v, w)$
- pressure p
- temperature T
- density ρ

In addition, the humidity in the atmopshere (and salinity in the ocean) are often dynamically important. An extra equation is usually required to track the concentration of these trace species. We therefore require six equations (plus one for humidity) and an appropriate set of boundary conditions to solve for the evolution of the fluid. These equations are:

- mass continuity (conservation of mass)
- Newton's second law applied to the fluid (conservation of momentum; 3 equations)
- the first law of thermodynamics (conservation of total energy)
- an equation of state (functional relationship between pressure, density and temperature)

In the next section, we sketch out heuristic derivations of these equations. For a more complete treatment, the reader is referred to sections 1.1-1.5 of Vallis (2017). For a fundamental derivation of the equations from Hamilton's principle, see the excellent lecture notes on geophysical fluid dynamics by Salmon (1998).

2.1.2 Conservation of mass

The equation of conservation of mass expresses the physical principle that mass cannot be created or destroyed. Consider a control volume, fixed in space, denoted by V_0. Conservation of mass requires that the net rate of change of the mass within the control volume must be equal to the flux of mass across its boundaries. Mathematically we may write,

$$\frac{\mathrm{d}}{\mathrm{d}t} \oint_{V_0} \rho \, \mathrm{d}V = - \oint_{\partial V_0} \rho \mathbf{u} \cdot \hat{\mathbf{n}} \, \mathrm{d}S, \tag{2.1}$$

where ρ is the fluid density, $\mathbf{u}$ is the vector fluid velocity, ∂V_0 represents the boundary of V_0, and $\hat{\mathbf{n}}$ is an outward unit normal. The left-hand side represents the change in mass within the control volume, and the right-hand side represents the flux of mass into the control volume. Since the volume is fixed, we can take the time derivative inside the integral as a partial derivative. Additionally applying the divergence theorem to the right-hand side gives,

$$\oint_{V_0} \frac{\partial \rho}{\partial t} \, \mathrm{d}V = - \oint_{V_0} \nabla \cdot (\rho \mathbf{u}) \, \mathrm{d}V, \tag{2.2}$$

Since (2.2) is valid for an arbitrary control volume, it can only be satisfied if the integrand is identically zero. Hence we may write a differential equation for mass continuity given by,

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0. \tag{2.3}$$

The material parcel

It will prove useful to consider in addition to a control volume fixed in space, a volume that moves with the fluid, known as a *material parcel*, which we denote V_t. The material parcel can be thought of as a mass of fluid made up of fluid elements that have been tagged so that the volume enclosed by the parcel "follows the flow". In particular, this means

that the velocity of the boundary of the parcel ∂V_t is equal to the fluid velocity at each point. Furthermore, by definition, the mass of a material parcel cannot change[1], so we have that,

$$\frac{\mathrm{d}}{\mathrm{d}t} \oint_{V_t} \rho \, \mathrm{d}V = 0. \tag{2.4}$$

The material parcel viewpoint corresponds to a Lagrangian perspective on the flow, while the control volume approach takes an Eulerian perspective. To convert between the material parcel view and the control volume view, we therefore need to convert from an Eulerian to a Lagrangian perspective. To do so, consider the following coordinate transformation.

We define the traditional Eulerian coordinates $\mathbf{x} = (x, y, z)$ as fixed in an inertial reference frame. We then define the Lagrangian coordinates $\mathbf{x_0} = (x_0, y_0, z_0)$ which refer to the position of a fluid element at some time $t = 0$. As the fluid evolves, the Eulerian coordinates of each fluid element (x, y, z) may change, while the Lagrangian coordinates, (x_0, y_0, z_0) remain unchanged. On the other hand, at a fixed location, (x, y, z) are unchanging, but the coordinates (x_0, y_0, z_0) vary as fluid elements with different starting locations pass through that particular location.

Now, by definition, the velocity of a given fluid element is given by,

$$\mathbf{u} = \left.\frac{\partial \mathbf{x}}{\partial t}\right|_{\mathbf{x_0}}, \tag{2.5}$$

where the partial derivative is taken at constant values of the Lagrangian coordinate. Furthermore, we may use the chain rule to relate the Lagrangian and Eulerian time derivatives. Consider a scalar variable which may be expressed in either coordinate system: $\gamma = \gamma(x, y, z, t) = \gamma(x_0, y_0, z_0, t)$. Taking the Lagrangian time derivative, we have,

$$\begin{aligned}\left.\frac{\partial \gamma}{\partial t}\right|_{\mathbf{x_0}} &= \left.\frac{\partial \gamma}{\partial x}\frac{\partial x}{\partial t}\right|_{\mathbf{x_0}} + \left.\frac{\partial \gamma}{\partial y}\frac{\partial y}{\partial t}\right|_{\mathbf{x_0}} + \left.\frac{\partial \gamma}{\partial z}\frac{\partial z}{\partial t}\right|_{\mathbf{x_0}} + \left.\frac{\partial \gamma}{\partial t}\right|_{\mathbf{x}}\frac{\partial t}{\partial t} \\ &= \mathbf{u}\cdot\nabla\gamma + \left.\frac{\partial \gamma}{\partial t}\right|_{\mathbf{x}}\end{aligned}$$

where partial derivatives with respect to an Eulerian coordinate (x, y, z) are taken while holding other Eulerian coordinates constant.

Since equations are typically defined in Eulerian coordinates, we denote the expression above for the Lagrangian derivative a special operator,

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \mathbf{u}\cdot\nabla, \tag{2.6}$$

[1] This is not strictly true if molecular diffusion is considered, since mass transport by diffusion is separate from the mass transport by the velocity $\mathbf{u}$. See Salmon (1998) for a detailed discussion of this separation between organised fluid motions that make up $\mathbf{u}$ and the random motions that make up molecular diffusion.

and henceforth we assume all partial derivatives are taken with Eulerian coordinates held constant.

Using (2.6), the equation for mass continuity (2.7) may be rearranged in terms of the Lagrangian derivative to give an evolution equation for density,

$$\frac{\mathrm{D}\rho}{\mathrm{D}t} = -\rho\nabla\cdot\mathbf{u}. \tag{2.7}$$

2.1.3 Conservation of momentum

To derive the equation for conservation of momentum in the atmosphere, we apply Newton's second law to a material parcel of fluid,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_t}\rho\mathbf{u}\,\mathrm{d}V = \sum_i \mathbf{F}_i, \tag{2.8}$$

where the integral on the left-hand side represents the momentum of a material fluid parcel, and the right hand side represents the sum of the forces on that fluid parcel $\mathbf{F}_i$.

To evaluate the right-hand side we need to tally the forces acting on the fluid. In general, the forces acting on the fluid may be body forces, which act on all elements of the fluid and result from an externally imposed potential, or surface forces, which act at the interface between the parcel under consideration and its environment and arise from the interaction of the fluid internally or with rigid boundaries. A full account of how the various surface and body forces arise is beyond the scope of these notes, and we present only a sketch here (see Salmon, 1998).

The main forces of relevance to geophysical fluids are:

- gravitational force (body)
- pressure force (surface)
- viscous stresses (surface)

Consider first the gravitational force $\mathbf{F}_g$ acting on a material parcel. As a body force, the gravitational force acts at every point within the fluid. We can express the force per unit mass at each point in terms of a gravitational potential Φ_g, so that the force acting on the entire parcel may be written,

$$\mathbf{F}_g = -\oint_{V_t}\rho\nabla\Phi_g\,\mathrm{d}V\,. \tag{2.9}$$

In principle, the gravitational potential is calculated by integrating the potential from all objects in the universe, including the atmosphere itself, based on Newton's law's of universal gravitation. However, for geophysical fluids, the gravitational potential of the Earth is the

only significant contributor. We will return to the precise form of the gravitational potential for terrestrial problems below.

Now consider the pressure force acting on the same material parcel $\mathbf{F}_p$. Pressure represents a force per unit area exerted by the environment on the parcel that acts to compress the parcel. This means that the pressure force acts normal to the parcel's surface, and in the opposite direction to the outward normal $\hat{\mathbf{n}}$ defined above. We may therefore write the pressure force on the parcel as,

$$\mathbf{F}_p = -\oint_{\partial V_t} p \, \mathrm{d}S \,. \tag{2.10}$$

Applying the divergence theorem, this may be written in a form similar to the gravitational force,

$$\mathbf{F}_p = -\oint_{V_t} \nabla p \, \mathrm{d}V \,. \tag{2.11}$$

The pressure represents the normal stress exerted by the fluid parcel's environment when there is no relative motion between the parcel and its environment. Under conditions where such motion is present, however, other stresses are produced, and these stresses need not be in the direction of the unit normal. The most general representation of such stresses is the viscous stress tensor $\mathbb{P}$. $\mathbb{P}$ is a rank two tensor (cf. vectors, which are rank one tensors). This means it can be described by specifying 9 components, three for each orthogonal direction. The components P_{ij} physically refer to a force in the ith direction acting on a surface oriented in the jth direction. In particular, this means the viscous force on the material parcel $\mathbf{F}_\nu$ may be written,

$$\mathbf{F}_\nu = -\oint_{\partial V_t} \hat{\mathbf{n}} \cdot \mathbb{P} \, \mathrm{d}S, \tag{2.12}$$

where the dot product reduces the rank-two tensor to a vector[2]. With the help of the divergence theorem, this may be written as,

$$\mathbf{F}_\nu = -\oint_{V_t} \nabla \cdot \mathbb{P} \, \mathrm{d}V \,. \tag{2.13}$$

Specifying the form of the viscous stress tensor is a difficult task. One approach is to assume that $\mathbb{P}$ at a given point in space is a linear function of velocity gradients at that point. Such a fluid is known as *Newtonian*. This approximation may be thought of as a local Taylor approximation for the stress tensor, and it is typically a reasonably good approximation for many applications in geophysical fluid dynamics[3]. Here, we will not

[2]For those familiar with index notation, the ith component of the term $\hat{\mathbf{n}} \cdot \mathbb{P}$ can be written $\sum_j n_j P_{ij}$.

[3]Notwithstanding the popular example of *Oobleck*, 1 part water to 1.5 parts corn starch. Google it. You will not regret it.

consider the stress tensor in detail, and generally we will only consider viscous forces in the atmosphere from a qualitative point of view.

Substituting the expressions for the forces into (2.8), we have,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_t}\rho\mathbf{u}\,\mathrm{d}V=\oint_{V_t}-\rho\nabla\Phi_g-\nabla p+\nabla\cdot\mathbb{P}\,\mathrm{d}V\,. \tag{2.14}$$

The remaining task is to move the time derivative into the integral on the left-hand side such that the region of integration becomes arbitrary. However, this is not trivial, as the integration region V_t is now a function of time. A way forward is to transform the integral into a set of coordinates for which the integration region is fixed. What are such coordinates? Precisely the Lagragian coordinate system we developed above. Making the transformation $(x,y,z,t)\rightarrow(x_0,y_0,z_0,t)$, we may write,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_t}\rho\mathbf{u}\,\mathrm{d}V=\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_0}\rho\mathbf{u}J(\mathbf{x},\mathbf{x}_0)\,\mathrm{d}V, \tag{2.15}$$

where V_0 is the volume of the parcel measured in the Lagragrian coordinate, and J is the Jacobian relating volume elements in the (x,y,z) coordinate to those in the (x_0,y_0,z_0) coordinate. Since V_0 is independent of time, the time derivative may be moved inside the integral so that,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_t}\rho\mathbf{u}\,\mathrm{d}V=\oint_{V_0}\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}\rho J(\mathbf{x},\mathbf{x}_0)\,\mathrm{d}V+\oint_{V_0}\mathbf{u}\frac{\mathrm{D}}{\mathrm{D}t}\left[\rho J(\mathbf{x},\mathbf{x}_0)\right]\mathrm{d}V, \tag{2.16}$$

where a time derivative at fixed $\mathbf{x_0}$ is denoted by $\frac{\mathrm{D}}{\mathrm{D}t}$, and we have used the product rule to split the integral into two terms.

The first term on the right-hand side of (2.16) may be simply transformed back into Eulerian coordinates since the Jacobian is unaffected. To evaluate the second term on the right-hand side, we transform the integral equation for conservation of mass of a material parcel (2.4) into Lagrangian coordinates so that,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_0}\rho J(\mathbf{x},\mathbf{x}_0)\,\mathrm{d}V=0. \tag{2.17}$$

Bringing the time derivative inside the integral as above, we have,

$$\oint_{V_0}\frac{\mathrm{D}}{\mathrm{D}t}\left[\rho J(\mathbf{x},\mathbf{x}_0)\right]\mathrm{d}V=0. \tag{2.18}$$

Since the volume V_0 is arbitrary, this implies $\mathrm{D}(\rho J)/\mathrm{D}t=0$, and therefore that the second term on the right-hand side of (2.16) is also zero.

Bringing this together, we have what is known as Reynolds transport theorem,

$$\frac{\mathrm{d}}{\mathrm{d}t}\oint_{V_t}\rho\mathbf{u}\,\mathrm{d}V=\oint_{V_t}\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}\rho\,\mathrm{d}V\,. \tag{2.19}$$

Substituting this into (2.14), we have,

$$\oint_{V_t}\rho\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}\,\mathrm{d}V=\oint_{V_t}-\rho\nabla\Phi_g-\nabla p-\nabla\cdot\mathbb{P}\,\mathrm{d}V\,. \tag{2.20}$$

Since the integration region V_t is arbritrary, this can only be satisfied if the integrand is exactly zero, giving,

$$\frac{\mathrm{D}\mathbf{u}}{\mathrm{D}t}=-\nabla\Phi_g-\frac{1}{\rho}\nabla p+\mathbf{f}_\nu, \tag{2.21}$$

where we have denoted the viscous stress term by the viscous force per unit mass $\mathbf{f}_\nu$ for simplicity. Expressing the above equation in purely in an Eulerian frame, we have the momentum equation for the atmosphere,

$$\frac{\partial\mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}=-\nabla\Phi_g-\frac{1}{\rho}\nabla p+\mathbf{f}_\nu, \tag{2.22}$$

Effects of rotation

In the atmosphere and ocean sciences, velocities are typically measured relative to the Earth's surface. It is therefore helpful to write the equations from this perspective. However, the rotating Earth is not an inertial reference frame, and there are therefore apparent forces that appear in the equations when viewed in this frame. In particular, rotation gives rise to two extra forces, the *centrifugal* force and the *Coriolis* force[4].

We do not derive the form of these additional forces here; the reader is referred to e.g., Houghton (2002), section 7.2 or Vallis (2017), section 2.1 for derivations of the equations of motion in a rotating frame. Here, we simply state that, when viewed from the rotating frame of the Earth, the momentum equation contains two extra terms given by $\mathbf{f}_{\text{cent}}=-\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{x})$ and $\mathbf{f}_{\text{cor}}=-2\boldsymbol{\Omega}\times\mathbf{u}$ relative to the equations in an intertial frame. The equation of motion relative to the rotating planet may therefore be written,

$$\frac{\partial\mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}=-\nabla\Phi_g-\boldsymbol{\Omega}\times(\boldsymbol{\Omega}\times\mathbf{x})-2\boldsymbol{\Omega}\times\mathbf{u}-\frac{1}{\rho}\nabla p+\mathbf{f}_\nu, \tag{2.23}$$

where all quantities are taken with respect to a coordinate system that rotates at angular velocity $\boldsymbol{\Omega}$.

[4]Despite sometimes being called fictitious forces, the Coriolis and centrifugal forces have important influences on the flow as we shall see throughout these notes. Indeed, gravity itself is a fictitious force when viewed from Einstein's theory of general relativity.

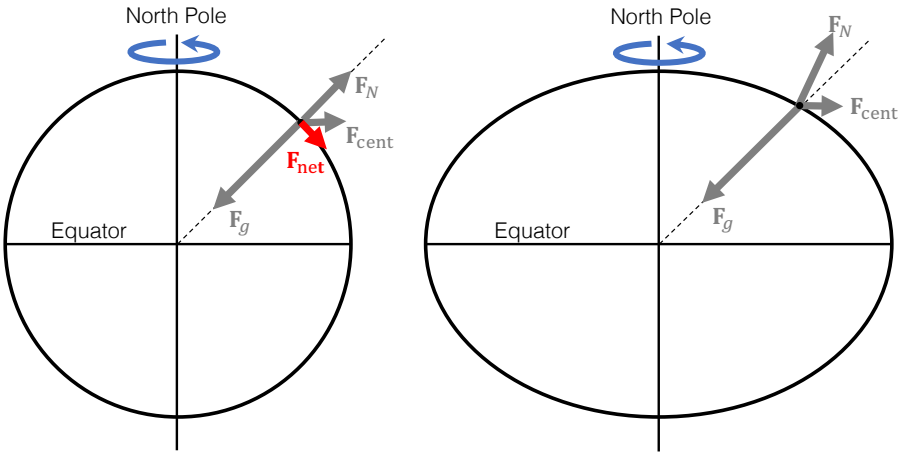

Figure 2.1: Force balance for a stationary object on the surface of the Earth. For a spherical planet (left), the combination of the gravitational force $\mathbf{F}_g$, centrifugal force $\mathbf{F}_{\text{cent}}$, and normal force $\mathbf{F}_N$ must result in a net force. If the planet bulges at the equator (right), the three forces can be balanced, and an object stationary on the surface will remain stationary.

The term $\mathbf{f}_{\text{cent}} = -\boldsymbol{\Omega} \times (\boldsymbol{\Omega} \times \mathbf{x})$ is the centrifugal force per unit mass. It acts on all objects in the rotating frame, whether they are stationary or moving. It is directed away from the Earth's rotation axis, with a magnitude equal to $\Omega^2 r_\perp$, where Ω is the Earth's rotation rate, and $r_\perp$ is the distance from the Earth's axis.

The term $\mathbf{f}_{\text{cor}} = -2\boldsymbol{\Omega} \times \mathbf{u}$ is the Coriolis force per unit mass. Since it depends on $\mathbf{u}$, it acts only on objects that are moving in the rotating frame. Furthermore, the Coriolis force acts at right angles to the motion, so it does not affect the speed of motion, only its direction.

The bulging of the Earth at the equator

Consider an object at rest on a flat surface on Earth. The forces exerted on this object are the gravitational force, which acts towards the centre of the Earth, the normal force, which acts normal to the surface, and the centrifugal force pointing away from the Earth's axis (Fig. 2.1). For a spherical Earth, the gravitational force and normal force would be parallel, but, apart from at the Equator, the centrifugal force would have a component parallel to the surface of the Earth (locally horizontal). Thus, regardless of the magnitudes of the gravitational force and normal force, the net force could not be balanced. Why then do objects on the surface of the Earth not spontaneously accelerate sideways?

The answer is that Earth is not a sphere, but bulges at the equator. This bulge is such that the combined effect of the gravitational force and the centrifugal force are perpendicular to what we think of as a "flat" surface (Fig. 2.1). At long enough timescales, the Earth may be

considered a fluid, and it has therefore attained a mass distribution that is approximately in equilibrium with its rotation. The net result is that what we consider as "down", as measured by a plumb bob for example, does not point to the centre of the Earth, but points in the direction of the sum of the centrifugal and gravitational forces.

What does all this mean for the equations of motion? Since the gravitational and centrifugal forces act together, it makes sense to combine them into a single entity. In particular, we can define a potential for the centrifugal force so that,

$$\mathbf{f}_{\text{cent}} = -\nabla\Phi_{\text{cent}} = -\nabla\left(\frac{-\Omega^2 r_\perp^2}{2}\right). \tag{2.24}$$

Defining a combined potential $\Phi = \Phi_g + \Phi_{\text{cent}}$, we may write the equation of motion as,

$$\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u}\cdot\nabla\mathbf{u} = -\nabla\Phi + 2\boldsymbol{\Omega}\times\mathbf{u} - \frac{1}{\rho}\nabla p + \mathbf{f}_\nu. \tag{2.25}$$

The potential Φ increases as one moves away from the Earth's oblate spheroidal surface, consistent with what we usually think of as the gravitational potential. For this reason we refer to Φ as the geopotential.

Spherical coordinates

So far, we have only considered the equations of motion in vector form. To be useful, however, we must express the equations in a set of coordinates. For the Earth, it is natural to use spherical coordinates, defined by λ the longitude, ϕ the latitude, and r the distance from the centre of the planet, with corresponding unit vectors as shown in Fig. 2.2. However, as described above, the Earth is not a true sphere, so applying spherical coordinates strictly would imply the existence of strong geopotential gradients on surfaces of constant r, making the equations rather awkward to use.

A solution to this problem, and the solution that is almost universally used in atmospheric science, is to make the approximation that the Earth, as well as surfaces of constant Φ, *are* true spheres[5]. Under this approximation, the sum of the gravitational and centrifugal acceleration $-\nabla\Phi$ is parallel to the r coordinate and so only appears in one of the component equations, which considerably simplifies the algebra.

The primitive equations

Rather than derive the coordinate forms of the momentum equation here, we will simply show an approximate form of these component equations known as the *primitive equations*.

[5] An alternative would be to define an oblate spheroidal coordinate system of latitude, longitude, and a third coordinate with level sets that are oblate spheroids. I am not aware of any model, textbook or research paper that does this for Earth, but it is potentially required for very rapidly rotating bodies (e.g., Pulsars).

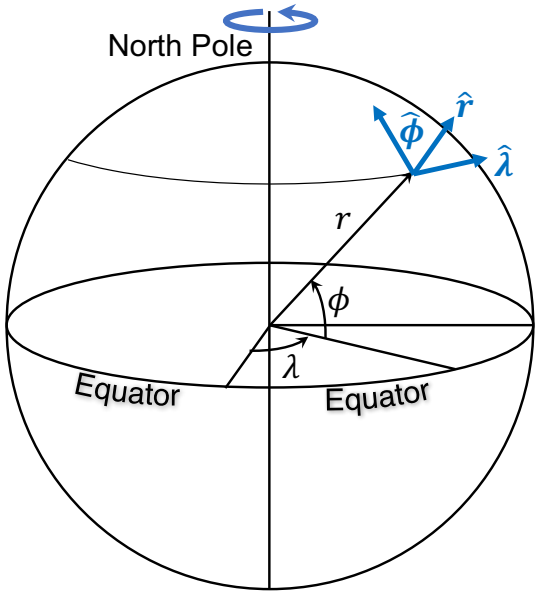

Figure 2.2: Spherical coordinate system commonly used in atmospheric science. The three coordinates are longitude λ, latitude ϕ and distance from the centre of the planet r, with corresponding unit vectors $\hat{\lambda}$, $\hat{\phi}$, and $\hat{\mathbf{r}}$. Adapted from Vallis (2017).

The primitive equations are constructed upon three approximations:

1. *The shallow fluid approximation:* Neglect the increase in surface area of the atmosphere as r increases, and use the distance from mean-sea level z for the radial coordinate (which we refer to as the vertical coordinate). This implies

$$z = r - R_e$$

   where R_e is the radius of the Earth (which is a true sphere under our approximations), and that the geopotential is given by $\Phi = gz$, where g is the (assumed constant) gravitational acceleration.

2. *The traditional approximation:* Neglect the components of the Coriolis force associated with vertical motion and neglect certain metric terms[6] that contain the vertical velocity.

3. Hydrostatic balance: assume the vertical momentum balance is between gravity and the pressure gradient.

[6]Metric terms are those terms in the component equations that arise from the curvilinear nature of the coordinate system. See Vallis (2017), section 2.2.3.

The first two assumptions above must be taken together in order to ensure the resultant equations retain a principle of conservation of angular momentum. Hydrostatic balance is not strictly required, but it is a very good approximation for large-scale flow in Earth's atmopshere, and it often assumed in global climate models.

Given these approximations, the primitive equations for momentum may be written,

$$\frac{\mathrm{D}u}{\mathrm{D}t} = 2\Omega \sin\phi v + \frac{uv}{R_e}\tan\phi - \frac{1}{R_e\rho\cos\phi}\frac{\partial p}{\partial\lambda} + \mathcal{F}_{\nu\lambda}, \tag{2.26a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} = -2\Omega \sin\phi u - \frac{u^2}{R_e}\tan\phi - \frac{1}{R_e\rho}\frac{\partial p}{\partial\phi} + \mathcal{F}_{\nu\phi}, \tag{2.26b}$$

$$\frac{\partial p}{\partial z} = -\rho g. \tag{2.26c}$$

Here u, v and w are the velocities in the zonal, meridional, and vertical directions, respectively, and $\mathcal{F}_{\nu\lambda}$ and $\mathcal{F}_{\nu\phi}$ are the frictional forces per unit mass in the zonal and meridional directions, respectively. The terms involving $\tan\phi$ are metric terms that arise from the curvilinear coordinate system.

Some geometry leads to the following expressions for the velocities,

$$u = R_e\cos\phi\frac{\mathrm{D}\lambda}{\mathrm{D}t}, \tag{2.27}$$

$$v = R_e\frac{\mathrm{D}\phi}{\mathrm{D}t}, \tag{2.28}$$

$$w = \frac{\mathrm{D}z}{\mathrm{D}t}. \tag{2.29}$$

While the Lagrangian derivative may be expressed,

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + \frac{u}{R_e\cos\phi}\frac{\partial}{\partial\lambda} + \frac{v}{R_e}\frac{\partial}{\partial\phi} + w\frac{\partial}{\partial z}. \tag{2.30}$$

Note also that, under the shallow fluid approximation, the divergence in spherical coordinates may be written,

$$\nabla\cdot\mathbf{u} = \frac{1}{R_e\cos\phi}\left[\frac{\partial u}{\partial\lambda} + \frac{\partial v\cos\phi}{\partial\phi}\right] + \frac{\partial w}{\partial z}. \tag{2.31}$$

2.1.4 Equation of state

The equation of state represents the relationship between the state variables of the fluid under consideration. For the atmosphere, the relevant equation of state is the ideal gas law, which in its most general form may be written,

$$p = \frac{n}{V}R^*T. \tag{2.32}$$

Here p is the pressure, T is the temperature, n is the number of moles of gas, V is the volume of the gas, and $R^* = 8.314$ J mol^{-1} K^{-1} is the universal gas constant. The ideal gas law may be derived from the kinetic theory of gases by assuming that interactions between individual gas molecules are weak (which is a good approximation at atmospheric pressures and temperatures).

For a gas composed of a single constituent chemical, we have that the mass of gas m is related to the number of moles by $n = m/M$, where M is the molar mass of the molecules making up the gas. We may therefore write,

$$p = \rho \frac{R^*}{M} T, \tag{2.33}$$

where we have noted that the density $\rho = m/V$. For a mixture of gases, we may use Dalton's law of partial pressures to derive an equation of state for the mixture,

$$p = \sum_i p_i = \rho R_{\text{mix}} T, \tag{2.34}$$

where

$$R_{\text{mix}} = \sum_i q_i \frac{R^*}{M_i}. \tag{2.35}$$

Here p_i and M_i are, respectively, the partial pressure and molecular mass of consituent i within the mixture, and q_i is the mass of consituent i per unit mass of the mixture.

In the atmosphere, the water vapour concentration is highly variable, and it therefore makes sense to consider the water vapour separately. In particular, we may write,

$$p_d = \rho_d R_d T \tag{2.36}$$

$$e = \rho_v R_v T, \tag{2.37}$$

where p_d and ρ_d are the pressure and dry air density, respectively, and e and ρ_v are the vapour pressure and vapour density, respectively. The gas constants are calculated using (2.35), and turn out to be $R_d = 287$ J kg^{-1} K^{-1} and $R_v = 461.5$ J kg^{-1} K^{-1}. Combining the above equations, we may write an equation of state for moist air given by,

$$p = \rho R_m T \tag{2.38}$$

where $R_m = (1-q)R_d + qR_v$ and $q = \frac{\rho_v}{\rho}$ is the specific humidity. Since $q << 1$ in Earth's atmosphere, an approximation that is suitable for large-scale flow is to simply neglect the effect of water vapour on density so that we have,

$$p \approx \rho R_d T. \tag{2.39}$$

This approximation is sufficient for our purposes, but can become dubious in situations where small-scale horizontal gradients of density are important (e.g., estimating the density of clouds relative to their environment).

2.1.5 The thermodynamic equation

The thermodynamic equation represents conservation of energy associated with the micro scale motions of air molecules[7]. The thermodynamic equation for a parcel of air of unit mass may be approximately written,

$$\frac{\mathrm{D}i}{\mathrm{D}t} = Q - p\frac{\mathrm{D}\alpha}{\mathrm{D}t}, \tag{2.40}$$

where Q is the diabatic heating rate, i is the internal energy of the air, and $\alpha = 1/\rho$ is the specific volume. This equation expresses the physical principle that the internal energy of a fluid parcel may be changed by external heat sources or by work done by the environment on the parcel.

For an ideal gas, gas particles are assumed to interact only weakly with one another, and the internal energy is therefore simply a mean of the kinetic energy of each individual molecule. Since the temperature is a measure of the mean kinetic energy of molecules, this means the internal energy $i = i(T)$ is a function of temperature only. Defining the specific heat capacity at constant volume $c_v = \frac{\mathrm{d}i}{\mathrm{d}T}$, we have,

$$Q = c_v\frac{\mathrm{D}T}{\mathrm{D}t} + p\frac{\mathrm{D}\alpha}{\mathrm{D}t}. \tag{2.41}$$

Note that the specific heat capacity c_v may not be constant with temperature, but empirically it is found that its variations with temperature are relatively weak, and it may be approximated as such.

It is useful to use the ideal gas law to transform the thermodynamic equation into a more convenient form. Taking the Lagrangian derivative of the ideal gas law $p\alpha = R_dT$ with respect to time, we have,

$$\alpha\frac{\mathrm{D}p}{\mathrm{D}t} + p\frac{\mathrm{D}\alpha}{\mathrm{D}t} = R_d\frac{\mathrm{D}T}{\mathrm{D}t}. \tag{2.42}$$

Substituting this into the thermodynamic equation (2.41), we have

$$c_p\frac{\mathrm{D}T}{\mathrm{D}t} = Q - \alpha\frac{\mathrm{D}p}{\mathrm{D}t}, \tag{2.43}$$

where we have defined $c_p = c_v + R_d$ to be the the isobaric specific heat capacity. This form of the thermodynamic equation is useful because it contains the Lagrangian derivative of pressure, which will be identified as the vertical velocity in pressure coordinates below. Note, however, that the left-hand side of the thermodynamic equation in the form (2.43)

[7] cf. the energy equation for macro-scale motion that may be derived by dotting the momentum equation with the velocity. These two energy equations interact through viscous dissipation, which converts macro-scale energy to micro-scale energy. We will not discuss the details of this *energy cascade* in these notes.

does not correspond to the internal energy, and the second term on the right-hand side does not correspond to the work done.

Finally, we simplify the thermodynamic equation further by defining the potential temperature θ so that

$$\theta = \frac{T}{\pi(p)} \tag{2.44}$$

where $\pi(p) = \left(\frac{p}{p_0}\right)^{\frac{R_d}{c_p}}$ is the Exner function, and p_0 is a reference pressure, generally taken to be 1000 hPa. Taking the Lagrangian derivative of the above equation, we have that,

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{1}{\pi(p)}\frac{\mathrm{D}T}{\mathrm{D}t} - \frac{R_d T}{c_p p \pi(p)}\frac{\mathrm{D}p}{\mathrm{D}t} \tag{2.45}$$

Substituing this into the thermodynamic equation, using the ideal gas law and rearranging, we have that,

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{Q}{c_p \pi(p)}. \tag{2.46}$$

2.1.6 Useful approximations and transformations

Equations (2.7), (2.26), (2.39) and (2.43), along with equations for additional tracers of interest (such as water substance) and specifications of appropriate boundary conditions and diabatic heating Q provide a complete set of equations for determining the evolution of the atmospheric flow. However, often for analysis a simpler set is more useful fo developing understanding. In this section we outline some approximations and transformations of the equations that will be used in these notes.

Pressure coordinates

A particularly useful form of the equations that we will use throughout these notes are those expressed with pressure as the vertical coordinate. In particular, we make the coordinate transformation $(\lambda, \phi, z, t) \rightarrow (\lambda', \phi', p, t')$, where $\lambda = \lambda'$, $\phi = \phi'$, $t = t'$, and p is the pressure field. Note that this change of coordinates does not involve a change in the direction of the basis vectors; for example, the definitions of u and v do not change. Rather, we take each component equation and recast them using the new coordinates.

Using a similar procedure to that in section 2.1.2, we may derive an expression for the Lagrangian derivative,

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t'} + \frac{u}{R_e \cos\phi}\frac{\partial}{\partial \lambda'} + \frac{v}{R_e}\frac{\partial}{\partial \phi'} + \omega\frac{\partial}{\partial p}. \tag{2.47}$$

where $\omega = \frac{\mathrm{D}p}{\mathrm{D}t}$. This expression may be compared to (2.30).

We also must transform derivatives at constant height into those at constant pressure. For example, consider a variable γ. Applying the chain rule, we have,

$$\frac{\partial \gamma}{\partial \lambda'} = \frac{\partial \gamma}{\partial \lambda}\frac{\partial \lambda}{\partial \lambda'} + \frac{\partial \gamma}{\partial z}\frac{\partial z}{\partial \lambda'} \tag{2.48}$$

where partial derivatives with respect to primed coordinates hold all other primed coordinates constant, and we have used the fact that the partial derivative of ϕ and t with respect to λ' is zero. Applying the above equation to the pressure itself, we have,

$$\frac{\partial p}{\partial \lambda'} = 0 = \frac{\partial p}{\partial \lambda} + \frac{\partial p}{\partial z}\frac{\partial z}{\partial \lambda'}. \tag{2.49}$$

Applying hydrostatic balance, we have that,

$$\frac{\partial p}{\partial \lambda} = \rho \frac{\partial \Phi}{\partial \lambda'}. \tag{2.50}$$

where $\Phi = gz$ is the geopotential. Substituting (2.50) and its equivalent for latitudinal derivatives into (2.26) and rearranging the hydrostatic relation, we may write the momentum equations in pressure-coordinates as,

$$\frac{\mathrm{D}u}{\mathrm{D}t} = 2\Omega \sin\phi v + \frac{uv}{R_e}\tan\phi - \frac{1}{R_e \cos\phi}\frac{\partial \Phi}{\partial \lambda} + \mathcal{F}_{\nu\lambda}, \tag{2.51a}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} = -2\Omega \sin\phi u - \frac{u^2}{R_e}\tan\phi - \frac{1}{R_e}\frac{\partial \Phi}{\partial \phi} + \mathcal{F}_{\nu\phi}, \tag{2.51b}$$

$$\frac{\partial \Phi}{\partial p} = -\alpha, \tag{2.51c}$$

where we have dropped the primes for clarity.

Finally, we must transform the mass continuity equation into pressure coordinates. This may be done directly using the mass continuity equation in height coordinates (2.7), but it is more convenient, although somewhat more heuristic, to use physical reasoning directly.

Consider the mass of a Lagrangian parcel, which we may write in spherical coordinates,

$$M = \oint_{V_t} \rho \,\mathrm{d}V = \iiint \rho R_e^2 \cos\phi \,\mathrm{d}\lambda \,\mathrm{d}\phi \,\mathrm{d}z \tag{2.52}$$

where the $R_e^2 \cos\phi$ factor is a result of the volume elements in spherical coordinates[8]. We may transform the integral using hydrostatic balance, so that,

$$M = \iiint \frac{1}{g} R_e^2 \cos\phi \,\mathrm{d}\lambda \,\mathrm{d}\phi \,\mathrm{d}p \tag{2.53}$$

[8]In full spherical coordinates, this factor would be $r^2 \cos\phi$, but under the shallow fluid approximation, we neglect the increase in the surface area of the atmosphere as one moves away from the surface.

The above equation implies that the "density" in pressure coordinates is constant and equal to the reciprocal of the gravitational acceleration. When viewed in pressure coordinates (and under hydrostatic balance), the atmosphere therefore behaves as an incompressible fluid. The relevant continuity equation for such a fluid is that the divergence of the velocity is zero. This implies that the relevant continuity equation in pressure coordinates is,

$$\nabla_p \cdot \mathbf{u_h} + \frac{\partial \omega}{\partial p} = 0. \tag{2.54}$$

Where ∇_p is the horizontal gradient operator at constant pressure, and $\mathbf{u_h}$ is the horizontal velocity vector.

Tangent-plane approximation

For many applications, it is sufficient to adopt a local cartesian coordinate system, rather than retain the full complexity of spherical geometry. In particular, if we define the horizontal coordinates,

$$x = R_e \cos\phi_0 \lambda, \tag{2.55}$$
$$y = R_e(\phi - \phi_0), \tag{2.56}$$

for some reference latitude ϕ_0, we recover the familiar cartesian equations (in pressure coordinates),

$$\frac{\mathrm{D}u}{\mathrm{D}t} = fv - \frac{\partial \Phi}{\partial x} + \mathcal{F}_{\nu x}, \tag{2.57}$$
$$\frac{\mathrm{D}v}{\mathrm{D}t} = -fu - \frac{\partial \Phi}{\partial y} + \mathcal{F}_{\nu y}, \tag{2.58}$$

where $f = f_0 = 2\Omega \sin\phi_0$ (f-plane) or $f = f_0 + \beta y$ (β-plane). Here (u, v) are the velocities in the (x, y) directions, and the Lagrangian derivative may be written,

$$\frac{\mathrm{D}}{\mathrm{D}t} = \frac{\partial}{\partial t} + u\frac{\partial}{\partial x} + v\frac{\partial}{\partial y} + \omega\frac{\partial}{\partial p}. \tag{2.59}$$

This approximation is reasonable when the region of interest is a limited range of latitudes, but it can break down for problems that involve the circulation on planetary scales.

Definition of the streamfunction

We close this section with a short discussion of the streamfunction and its definition. We define the streamfunction Ψ such that the difference between the value of the streamfunction at two height levels is equal to the total meridonal mass flux between those levels. In particular,

$$\Psi(\phi, z) = R_e \cos\phi \int_0^{2\pi} \int_0^z \rho v \,\mathrm{d}z \,\mathrm{d}\lambda, \tag{2.60}$$

where we have taken the streamfunction to be zero at the surface ($z = 0$). Since under the hydrostatic approximation, $\mathrm{d}p = -\rho g \, \mathrm{d}z$, we have,

$$\Psi(\phi, p) = \frac{1}{g} R_e \cos\phi \int_0^{2\pi} \int_p^{p_s} v \, \mathrm{d}p \, \mathrm{d}\lambda, \tag{2.61}$$

where p_s is the surface pressure. Differentiating in pressure, we have,

$$\frac{\partial \Psi}{\partial p} = -\frac{1}{g} R_e \cos\phi \int_0^{2\pi} v \, \mathrm{d}\lambda \,. \tag{2.62}$$

Rearranging,

$$[v] = -\frac{g}{2\pi R_e \cos\phi} \frac{\partial \Psi}{\partial p}, \tag{2.63}$$

where $[v]$ represents the zonal-mean meridional velocity (see the next section).

2.2 Decomposing the circulation

A large part of these notes are devoted to understanding the zonal- and time-mean atmospheric circulation. In this section, we develop a few basic techniques for analysing the time-and zonal-mean circulation and the deviations from this mean (the eddies), and how these two components of the circulation interact. The notation and development of this section follows that of Peixoto and Oort (1992).

2.2.1 Spatial and temporal averages

We define the time mean of a given variable $a = a(\lambda, \phi, p, t)$ as,

$$\overline{a} = \frac{1}{\tau} \int_0^{\tau} a \, \mathrm{d}t, \tag{2.64}$$

where τ is a suitable temporal averaging period (e.g., 20 years for a climatological mean). Furthermore, we denote the deviations of a from the time mean by a prime, so that,

$$a' = a - \overline{a}. \tag{2.65}$$

Note that this definition guarantees that $\overline{a'} = 0$.

In a similar fashion, we define the zonal average,

$$[a] = \frac{1}{2\pi} \int_0^{2\pi} a \, \mathrm{d}\lambda, \tag{2.66}$$

and deviations

$$a^* = [a] - a, \tag{2.67}$$

which similarly guarantees that $[a^*] = 0$.

The averaging operators have a few useful properties worth noting. Firstly, it is easily shown that the averaging operators are linear, so that $\overline{a+b} = \overline{a} + \overline{b}$, and that $\overline{ka} = k\overline{a}$ for any scalar k. Secondly, the zonal and time averaging operators commute with each other (since integration is commutative for sufficiently well behaved functions). Finally, the zonal mean of a zonal derivative vanishes:

$$\begin{aligned}\left[\frac{\partial a}{\partial \lambda}\right] &= \frac{1}{2\pi}\int_0^{2\pi} \frac{\partial a}{\partial \lambda}\,\mathrm{d}\lambda \\ &= \frac{1}{2\pi}\, a|_{\lambda=0}^{\lambda=2\pi} \\ &= 0,\end{aligned}$$

since the longitude 0 and 2π correspond to the same location. Furthermore, if we assume that temporal trends are weak and the averaging period τ is long enough,

$$\overline{\frac{\partial a}{\partial t}} = 0.$$

The above properties are true provided the pressure or height level under consideration does not intersect the surface. If, however, a given pressure surface *outcrops*, the integral must be split into a series of integrals over all regions for which the level under question is defined, and, for example, the zonal mean of a zonal derivative may be non zero. We shall see this behaviour in our discussion of *form drag* in section 5.3.

The utility of the averaging operators arises when we consider the average of a product of more than one field. For example, consider the produce of two fields a and b. Taking the time mean, we have,

$$\begin{aligned}\overline{ab} &= \overline{(\overline{a}+a')(\overline{b}+b')} \\ &= \overline{a}\,\overline{b} + \overline{\overline{a}b'} + \overline{a'\overline{b}} + \overline{a'b'} \\ &= \overline{a}\,\overline{b} + \overline{a'b'}.\end{aligned} \tag{2.68}$$

The time-mean value of the quantity ab does not just have contributions from the time-mean value of a and the time-mean value of b, but there is an additional contribution from the *covariance* between a and b. Similarly we may write for the zonal mean,

$$[ab] = [a][b] + [a^*b^*]. \tag{2.69}$$

Why is this important? Let's consider a more physical example and study the evolution of the specific humidity q. Absent precipitation or evaporation, the specific humidity is a

conserved quantity. For simplicity, we neglect evaporation and precipitation here, and we treat the specific humidity as a passive tracer with a governing equation of the form,

$$\frac{\mathrm{D}q}{\mathrm{D}t} = 0. \tag{2.70}$$

We may use continuity to write the above equation in a form more convenient for the present analysis called *flux form*. Expanding the Lagrangian derivative in pressure coordinates and using the tangent plane approximation for simplicity, we have,

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \frac{\partial q}{\partial t} + u\frac{\partial q}{\partial x} + v\frac{\partial q}{\partial y} + \omega\frac{\partial q}{\partial p} \tag{2.71}$$

By the product rule, we have that

$$\frac{\partial uq}{\partial x} = u\frac{\partial q}{\partial x} + q\frac{\partial u}{\partial x} \tag{2.72}$$

Applying this to (2.71) above, we have

$$\frac{\mathrm{D}q}{\mathrm{D}t} = \frac{\partial q}{\partial t} + \frac{\partial uq}{\partial x} + \frac{\partial vq}{\partial y} + \frac{\partial \omega q}{\partial p} - q\left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial \omega}{\partial p}\right) \tag{2.73}$$

The last term in the above equation is zero by continuity, and we may therefore write the flux form of the equation for q as

$$\frac{\partial q}{\partial t} + \nabla_p \cdot (q\mathbf{u_h}) + \frac{\partial q\omega}{\partial p} = 0. \tag{2.74}$$

Any equation with a Lagrangian derivative can be transformed into flux form, either in pressure coordinates or in height coordinates (in which case the resultant flux-form equation includes the density). We will often switch between the flux form and Lagrangian form of equations rapidly, and so the reader is advised to become comfortable with this transformation.

Now, let us use the flux form (2.75) to derive an equation for the evolution of the *zonal-mean* specific humidity $[q]$. Taking the zonal average, we have,

$$\frac{\partial [q]}{\partial t} + \frac{\partial [qv]}{\partial y} + \frac{\partial [q\omega]}{\partial p} = 0. \tag{2.75}$$

Using (2.69) and rearranging, we have,

$$\frac{\partial [q]}{\partial t} + \frac{\partial [q][v]}{\partial y} + \frac{\partial [q][\omega]}{\partial p} = -\frac{\partial [q^*v^*]}{\partial y} - \frac{\partial [q^*\omega^*]}{\partial p}. \tag{2.76}$$

Finally, expanding the derivatives on the left-hand side using the product rule and using the zonal-mean continuity equation,

$$\frac{\partial[v]}{\partial y} + \frac{\partial[\omega]}{\partial y} = 0,$$

we may write the equation for the zonal-mean specific humidity in Lagrangian form,

$$\frac{\mathrm{D}[q]}{\mathrm{D}t} = -\frac{\partial[q^* v^*]}{\partial y} - \frac{\partial[q^* \omega^*]}{\partial p}. \tag{2.77}$$

Equation 2.77 is rather remarkable, as it states that the zonal-mean specific humidity $[q]$ is governed by the same equation as q itself but with additional forcing terms on the right-hand side that depend on the deviations q^*. More specifically, these additional terms represent the meridional and vertical eddy flux of q produced by correlations between the specific humidity and the winds. These fluxes do not directly depend on the zonal-mean flow; even if the zonal-mean flow $[v]$ is zero, there may still be a non-zero meridional flux of tracer through the term $[v^* q^*]$. This fact will be of central importance to our discussion in later chapters.

2.2.2 Combining the spatial and temporal averaging operators

Often when studying the general circulation, we will be interested in the time- and zonal-mean circulation. To deal with this, we need to combine the zonal and temporal averaging operators. Consider again the product ab. Taking the time mean as in (2.68), and then applying a zonal mean, we have,

$$[\overline{ab}] = [\overline{a}\,\overline{b}] + [\overline{a'b'}]. \tag{2.78}$$

Applying the zonal decomposition to the first term, we have

$$[\overline{ab}] = [\overline{a}][\overline{b}] + [\overline{a}^* \overline{b}^*] + [\overline{a'b'}]. \tag{2.79}$$

This three-term decomposition will turn out to be particularly useful for analysing the general circulation.

Consider the decomposition applied to the meridional flux of specific humidity,

$$[\overline{vq}] = [\overline{v}][\overline{q}] + [\overline{v}^* \overline{q}^*] + [\overline{v'q'}]. \tag{2.80}$$

This flux represents the mean transport of humidity across a given latitude circle at a given level in the atmosphere. On the right-hand side, the first term represents the product of the zonal- and time-mean meridional velocity with the mean humidity; this is the flux due to the zonal-and time-mean circulation, which we will refer to as the mean flux. The

second term represents the flux owing to motions that are steady in time, but correspond to deviations from the zonal mean; this is the flux owing to stationary eddies. The third term represents the flux owing to motions that vary in time, and it is therefore the flux owing to transient eddies. This decomposition places a precise mathematical definition on the concepts of 'stationary eddies' and 'transient eddies' as used in, for example, Fig. 1.10.

2.3 State estimation techniques

The analysis tools developed in the previous section assume we have enough information to produce well-defined estimates of zonal averages $[a]$ and time averages $\overline{a}$ and their deviations. But observations of the atmosphere are usually scattered in space or time, and it is not obvious that producing such estimates is possible. In this chapter, we briefly introduce two methods for transforming direct observations of the atmosphere into a gridded estimate of the time- and zonal-mean fields and their deviations.

2.3.1 The Peixoto-Oort method

Peixoto and Oort (1992) present estimates of the general circulation of the atmosphere, include time- and zonal-mean fields and the mean and eddy fluxes of various quantities based solely on radiosonde observations. Radiosondes are typically launched twice a day at locations around the world by the Meteorologial agencies in various countries. Radiosonde data is therefore relatively regular in time, but, because of the irregular spacing of radiosonde launching sites, it is highly irregular in space.

The procedure of Peixoto and Oort (1992) proceeds in three steps.

Step 1: Prepare the station data

The raw radiosonde data is first quality controlled by removing obvious outliers. These outliers are found by examining the distribution of a given variable observed by a given radiosonde site over time. The quality controlled data is then used to calculate (monthly) time means (e.g., $\overline{v}$, $\overline{T}$, $\overline{vT}$, etc.). From these means, the time-averaged covariances (e.g., $\overline{T'^2}$, $\overline{v'T'}$) may also be calculated, using, for example,

$$\overline{v'T'} = \overline{vT} - \overline{v}\,\overline{T}. \tag{2.81}$$

Note that time-mean quantities are relatively easy to calculate because of the high time resolution and relative regularity of the input radiosonde obsverations.

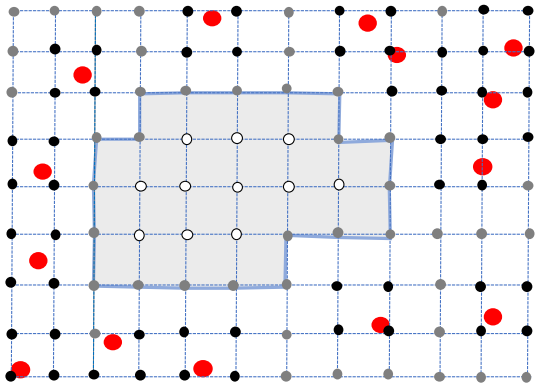

Figure 2.3: Example of a spatial grid used in the Peixoto and Oort (1992) analysis. Radiosonde stations are shown in red, class 1 gridpoints in black, class 2 gridpoints in grey, and class 3 gridpoints in white. The analysis at class 3 points is found by solving a Poisson equation (2.82) over the shaded region with boundary conditions taken from the surrounding class 2 points.

Step 2: Define the grid and initial guess

In the second step, a regular grid on which the atmospheric state is to be estimated is defined. Peixoto and Oort (1992) use a $2.5° \times 5°$ grid. This is considerably lower resolution than many modern reanalyses, but it is sufficient from the point of view of the general circulation.

An initial guess of the atmospheric state is then defined on this grid. Peixoto and Oort (1992) use the climatological and zonal mean within 10° latitude bins as the initial guess. This implies that in the initial guess $[\overline{v^*T^*}] = 0$, and similar for all stationary eddy components.

Step 3: Update the initial guess using the observations

The initial guess is updated at each gridpoint using a set of rules based on the proximity of the gridpoint to a radiosonde station. Specifically, three classes of gridpoints are defined. Class 1 gridpoints are directly adjacent to a radiosonde site; class 2 gridpoints are adjacent to a class 1 gridpoint; and all other gridpoint are denoted class 3 (Fig. 2.3). The rules for each class are then,

Class 1: Use linear interpolation to calculate the required correction to the initial guess.

Class 2: Apply the same correction to the initial guess as for the nearest class 1 gridbox.

Class 3: For gridboxes far away from available radiosonde sites, solve the following boundary value problem:

$$\nabla^2 T = F(\lambda, \phi), \tag{2.82}$$

where the forcing function F is chosen to be the Laplacian of the initial guess, and the boundary conditions correspond to the nearest gridboxes that have had a correction applied.

Step 3 of the above procedure is repeated until a stable analysis is found. Once the time-mean quantities have been estimated on the grid, zonal averages may be taken, and the stationary eddy component may be estimate using, for example,

$$[\overline{v^*T^*}] = [\overline{vT}] - [\overline{v}][\overline{T}] - [\overline{v'T'}]. \tag{2.83}$$

It is useful to consider why the method above is superior to a simple interpolation procedure. Consider an analysis of surface temperature which includes two adjacent gridpoints, one at sea level and one in a coastal mountain range. The interpolation of surface temperature across such a large difference in elevation is likely to be rather inaccurate. However, if one knows something of the climatology of temperature at both locations, the anomalies from this climatology are likely to be much better behaved when interpolated. The Peixoto-Oort analysis applies the same reasoning over latitude rather than elevation. The challenge is to optimally combine the information provided by the climatology at a given location with the information provided by the weather at nearby locations.

2.3.2 Modern analysis

While the Peixoto-Oort analysis has the advantage of simplicity; it uses only radiosonde observations and is relatively transparent in its treatment of the input data, modern analysis methods employ far more sophisticated techniques with a much wider range of data. Such methods came about from the need to provide an initial state for numerical weather prediction (operational analysis). However, there are also a number of "reanalysis" products available from different groups around the world which apply consistent methods to estimate the state of the atmosphere at regular intervals over a number of decades which are primarily used for research processes.

Data assimilation

The primary method of creating operational analyses and reanalyses is called data assimilation. As in the Peixoto-Oort method, an initial guess of the atmospheric state is required. In this case, the initial guess comes from the 6-hourly forecast of a numerical weather prediction model. The initial guess is then corrected by minimising a cost function that depends on the difference between the analysis and various observed fields, as well as the size of the corrections that are required.

Data assimilation has the advantages that many different types of observations are able to be used, and that the resultant analysis is based on a physical model of the system (which respects the laws of physics). Nevertheless, such analyses typically do not satisfy budget

constraints (e.g., water, momentum, mass). Furthermore, while, in the case of reanalyses, the methodology (including the numerical model) used is fixed in time, the density and type of observations vary. This can make identifying climatic trends in reanalysis rather challenging. Finally, different reanalysis products are known to disagree on aspects of the atmospheric circulation. This is not likely to be important for our purposes in this unit, but it is important to keep in mind if you use reanalysis in your research.

Chapter 3

Radiative-convective equilibrium & Hide's theorem

3.1 Why is there an atmospheric circulation?

In this chapter we consider the most fundamental question possible about the atmospheric circulation. Why does it exist? Sometimes it is said that the atmospheric circulation exists because there is a net energy input into the equatorial latitudes and a net energy deficit at polar latitudes. But this is incorrect. The existence of energy imbalances in the tropics and polar latitude implies the existence of a circulation, but it cannot be used as a causal mechanism any more than the existence of an atmospheric circulation can be thought of as causing energy imbalances at different latitudes.

3.1.1 Radiative equilibrium

To determine the causes of the atmospheric circulation, it is helpful to imagine how the atmosphere might look if such a circulation did not exist. In particular, if we consider the case where all atmospheric motion ceases, the thermodynamic equation reduces to,

$$c_p \frac{\partial T}{\partial t} = Q_{\text{rad}}, \tag{3.1}$$

where Q_{rad} is the diabatic heating, which in the case of no motion can only occur via radiation[1]. If we take a sufficiently long time average, the temperature tendency on the left-hand side of (3.1) vanishes, and, expressing the radiative heating rate in terms of the

[1] In principle there would also be conduction from the surface to the atmosphere, and throughout the atmosphere, but this is only efficient very close to the surface

divergence of the radiative fluxes, we have,

$$\frac{\partial}{\partial z}\left(F_{\uparrow} - F_{\downarrow}\right) = 0, \tag{3.2}$$

where $F_{\uparrow}$ and $F_{\downarrow}$ are the upward and downward radiative fluxes (both shortwave and longwave) through the atmosphere, respectively. This is the equation for *radiative equilibrium*. Note that (3.2) depends only on the vertical coordinate, and so radiative equilibrium can be solved for independently in each column of the atmosphere.

With a suitable radiative transfer code, one may solve (3.2) for the radiative-equilibrium temperature profile by providing the solar radiation flux at the top of the atmosphere and the distribution of radiatively active gases in the atmosphere, including water vapour and clouds[2]. One of the first such calculations was performed by Manabe and Strickler (1964) and is shown in the solid curve on Fig. 3.1.

By definition, in the radiative equilibrium state there is no net energy input or output at the top of the atmosphere. If each column of the atmosphere was in radiative equilibrium, the tropics and polar latitudes would each have a balanced heat budget locally, and there would be no circulation. This clearly demonstrates that energy imbalances in the tropics and polar latitudes are not the ultimate cause of the atmospheric circulation.

Why is Earth's atmosphere out of radiative equilibrium? One reason can be seen by examining Fig. 3.1. The radiative-equilibrium temperature profile is broadly similar to the observed temperature profile, decreasing in the troposphere and increasing in the stratosphere above (e.g., Fig. 1.5), but there are some important differences. In particular, the radiative equilibrium is warmer at the surface and cooler at the tropopause than Earth's observed temperature profiles. As a result, the radiative-equilibrium lapse rate in the troposphere is higher than in observations, particularly near the surface. In fact, one can show that there is a substantial discontinuity in the radiative-equilibrium temperature profile in which the surface is a few kelvin warmer than the atmosphere immediately above.

It may be shown that the radiative-equilibrium temperature shown in Fig. 3.1 is in fact gravitationally/convectively unstable. If an air parcel in contact with the surface is perturbed upward, it experiences positive buoyancy and continues to rise. An initial state in radiative equilibrium would therefore lead to convection, and the atmosphere would no longer be stationary.

3.1.2 Radiative-convective equilibrium

We argued above that the atmosphere cannot remain in radiative equilibrium because this state is gravitationally unstable to perturbations. But does this directly imply there should be a large-scale circulation akin to the Hadley Cell? Gravitational instability refers to an

[2]Strictly speaking, if there were no atmospheric circulation there would be no clouds

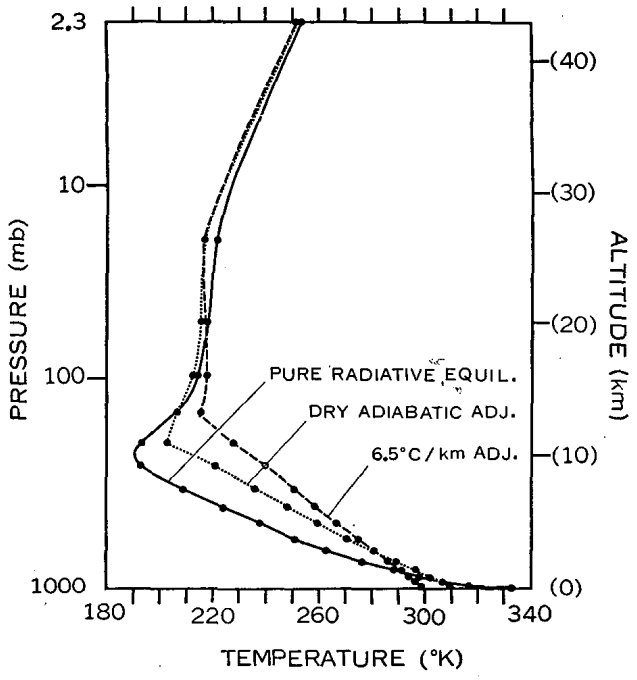

Figure 3.1: Thermal equilibrium profiles calculated using hemispherically averaged insolation and assuming a vertical distribution of gaseous absorbers typical of 35°N in April with no clouds. Solid curve gives radiative equilibrium and dashed and dotted curves give equilibrium calculated with a convective adjustment to the dry adiabatic lapse rate and a lapse rate of 6.5 K km^{-1}, respectively. From Manabe and Strickler (1964).

instability to vertical perturbations, and it may be removed by vertical rearrangements of air parcels. We may therefore imagine a circumstance in which convection acts to locally remove gravitational instability within each column while producing no "large-scale" circulation. Such a state is known as radiative-convective equilibrium (RCE).

What does the RCE state look like? In RCE, there is a balance between the net radiative cooling of the troposphere and net convective heating (both sensible and through latent heat release in clouds) within each column. Since convective clouds develop over a timescale of a few minutes, whereas the radiative relaxation timescale of the atmosphere is on the order of weeks, we may think of convection as a "fast" process relative to radiation. We therefore expect convection to rapidly remove instability produced by "slow" radiative processes, implying that the the RCE state is one that is close to *convective neutrality.*

What is meant by convective neutrality? In an atmosphere with no latent heat release, convective neutrality is a state in which the temperature follows a dry adiabat, and potential

temperature is invariant with height. That is, we have that,

$$\frac{\partial \theta}{\partial z} = 0, \tag{3.3}$$

where $\theta = T\left(\frac{p_0}{p}\right)^{R_d/c_p}$ is the potential temperature, T is the temperature, p is the pressure and $p_0 = 1000$ hPa is a reference pressure. For reasons that will become clear in later sections, it is useful to express convective neutrality conditions in terms of the *entropy* rather than potential temperature. For a dry atmosphere, the entropy is defined $s_d = c_p \ln \theta$, and convective neutrality is described by,

$$\frac{\partial s_d}{\partial z} = 0. \tag{3.4}$$

While the entropy has some interesting properties relating to irreversible processes (see Pauluis and Held, 2002), for our purposes, it is sufficient to treat it as a state variable related to potential temperature.

In a convectively-neutral dry atmosphere, a parcel of air lifted adiabatically will remain neutrally buoyant with respect to its environment and will feel no net force. Manabe and Strickler (1964) found solutions for a dry RCE state using a simple "convective adjustment" scheme. Under this approximation, convection is assumed to act wherever the atmosphere becomes unstable to enforce a dry adiabatic lapse rate (Fig. 3.1). The dry-adiabatic RCE solution avoids the problem of a surface discontinuity in temperature, but it still has a substantially higher lapse rate than is observed in Earth's tropics.

In a moist atmosphere, we expect the lapse rate in RCE to be smaller than that of a dry adiabat because of the latent heat release within clouds. Manabe and Strickler (1964) used convective adjustment to a lapse rate of 6.5 K km^{-1} to provide an approximate solution to moist RCE (Fig. 3.1). More generally, we might think of an atmosphere that is neutral to moist convection as one in which a parcel of air lifted from the boundary layer remains neutrally buoyant as it rises through its lifted condensation level and into the upper troposphere. This corresponds to a temperature that follows a dry adiabat up to the LCL and a moist adiabat thereafter[3]. Mathematically, we express this state as one in which

1. the saturation equivalent potential temperature θ_e^* is constant vertically within the free troposphere;
2. the free-tropospheric saturation equivalent potential temperature θ_e^* is equal to the actual equivalent potential temperature θ_e within the boundary layer.

[3] Singh and O'Gorman (2013) showed that, if entrainment is considered, the correct equilibrium profile is actually slightly more unstable than a moist adiabat, but we will neglect this difference here.

Equivalently, we may express the moist neutrality condition above using the moist entropy s

$$s = c_p \ln \theta_e \approx c_p \ln(\theta) + \frac{L_v q}{T_{\mathrm{LCL}}}, \tag{3.5}$$

and the saturation moist entropy s^*

$$s^* = c_p \ln \theta_e^* \approx c_p \ln(\theta) + \frac{L_v q^*}{T}. \tag{3.6}$$

Here, q is the specific humidity, L_v is the latent heat of vaporisation, and T_{LCL} is the temperature of an air parcel after it has been brought to saturation by adiabatic expansion [see appendix D of Holton (2004) for a derivation, or Emanuel (1994) for a more detailed treatment]. Note that the saturation moist entropy is a function of temperature and pressure only; the two conditions above imply that the temperature profile of the atmosphere is fully determined if one knows the boundary layer temperature and moisture content.

3.1.3 The tropical thermal structure

As we shall see below, Earth's atmosphere is typically in a state far from RCE. Nevertheless, the thermal structure within the tropics is well approximated by our RCE solution above, particularly above the boundary layer.

Fig. 3.2 compares the mean temperature profile within tropical regions (20°S-20°N) to that of a moist adiabat lifted from saturation from 850 hPa. The two curves differ by only 1-2 K within the troposphere, indicating that constant s^* is a rather good approximation to the large-scale temperature structure of the tropical troposphere (see also Fig. 1.5). While Fig. 3.2 is based on reanalysis data, other studies have examined the tropical thermal structure from radiosondes and found broadly similar results (Xu and Emanuel, 1989; Singh and O'Gorman, 2013).

The close proximity of the tropical troposphere to a state in which s^* is invariant with height may be understood with the help of two conceptual models: *convective quasi-equilibrium (QE)* and the *weak temperature gradient (WTG) approximation*. According to the QE hypothesis, the effect of moist convection on the large-scale state of the atmosphere is to rapidly relax the temperature profile to one that is moist neutral. As a result, in regions of active deep convection, the atmosphere may be thought of as being in a series of quasi steady states in which convective heating balances destabilisation by non-convective processes (which includes radiation as well as large-scale dynamical uplift), and the same arguments for moist neutrality used for RCE are applicable. We therefore expect that, in convecting regions, the boundary-layer entropy s is roughly equal to saturation entropy s^* within the free troposphere, and s^* itself is roughly constant in the vertical.

The WTG approximation states that, because of the smallness of the Coriolis parameter in tropical regions, temperature gradients in the free troposphere are rapidly removed by

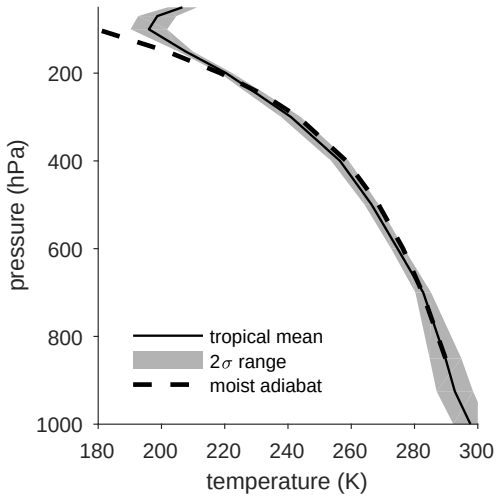

Figure 3.2: Mean temperature profile of the atmosphere for the tropical region (20°S-20°N) for the years 1981-2010 according to the NCEP-DOE reanalysis (solid) and temperature of a pseudoadiabatic parcel ascent lifted from saturation at 850 hPa and initialised at the tropical mean temperature (dashed). Shading represents the $\pm 2\sigma$ range of monthly temperatures for all months and all gridpoints in the tropical belt.

the action of gravity waves. As a result, the tropical atmosphere cannot maintain strong temperature gradients within the free troposphere. This implies that the temperature profile in non-convecting regions of the tropics is strongly constrained by the temperature profile within convecting regions.

Combining QE with the WTG approximation implies that

1. the thermal strcuture in convecting regions is constrained by the QE hypothesis to remain close to moist neutral.

2. by the WTG approximation, the temperature profile within convective regions is communicated to non-convective regions of the tropics, and the entire tropical troposphere is maintained in a state where the saturation entropy s^* is roughly constant.

Fig. 3.2 shows that these conclusions are well justified; not only is the mean tropical thermal structure one in which s^* is almost constant with height, the spatial and temporal variations in temperature within the troposphere are on the order of a few kelvin, certainly much smaller than the variations in mid-tropospheric temperature in midlatitude regions.

The preceding discussion has some far reaching consequences for the tropical precipitation distribution. In particular, it suggests that regions of strong precipitation are characterised by a value of boundary-layer entropy s_b equal to the free tropospheric saturation entropy

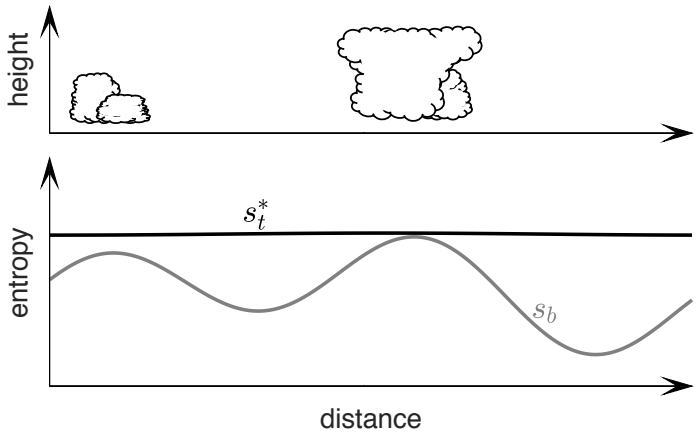

Figure 3.3: (bottom) Idealised distribution of the boundary-layer entropy (s_b; grey) and saturation entropy of the free troposphere (s_t^*; black) and (top) the associated distribution of convective clouds according to the quasi-equilibrium view of tropical precipitation. Deep convection occurs at the s_b maximum, where $s_b \approx s_t^*$. Shallower convection (e.g., congestus) occurs at a secondary, weaker local maximum of s_b.

s_t^*, while in other regions s_b is lower than s_t^*, and the boundary layer is decoupled from the free troposphere above (Fig. 3.3). This implies that the value of s_b is a key determinant of whether a region is able to support deep convection.

The viewpoint described above has become known as the quasi-equilibrium view of tropical precipitation, and there is a vast literature devoted to studying the extent to which it is useful in understanding the tropical general circulation (see e.g., Neelin and Held, 1987; Emanuel et al., 1994; Emanuel, 2007; Privé and Plumb, 2007a,b; Nie et al., 2010; Singh, 2019). While this view has its limits, there is little doubt that the low-level entropy or equivalent potential temperature distributions are more relevant for understanding large-scale circulations than the temperature itself (e.g., Hurley and Boos, 2013).

3.2 Hide's theorem

Let us return to the radiative-convective equilibrium solution to the thermodynamic equation. Can this be used to construct a full solution to the governing equations? And if so, why is this state not observed? More generally, under what conditions is such a solution physically realisable?

3.2.1 Column-by-column RCE with balanced zonal flow

As we have shown above, we may construct a radiative-convective equilibrium state for the atmosphere in which the thermodynamic balance within each column is between radiative cooling and convective heating. Under certain simplifying assumptions, this state can be used to construct a complete solution to the governing equations. In particular, if we assume that frictional forces are negligible except at the surface, we may then construct a column-by-column RCE solution as follows:

1. The local thermodynamic balance is that of RCE.
2. Each column is in a state of hydrostatic balance
3. There is no large-scale meridional overturning: $v = 0$ and $\omega = 0$.
4. The solution is axisymmetric; all quantities are independent of longitude.
5. The winds at the surface are identically zero.

Condition (i) implies that our solution satisfies the thermodynamic equation. Condition (ii) ensures that the vertical momentum equation is satisfied. Condition (iii) and (iv) imply that the flow is non-divergent, and so the mass conservation equation is also satisfied. Condition (v) implies that there is no frictional stress between the surface and the atmosphere, and therefore we may neglect friction everywhere. Applying conditions (iii-v) to the zonal momentum equation, we find that it is trivially satisfied. The final equation to satisfy is the meridional momentum equation, which may be written,

$$2\Omega \sin\phi u + \frac{u^2}{R_e}\tan\phi = -\frac{1}{R_e}\frac{\partial \Phi}{\partial \phi}, \tag{3.7}$$

where we have neglected friction and terms involving v and ω. Equation (3.7) represents a nonlinear balance equation for the zonal wind. Under the Cartesian approximation it reduces to geostrophic balance.

Differentiating (3.7) with respect to pressure, we have the thermal wind relation,

$$\frac{\partial}{\partial p}\left\{2\Omega \sin\phi u + \frac{u^2}{R_e}\tan\phi\right\} = -\frac{1}{R_e}\frac{\partial}{\partial \phi}\frac{\partial \Phi}{\partial p}. \tag{3.8}$$

Applying hydrostatic balance and using the ideal gas law, this may be written in terms of the temperature,

$$\frac{\partial}{\partial p}\left\{2\Omega \sin\phi u + \frac{u^2}{R_e}\tan\phi\right\} = \frac{R_d}{pR_e}\frac{\partial T}{\partial \phi}. \tag{3.9}$$

Integrating with respect to pressure, we have,

$$2\Omega R_e \sin\phi u + u^2 \tan\phi = -R_d \ln\left(\frac{p_s}{p}\right)\frac{\partial \widehat{T}}{\partial \phi}, \tag{3.10}$$

where we have assumed the surface zonal wind is zero, and we have defined an average temperature,

$$\widehat{T} = \frac{1}{\ln\left(\frac{p_s}{p}\right)} \int_{\ln p}^{\ln p_s} T \, \mathrm{d}\ln(p), \tag{3.11}$$

where p_s is the (assumed constant) surface pressure.

The temperature in RCE may be calculated for each latitude using a procedure similar to that of Manabe and Strickler (1964). The zonal wind can then be evaluated using (3.10). Combined with our initial assumption that $v = \omega = 0$, we now have a complete RCE solution to the governing equations!

The RCE solution has no large-scale overturning, and the solar energy absorbed at each latitude is emitted back to space as long-wave radiation; there are no meridional energy transports by atmospheric motions. Apart from the convective motions that transport heat vertically (and which we assume remain local), there is no weather to speak of!

Why does Earth's general circulation look nothing like this RCE solution? There are a number of possible answers, including that we have neglected the seasonal cycle in our discussion, that the RCE solution may be baroclinically unstable, that our assumption that convective motions remain local may be unfounded (Raymond, 2000; Emanuel et al., 2014), or indeed that an axisymmetric solution cannot be realised on a planet with continents. But there is a more fundamental reason that the RCE solution cannot exist that has to do with the unattainability of the zonal wind distribution implied by (3.10). The relevant constraint is known as Hide's theorem and is discussed in the next section.

3.2.2 Constraints on the RCE state

Hide's theorem (Hide, 1969) discusses the range of possible planetary circulations attainable for flows of a certain class. It is based on the concept of angular momentum, which we now introduce.

3.2.3 Angular momentum of the atmosphere

The angular momentum of an object is equal to its momentum crossed with a moment arm. For a fluid, it is useful to consider quantities per unit mass, and so we write the angular-momentum per unit mass $\mathbf{M}$ as,

$$\mathbf{M} = \mathbf{x} \times \mathbf{u}_i, \tag{3.12}$$

where $\mathbf{u}_i$ is the velocity measured in an inertial reference frame. The moment arm can be taken with respect to any point, and here we assume this point is the origin, so that

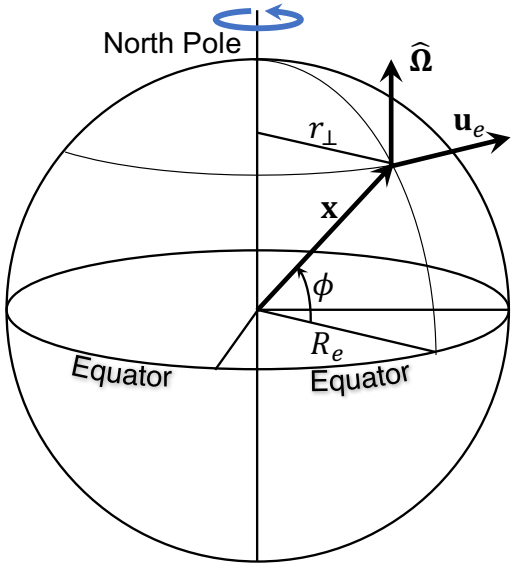

Figure 3.4: Schematic of the position vector $\mathbf{x}$ the velocity of the Earth's surface $\mathbf{u}_e$ and the unit vector in the direction of Earth's rotation axis $\hat{\Omega}$. Additionally the scalar distances R_e and $r_\perp$ are depicted. Adapted from Vallis (2017).

the moment arm is simply the position vector. Taking the Lagrangian time derivative, we have,

$$\frac{\mathrm{D}\mathbf{M}}{\mathrm{D}t} = \frac{\mathrm{D}\mathbf{x}}{\mathrm{D}t} \times \mathbf{u}_i + \mathbf{x} \times \frac{\mathrm{D}\mathbf{u}_i}{\mathrm{D}t}. \tag{3.13}$$

The first term on the right-hand side is the velocity crossed with itself, which is zero. The second term corresponds to the moment arm crossed with the acceleration of the air parcel, which by Newton's Second Law is the force per unit mass. We therefore have that,

$$\frac{\mathrm{D}\mathbf{M}}{\mathrm{D}t} = \mathbf{x} \times \mathbf{f} = \boldsymbol{\tau}, \tag{3.14}$$

where τ is known as the *torque*.

Angular momentum of objects on Earth (planetary angular momentum)

Let us consider the angular momentum of an object stationary on the surface of the Earth. It is convenient to take the moment arm with respect to the centre of the Earth. We may then write,

$$\mathbf{M}_e = \mathbf{x} \times \mathbf{u}_e \tag{3.15}$$

where $\mathbf{u}_e$ is the velocity of the Earth's surface as it rotates. Of particular interest will be the angular momentum in the direction of the Earth's rotation axis, which we may

write,

$$M_e = \hat{\mathbf{\Omega}} \cdot (\mathbf{x} \times \mathbf{u}_e)\,, \tag{3.16}$$

where $\hat{\mathbf{\Omega}}$ is a unit vector in the direction of the Earth's rotation vector (Fig. 3.4). While this is only one component of the angular-momentum vector, we will refer to it as the angular momentum for convenience.

Using the "scalar triple product" identity, we may write the angular momentum as,

$$M_e = \mathbf{x} \cdot \left(\hat{\mathbf{\Omega}} \times \mathbf{u}_e\right). \tag{3.17}$$

Since $\hat{\mathbf{\Omega}}$ and $\mathbf{u}_e$ are perpendicular, we therefore have,

$$M_e = |\mathbf{x}||\mathbf{u}_e| \cos\xi, \tag{3.18}$$

where ξ is the angle between the position vector $\mathbf{x}$ and the cross product $\hat{\mathbf{\Omega}} \times \mathbf{u}_e$. Examining Fig 3.4 and performing some geometry, it may be seen that the angular momentum of the Earth's surface may be written,

$$M_e = |\mathbf{x}||\mathbf{u}_e| \cos\phi. \tag{3.19}$$

It may also be seen that $|\mathbf{x}| = R_e$, and that $|\mathbf{u}_e| = \Omega R_e \cos\phi$, so that the angular momentum of an object on Earth is given by,

$$M_e = \Omega R_e^2 \cos^2\phi. \tag{3.20}$$

All objects on Earth have angular momentum M_e as a result of the Earth's rotation. This component of angular momentum will be referred to as *planetary angular momentum.*

Angular momentum of the winds (relative angular momentum)

In addition to planetary angular momentum, the atmosphere has angular momentum because of its motion relative to that of the Earth, which we refer to as *relative angular momentum*, M_r. The total angular momentum is simply the sum of the planetary and relative components,

$$M = \hat{\mathbf{\Omega}} \cdot \{\mathbf{x} \times (\mathbf{u}_e + \mathbf{u})\} \tag{3.21}$$

where $\mathbf{u}$ is the velocity in the frame of reference of the Earth. The only component of the velocity that contributes to angular momentum in the direction of the Earth's axis is that in the zonal direction. We therefore have that,

$$M = |\mathbf{x}|\,(|\mathbf{u}_e| + u)\cos\phi = R_e \cos\phi\,(\Omega R_e \cos\phi + u)\,. \tag{3.22}$$

Note that our definition for M contains the radius of the Earth R_e rather than the distance of the air parcel from the centre of the Earth. This is to be consistent with our use of the shallow fluid approximation in deriving the governing equations in the previous chapter.

Conservation of angular momentum

The utility of angular momentum comes about because it is conserved in the absence of torques. In particular, the angular momentum M is conserved for axisymmetric, frictionless flow. This may be easily seen by substituting (3.22) into the zonal momentum equation in height coordinates (2.26a), which, with a little algebra gives,

$$\frac{\mathrm{D}M}{\mathrm{D}t} = -\frac{1}{\rho}\frac{\partial p}{\partial \lambda} + R_e \cos\phi \mathcal{F}_{\nu\lambda}. \tag{3.23}$$

If the flow is axisymmetirc, all variables are independent of longitude and the first term on the right-hand side is zero. If the flow is frictionless, the second term on the right-hand side is also zero, and angular momentum is conserved.

Multiplying the above equation by ρ and assuming axisymmetry, we have,

$$\rho\frac{\partial M}{\partial t} + \rho\mathbf{u}\cdot\nabla M = R_e\rho\cos\phi\mathcal{F}_{\nu\lambda}. \tag{3.24}$$

Recall also that the continuity equation may be written,

$$\frac{\partial \rho}{\partial t} + \nabla\cdot(\rho\mathbf{u}) = 0. \tag{3.25}$$

Summing the two previous equations, we can derive the angular momentum conservation equation in flux form,

$$\frac{\partial \rho M}{\partial t} + \nabla\cdot(\rho\mathbf{u}M) = R_e\rho\cos\phi\mathcal{F}_{\nu\lambda}, \tag{3.26}$$

valid for axisymmetric flow.

3.2.4 Angular-momentum extrema in axisymmetric steady flow

We now consider constraints on the range of possible flows placed on the atmosphere by conservation of angular momentum. Consider an atmosphere with a steady, axisymmetric circulation. Further, suppose that the frictional stress within this atmosphere acts as a downgradient diffusion in angular momentum with some viscosity ν. Under these assumptions, the angular momentum conservation equation (3.26) may be written,

$$\nabla\cdot(\rho\mathbf{u}M) = \nabla\cdot(\rho\nu\nabla M)\,. \tag{3.27}$$

If the atmosphere is stationary, the angular momentum is a maximum on the equator, and decreases monotonically poleward. Suppose for this atmosphere, however, the zonal winds are westerly in the equatorial upper troposphere, so that there is a local maximum of angular momentum there, as shown in Fig. 3.5.

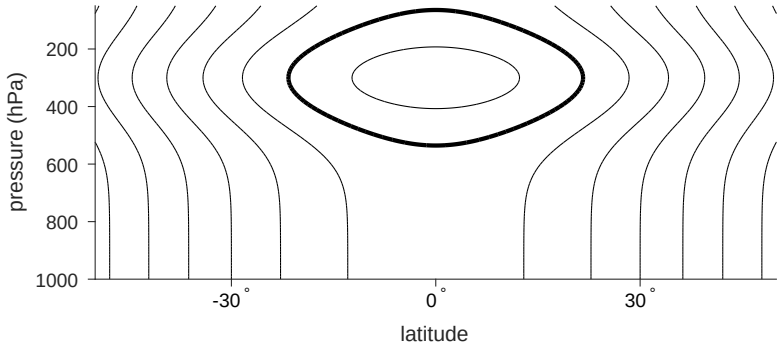

Figure 3.5: Idealised angular-momentum distribution that violates Hide's theorem. Lines represent contours of angular momentum with vales increasing toward the equator. The thick contour corresponds to a closed boundary on which $M = M_0$ is constant as in (3.29).

Since we assume a local maximum in M exists, we must be able to find a closed contour surrounding the maximum on which $M = M_0$ is a constant. Integrating (3.28) over the region bounded by this contour, we have,

$$\oint_{M>M_0} \nabla \cdot (\rho \mathbf{u} M)\, \mathrm{d}V = \oint_{M>M_0} \nabla \cdot (\rho \nu \nabla M)\, \mathrm{d}V\,. \tag{3.28}$$

By the divergence theorem, these integrals may be written as surface integrals over the closed surface on which $M = M_0$,

$$\oint_{M=M_0} \rho M \mathbf{u} \cdot \hat{\mathbf{n}}\, \mathrm{d}S = \oint_{M=M_0} \rho \nu \nabla M \cdot \hat{\mathbf{n}}\, \mathrm{d}S\,. \tag{3.29}$$

Consider the left-hand side of the above equation. Since $M = M_0$ on the integration surface, we may remove it from the integral so that,

$$\oint_{M=M_0} \rho M \mathbf{u} \cdot \hat{\mathbf{n}}\, \mathrm{d}S = M_0 \oint_{M=M_0} \rho \mathbf{u} \cdot \hat{\mathbf{n}}\, \mathrm{d}S\,. \tag{3.30}$$

The integral on the right-hand side corresponds to the net mass flux out of the region, which by steady mass conservation must be zero. Now consider the right-hand side of (3.29). ∇M is perpendicular to level sets of M and points towards increasing M so that, on the integration surface, $\nabla M \cdot \hat{\mathbf{n}} = -|\nabla M| < 0$. Since ρ and ν are both positive, this implies that the right-hand side is negative definite!

We have therefore arrived at a contradiction; the left-hand side of (3.29) is zero, but for any non-zero viscosity, the right-hand side is negative definite. The conclusion we reach is known as Hide's theorem:

> *A steady, axisymmetric flow cannot develop maxima in angular momentum away from boundaries; such maxima will always be eroded by diffusion.*

Hide's theorem applies to any steady, axisymmetric flow with non-zero viscosity. This means that, strictly speaking, it does not apply to inviscid flows. However, since even a negligible amount of viscosity produces a contradiction in (3.29), it applies to any real flow no matter how small the effect of friction may be. Thus while an inviscid solution that violates Hide's theorem is mathematically sound (in that both sides of (3.29) are zero), it is singular, in that it is not the correct solution in the limit in which the viscosity is reduced to zero from a positive value. Thus, any physically realisable solution of the atmospheric circulation cannot violate Hide's theorem.

3.2.5 The violation of Hide's theorem

Let us now consider the RCE solution developed in the previous section from the point of view of Hide's theorem. Since the RCE solution is assumed steady and axisymmetric, Hide's theorem is applicable, and such a solution must not produce an angular-momentum maximum above the surface in order to be attainable. How can we tell if the RCE solution produces such a maximum?

Consider the thermal wind equation (3.8),

$$\frac{\partial}{\partial p}\left\{2\Omega\sin\phi u + \frac{u^2}{R_e}\tan\phi\right\} = \frac{1}{R_e}\frac{\partial\alpha}{\partial\phi}, \tag{3.31}$$

where we have used hydrostatic balance. Now, usually, we rewrite this equation in terms of temperature. However, in this case, we will use some thermodynamics, in combination with the assumption of moist neutrality, to instead put it in terms of the boundary-layer entropy.

Neglecting the effects of condensed water, we may write the specific volume as a function of pressure and temperature $\alpha = \alpha(p, T)$. But since the saturation entropy is also a function of temperature and pressure $s^* = s^*(p, T)$, we may, without loss of generality, write, $\alpha = \alpha(p, s^*)$. By the chain rule, we also have that,

$$\left.\frac{\partial\alpha}{\partial\phi}\right|_p = \left.\frac{\partial\alpha}{\partial s^*}\right|_p \left.\frac{\partial s^*}{\partial\phi}\right|_p. \tag{3.32}$$

The derivative of the specific volume with respect to saturation entropy is a thermodynamic function of pressure and temperature. In fact, we can derive a Maxwell relation that states,

$$\left.\frac{\partial\alpha}{\partial s^*}\right|_p = \left.\frac{\partial T}{\partial p}\right|_{s^*}, \tag{3.33}$$

where the right-hand side corresponds to the moist adiabatic lapse rate (see Emanuel, 1994, for an introduction to Maxwell relations) . Substituting (3.32) and (3.33) into (3.31), we have

$$\frac{\partial}{\partial p}\left\{2\Omega\sin\phi u + \frac{u^2}{R_e}\tan\phi\right\} = \frac{1}{R_e}\frac{\partial s^*}{\partial\phi}\left.\frac{\partial T}{\partial p}\right|_{s^*}, \tag{3.34}$$

Now, by the assumption that the RCE solution is moist neutral, we may assume that s^* is independent of pressure and equal to the boundary-layer entropy s_b. We may therefore directly integrate the above equation from the boundary layer to the tropopause to give,

$$2\Omega\sin\phi u_t + \frac{u_t^2}{R_e}\tan\phi = -\frac{(T_b - T_t)}{R_e}\frac{\partial s_b}{\partial\phi}, \tag{3.35}$$

where the subscript b and t refer to variables evaluated in the boundary layer and at the tropopause, respectively, and we have used that the zonal wind in the RCE solution is zero at the surface.

Equation (3.34) is a quadratic for the zonal wind at the tropopause. Applying the quadratic formula, we may solve this equation for the tropopause zonal velocity u_t,

$$\frac{u_t}{\Omega R_e} = \left\{\left(1 - \frac{(T_b - T_t)}{\Omega^2 R_e^2\sin\phi\cos\phi}\frac{\partial s_b}{\partial\phi}\right)^{\frac{1}{2}} - 1\right\}\cos\phi, \tag{3.36}$$

where we have taken the root that ensures that $u_t \to 0$ as $\partial_\phi s_b \to 0$.

The tropopause zonal wind distribution (and therefore the distribution of angular momentum) within the RCE state is dependent on the meridional gradient of boundary layer entropy s_b. A detailed evaluation of s_b in the RCE state requires a full radiative-convective calculation with a numerical model. Nevertheless, some insight may be gained by thinking about hypothetical s_b distributions.

As a starting point for thinking about the RCE distribution of s_b, we note that, to a very good approximation, the annual-mean solar insolation as a function of latitude has a dependence on latitude proportional to $\sin^2\phi$ (see Hartmann, 1994, and Fig. 1.1). Based on this, we consider an idealised boundary-layer entropy distribution that has the functional form,

$$s_b^{\mathrm{RCE}} = s_{b0}^{\mathrm{RCE}} - \delta s_b^{\mathrm{RCE}}\sin^2\phi. \tag{3.37}$$

where s_{b0}^{RCE} and $\delta s_b^{\mathrm{RCE}}$ are constants. Substituting this distribution into (3.36), we have

$$\frac{u_t^{\mathrm{RCE}}}{\Omega R_e} = \left\{\left(1 + \frac{2(T_b - T_t)\delta s_b^{\mathrm{RCE}}}{\Omega^2 R_e^2}\right)^{\frac{1}{2}} - 1\right\}\cos\phi, \tag{3.38}$$

where the superscript RCE refers to a quantity in the RCE state. It may be easily seen from the above equation that the zonal wind at the tropopause in the RCE solution is

westerly (positive) at the equator for any positive value of $\delta s_b^{\mathrm{RCE}}$. This implies that the angular momentum of the equatorial tropopause is greater than the angular momentum of Earth's surface at any point on the planet. In particular, this implies that the maximum value of the angular momentum in the atmosphere occurs above the surface. That is, the RCE solution violates Hide's theorem.

More generally, one can show that any RCE distribution s_b^{RCE} will violate Hide's theorem if it has non-zero latitudinal curvature at the equator. Given the annual-mean distribution of solar insolation on Earth, any reasonable estimate of s_b^{RCE} is likely to include such curvature.

The above discussion provides an answer to the main question of this chapter. Why does the atmospheric circulation exist? It is because the relevant solution with no large-scale circulation is unphysical, it produces an angular-momentum distribution that cannot be maintained by any real flow. In order to produce a physically attainable angular momentum distribution, the maximum in angular momentum at the equatorial tropopause must be removed. By thermal wind balance, this requires a reduction in the meridional temperature gradients. Such a reduction can be produced by a large-scale overturning circulation that transports energy poleward.

3.3 Summary

In this chapter we investigated the fundamental reasons for the existence of a large-scale overturning circulation in the atmosphere by attempting to construct solutions to the governing equations that had no circulation. Our line of reasoning was as follows:

- We showed that an atmosphere in radiative equilibrium is gravitationally unstable, and any small perturbation to this state would lead to the onset of convection.
- We constructed a radiative-convective equilibrium solution in which gravitational instability is assumed to be released locally and thus there is no large-scale circulation.
- We showed that the RCE solution required, by thermal wind balance, an unphysical distribution of angular momentum within the atmosphere. This implies that the RCE solution requires horizontal temperature gradients that cannot be maintained; a large-scale circulation must exist in order to reduce these temperature gradients.

In the next chapter, we will investigate the properties of the atmospheric circulation in an axisymmetric atmosphere as a first step toward considering the atmosphere in its full complexity.

Chapter 4

Axisymmetric Hadley Cells

In the previous chapter, we constructed a radiative-convective equilibrium solution for the atmosphere in which there was no large-scale circulation. Such a solution is valid for a strictly inviscid fluid, but it is not the correct solution when one takes the limit in which the viscosity is reduced to zero from a positive value. In this chapter we will describe this "nearly-inviscid" solution to the governing equations. For the time being, we remain within the axisymmetric framework (thereby neglecting zonal variations) and take an annual-mean view (thereby neglecting seasonal variations). These assumptions will be reexamined in chapter 8.

What does the axisymmetric nearly-inviscid circulation look like? Assuming steady, axisymmetric, nearly-inviscid flow, the equation for angular momentum conservation becomes,

$$\mathbf{u} \cdot \nabla M = 0. \tag{4.1}$$

Assuming drag near to the surface is the only important form of friction, this equation applies everywhere above the planetary boundary layer. Equation (4.1) states that the vector velocity and the gradient ∇M are perpendicular. Since ∇M is also perpendicular to the surfaces of constant angular momentum, this implies that the velocity is parallel to angular momentum contours. That is, angular momentum is conserved along streamlines of the flow, and streamlines and angular momentum contours are parallel. A complete model of such a circulation was presented by Held and Hou (1980).

4.1 The Held & Hou model

The model of Held and Hou (1980) was originally framed in a Boussinesq, stably stratified fluid with rigid top and bottom boundaries. Here, we will give it a slightly more atmospheric

flavour, by framing the model as an ideal gas in which the thermal structure is maintained by convection through the assumption of convective quasi-equilibrium.

4.1.1 Conservation of angular momentum along streamlines

We begin by assuming, on physical grounds, that the zonal-mean circulation corresponds to rising air at the equator which then reaches the tropopause and flows poleward. We will refer to this circulation as the Hadley Cell, although at this stage we do not know its properties. We assume that, in the boundary layer, friction is strong, and the boundary-layer zonal winds are everywhere weak, so that the angular momentum of boundary-layer air M_b is roughly equal to the planetary angular momentum $M_b \approx M_p$. This implies that air rising at the equator will initially have an angular-momentum equal to that of the Earth at the equator $M_p(\phi = 0) = \Omega R_e^2$. By our discussion above, angular momentum is conserved along this streamline when it is above the boundary layer, and the air will maintain this angular momentum as it reaches the tropopause and flows poleward (Fig. 4.1). We may therefore write the angular momentum at the tropopause within the Hadley cell $M_t(\phi)$ as,

$$M_t(\phi) = \Omega R_e^2. \tag{4.2}$$

Expressing the angular momentum in terms of the zonal wind, we have,

$$R_e \cos\phi \left(u_t + \Omega R_e \cos\phi\right) = \Omega R_e^2. \tag{4.3}$$

This may be rearranged to give an explicit expression for the zonal velocity at the tropopause,

$$u_t^{\mathrm{AMC}} = \Omega R_e \sin\phi \tan\phi. \tag{4.4}$$

where we use the superscript AMC to refer to the angular momentum conserving wind, that is, the wind distribution associated with an air parcel initially stationary at the equator moving poleward and conserving its angular momentum.

Assuming that the atmosphere is in convective quasi equilibrium, and therefore s^* is invariant with height, the thermal wind equation (3.34) is valid. Further assuming that the surface zonal winds are weak relative to the zonal wind at the tropopause, the integral form (3.35) is also valid. This provides a relationship between the boundary-layer entropy and the zonal winds at the tropopause. Substituting u^{AMC} into (3.35), we have,

$$2\Omega^2 R_e \sin^2\phi \tan\phi + \Omega^2 R_e \sin^2\phi \tan^3\phi = -\frac{(T_b - T_t)}{R_e}\frac{\partial s_b^{\mathrm{AMC}}}{\partial \phi}, \tag{4.5}$$

where s_b^{AMC} is the boundary-layer entropy associated with the AMC solution. After some rearrangement, this gives an equation for the meridional gradient of boundary-layer entropy,

$$\frac{(T_b - T_t)}{\Omega^2 R_e^2}\frac{\partial s_b^{\mathrm{AMC}}}{\partial \phi} = -\tan^3(\phi)(1 + \cos^2\phi). \tag{4.6}$$

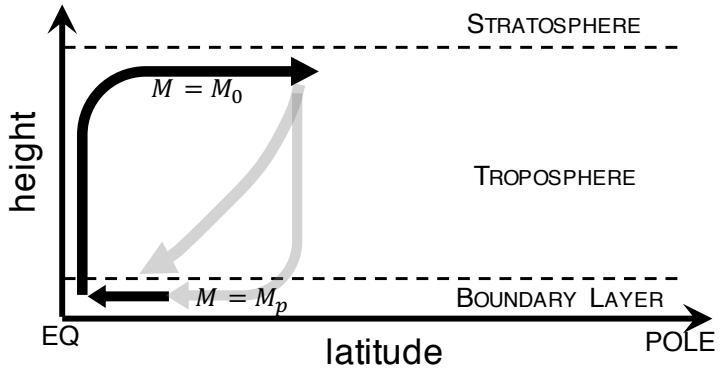

Figure 4.1: Schematic of the axisymmetric Hadley Cell. Angular momentum is conserved along the rising and poleward branches ($M = M_0$), while the zonal surface winds in the region of the return flow are assumed to be weak ($M = M_p$) (black arrows). The descending branch is either vertical or slanted depending on the mechanisms that bring the angular momentum in the descending branch to the surface value (grey arrows).

Finally, assuming the depth of the troposphere in temperature coordinates $T_b - T_t$ is constant, we may integrate the above equation with respect to latitude to give an explicit equation for s_b^{AMC},

$$s_b^{\mathrm{AMC}}(\phi) = s_0 - \frac{\Omega^2 R_e^2}{2(T_b - T_t)} \frac{\sin^4 \phi}{\cos^2 \phi} \tag{4.7}$$

where s_0 is the value of boundary-layer entropy at the equator, which remains a free parameter.

At this point, we must stress that the only assumptions being made about the circulation is that it is in convective quasi-equilibrium and angular momentum is conserved along streamlines. And yet we have discovered a very strong constraint on the thermal structure of the atmosphere. In particular, we have not assumed anything about the strength or energy fluxes associated with the Hadley Cell, and yet we can confidently state the horizontal temperature gradients that must result. Clearly angular momentum conservation is a powerful principle.

Equation (4.7) is characterised by a very flat distribution of entropy near the equator (Fig. 4.2). Indeed, as required by Hide's theorem, the functional form embodied in (4.7) has no curvature at the equator. On the other hand, the gradient of s_b^{AMC} increases rapidly outside the tropics, and both the entropy gradient and the angular momentum conserving wind u^{AMC} approach infinity at the pole. This implies that the angular momentum conserving solution cannot continue to the pole; the theory predicts that the Hadley Cell has a latitudinal extent! Beyond the Hadley Cell, the only available solution within the axisymmetric

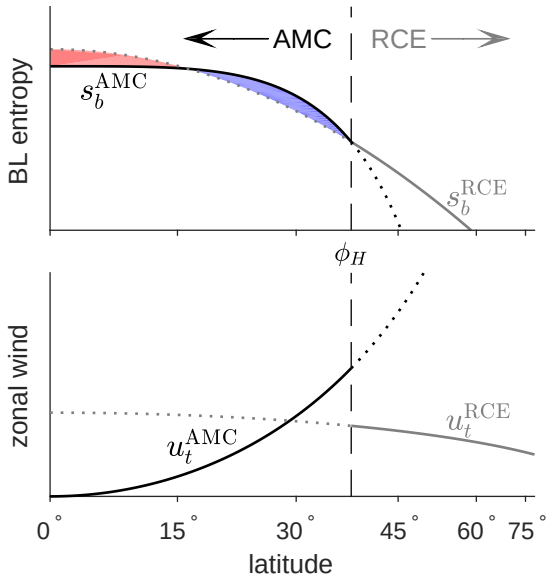

Figure 4.2: Solution to the Held-Hou model of the Hadley Cell plotted against the sine of latitude. (top) Boundary-layer entropy and (bottom) tropopause zonal wind according to the angular-momentum conserving (AMC) solution (black) and according to the RCE solution (grey). The AMC solution is valid equatorward of the vertical dashed line, while the RCE solution is valid poleward of the vertical dashed line. The red and blue areas in the top panel are of equal size.

framework is that of RCE. The theory of Held and Hou (1980) therefore predicts a well defined Hadley Cell beyond which the atmosphere is in RCE (Fig. 4.2). Specifically, the boundary layer entropy s_b equals s_b^{AMC} within the Hadley Cell, while $s_b = s_b^{\mathrm{RCE}}$ further poleward. In the next section we seek to determine what sets the Hadley Cell extent.

4.1.2 Thermodynamic constraints on the Hadley Cell extent

To estimate the Hadley Cell extent, we need to make some appeal to thermodynamic considerations. On physical grounds, we expect the Hadley Cell to transport energy polewards. In an steady-state axisymmetric atmosphere, the energy flux divergence owing to the Hadley Cell heat flux must be balanced by the net top-of-atmosphere (TOA) flux of radiation at each latitude[1] . As for the RCE solution, a detailed estimate of these fluxes and their dependence on the Hadley Cell itself requires a numerical calculation with a radiative-convective model. However, we can advance or understanding by considering a very simple parameterisation for the energy fluxes.

[1] convective heating/cooling cannot affect the top-of-atmosphere energy balance because it only transports energy within the column.

Suppose the net radiative flux at the top of the atmosphere $F_{\text{net}}^{\text{TOA}}$ is a linear function of the boundary-layer entropy, so that,

$$F_{\text{net}}^{\text{TOA}} = -k(s_b - s_b^{\text{RCE}}) \tag{4.8}$$

where k is a constant. This parameterisation is doubtless oversimplified, but it has the useful properties that the net flux is zero when the atmosphere is in RCE, and that the net flux increases the farther away the atmosphere is from RCE.

For a steady atmosphere, the global average TOA flux of radiation must be zero, and this implies,

$$\oint_{\text{globe}} s_b - s_b^{\text{RCE}} \, \mathrm{d}A = 0. \tag{4.9}$$

Since poleward of the Hadley Cell, $s_b = s_b^{\text{RCE}}$, we may write,

$$\int_0^{\phi_H} s_b^{\text{AMC}} \cos\phi \, \mathrm{d}\phi = \int_0^{\phi_H} s_b^{\text{RCE}} \cos\phi \, \mathrm{d}\phi, \tag{4.10}$$

where we have defined ϕ_H as the latitudinal extent of the Hadley Cell, and we have used the zonal and hemispheric symmetry of the problem to write it in terms of an integral in longitude within one hemisphere.

In addition to the above integral constraint, we also have a continuity constraint on the entropy distribution where we assume that the boundary-layer entropy is continuous at the Hadley Cell boundary,

$$s_b^{\text{AMC}}(\phi_H) = s_b^{\text{RCE}}(\phi_H). \tag{4.11}$$

Graphically, the two constraints above may be interpreted as the requirement that the red and blue areas on Fig. 4.2 are equal. In principle, one may use (4.7) for the AMC solution and our simplified expression (3.37) for the RCE solution to solve (4.10) and (4.11) for ϕ_H. In practice, this must be done numerically. However, if one makes the small angle approximation, in which $\sin\phi \approx \phi$, and $\cos\phi \approx 1$, and analytic solution for ϕ_H may be obtained,

$$|\phi_H| = \left(\frac{5}{3} \frac{(T_b - T_t)}{\Omega^2 R_e^2} \delta s_b^{\text{RCE}} \right)^{\frac{1}{2}}. \tag{4.12}$$

To obtain an estimate of the Hadley Cell extent according to Held-Hou theory, we substitute reasonable values into the above equation. Using (3.5), we have, for δs_b^{RCE},

$$\delta s_b^{\text{RCE}} \approx \frac{c_p \delta T_b^{\text{RCE}}}{T} + \frac{L_v}{T} \delta q^{*\text{RCE}} \tag{4.13}$$

where δT^{RCE} and $\delta q^{*\text{RCE}}$ are the pole to equator differences in temperature and saturation specific humidity within the RCE state, and T is a temperature scale. Supposing that

the tropospheric depth in temperature coordinates $(T_b - T_t) = 100\text{K}$, $\delta T_b^{\text{RCE}} = 100$ K, $\delta q^{*\text{RCE}} = 0.02$ and the temperature scale $T = 270$ K, we get an estimate of $\phi_H = 38°$. The precise value obviously depends on the exact parameters used, but it is nonetheless roughly similar to the observed Hadley Cell extent, which is about 30°.

4.1.3 Tropopause zonal wind distribution

The tropopause level zonal wind in the Held-Hou solution is given by the AMC wind (4.4) for latitudes equatorward of ϕ_H and the RCE zonal wind (3.38) further poleward (Fig. 4.2). The AMC wind increases with latitude, roughly quadratically for small angles, while the RCE solution weakly decreases with latitude.

Unlike the boundary-layer entropy, the tropopause zonal wind is discontinuous at the boundary ϕ_H. This results in the wind distribution being characterised by a sharp jet at the poleward edge of the Hadley Cell, just as is observed! However, for Earth-like parameters, u_t^{AMC} increases to close to 200 m s^{-1} as the latitude approaches the estimated Hadley Cell extent of roughly 38°. Clearly, the zonal winds predicted by the Held-Hou model are considerably stronger than those observed.

4.1.4 Energy fluxes and the strength of the Hadley cell

To determine the strength of the Hadley Cell, we will once again turn to the atmospheric energy budget. In steady state, the divergence of the meridional energy flux by the Hadley Cell must be equal to the net radiative flux at the top of the atmosphere $F_{\text{net}}^{\text{TOA}}$. Applying the simple parameterisation of radiative heating (4.8),

$$\frac{1}{R_e \cos\phi} \frac{\partial F_H \cos\phi}{\partial \phi} = k(s_b^{\text{AMC}} - s_b^{\text{RCE}}) \tag{4.14}$$

where F_H is the zonal-mean meridional flux of energy by the Hadley Cell and may be written,

$$F_H = \int_0^{p_s} [v][h] \frac{\mathrm{d}p}{g}, \tag{4.15}$$

where h is the relevant energy variable discussed further below.

Held and Hou (1980) assumed that the meridional mass flux within the Hadley Cell is confined to layers of pressure depth δp at the surface and tropopause, so that we may write,

$$F_H = h_b \int_{p_s - \delta_p}^{p_s} [v] \frac{\mathrm{d}p}{g} + h_t \int_{p_t - \delta p}^{p_t} [v] \frac{\mathrm{d}p}{g}. \tag{4.16}$$

By mass conservation, the two integrals in the above equation are equal in magnitude and opposite in sign. We therefore have,

$$F_H = (h_t - h_b) \frac{\Psi_{\max}}{2\pi R_e \cos\phi}, \tag{4.17}$$

where $\Psi_{\max}$ is the maximum value of the streamfunction at each latitude, with the streamfunction being defined in the usual way following (2.61). The strength of the Hadley cell is therefore related to the energy flux by the Hadley cell and the difference $\Delta h = h_t - h_b$ between the poleward and equatorward flowing layers. But what exactly is h?

A first look at energetics

Held and Hou (1980) considered the Hadley cell in a dry stably stratified fluid. In the absence of moisture, the relevant energy variable is the dry static energy[2] $h = h_d = c_p T + \Phi$. The difference in h between the surface and tropopause in such a fluid depends on the stratification, which Held and Hou (1980) took as an external parameter. Therefore, in the original Held-Hou model, the energy and strength of the Hadley Cell are related by an external parameter. Given a parameterisation of the top-of-atmosphere radiative fluxes, one may calculate the required heat transport, and this fixes the Hadley Cell strength for a given stratification.

In our case, we are considering a moist atmosphere in convective quasi-equilibrium. The relevant energy variable is the moist static energy $h = h_m = c_p T + \Phi + L_v q$. Moist static energy is a function of both temperature and humidity, and it is closely related to the entropy. In fact, one may define a saturation moist static energy $h_m^* = c_p T + \Phi + L_v q^*$, with the property that, for a moist neutral atmosphere,

$$\frac{\partial h_m^*}{\partial z} \approx 0, \tag{4.18}$$

above the boundary layer.

Now, let us consider the difference between the saturation moist static energy and the actual moist static energy. By their respective definitions, we may write,

$$h_m^* - h_m = L_v(q^* - q). \tag{4.19}$$

The function in brackets on the right-hand side is known as the saturation deficit, and it has the following properties:

1. In the boundary layer, the relative humidity is high, and $q^* - q$ is small.
2. In the mid troposphere, the relative humidity is typically lower, and $q^* - q$ is large.
3. In the upper troposphere, the temperature is low, and therefore q^* is small, implying that $q^* - q$ is also small.

The properties above suggest that, for an atmosphere in convective quasi-equilibrium, h_m has a minimum in the midtroposphere. Indeed, this is also the case in Earth's tropics (Fig. 4.3).

[2] We also neglect here the transport of kinetic energy, but this is generally two orders of magnitude smaller than the transport of dry static energy.

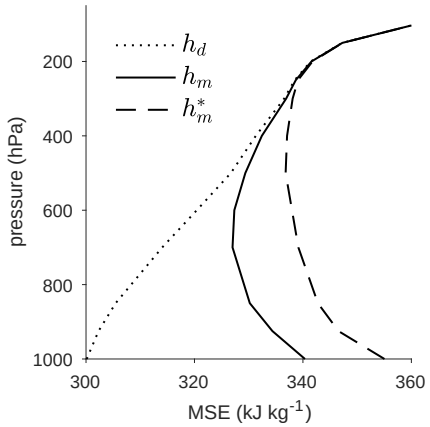

Figure 4.3: Profiles of the tropical-mean (20°S-20°N) and time-mean dry static energy (dotted), moist static energy (solid) and saturation moist static energy (dashed) according to the NCEP-DOE reanalysis for the years 1981-2010.

The implications of the non-monotonic profile of h_m is that the Hadley Cell meridional energy flux,

$$F_H = \int_0^{p_s} [v][h_m] \frac{\mathrm{d}p}{g}, \tag{4.20}$$

is highly sensitive to both the profile of the meridional velocity $[v]$ and the profile of the saturation deficit $(q^* - q)$. Indeed, it is not obvious *a priori* that the Hadley Cell should even transport energy poleward!

The above complications in the energetics of moist atmospheres are usually described in terms of the *gross moist stability (GMS)* first introduced by Neelin and Held (1987). The GMS effectively plays the role of Δh in a moist atmosphere, but it depends on both the profile of h_m, and the vertical structure of the overturning circulation. A full understanding of the behaviour of GMS is still lacking and is the subject of active research (see e.g., Raymond et al., 2009; Inoue and Back, 2015, 2017).

A further complication in the energetics of the Hadley Cell that we have not considered yet is the ocean. At low latitudes, most of the poleward energy transport through the climate system occurs in the ocean rather than in the atmosphere. In fact, it may be shown that the response of the oceanic circulation to the Hadley Cell wind stress results in a poleward flux of energy a factor of three larger than that of the Hadley Cell itself (Held, 2001). Neglect of oceanic processes in considering poleward energy transport in the climate system is therefore of dubious utility.

It is clear that a detailed analysis of the energetics of the Hadley Cell, and the energetics of the low-latitude climate system in general, requires a consideration of a number of aspects

not present in the simple Held-Hou model. We therefore postpone further discussion of atmospheric energetics to chapter 9 when we discuss the energy budget of the atmosphere in more detail.

4.1.5 Summary of the Held & Hou model

The key features of the Held-Hou model may be summarised as follows:

- The Held-Hou model predicts a latitudinally confined Hadley Cell, with rising motion at the Equator and descending motion in the subtropics. Further poleward, there is no large-scale circulation.
- Associated with the Hadley Cell, there is a sharp subtropical jet at the cell edge.
- The Hadley Cell transports energy from the tropics to the subtropics, but there is no energy transport beyond the poleward edge of the Hadley Cell.
- The boundary-layer entropy distribution becomes very flat in the tropics, but remains at its RCE value poleward of the Cell edge.

These features share much in common with the observed zonal-mean circulation on Earth; The Hadley Cell is the dominant meridional overturning (Fig. 1.9), it terminates within the subtropics and has an associated jet (Fig. 1.6), and the meridional gradient of temperature in the tropical troposphere is close to zero (Fig. 1.5). Obviously the detailed predictions of the Held-Hou theory are inaccurate given the simplicity of the model. But given its similarity to the observed circulation, it is tempting to conclude that it captures much of the essence of the dynamics of the Hadley Cell. In the next section, we discuss some limitations of the Held-Hou theory as a representation of Earth's Hadley Cell, and this will lead us to consider the role played by eddies in the general circulation in much more detail.

4.2 Limitations of the Held & Hou model

4.2.1 Numerical simulations

To motivate this section, we present the original numerical simulations of the Held-Hou model in Fig. 4.4. The figure shows three simulations, each with different values of viscosity. As expected, the subtropical jet becomes sharper as the viscosity decreases, more closely approximating the nearly-inviscid analytical solution. However, the strength of the Hadley Cell, measured by the maximum value of the streamfunction, is strongly dependent on the viscosity, and in fact increases with increasing viscosity! This is rather surprising, as we do not generally expect that increasing the importance of friction within a flow should lead to a more energetic circulation. How can we understand these results?

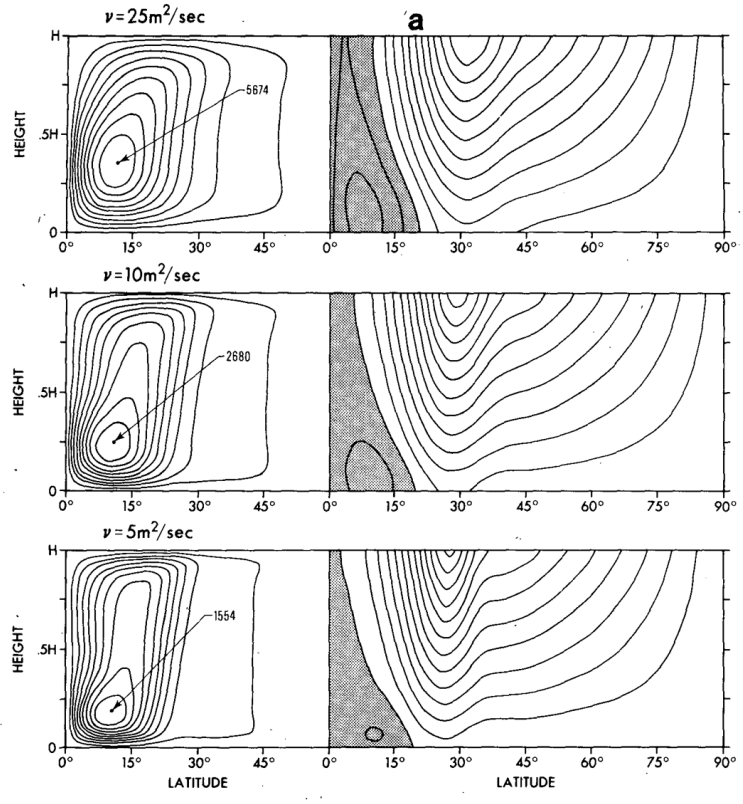

Figure 4.4: Numerical solutions of the Held-Hou model with decreasing viscosity (top to bottom). (left) Overturning streamfunction Ψ with contour interval of $0.1\Psi_{\max}$ and (right) zonal wind speed with contour interval 5 m s^{-1} and easterly winds shaded. The Hadley Cell strength $\Psi_{\max}$ is 5674, 2680 and 1554 m^2 s^{-1} in the top, middle, and bottom panels, respectively. From Held and Hou (1980).

4.2.2 Angular-momentum sources and sinks

Let us return to our original conception of the Held-Hou model (Fig. 4.1). Air rises at the equator with angular momentum $M_t(\phi) = M_p(0)$ and conserves its angular momentum as it reaches the tropopause and flows polewards. In the return flow within the boundary layer, the winds are weak, and the angular momentum is close to that of the Earth $M_b(\phi) = M_p(\phi)$. But how does the air at the tropopause with angular momentum $M_p(0)$ return to the boundary layer having angular momentum $M_p(\phi)$?

There are two possible solutions (Fig. 4.1):

1. The air takes a slanted path on its descent, such that it enters the boundary layer close to the equator where its angular momentum matches that of the Earth.

2. The air loses its angular momentum on its descent by some process.

The first solution is the only option in axisymmetric flows with very low viscosity, and slanted descending branches have been found in a number of numerical (Satoh et al., 1995; Singh and Kuang, 2016) and theoretical (Fang and Tung, 1996) models of axisymmetric Hadley Cells.

If the viscosity is somewhat higher, friction can act as a sink of angular momentum in order to reduce it as air slowly descends from the tropopause to the boundary layer in the Hadley Cell's descending branch. The smaller the viscosity, the more time required to reduce the tropopause angular momentum to boundary-layer values, and the weaker the descent rate within the Hadley Cell. This heuristic argument provides an explanation for the viscosity dependence of the Hadley Cell strength found in the numerical solutions of Held and Hou (1980).

In Earth's atmosphere, angular momentum can also be transported by eddy fluxes, and this provides another pathway through which the required sink of angular momentum may come about. This provides a hint that a purely axisymmetric model may not be able to account for the observed behaviour of Earth's Hadley Cell. In the next few chapters, we will explore the role of non-axisymmetric motions on Earth's angular momentum budget, first from the perspective of the midlatitude circulation, in chapter 5-7, and then through the effect of eddy fluxes of angular momentum on the Hadley Cell in chapter 8.

Chapter 5

The angular-momentum budget of the atmosphere

In the previous two chapters, we investigated the circulation of an idealised axisymmetric atmosphere, in which we neglected longitudinal variations in the flow. We saw that this produced a somewhat realistic model for the tropical overturning, but it implied there could be no large-scale circulation in the extratropics. This suggests that longitudinal variations – eddies – are of first order importance to the large-scale circulation in extratropical regions. In this chapter we will demonstrate this importance from the point of view of the angular-momentum budget.

5.1 Motivation

In chapter 1, we showed some of the key features of the observed atmospheric circulation. In particular we saw that Earth's midlatitude circulation was characterised by,

- strong upper-level westerly winds (extending further poleward than the subtropical jet discussed in the previous chapter)
- strong latitudinal temperature gradients
- westerly surface winds
- a thermally indirect meridional overturning (the Ferrel Cell)
- a strong upper-tropospheric maximum in eddy kinetic energy

Our aim in this chapter is to understand why these features exist.

Thermal wind balance

A key concept that relates the first two points above is that of thermal wind balance. We already touched upon this concept in chapters 2 and 3, but here we derive it in a simpler form. Neglecting the time derivative, advection, metric terms, and viscous forces from the meridional momentum equation (2.51b), we have geostrophic balance,

$$\frac{1}{R_e}\frac{\partial \Phi}{\partial \phi} = -2\Omega \sin\phi u. \tag{5.1}$$

Taking the vertical (pressure) derivative and using hydrostatic balance, we have,

$$\frac{1}{R_e}\frac{\partial \alpha}{\partial \phi} = 2\Omega \sin\phi \frac{\partial u}{\partial p}. \tag{5.2}$$

Finally, using the ideal gas law, we may write,

$$\frac{R_d}{pR_e}\frac{\partial T}{\partial \phi} = 2\Omega \sin\phi \frac{\partial u}{\partial p}. \tag{5.3}$$

Or, in more familiar Cartesian notation,

$$\frac{R_d}{p}\frac{\partial T}{\partial y} = f\frac{\partial u}{\partial p}. \tag{5.4}$$

The thermal wind equation relates meridional temperature gradients to vertical shear in the zonal wind. In particular, where the temperature decreases rapidly toward the pole, we expect westerly wind shear. To the extent that the surface winds are constrained by friction to remain relatively weak, this implies that where temperature decreases strongly toward the pole in the troposphere, there will be a westerly jet aloft.

The thermal wind relation therefore shows that the first two points above are strongly related. However, it does not provide any insight as to *why* there should be either strong temperature gradients or strong wind shear. This is particularly true when one examines snapshots of the wind distribution rather than time means (Fig. 1.8), which reveals the existence of very localised upper-tropospheric wind maxima, associated with regions of sharp temperature gradients. Such features have have no obvious cause in the relatively smoothly varying solar forcing, and they must be produced by the circulation itself. It turns out that the angular momentum budget can provide considerable insight into the relevant dynamics.

5.2 Angular momentum of Earth's atmosphere

5.2.1 The angular-momentum budget

As discussed in the previous chapter, angular momentum is conserved for axisymmetric, frictionless flow in the atmosphere. More specifically, we can use the primitive zonal-

momentum equation in pressure coordinates (2.51a) to write an equation for angular momentum valid under general conditions given by,

$$\frac{\mathrm{D}M}{\mathrm{D}t} = -\frac{\partial \Phi}{\partial \lambda} + R_e \cos\phi \mathcal{F}_{\nu\lambda}, \tag{5.5}$$

where $M = R_e \cos\phi(\Omega R_e \cos\phi + u)$ is the angular momentum. It will also prove useful to express this in terms of the relative angular momentum $M_r = R_e \cos\phi u$ and the planetary angular momentum $M_p = \Omega R_e^2 \cos^2\phi$, which are governed by the equations,

$$\frac{\mathrm{D}M_r}{\mathrm{D}t} = -\frac{\partial \Phi}{\partial \lambda} + R_e \cos\phi(fv + \mathcal{F}_{\nu\lambda}), \tag{5.6a}$$

$$\frac{\mathrm{D}M_p}{\mathrm{D}t} = -R_e \cos\phi(fv), \tag{5.6b}$$

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

Also, recall from chapter 2 that we can rewrite Lagrangian conservation equations like those above in flux form. In particular, we may write the flux form of the relative angular momentum equation as,

$$\frac{\partial M_r}{\partial t} + \frac{1}{R_e \cos\phi}\left\{\frac{\partial (uM_r)}{\partial \lambda} + \frac{\partial (vM_r \cos\phi)}{\partial \phi}\right\} + \frac{\partial (\omega M_r)}{\partial p} = -\frac{\partial \Phi}{\partial \lambda} + R_e \cos\phi(fv + \mathcal{F}_{\nu\lambda}). \tag{5.7}$$

where we have used the continuity equation in pressure coordinates,

$$\frac{1}{R_r \cos\phi}\left\{\frac{\partial u}{\partial \lambda} + \frac{\partial (v\cos\phi)}{\partial \phi}\right\} + \frac{\partial \omega}{\partial p} = 0. \tag{5.8}$$

5.2.2 Angular-momentum budget of the upper troposphere

Let us first consider the angular momentum budget in the upper troposphere. In this region, frictional torques are weak, and so we may neglect $\mathcal{F}_{\nu\lambda}$. Taking the zonal and time mean of (5.7), we have,

$$\frac{1}{R_e \cos\phi}\frac{\partial [\overline{vM_r}]\cos\phi}{\partial \phi} + \frac{\partial [\overline{\omega M_r}]}{\partial p} = R_e \cos\phi f[\overline{v}]. \tag{5.9}$$

Here, the time average of a time derivative is assumed to be zero, and a zonal average of a zonal derivative is zero by periodicity. Using the definition of M_r, and rearranging,

$$f[\overline{v}] = \frac{1}{R_e \cos^2\phi}\frac{\partial [\overline{uv}]\cos^2\phi}{\partial \phi} + \frac{\partial [\overline{u\omega}]}{\partial p}. \tag{5.10}$$

This gives an equation for the mean meridional flow in the upper troposphere. The right hand side is proportional to the divergence of the flux of relative angular momentum out

of a given latitude band. Incredibly, considering the zonal (angular) momentum budget has given us a constraint on the meridional mean flow!

To make sense of the above equation, we can divide the transport terms on the right hand side into mean and eddy components. Recall that for any variable q, we may write its time- and zonal-mean latitudinal flux $[\overline{vq}]$ as the sum of three terms,

$$[\overline{vq}] = [\overline{v}][\overline{q}] + [\overline{v}^*\overline{q}^*] + [\overline{v'q'}], \tag{5.11}$$

with a similar equation applying for the vertical flux $[\overline{\omega q}]$. Here, the terms on the right-hand side of the above equation give the mean flux, the stationary eddy flux, and the transient eddy flux, respectively. For brevity, we will define a "total" deviation $q^\dagger = q - [\overline{q}]$, so that we may write the sum of the transient and stationary eddies as,

$$[\overline{v}^*\overline{q}^*] + [\overline{v'q'}] = [\overline{v^\dagger q^\dagger}]. \tag{5.12}$$

Applying this decomposition to (5.10), we have,

$$f[\overline{v}] = \frac{1}{R_e \cos^2\phi}\frac{\partial [\overline{u}][\overline{v}]\cos^2\phi}{\partial\phi} + \frac{\partial [\overline{u}][\overline{\omega}]}{\partial p} + \frac{1}{R_e\cos^2\phi}\frac{\partial [\overline{u^\dagger v^\dagger}]\cos^2\phi}{\partial\phi} + \frac{\partial [\overline{u^\dagger\omega^\dagger}]}{\partial p}. \tag{5.13}$$

With the help of the zonal- and time-mean continuity equation,

$$\frac{1}{R_r\cos\phi}\frac{\partial [\overline{v}]\cos\phi}{\partial\phi} + \frac{\partial [\overline{\omega}]}{\partial p} = 0, \tag{5.14}$$

we may transform (5.13) back into advective form,

$$f[\overline{v}] - \frac{[v]}{R_e\cos\phi}\frac{\partial [\overline{u}]\cos\phi}{\partial\phi} - \omega\frac{\partial [\overline{u}]}{\partial p} = \frac{1}{R_e\cos^2\phi}\frac{\partial [\overline{u^\dagger v^\dagger}]\cos^2\phi}{\partial\phi} + \frac{\partial [\overline{u^\dagger\omega^\dagger}]}{\partial p}. \tag{5.15}$$

Defining the vorticity,

$$\zeta = \frac{1}{R_e\cos\phi}\left\{\frac{\partial v}{\partial\lambda} - \frac{\partial u\cos\phi}{\partial\phi}\right\}, \tag{5.16}$$

and taking its zonal mean, we may write (5.15) as

$$(f + [\overline{\zeta}])[\overline{v}] - \omega\frac{\partial [\overline{u}]}{\partial p} = \frac{1}{R_e\cos^2\phi}\frac{\partial [\overline{u^\dagger v^\dagger}]\cos^2\phi}{\partial\phi} + \frac{\partial [\overline{u^\dagger\omega^\dagger}]}{\partial p}. \tag{5.17}$$

In the extratropics, the large-scale flow tends to be close to geostrophic balance[1], and this may be shown to imply that (i) the advection of angular momentum is dominated by its

[1] In particular, the flow is in what is known as the *quasigeostrophic* regime, implying that the Rossby number, equal to the ratio of the Coriolis acceleration to advection, is small, and the Burger number, equal to the ratio of the Rossby radius the horizontal scale of the motions in question is of order unity or smaller.

horizontal component, and (ii) $|\zeta| << |f|$. Outside the deep tropics, (5.17) therefore takes a particularly simple form,

$$f[\overline{v}] = -\mathcal{S}, \tag{5.18}$$

where $\mathcal{S}$ is known as the eddy momentum flux convergence[2],

$$\mathcal{S} = -\frac{1}{R_e \cos^2\phi}\frac{\partial[\overline{u^\dagger v^\dagger}]\cos^2\phi}{\partial\phi}. \tag{5.19}$$

Equation (5.18) relates the time- and zonal-mean meridional flow directly to the eddy fluxes of angular momentum. In regions where there is a convergence of angular momentum by eddies, the flow is equatorward and in regions where there is an eddy momentum flux divergence, there is poleward flow.

Observational estimates of the angular-momentum fluxes are consistent with this picture (Fig. 5.1). Outside the deep tropics, the angular momentum fluxes are dominated by the transient eddy component, and, in some regions of the northern hemisphere extratropics, the stationary eddy component. The angular-momentum fluxes are dominated by a large-scale poleward transport from the subtropics to the subpolar regions in the upper troposphere of each hemisphere, and a weaker transport from the polar regions towards the equator. As a result, the eddy momentum flux divergence $\mathcal{S}$ has a tripole structure, with maxima of divergence in the subtropics and polar latitudes, and a convergence maximum in between.

The tripole strcuture in $\mathcal{S}$ is closely associated with the sign of the upper-tropospheric meridional flow (i.e., the upper branches of the Hadley, Ferrel and Polar Cells), as expected from (5.18). The thermally indirect Ferrel Cell may therefore be seen to be associated with convergence of angular momentum by midlatitude eddies. Indeed, Fig. 5.1 indicates that such eddies contribute to the upper-tropospheric poleward flow of the Hadley Cell; we will discuss the influence of eddy processes on the Hadley Cell further in Chapter 8.

The above argument relates eddy fluxes of angular-momentum to the meridional flow in the upper troposphere. But what about the lower troposphere? As shown in Fig. 5.1, the relationship between the eddy angular momentum fluxes and the overturning appears to be rather different in the lower troposphere. To see why this is, consider again (5.18) and integrate from the surface to the top of the atmosphere with mass weighting,

$$f\int_0^{p_s}[\overline{v}]\frac{\mathrm{d}p}{g} = -\int_0^{p_s}\mathcal{S}\frac{\mathrm{d}p}{g}. \tag{5.20}$$

The right-hand side corresponds to the divergence of the eddy angular momentum flux within the column, while the left-hand side is the total meridional mass flux within the

[2]Strictly speaking $R_e\mathcal{S}$ is the horizontal eddy angular momentum flux covergence, but $\mathcal{S}$ is usually referred to as the eddy momentum flux convergence in the literature.

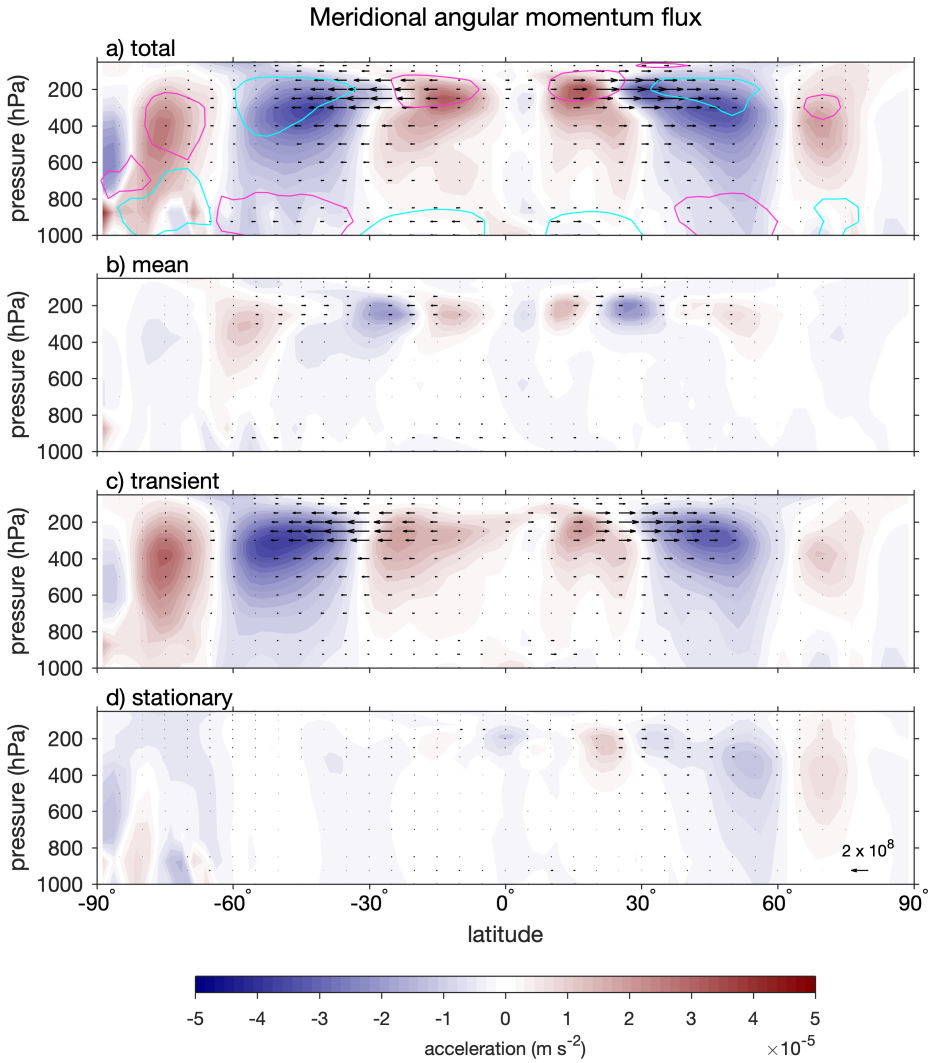

Figure 5.1: Zonal- and time-mean meridional relative angular momentum flux $[\overline{vM_r}]$ (arrows; $\mathrm{m}^3\ \mathrm{s}^{-2}$) and the acceleration associated with the divergence of this flux (colors) broken down into components associated with (a) total flow, (b) mean flow, (c) transient eddies, and (d) stationary eddies according to the NCEP-DOE reanalysis. Contours in panel (a) show regions in which the Coriolis acceleration $f[\overline{v}]$ is greater than 2×10^{-5} m s^{-2} (pink) and less than -2×10^{-5} m s^{-2} (cyan).

column. In steady-state, one cannot maintain a net mass flux across a latitude circle, and the left-hand side must be zero. But there is no such restriction on the right-hand side, and casual inspection of Fig. 5.1 shows that the right-hand side is not zero in the extratropics.

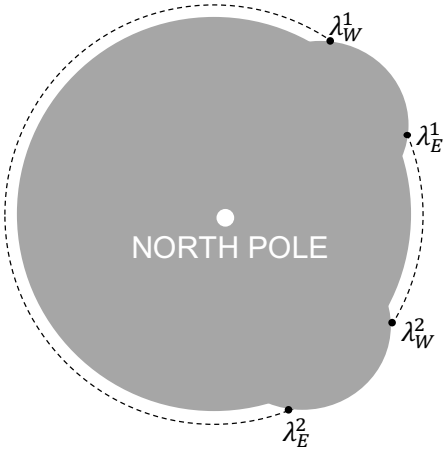

Figure 5.2: Isobar (dashed) that intersects the surface of the Earth (grey) as viewed from directly above the North Pole. Intersections of the isobar and the surface at the western slopes (λ_W) and eastern slopes (λ_E) are also shown. Adapted from Peixoto and Oort (1992).

What is going on here?

Clearly, our assumptions are breaking down, and (5.18) is not valid throughout the column. The key physical process we are missing is drag on the atmosphere by the Earth. This drag allows for a different balance of the angular momentum budget in the planetary boundary layer, and we shall see that, near to the surface, we cannot neglect frictional torques and nor can we neglect the geopotential gradient term as was done to derive (5.18).

5.3 Drag and the angular-momentum budget

Let us return to the un-approximated angular momentum budget (5.7). Taking the time and zonal mean and assuming a steady state, we have,

$$\frac{1}{R_e \cos\phi}\left\{\left[\frac{\partial \overline{uM_r}}{\partial \lambda}\right] + \frac{\partial [\overline{vM_r}]\cos\phi}{\partial \phi}\right\} + \frac{\partial [\overline{\omega M_r}]}{\partial p} = -\left[\frac{\overline{\partial \Phi}}{\partial \lambda}\right] + R_e \cos\phi \left(f[\overline{v}] + [\overline{\mathcal{F}_{\nu\lambda}}]\right). \tag{5.21}$$

Ordinarily, we would be justified in assuming that the zonal mean of a zonal derivative is zero. However, near the surface, we must be careful, because if the isobar that we are considering intersects the surface, the term in question will not be defined at all longitudes (Fig. 5.2). Under such conditions, the zonal mean must be evaluated as a series of integrals over all regions for which $p < p_s$. For example, if we imagine a given isobar intersecting

the surface at say, $\lambda = \lambda_W$, and reemerging at $\lambda = \lambda_E$, we would have that,

$$\begin{aligned}\left[\frac{\partial a}{\partial \lambda}\right] &= \frac{1}{2\pi}\int_{\lambda_E}^{2\pi} \frac{\partial a}{\partial \lambda}\,\mathrm{d}\lambda + \frac{1}{2\pi}\int_{0}^{\lambda_W} \frac{\partial a}{\partial \lambda}\,\mathrm{d}\lambda \\ &= a(\lambda = \lambda_W) - a(\lambda = \lambda_E).\end{aligned}$$

where we have used that $a(\lambda = 2\pi) = a(\lambda = 0)$. If there is more than one region in which $p > p_s$, then the integral must be broken into multiple pieces. But it is easy to see that the procedure above will lead to a contribution at each western outcrop of the pressure surface and a negative contribution from each eastern outcrop. Therefore, we may write in general,

$$\left[\frac{\partial a}{\partial \lambda}\right] = \sum_i a(\lambda = \lambda_W^i) - a(\lambda = \lambda_E^i), \tag{5.22}$$

for all regions i in which $p > p_s$.

Applying this to (5.21), we note that $\mathbf{u} = 0$ at the surface, and so the zonal derivative on the left-hand side does not contribute to the zonal average. However, the contribution of the geopotential gradient on the right-hand side is in general nonzero, and we have,

$$\frac{1}{R_e \cos\phi}\frac{\partial [\overline{vM_r}]\cos\phi}{\partial \phi} + \frac{\partial [\overline{\omega M_r}]}{\partial p} = \sum_i \overline{\Phi_E^i - \Phi_W^i} + R_e \cos\phi \left(f[\overline{v}] + [\overline{\mathcal{F}_{\nu\lambda}}]\right). \tag{5.23}$$

Integrating vertically over the atmospheric column, we have,

$$\frac{1}{R_e \cos\phi}\frac{\partial}{\partial \phi}\int_0^{p_s} [\overline{vM_r}]\cos\phi\,\mathrm{d}p = \int_0^{p_s}\sum_i \overline{\Phi_E^i - \Phi_W^i}\,\mathrm{d}p + \int_0^{p_s} R_e\cos\phi[\overline{\mathcal{F}_{\nu\lambda}}]\,\mathrm{d}p\,. \tag{5.24}$$

where we have assumed u and v are zero at the surface, and $\int_0^{p_s}[\overline{v}]\,\mathrm{d}p$ is zero by mass conservation.

The above equation is the more general counterpart of (5.20). It states that, in steady state, the divergence of the flux of angular momentum out of any latitude band must be equal to the two terms on the right-hand side. These two terms are both angular momentum transports into the Earth associated with, from left to right, form drag and frictional drag,

$$\frac{1}{R_e \cos\phi}\frac{\partial}{\partial \phi}\int_0^{p_s} [\overline{vM_r}]\cos\phi\,\mathrm{d}p = D_{\text{form}} + D_{\text{friction}}. \tag{5.25}$$

We discuss each of these angular momentum transports below.

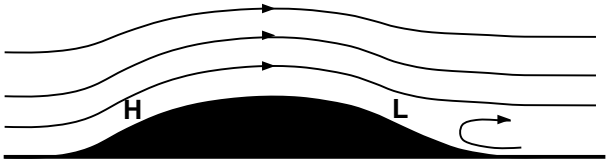

Figure 5.3: Schematic showing streamlines of flow over a mountain. The surface pressure on the upstream slope is higher than that of the downstream slope. This results on a net force on the mountain in the downstream direction. By Newton's Third Law, this is associated with a net force, termed form drag, on the atmosphere toward the upstream direction.

5.3.1 Surface friction

To evaluate the surface friction term, we recall that the surface friction per unit mass $\mathcal{F}_{\nu\lambda}$ may be written more generally using the viscous stress tensor,

$$\begin{aligned}\mathcal{F}_{\nu\lambda} &= \frac{1}{\rho}\nabla\cdot\mathbb{P}\cdot\hat{\boldsymbol{\lambda}} \\ &= -\frac{1}{\rho R_e\cos\phi}\left\{\frac{\partial P_{\lambda\lambda}}{\partial\lambda}+\frac{\partial P_{\lambda\phi}\cos\phi}{\partial\phi}\right\}-\frac{1}{\rho}\frac{\partial P_{\lambda z}}{\partial z}.\end{aligned}$$

At large-scales in the atmosphere, friction is dominated by vertically oriented stresses, and we may therefore write,

$$\mathcal{F}_{\nu\lambda} = g\frac{\partial P_{\lambda z}}{\partial p} \tag{5.26}$$

where we have used hydrostatic balance to convert to pressure coordinates. Integrating vertically and taking a time and zonal mean, we therefore have that,

$$D_{\text{friction}} = gR_e\cos\phi[\overline{P_s}] \tag{5.27}$$

where P_s is the frictional stress at the surface.

5.3.2 Form drag

To express the form drag term more simply, it is useful to consider this term as it appears in (5.21). Taking the vertical integral, we have,

$$D_{\text{form}} = -\int_0^{p_s}\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial\Phi}{\partial\lambda}\mathcal{H}(p_s-p)\,\mathrm{d}\lambda\,\mathrm{d}p \tag{5.28}$$

where $\mathcal{H}$ is the Heaviside step function, equal to unity when its argument is positive and zero otherwise, and we have omitted the time average for clarity. Interchanging the order of integration, we see that $\mathcal{H}(p_s-p)$ is equal to unity over the interval containing the integral in pressure, and we may therefore omit it,

$$D_{\text{form}} = -\frac{1}{2\pi}\int_0^{2\pi}\int_0^{p_s}\frac{\partial\Phi}{\partial\lambda}\,\mathrm{d}\lambda\,\mathrm{d}p \tag{5.29}$$

Using the Leibniz integral rule[3], we have,

$$D_{\text{form}} = -\frac{1}{2\pi}\int_0^{2\pi}\frac{\partial}{\partial\lambda}\left\{\int_0^{p_s}\Phi\,\mathrm{d}p\right\} - \Phi_s\frac{\partial p_s}{\partial\lambda}\,\mathrm{d}\lambda\,. \tag{5.30}$$

where Φ_s is the surface geopotential. The first term in the above equation is zero by periodicity, so that we may write,

$$D_{\text{form}} = \left[\Phi_s\frac{\partial p_s}{\partial\lambda}\right] \tag{5.31}$$

Finally, using the product rule, we may write this in the slightly more suggestive form,

$$D_{\text{form}} = -\left[p_s\frac{\partial\Phi_s}{\partial\lambda}\right] \tag{5.32}$$

From the above discussion, we may write the angular momentum budget for a latitude band as,

$$\frac{1}{R_e\cos\phi}\frac{\partial}{\partial\phi}\int_0^{p_s}[\overline{vM_r}]\cos\phi\,\mathrm{d}p = gR_e\cos\phi[\overline{P_s}] - \left[p_s\frac{\partial\Phi_s}{\partial\lambda}\right]. \tag{5.33}$$

This equation states that there is a balance between the net flux of angular momentum into a latitude band and its removal at the surface from drag. This drag is of two types.

The first term on the right-hand side corresponds to frictional drag and is proportional to the frictional stress at the surface. It represents a viscous force between the Earth's surface and the air immediately above, and it depends on the difference in velocity between the surface and the atmosphere immediately above. In general, we would expect the action of viscous stresses to tend to reduce the surface wind.

The second term on the right-hand side corresponds to form drag between the surface and atmosphere. Form drag is related to the difference in pressure between the western slopes and eastern slopes of the topography in a given latitude circle. If the pressure is higher on the western slopes, the atmosphere is pushing the Earth in the direction of its rotation (eastward), and, by Newton's Third Law, the Earth is imparting westward angular momentum to the atmosphere. In general, one would expect the pressure to be higher on the windward side of a mountain than the leeward side (Fig. 5.3), and form drag therefore tends to act against the near surface wind.

5.3.3 Angular momentum budget and the midlatitude westerlies

As already hinted at above, drag between the atmosphere and surface is a function of the relative motion between these two bodies, that is, it is a function of the surface wind.

[3]This rule states that $\frac{\mathrm{d}}{\mathrm{d}t}\int_0^{b(x)} f(x,t)\,\mathrm{d}t = f(x,b(x))b'(x) + \int_0^{b(x)} f'(x,t)\,\mathrm{d}t$.

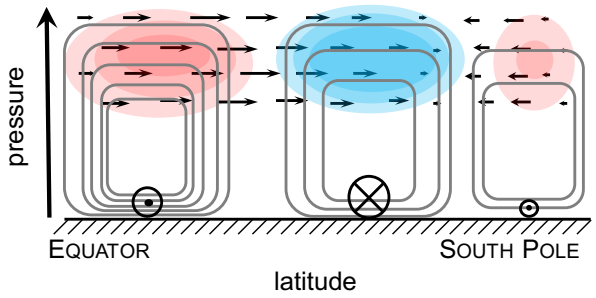

Figure 5.4: Latitude-pressure schematic of the extratropical angular momentum budget and its influence on the circulation. Arrows represent angular momentum fluxes and colours represent contours of the angular momentum flux divergence (red) and convergence (blue). Grey lines represent the overturning streamfunction, and $\otimes$ and $\odot$ symbols represent surface winds into and out of the page, respectively.

Assuming the viscous stress tends to reduce this relative motion, then, in regions where the surface wind is westerly, the Earth is exerting an eastward frictional torque on the atmosphere, and the atmosphere is exerting a westward frictional torque on the Earth. There is therefore a transport of angular momentum from the atmosphere to the surface. Similarly, when the background flow is westerly, one would expect the surface pressure to be higher on the western slopes of topography than on the eastern slopes, also implying a westward acceleration of the atmosphere and a transport of angular momentum from the atmosphere to the Earth.

The above discussion suggests that westerly winds can only be maintained against drag in the presence of a flux of angular momentum from the atmosphere to the surface. By (5.33), this requires a convergence of angular momentum within the atmosphere. We have thus constructed an explanation for the pattern of surface winds plotted in 1.7: where there is a net convergence of angular momentum in the atmospheric column, there will be surface westerlies. Where there is a net divergence, there will be easterlies.

Combining the results of this section, we have a powerful set of relationships tying the general circulation in the extratropics to the angular momentum budget (Fig. 5.4). Regions of convergence of angular momentum (midlatitudes) correspond to equatorward zonal-mean flow in the free troposphere. This implies an overturning circulation (the Ferrel Cell) that can only be closed in the boundary layer, where frictional torques are able to balance the Coriolis acceleration. The implied frictional torques require zonal-mean surface westerlies in regions of angular momentum convergence and easterlies elsewhere. The atmospheric transport of angular momentum producing these circulations is primarily accomplished by eddies, both transient, and to a lesser extent, stationary. To understand the Ferrel Cell, and the pattern of surface winds we must therefore understand the mechanisms that

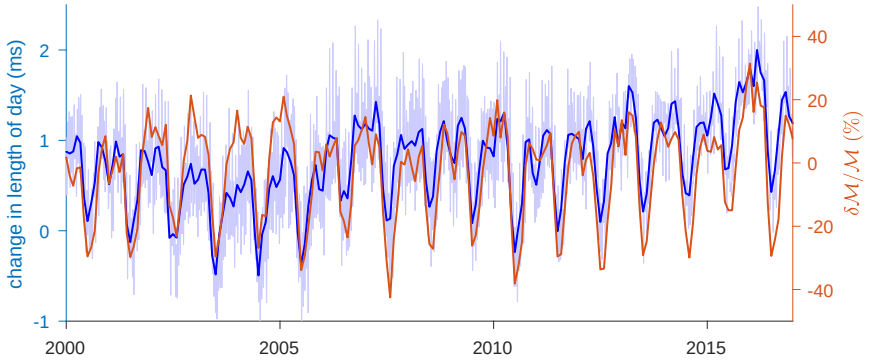

Figure 5.5: Monthly fractional variations in relative angular momentum $\mathcal{M}$ estimated from the NCEP-DOE reanalysis (red; right axis), and daily (light blue) and monthly (dark blue) variations in the length of the day (relative to 86 400 seconds) taken from the International Earth Rotation and Reference Systems Service (IERS) Earth Orientation Parameter C01 dataset (left axis).

produce these angular momentum transports. This will be the subject of the next two chapters.

5.4 The global angular momentum cycle

To finish this chapter, we will consider the angular momentum budget of the entire global atmosphere. Taking (5.7), integrating over the entire globe and following a procedure similar to that above, we may derive the global angular momentum budget,

$$\frac{\mathrm{d}\mathcal{M}}{\mathrm{d}t} = \oint_{\text{globe}} R_e \cos\phi \left\{ gP_s - \frac{p_s}{R_e \cos\phi} \frac{\partial \Phi_s}{\partial \lambda} \right\} \mathrm{d}S \,. \tag{5.34}$$

where

$$\mathcal{M} = \oint_{\text{atmosphere}} M_r \, \mathrm{d}V \,. \tag{5.35}$$

Equation (5.34) simply expresses a balance between the rate of change of angular momentum of the atmosphere and the frictional torques that transfer angular momentum between the Earth and the atmosphere. In particular, the global angular momentum budget states that, in steady state, the net frictional torque on the atmosphere must be zero. This implies, for example, that the surface winds over the Earth cannot be everywhere westerly or easterly. Moreover, if we assume that the sum of frictional drag and form drag are linearly related to the surface winds, the mean surface wind, averaged over the globe, must be zero. In reality, the frictional torques on the atmosphere are not a linear function of the winds, but nevertheless, the steady-state angular momentum budget implies that regions

of westerly winds must be roughly compensated by regions of easterlies over the Earth's surface.

In actual fact, the atmospheric angular momentum budget is not in perfect balance, but it is exchanging angular momentum with the solid Earth. This exchange is expressed by a speeding up or slowing down of the Earth's rotation that causes a change in the length of the day of a few milliseconds–large enough to be measurable. Such exchanges of angular momentum between the Earth and atmosphere occur on timescales from days, to seasonal and interannual timescales (associated with, e.g., the El Niño-Southern Oscillation). Indeed, estimates of the length of day and atmospheric angular momentum are highly correlated on timescales from months to years (Fig. 5.5).

Note that, since the northern hemisphere has the larger seasonal cycle in winds and temperature, the angular momentum of the atmosphere peaks in Boreal winter, when the northern hemisphere jet is strongest. This means that, for citizens of the southern hemisphere, not only are there more daylight hours in summer, the days themselves are slightly longer!

The length of day changes associated with exchanges of angular momentum between the Earth and the atmosphere occur in addition to the much smaller secular change in the length of the day that is a result of the Moon's tidal pull on the Earth.

Chapter 6

The maintenance of a barotropic jet

In the previous chapter we highlighted the importance of the angular momentum budget for understanding relationships between the eddies and the mean flow in the extratropical atmosphere. In particular, we described how the pattern of eddy momentum fluxes is a key determinant of the meridional overturning circulation and zonal-mean surface winds. But what determines this pattern of eddy fluxes? This is the question we attempt to answer in this chapter. To do so, we consider, rather than angular momentum flux, the closely related quantity of the vorticity flux.

6.1 Vorticity and circulation

The vorticity $\boldsymbol{\zeta}$ of the flow is a vector quantity defined as the curl of the velocity,

$$\boldsymbol{\zeta} = \nabla \times \mathbf{u} \tag{6.1}$$

In geophysical fluid dynamics, the component of the vorticity in the local vertical direction is of particular interest, which we simply refer to as ζ,

$$\zeta = \boldsymbol{\zeta} \cdot \hat{\mathbf{z}} = \frac{1}{R_e \cos\phi} \left\{ \frac{\partial v}{\partial \lambda} - \frac{\partial u \cos\phi}{\partial \phi} \right\} \tag{6.2}$$

6.1.1 The vorticity equation

We may derive a governing equation for vorticity (called the vorticity equation) using the u- and v-momentum equations (2.26a) & (2.26b). For simplicity, we will initially consider

an inviscid, homogenous fluid (in which the density $\rho = \rho_0$ is constant). Under these conditions, the momentum equations may be written,

$$\frac{\mathrm{D}u}{\mathrm{D}t} = fv + \frac{uv}{R_e}\tan\phi - \frac{1}{\rho_0 R_e \cos\phi}\frac{\partial p}{\partial \lambda}, \tag{6.3}$$

$$\frac{\mathrm{D}v}{\mathrm{D}t} = -fu - \frac{u^2}{R_e}\tan\phi - \frac{1}{\rho_0 R_e}\frac{\partial p}{\partial \phi}, \tag{6.4}$$

$$\tag{6.5}$$

where $f = 2\Omega\sin\phi$. To derive the vorticity equation, we will first rearrange the momentum equations in a more convenient form.

Consider the meridional flux of vorticity,

$$\begin{aligned} v\zeta &= \frac{v}{R_e\cos\phi}\left\{\frac{\partial v}{\partial\lambda} - \frac{\partial u\cos\phi}{\partial\phi}\right\} \\ &= \frac{1}{R_e\cos\phi}\frac{\partial}{\partial\lambda}\left(\frac{v^2}{2}\right) + \frac{uv\sin\phi}{R_e\cos\phi} - \frac{v}{R_e}\frac{\partial u}{\partial\phi}. \end{aligned}$$

Rearranging the above equation for the meridional advection (the last term on the right-hand side) and substituting this into the defintion of $\frac{\mathrm{D}u}{\mathrm{D}t}$, we may write the Lagrangian derivative of u as,

$$\frac{\mathrm{D}u}{\mathrm{D}t} = \frac{\partial u}{\partial t} + \frac{1}{R_e\cos\phi}\frac{\partial}{\partial\lambda}\left(\frac{u^2+v^2}{2}\right) - v\zeta + \frac{uv\tan\phi}{R_e}. \tag{6.6}$$

Substituting this into the u-momentum equation, we have,

$$\frac{\partial u}{\partial t} - v\left(f+\zeta\right) = -\frac{1}{R_e\cos\phi}\frac{\partial}{\partial\lambda}\left(\frac{u^2+v^2}{2} + \frac{p}{\rho_0}\right). \tag{6.7}$$

By a similar procedure, we have for v,

$$\frac{\partial v}{\partial t} + u\left(f+\zeta\right) = -\frac{1}{R_e}\frac{\partial}{\partial\phi}\left(\frac{u^2+v^2}{2} + \frac{p}{\rho_0}\right). \tag{6.8}$$

Taking

$$\frac{1}{R_e\cos\phi}\left\{\frac{\partial}{\partial\lambda}(6.8) - \frac{\partial}{\partial\phi}\cos\phi(6.7)\right\},$$

we have, with some rearrangement of terms,

$$\frac{\partial\zeta}{\partial t} + \frac{u}{R_e\cos\phi}\frac{\partial}{\partial\lambda}\left(f+\zeta\right) + \frac{v}{R_e}\frac{\partial}{\partial\phi}\left(f+\zeta\right) + \frac{f+\zeta}{R_e\cos\phi}\left\{\frac{\partial u}{\partial\lambda} + \frac{\partial v\cos\phi}{\partial\phi}\right\} = 0. \tag{6.9}$$

Which gives us the vorticity equation for a inviscid, homogenous fluid,

$$\frac{\mathrm{D}(f+\zeta)}{\mathrm{D}t} = -(f+\zeta)\nabla_h \cdot \mathbf{u}_h. \tag{6.10}$$

The quantity $\zeta_a = f + \zeta$ is the absolute vorticity, equal to the sum of the planetary vorticity f and the relative vorticity ζ. Consider a single layer of fluid of depth H. By mass continuity, the flow must be horizontally non divergent $\nabla_h \cdot \mathbf{u}_h = 0$, and the absolute vorticity is conserved,

$$\frac{\mathrm{D}(f+\zeta)}{\mathrm{D}t} = 0. \tag{6.11}$$

Note that, for a fluid at rest, $\zeta_a = f$, which increases monotonically with latitude. Furthermore, at mid and high latitudes, $f >> \zeta$, and thus for any realistic flow we expect $\partial_\phi \zeta > 0$.

6.1.2 Kelvin's circulation theorem

The final piece of machinery we need for our analysis is Kelvin's circulation theorem. Consider the integral of the vorticity over a material volume V_t for our single layer, homogenous fluid,

$$\hat{\Gamma}(t) = \oint_{V_t} (f+\zeta)\,\mathrm{d}V \tag{6.12}$$

Since the density is assumed to be constant, the volume of fluid parcels is conserved, and we may apply Reynold's transport theorem (2.19), to move the time derivative inside the integral,

$$\frac{\mathrm{d}\hat{\Gamma}}{\mathrm{d}t} = \oint_{V_t} \frac{\mathrm{D}(f+\zeta)}{\mathrm{D}t}\,\mathrm{d}V = 0, \tag{6.13}$$

where the last equality arises from the vorticity equation. For our fluid of constant depth, we may also write,

$$\hat{\Gamma}(t) = H \oint_{A_t} (f+\zeta)\,\mathrm{d}A, \tag{6.14}$$

where H is the depth of the fluid and A_t is the horizontal area of the material volume. By Stokes' theorem, we then have,

$$\hat{\Gamma}(t) = H \oint_{\partial A_t} \mathbf{u}_i \cdot \mathrm{d}\mathbf{r}, \tag{6.15}$$

where ∂A_t is the boundary of A_t, and $\mathbf{u}_i$ is the velocity in an inertial frame. The quantity $\Gamma = \hat{\Gamma}/H$ is an integral of the velocity along a closed material path of the flow and is known as the circulation. By (6.13), the circulation is conserved for our single layer homogenous fluid.

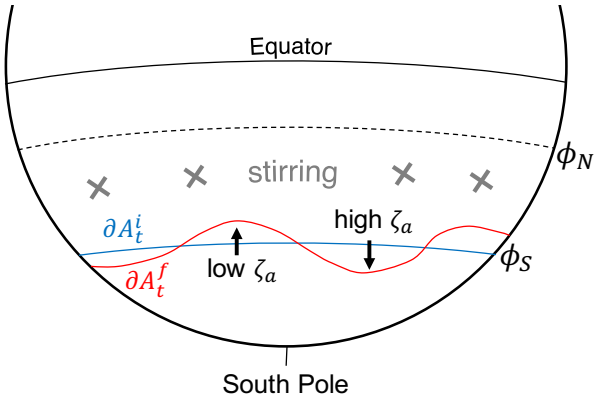

Figure 6.1: Schematic showing material contour ∂A_t that begins as a latitude circle at ϕ_S bounding the polar cap (blue line) and at some later time becomes deformed by disturbances that propagate from the stirred region (red line). The deformation results in a net flux of vorticity into the polar cap, reducing the zonal-mean wind at ϕ_S. A similar argument can be made for a latitude north of the stirred region ϕ_N. Adapted from Vallis (2017).

6.2 Irreversible mixing and vorticity fluxes

Let's apply Kelvin's circulation theorem to a material surface A_t that initially corresponds to a "polar cap" up to a latitude ϕ_S as shown in Fig. 6.1. Since the material path ∂A_t is initially simply a latitude circle, we have that the initial circulation $\Gamma^i_{\phi_S}$ is given by,

$$\begin{aligned}\Gamma^i_{\phi_S} = \int_{\phi<\phi_S} (f+\zeta)\,\mathrm{d}A &= -\int_0^{2\pi} (u_{\phi_S} + \Omega R_e \cos\phi) R_e \cos\phi\,\mathrm{d}\lambda \\ &= 2\pi R_e \cos\phi [u_{\phi_S}] + \Omega R_e^2 \cos^2\phi,\end{aligned}$$

where we have once again used Stokes' theorem. The negative sign comes about because, in the southern hemisphere, the boundary integral is performed east-to-west following the right-hand rule for relating surface integrals to line integrals.

Suppose that the atmosphere is initially at rest, and at some time we begin to stir the fluid at some latitude equatorward of ϕ_S in such a way that no net vorticity is supplied to the fluid. The stirring will create disturbances in the flow which will propagate to the latitude ϕ_S, deforming the material contour ∂A_t. In our non-divergent single-layer fluid, we have that:

1. the total area encompassed by the contour must remain fixed by mass conservation; at some longitudes the contour will move equatorwards of ϕ_S, at others it will move poleward;

2. air parcels conserve their vorticity as they are advected by the flow;

3. at the initial time, the fluid is at rest, and therefore absolute vorticity increases toward the north.

These three properties imply that the deformation of the material contour advects high vorticity air into the polar cap and low vorticity air out of it. That is, if we calculate a new path integral around a latitude circle, $\Gamma^f_{\phi_S}$, we have that,

$$\Gamma^f_{\phi_S} > \Gamma^i_{\phi_S}, \tag{6.16}$$

which directly implies that,

$$[u^f_{\phi_S}] < [u^i]_{\phi_S}. \tag{6.17}$$

That is, the zonal-mean velocity at the latitude ϕ_S has been reduced!

Some time later, we might imagine that the flow returns to its original position reversibly. in this case, the zonal velocity would revert to its initial value. However, if instead we suppose that there some irreversible mixing occurs (e.g., the vorticity contour ∂A_t overturns), we might expect that the vorticity transport is also irreversible, and that there is a net flux of vorticity through the latitude ϕ_S.

In summary, the stirring equatorward of ϕ_S has produced a deceleration of the flow over the polar cap. Nothing in this argument is specific to the latitude ϕ_S. In particular, we could pick a different latitude, ϕ_N that is to the north of the stirred region. Applying Kelvin's circulation theorem to this latitude (Fig. 6.1), and with a careful application of the right-hand rule, one finds that the zonal wind north of the stirred region is also reduced.

We have shown that stirring that produces disturbances that propagate away from their source and then decay irreversibly tend to decelerate the flow in the regions in which they produce irreversible mixing. By conservation of angular momentum, we may further argue that the stirred region itself must experience a westerly acceleration. That is, localised stirring produces westerly winds!

The argument above shows how a jet can form from eddy motions on a rotating planet. Eddies forming in one region produce a momentum flux that converges momentum towards their source. This model may be thought of as analogous to the midlatitude upper troposphere. In our barotropic example, however, these eddies are externally imposed, while in Earth's atmosphere, the eddies are driven by baroclinic instability as a result of the strong temperature gradients present in midlatitudes. In turn these temperature gradients are related to the eddy fluxes in a way that we will attempt to make sense of in chapter 7.

In the meantime, we continue within a barotropic framework, but we analyse in more detail how the momentum fluxes implied by our vorticity-mixing example are actually produced.

6.3 Momentum transport by Rossby waves

6.3.1 Dynamics of Rossby wave momentum transport

We now consider the momentum transport produced by propagating Rossby waves. For simplicity, we stick with the horizontally non-divergent single-layer fluid above. Consider again the vorticity equation, this time adopting the tangent-plane approximation for simplicity,

$$\frac{\partial \zeta}{\partial t} + u\frac{\partial \zeta}{\partial x} + v\left(\frac{\partial \zeta}{\partial y} + \beta\right) = 0, \tag{6.18}$$

where $\beta = \mathrm{d}f\,/\,\mathrm{d}y$ is the meridional gradient of the Coriolis parameter. Our discussion in the previous section was rather general, and we considered the vorticity equation in its full nonlinear form. However, it is useful when examining mechanisms to consider linearised dynamics. To this end, we linearise the above equation about a base state $[\overline{\mathbf{u}}] = ([\overline{u}], 0)$, where we initially assume that the zonal- and time-mean zonal wind is independent of latitude. The linearised vorticity equation may be written,

$$\frac{\partial \zeta^\dagger}{\partial t} + [\overline{u}]\frac{\partial \zeta^\dagger}{\partial x} + \beta v^\dagger = 0, \tag{6.19}$$

where the superscript † refers to deviations from the time and zonal mean. Since the flow is non divergent, we may define a perturbation streamfunction Ψ, where,

$$u^\dagger = -\frac{\partial \Psi}{\partial y},$$
$$v^\dagger = \frac{\partial \Psi}{\partial x},$$

The vorticity may be written in terms of the streamfunction,

$$\zeta^\dagger = \frac{\partial v^\dagger}{\partial x} - \frac{\partial u^\dagger}{\partial y} = \nabla^2\Psi, \tag{6.20}$$

allowing the linearised vorticity equation to be written purely in terms of the streamfunction,

$$\frac{\partial}{\partial t}\nabla^2\Psi + [\overline{u}]\frac{\partial}{\partial x}\nabla^2\Psi + \beta\frac{\partial \Psi}{\partial x} = 0. \tag{6.21}$$

We now search for wave-like solutions to the above equation of the form,

$$\begin{aligned}\Psi(x, y, t) &= A\Re\left\{\exp(i(kx + ly - \omega t))\right\}, \\ &= A\cos(kx + ly - \omega t)\end{aligned} \tag{6.22}$$

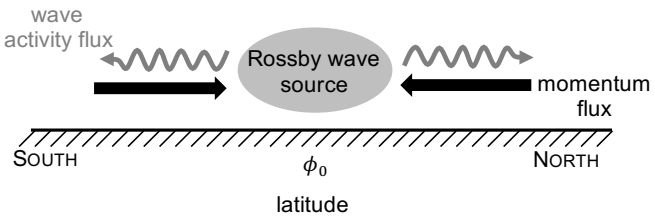

Figure 6.2: Fluxes of momentum (grey arrows) and wave activity (black arrows) produced by a localised source of Rossby waves.

where A is the wave amplitude, k and l are the zonal and meridional wavenumbers, ω is the frequency and $\Re\{\}$ refers to the real component. For this plane wave form, we have that,

$$\begin{aligned} u^\dagger v^\dagger &= A^2 \Re\left\{-il\exp(i(kx+ly-\omega t))\right\}\Re\left\{ik\exp(i(kx+ly-\omega t))\right\} \\ &= -A^2 kl\sin^2(kx+ly-\omega t). \end{aligned}$$

Taking the zonal or time average over one period, we have,

$$[\overline{u^\dagger v^\dagger}] = -\frac{A^2 kl}{2}. \tag{6.23}$$

Thus the momentum flux averaged over the wave is nonzero and depends on the meridional and zonal wavenumbers. Note that we have not used the vorticity equation to derive (6.23), rather, (6.23) is a general property of plane waves described by (6.22).

Now, let us calculate the dispersion relation for waves governed by (6.21). Substituting the plane wave solution into (6.21), we have,

$$-(k^2+l^2)(-\omega + k[\overline{u}]) + k\beta = 0. \tag{6.24}$$

Rearranging for the frequency, we have,

$$\omega = k[\overline{u}] - \frac{k\beta}{(k^2+l^2)}, \tag{6.25}$$

which is the familiar barotopic Rossby wave dispersion relation in a flow with mean zonal wind $[\overline{u}]$. We may also calculate the meridional group velocity c_{gy} as,

$$c_{gy} = \frac{\partial\omega}{\partial l} = kl\left\{\frac{2\beta}{(k^2+l^2)^2}\right\}. \tag{6.26}$$

The group velocity gives the velocity at which energy (actually wave activity) is propagated by a wave. Since the term in curly braces in the above equation is positive definite (β is

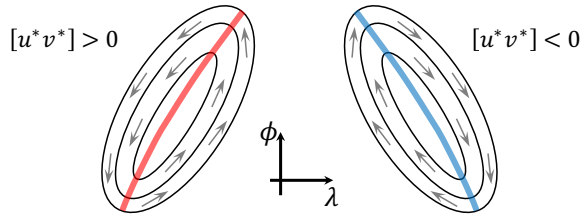

Figure 6.3: The momentum transport associated with eddies that have "tilted troughs/ridges". Black lines represent contours of geopotential anomaly Φ^* and grey arrows represent anomalies of the horizontal winds (u^*, v^*) for idealised eddies in which the trough axes are tilted positively (red) or negatively (blue). Considering the sign of u^* and v^* for each eddy leads to the conclusion that positively tilted trough/ridge axes produce northward eddy momentum flux and negatively tilted trough/ridge axes produce southward eddy momentum flux.

positive everywhere), the group velocity is proportional to, but opposite in sign to the momentum flux. This means that the flux of momentum associated with a Rossby wave is in the opposite direction to the flux of wave activity!

In particular, we might imagine a situation in which a group of Rossby waves form at a latitude ϕ_0 and propagate both northwards and southwards (Fig. 6.2). As wave energy moves away from the Rossby wave source, momentum is converged toward it. As in our nonlinear example above, momentum is converged in the source region of the disturbances! A localised source of Rossby waves will therefore produce a convergence of (angular) momentum toward the source. Such a momentum convergence provides a mechanism for maintaining westerly winds through the depth of the atmosphere, that is, a mechanism to maintain a barotropic jet.

6.3.2 Kinematics of momentum transport

We have shown, using both linearised arguments and a rather general argument relating to vorticity mixing, that localised stirring produces a convergence of momentum within rapidly rotating flows. But how is this momentum flux effected?

Consider an isolated upper tropospheric eddy characterised by a local extremum in the geopotential anomaly from the zonal mean Φ^*. Associated with this height extremum, there is a set of closed contours of Φ^*, and by geostrophic balance, anomalous winds flowing roughly parallel to these contours. What is the net momentum flux by these winds? If the eddy is circular, the symmetry of the problem suggests that there will be no net flux of momentum in any direction. However, if the eddy is elongated to form a trough or ridge axis, the net momentum flux then depends on the tilt of this axis relative to the meridians. To see this, consider the left panel in Fig. 6.3, showing an eddy in which the trough axis is "positively" tilted. West of the trough axis, $u* < 0$ and $v^* < 0$ implying $u^*v^* > 0$. To

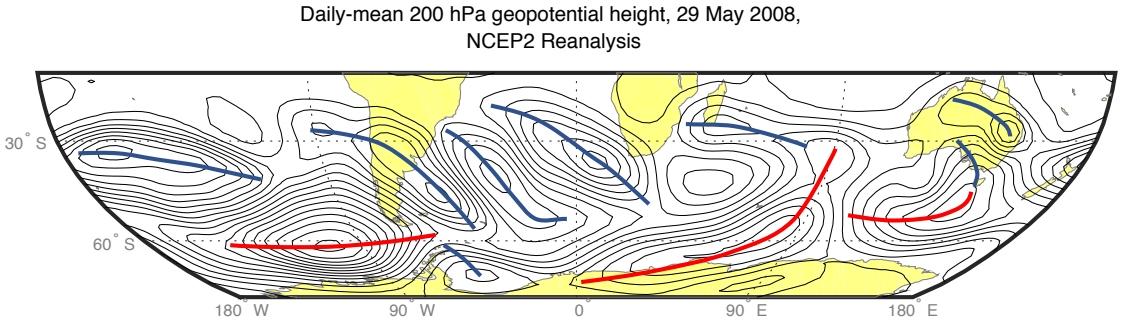

Figure 6.4: Daily-mean 200 hPa height anomaly for 29 May 2008 over the southern hemisphere. Anomaly is taken from the annual and zonal mean and the contour interval is 50 m. Trough and ridge axes with positive (red) and negative (blue) tilt are marked subjectively to guide the eye.

the east of the trough axis, $u^* > 0$ and $v^* > 0$ also implying that $u^*v^* > 0$. Averaged over the eddy, we therefore have that $[u^*v^*] > 0$, there is a net northward flux of momentum by the eddy!

Similar reasoning can be used to show that eddies with negative tilt have a southward associated momentum flux (Fig. 6.3, right panel). Note that this does not depend on the sign of the geopotential height anomaly; whichever direction the winds rotate, it is the orientation of the trough/ridge axis that determines their meridional momentum flux.

Given the observed eddy momentum fluxes in the atmosphere are dominated by poleward transport except in polar regions (Fig. 5.1), one would expect most eddies in the atmosphere to have negative tilted trough/ridge axes, except perhaps near the poles, where positive tilted trough/ridge axes may dominate. Indeed, these tilts are even observed in single daily-mean snapshots of the upper tropospheric geopotential height (Fig. 6.4). Most of the trough/ridge axes identified on Fig. 6.4 are negatively tilted, and those that are positively tilted tend to be poleward of 60°S.

Of particular interest on Fig. 6.4 are the three anomalies of geopotential height over the south Atlantic, extending from Argentina to southern Africa. This pattern is characteristic of a Rossby wave train. Given the northward extension of these anomalies as one moves east, we may identify this wavetrain as having positive k and l, and therefore northward group velocity, and a southward momentum flux. This connects our kinematic discussion with the dynamic one of the previous section; a Rossby wave produces northward momentum fluxes by developing a negative tilt in the trough axes of its associated anomalies.

6.4 Rossby wave propagation

In the previous section we saw that a barotropic jet may be maintained when there exists a localised region of eddy generation from which eddies propagate meridionally, decaying remotely from their source. We now consider what governs the propagation of these eddies, by again appealing to linear Rossby wave dynamics.

Consider again the dispersion relation for barotropic Rossby waves (6.3.1). Solving for the meridional wavenumber l, we have,

$$l^2 = \frac{\beta}{[\overline{u}] - c} - k^2, \tag{6.27}$$

where $c = \omega/k$ is the zonal phase speed. When $l^2 > 0$, l is real, and the vorticity equation admits wave-like solutions in the y-direction. When $l^2 < 0$, l is imaginary, and the solutions are said to be evanescent. Meridional propagation of Rossby waves therefore requires that $l^2 > 0$. A necessary condition for this is that,

$$[\overline{u}] - c > 0. \tag{6.28}$$

This implies that the Doppler-shifted phase speed of the waves must be to the west. Or, put another way, individual troughs and crests of the wave must propagate to the west relative to the mean flow.

The condition (6.28) for meridional Rossby wave propagation was derived based on the assumption that $[\overline{u}]$ is constant in latitude. However, the principal result carries over to the case in which $[\overline{u}]$ varies with latitude. In particular, for slowly varying $[\overline{u}]$, we may construct an approximate solution[1] to (6.21), of the form,

$$\Psi(x, y, t) = A(y) \exp\left\{ i \left(kx + \int l(y) dy - \omega t \right) \right\}, \tag{6.29}$$

where $A(y) \propto l^{-\frac{1}{2}}$ is a slowly varying wave amplitude, and $l(y)$ is a local solution to the Rossby wave dispersion relation,

$$l^2(y) = \frac{\tilde{\beta}(y)}{[\overline{u}](y) - c} - k^2, \tag{6.30}$$

with $\tilde{\beta}(y) = \partial_y(f + [\overline{\zeta}])$.

As $[\overline{u}] - c \to 0$, the above solution predicts that, consistent with the constant $[\overline{u}]$ case, the wave group velocity slows down – the wave cannot propagate any further. At this point, the linearity assumption breaks down, and we expect nonlinear wave breaking.

[1] This approximation is referred to as the WKB method, after the initials of (some of) the mathematicians that developed it. Further details may be found in section 13.2 of Vallis (2017).

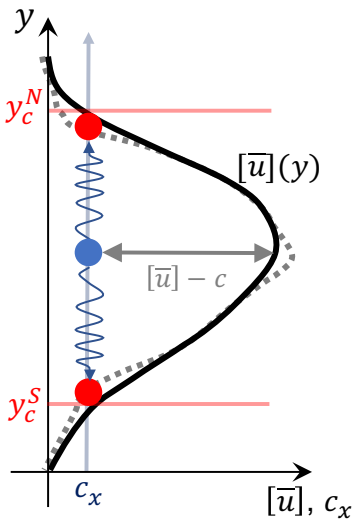

Figure 6.5: Schematic of the propagation of a single Rossby wave packet initiated near the centre of the jet in the phase speed-latitude plane. The wave propagates both north and south, conserving its zonal phase speed c_x until it reaches the critical latitudes y_c^N and y_c^S, where the phase speed equals that of the mean zonal wind, at which point it breaks and is absorbed. Black and grey dashed lines give zonal- and time-mean zonal wind before and after the Rossby wave event; the net effect of the wave is to converge momentum into the jet.

Locations where $l \to \infty$ are known as critical lines, and they correspond to the locations beyond which no Rossby wave propagation is possible. Waves that approach their critical latitude are expected to dissipate, either directly through drag or first through nonlinear wave breaking and eventually via drag.

The conclusion of the above discussion is that Rossby waves propagate meridionally until they approach a latitude at which the wave phase speed c is equal to the background flow $[\overline{u}]$, at which point they break. Given a localised source of waves, this gives us a powerful heuristic for understanding the distribution of wave activity fluxes (and therefore the eddy momentum fluxes) in the atmosphere.

6.4.1 Jet maintenance: a conceptual picture

Interaction between a Rossby wave packet and the jet

Let us consider the implications of (6.28). As we have seen, the midlatitude upper troposphere is characterised by strong eddy motions driven by baroclinic instability (Fig. 1.10) and a strong westerly jet. From the perspective of the upper troposphere, we may therefore consider the midlatitudes as a Rossby wave source near the region of strongest temperature

gradients, which, by thermal wind balance is the centre of the jet. The waves produced by this baroclinic generation will then propagate, both northward and southward, away from the jet core. As they do so, the Rossby wave packets will maintain their zonal phase speed (blue line)[2], but, by virtue of moving away from the jet core, the quantity $[\overline{u}] - c$ will decrease (Fig. 6.5). Eventually, a given wave packet may reach its critical latitude, where $[\overline{u}] = c$ (red lines), at which point it must break and dissipate. The wave activity associated with this wave packet is transported from the centre of the jet to its flanks, while the momentum is transported in the opposite direction – from the flanks to the centre of the jet. The net result of this wave is therefore to accelerate the jet core and decelerate its flanks – i.e., to sharpen the jet!

The above discussion describes the curious tendency of turbulence in rapidly rotating flows to transport momentum up the mean gradient. Historically, this was described as the flow having a negative viscosity (see e.g., Starr, 1968), although that term has largely fallen out of use. It also hints at a mechanism by which Rossby wave propagation can contribute to the maintenance of the jet; the wave-induced sharpening of the jet is, by thermal wind balance, associated with an increase in the meridional temperature gradient. This increase in the temperature gradient may in turn act to strengthen the baroclinic instability that was the original source of the Rossby waves.

While we have appealed to thermal wind balance and baroclinic instability in this section, we have nevertheless largely limited ourselves to a barotropic model of the midlatitude upper troposphere. To fully describe the feedback hinted at above, however, we must incorporate baroclinic processes more fully into our discussion. We do so in the next chapter through the powerful framework of quasigeostrophic theory.

The phase-speed spectrum

Figure 6.5 depicts the interaction of a single Rossby wave packet with the zonal-mean flow. But how well does this describe the complicated spectrum of real, nonlinear disturbances that exist in the atmosphere?

To answer this question, we calculate the phase speed spectrum of eddy momentum fluxes in the upper troposphere $K(c, \phi)$ following Randel and Held (1991) (Fig. 6.6). The integral $\int_{c_1}^{c_2} K(c, \phi) dc$ describes, at each latitude, the component of the covariance $[\overline{u'v'}]$ that is due to disturbances with phase speeds between c_1 and c_2. We may define the phase speed spectrum at any level, but we show the upper troposphere (300 hPa) in Fig. 6.6. The spectrum K is calculated by finding the space-time spectra of $[\overline{u'v'}]$, which divides the covariance into its frequency and wavenumber components, converting the frequency spectrum to one in terms of the zonal phase speed $c = \omega/k$, and integrating over all wavenumbers. Finally, the spectrum is multiplied by $\cos\phi$ so that it is proportional to the

[2] In spherical geometry, the quantity that is conserved is actually the angular phase speed $\frac{c}{\cos\phi}$.

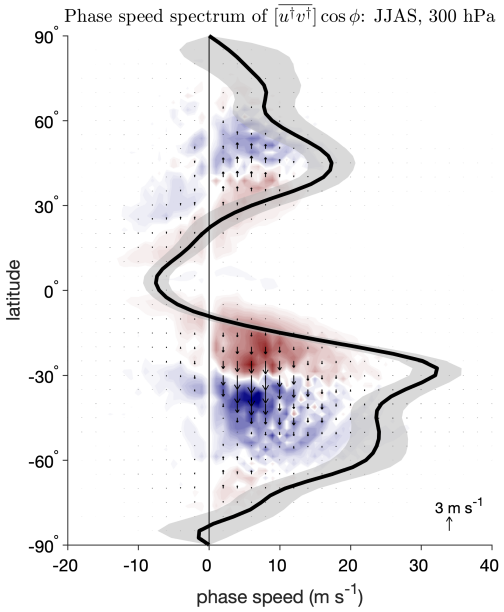

Figure 6.6: Phase speed spectrum of meridional eddy momentum flux at 300 hPa for Austral winter (JJAS) following (Randel and Held, 1991). Arrows give the phase speed spectrum of eddy momentum fluxes, $K(c,\phi)\cos\phi$, proportional to the flux of angular momentum at each latitude owing to eddies with phase speed c, and colours give the divergence of this flux, which is proportional to the torque exerted by eddies on the mean flow [see (5.7)]. Black line denotes the seasonal- and zonal-mean zonal wind at 300 hPa and shading denotes plus and minus one daily standard deviation (shaded).

eddy angular momentum flux.

The phase speed spectrum shows that disturbances responsible for transient eddy momentum fluxes in the atmosphere typically range in phase speed between 0 and 15 m s^{-1}. Such disturbances produce a strong momentum flux convergence at midlatitudes, and divergences either side. This implies that the midlatitudes are a source of wave activity, which then decays in the subtropics and subpolar regions. The time- and zonal-mean zonal wind is plotted on Fig. 6.6, and it marks the critical latitude for a barotropic Rossby wave. Consistent with linear theory, the momentum flux is close to zero for regions in which $[\overline{u}] < c$, and Rossby waves are evanescent, and the regions for which the momentum flux is divergent (corresponding to wave dissipation) are close to the critical latitudes.

The phase speed spectrum shows that the conceptual picture describing a single Rossby wave packet may be usefully applied to the entire spectrum of disturbances in the atmosphere. A range of disturbances with different phase speeds is produced in the core of the jet (which we presume is a result of baroclinic instability). These disturbances then

propagate to north and south in the upper troposphere. As they approach their critical latitudes (whch depends on the mean wind and their phase speed) they begin to break nonlinearly, eventually dissipating. The net result is a flux of momentum toward the source latitudes which accelerates and sharpens the jet. A sharper jet implies stronger meridional temperature gradients, which in turn results in increased baroclinic instability.

At first glance, the above feedback process appears as if it should continue indefinitely, producing frontal collapse at the jet axis. However, our barotropic picture in which eddies transport angular momentum is only half the story; eddies also transport heat in addition to momentum. This heat transport alters the temperature field, which itself influences the winds through thermal wind balance. Making sense of this strong coupling between the thermodynamic and dynamic fields in rapidly rotating atmospheres is the subject of the next chapter.

Chapter 7

Eliassen-Palm fluxes and the transformed Eulerian mean

In the last few chapters, we have focused on the angular-momentum budget and the constraints on the flow one may derive from considerations of meridional fluxes of angular momentum, either through axisymmetric dynamics or more general considerations. However, on large-scales away from the equator, the flow in Earth's atmosphere is known to be close to a state of geostrophic balance; as we have seen in previous chapters such balance implies a strong relationship between dynamic and thermodynamic fields embodied by the thermal wind relation. This implies that we cannot think of the zonal winds as independent of the temperature structure of the atmosphere; these two aspects are strongly coupled. In this chapter we will introduce a powerful set of methods for understanding how this coupled system behaves, and in particular, how eddies influence the mean flow within fluids that are close to a state of geostrophic balance.

We begin by introducing the framework we will use which is known as quasigeostrophic (QG) theory. QG theory is an extremely influential simplification of the governing equations that is commonly used to analyse extratropical dynamics. Entire books have been written on QG theory, and we will present a very brief introduction here; the reader is referred to chapter 5 of Vallis (2017) for further details.

7.1 Quasigeostrophic preliminaries

We begin by considering the equations of motion under the beta-plane approximation and in pressure coordinates,

$$\frac{\partial u}{\partial t} + \mathbf{u_h} \cdot \nabla_p u + \omega \frac{\partial u}{\partial p} = -\frac{\partial \Phi}{\partial x} + fv + \mathcal{F}_x, \tag{7.1a}$$

$$\frac{\partial v}{\partial t} + \mathbf{u_h} \cdot \nabla_p v + \omega \frac{\partial v}{\partial p} = -\frac{\partial \Phi}{\partial y} - fu + \mathcal{F}_y, \tag{7.1b}$$

$$\frac{\partial \theta}{\partial t} + \mathbf{u_h} \cdot \nabla_p \theta + \omega \frac{\partial \theta}{\partial p} = \frac{Q}{c_p \pi}, \tag{7.1c}$$

$$\frac{\partial \Phi}{\partial p} = -\alpha, \tag{7.1d}$$

$$\nabla_p \cdot \mathbf{u}_h + \frac{\partial \omega}{\partial p} = 0. \tag{7.1e}$$

where $f = f_0 + \beta y$. The first two equations above are the horizontal momentum equations, the third gives the thermodynamic equation, the fourth gives hydrostatic balance, and the final equation gives mass continuity.

The quasigeostrophic approximation is valid in fluids that may be considered to be close to a state of geostrophic balance at the appropriate scale. It is therefore useful to define the geostrophic velocity $\mathbf{u}_g = (u_g, v_g, 0)$, where,

$$u_g = -\frac{1}{f_0}\frac{\partial \Phi}{\partial y}, \tag{7.2}$$

$$v_g = \frac{1}{f_0}\frac{\partial \Phi}{\partial x}. \tag{7.3}$$

Note that, since we have used f_0 in the definitions above, the geostrophic wind is non-divergent,

$$\nabla \cdot \mathbf{u}_g = \frac{1}{f_0}\left\{\frac{\partial^2 \Phi}{\partial x \partial y} - \frac{\partial^2 \Phi}{\partial x \partial y}\right\} = 0. \tag{7.4}$$

The quasigeostrophic approximation involves two assumptions,

1. That the velocity is close to its geostrophic value, that is, $|\mathbf{u}_{ag}| << |\mathbf{u}_g|$, where $\mathbf{u}_{ag} = \mathbf{u} - \mathbf{u}_g$. This requires that the Rossby number,

$$\mathrm{Ro} = \frac{U}{f_0 L} << 1, \tag{7.5}$$

where U and L are characteristic horizontal velocity and length scales of the flow.

2. That the atmosphere is strongly stratified. More specifically, we require that the Burger number,

$$\mathrm{Bu} = \frac{NH}{f_0 L} \sim \mathcal{O}(1), \tag{7.6}$$

where H is the scale height of the atmosphere and $N = \frac{g}{\theta}\frac{\partial \theta}{\partial z}$ is the Brunt Väisälä frequency.

Based on these assumptions, the quasigeostrophic system of equations may be derived by a formal asymptotic expansion of (7.1) about the small parameter Ro. That is, we may define a series expansion for the velocity as,

$$\mathbf{u} = \mathbf{u}_0 + \mathrm{Ro}\mathbf{u}_1 + \mathcal{O}(\mathrm{Ro}^2), \tag{7.7}$$

where $\mathbf{u}_0$ is the geostrophic velocity, $\mathbf{u}_1$ is the first order correction, and we neglect higher order terms. Here we do not derive this expansion in detail, and instead we simply present the results of such a series expansion.

The QG zonal velocity equation is found by replacing all velocities in the equation by their geostrophic counterparts, except when multiplied by the Coriolis parameter f_0. That is, we have,

$$\frac{\partial u_g}{\partial t} + \mathbf{u}_g \cdot \nabla_p u_g = f_0 v_{ag} + \beta y v_g + \mathcal{F}_x, \tag{7.8}$$

where geostrophic balance has been used to remove the geopotential gradient. A similar equation is valid for the meridional momentum equation,

$$\frac{\partial v_g}{\partial t} + \mathbf{u}_g \cdot \nabla_p v_g = -f_0 u_{ag} - \beta y u_g + \mathcal{F}_y. \tag{7.9}$$

To derive the QG thermodynamic equation, we define a base state potential temperature $\Theta(p)$, that only depends on pressure, so that,

$$\theta(x, y, p, t) = \Theta(p) + \theta^{\times}(x, y, p, t) \tag{7.10}$$

where $\theta^{\times}$ refers to the deviations of the potential temperature from its base state. Our assumption that the Burger number is order unity implies that the atmospheric stratification is dominated by that of the base state. That is, we have that,

$$\left|\frac{\mathrm{d}\Theta}{\mathrm{d}p}\right| >> \left|\frac{\partial \theta^{\times}}{\partial p}\right|. \tag{7.11}$$

We then neglect ageostrophic advection of the potential temperature anomalies, so that we may write the thermodynamic equation as,

$$\frac{\partial \theta}{\partial t} + \mathbf{u_g} \cdot \nabla_p \theta + \omega\sigma = \frac{Q}{c_p \pi}, \tag{7.12}$$

where $\sigma = \mathrm{d}\Theta \,/\, \mathrm{d}p$.

Since the geostrophic velocity is non divergent, the QG continuity equation may be written,

$$\frac{\partial u_{ag}}{\partial x} + \frac{\partial v_{ag}}{\partial y} + \frac{\partial \omega}{\partial p} = 0. \tag{7.13}$$

Equations (7.8), (7.9), (7.12), (7.13) and hydrostatic balance give us the QG system of equations.

7.2 Forcing of the zonal mean

7.2.1 Zonal-mean QG equations

We now derive a compact set of equations that govern the zonal-mean structure of the atmosphere in the QG system. We first write the QG zonal momentum equation (7.8) in flux form,

$$\frac{\partial u_g}{\partial t} + \frac{\partial u_g^2}{\partial x} + \frac{\partial u_g v_g}{\partial y} = f_0 v_{ag} + \beta y v_g + \mathcal{F}_x, \tag{7.14}$$

where we have used the nondivergence of the geostrophic velocity. Taking the zonal mean of the above equation, we have that,

$$\frac{\partial [u_g]}{\partial t} = f_0 [v] + [\mathcal{F}_x] - \frac{\partial [u_g^* v_g^*]}{\partial y}. \tag{7.15}$$

Here we have divided the velocities into their zonal mean and eddy components, and we have used the fact that the meridional geostrophic velocity is an exact zonal derivative, and that the zonal mean of a zonal derivative is zero to note that $[v_g] = \frac{1}{f_0}[\frac{\partial \Phi}{\partial x}] = 0$ and $\left[\frac{1}{2}\frac{\partial u_g^2}{\partial x}\right] = 0$.

We may apply a similar procedure to the thermodynamic equation (7.12) to give,

$$\frac{\partial [\theta]}{\partial t} + \sigma [\omega] = \frac{[Q]}{c_p \pi} - \frac{\partial [v_g^* \theta^*]}{\partial y}. \tag{7.16}$$

The zonal mean continuity equation may be derived straightforwardly from (7.13) to give,

$$\frac{\partial [v]}{\partial y} + \frac{\partial [\omega]}{\partial p} = 0, \tag{7.17}$$

noting that $[v] = [v_{ag}]$.

Finally, we may connect the zonal wind u_g to the potential temperature deviation θ by deriving the thermal wind relation for this system. Consider the equation for hydrostatic balance,

$$\frac{\partial \Phi}{\partial p} = -\alpha \tag{7.18}$$

Using the ideal gas law and the definition of potential temperature, we have,

$$\frac{\partial \Phi}{\partial p} = -\frac{R_d \theta \pi}{p} \tag{7.19}$$

Taking the y-derivative of the above equation and substituting the definition of the zonal geostrophic velocity, we have,

$$\frac{\partial u_g}{\partial p} = \frac{R_d \pi}{f_0 p} \frac{\partial \theta}{\partial y} \tag{7.20}$$

Taking the zonal mean, we have,

$$\frac{\partial [u_g]}{\partial p} = \frac{R_d \pi}{f_0 p} \frac{\partial [\theta]}{\partial y} \tag{7.21}$$

Equations (7.15), (7.16), (7.17) and (7.21) gives a set of four equations in the four unknown zonal mean quantities $[u_g]$, $[\theta]$, $[v]$ and $[\omega]$. These equations may be solved given the eddy forcing terms $[u_g^* v_g^*]$ and $[v_g^* \theta^*]$, the zonal-mean frictional force $[\mathcal{F}_x]$ and diabatic heating $[Q]$, and the parameter σ.

Some remarks on the QG equation system

The zonal-mean QG system is a rather remarkable set of equations. First, note that the only prognostic variables are the zonal-mean zonal wind and potential temperature perturbation. That is, $[v]$ and $[\omega]$ are diagnostic variables within this framework. This is because $[v]$ and $[\omega]$ have no geostrophic component and are entirely ageostrophic. In particular, $[\omega]$ may be diagnosed by requiring that the vertical velocity is sufficient to adiabatically warm or cool the atmosphere to maintain thermal wind balance at all times. The zonal-mean meridional velocity may then be diagnosed from mass continuity.

In deriving the QG system, we have effectively filtered out the gravity waves that ordinarily act to adjust the atmosphere toward geostrophic balance (as in the Rossby adjustment problem). Rather, the QG system enforces thermal wind balance (and mass continuity) at all instants. It should also be noted that in our QG equation set, the vertical eddy fluxes do not appear. This is a consequence of QG scaling, the vertical velocity is small (of order Rossby number), and thus its effects may be neglected, except where it multiplies the mean stratification.

While useful, the QG system has a number of limitations that should be mentioned. Firstly, the assumptions of low Rossby number required for QG to be valid are not satisfied in the deep tropics; QG is a model of the extratropical circulation.

Secondly, the QG equations are not closed, in the sense that one must know the forcing terms on the right hand side, or parameterise them in terms of the zonal-mean flow. This is particularly challenging for the diabatic heating rate, since latent heating by clouds is highly coupled to the flow itself in a highly nonlinear way (upward motion is associated with latent heating, but downward motion is not associated with an equivalent amount of latent cooling).

Thirdly, a limitation of QG that is not often discussed is that one must specify the mean stratification *a priori*. In reality, the stratification of the atmosphere in extratoprical regions is determined by the effect of large-scale eddies[1], and so a full theory would include a prediction of the mean stratification σ.

Finally, we note that QG is simply the first order solution to small Rossby number dynamics in the atmosphere; one may derive higher-order solutions that take into account terms of order Ro^2. One such extension, called semi-geostrophy (see Hoskins, 1975), predicts the existence of discontinuities in the flow that appear in finite time; it predicts the existence of fronts! At such locations, the assumptions of small Rossby number flow break down. This implies that, given a flow that satisfies the assumptions of QG, evolution of the flow will eventually lead to a breakdown of this assumption.

Surprisingly, given the number of caveats above, QG theory is one of the most powerful conceptual tools in dynamical meteorology, and our understanding of the extratropical atmosphere and the interaction between the zonal-mean flow and eddies would be much poorer without it.

7.2.2 Perturbation from a balanced state

Let us consider how the zonal-mean QG system is affected by eddy forcing. Suppose we have a statistically steady state in which

$$\frac{\partial [u_g]}{\partial t} = \frac{\partial [\theta]}{\partial t} = 0. \tag{7.22}$$

[1] A general theory for the extratropical stratification is still lacking, although there are a number of ideas in the literature (see e.g., Juckes, 2000; Schneider and O'Gorman, 2008; Jansen and Ferrari, 2012). For the extratropics, we therefore have a well-developed theory for balanced dynamics, but no full theory for the mean thermodynamic state. This is somewhat of the opposite situation to that of the tropics, where, as mentioned in chapter 3, the thermal stratification is known to be controlled by convection, but we lack a compelling theory for balanced dynamics (see, e.g., Sobel et al., 2001; Raymond et al., 2015, for some attempts at constructing such a theory in the tropics).

Now let us perturb this state with a vertically propagating wave. We will show later that such a wave will produce a meridional heat flux, but no meridional momentum flux, that is, the perturbation has the properties

$$\frac{\partial[v_g^*\theta^*]}{\partial y} \neq 0,$$
$$\frac{\partial[u_g^* v_g^*]}{\partial y} = 0.$$

Examining the momentum equation, we see that the momentum equation is unaffected by this wave, and so we might expect no change in the zonal wind due to the wave, i.e., $\partial_t u_g = 0$. On the other hand, the meridional heat flux is a forcing to the potential temperature equation, and so we would expect $\partial_t \theta \neq 0$. But by thermal wind balance, if either u_g or θ changes, the other has to evolve in tandem to maintain thermal wind balance! What is going on here?

The mechanism by which the QG system maintains thermal wind balance is through the ageostrophic mean circulation ($[v], [\omega]$). A meridional eddy flux of heat with non-zero divergence $\partial_y[v_g^*\theta^*]$ induces a mean ageostrophic circulation that in turn affects the momentum equation through the Coriolis acceleration. More specifically, consider the case in which the heat flux through the midlatitudes is suddenly increased. All else being equal, we would expect such an increase to act to reduce the temperature gradient between subrtropical and subpolar latitudes. In order to maintain thermal wind balance, however, this reduction in the temperature gradient must be accompanied by a reduction in the wind shear. In the QG system, this reduction in wind shear is achieved by an induced equatorward flow in the upper atmosphere, and a poleward flow in the lower atmosphere, which produces a Coriolis acceleration that increases the zonal wind at low levels and decreases it at upper levels. The induced meridional circulation is driven mechanically by the requirement for the atmosphere to maintain thermal wind balance; it may or may not be thermally direct, and in the case of the Ferrel cell, the mean meridional circulation is indeed thermally indirect.

Clearly, in a baroclinic atmosphere, we cannot consider eddy heat and momentum fluxes separately as we did in the previous chapter; they act in tandem and each affects both the thermal structure and the zonal wind structure.

7.2.3 Non-acceleration theorem

How can we make sense of the eddy influence on the mean circulation within the QG framework? Consider steady, non-dissipative, adiabatic flow. Under such conditions, time derivatives are zero, and friction and diabatic heating are absent. The equations for the zonal-mean zonal wind and potential temperature perturbation (7.15) and (7.16) then

become diagnostic equations for the overturning circulation,

$$f_0[v] = \frac{\partial[u_g^* v_g^*]}{\partial y}, \tag{7.23}$$

$$[\omega] = -\frac{1}{\sigma}\frac{\partial[v_g^* \theta^*]}{\partial y}. \tag{7.24}$$

Substituting these expressions into the zonal-mean continuity equation, we have,

$$\frac{\partial}{\partial y}\left\{\frac{\partial}{\partial y}\left(-\frac{[u_g^* v_g^*]}{f_0}\right) + \frac{\partial}{\partial p}\left(\frac{[v_g^* \theta^*]}{\sigma}\right)\right\} = 0. \tag{7.25}$$

Integrating meridionally, we find that the term in curly brackets is independent of latitude. Consider the behaviour of the function in curly brackets at the pole. Since at the pole $v = [v] = v^* = 0$, we have that,

$$\frac{\partial}{\partial p}\left(\frac{[v_g^* \theta^*]}{\sigma}\right) = 0. \tag{7.26}$$

Furthermore, by (7.23), the term $\partial_y\left(-[u_g^* v_g^*]\right)$ is also zero. This implies that the term in curly brackets in (7.25) is zero at the pole, and, by (7.25) itself, this term must be zero at all latitudes.

The above conclusion may be expressed concisely as,

$$\nabla_p \cdot \mathbf{F}_{EP} = 0. \tag{7.27}$$

where $\mathbf{F}_{EP}$ is the Eliassen-Palm flux

$$\mathbf{F}_{EP} = -[u_g^* v_g^*]\mathbf{j} + \frac{f_0[v_g^* \theta^*]}{\sigma}\mathbf{k} \tag{7.28}$$

where $\mathbf{j}$ and $\mathbf{k}$ are unit vectors in the meridional (y) and vertical (pressure) directions, respectively[2].

The above derivation has shown that for non-dissipative, adiabatic flow, if the EP flux is zero, there is a trivial solution to the QG governing equations in which $\partial_t[u_g] = \partial_t[\theta] = 0$ – eddies do not accelerate the mean flow. This is known as the nonacceleration theorem.

Alternatively, we may state that, in order for an adiabatic non-dissipative flow to be steady, the divergence of the Eliassen-Palm flux must be zero. When these conditions are satisfied, eddy fluxes of heat and momentum are balanced by the meridional overturning circulation

[2]Note that here we are defining the pressure-coordinate version of the divergence of a vector field $\mathbf{F} = F_x\mathbf{i} + F_y\mathbf{j} + F_p\mathbf{k}$ as $\nabla_p \cdot \mathbf{F} = \partial_x F_x + \partial_y F_y + \partial_p F_p$. This is simply a notational convenience; in a proper coordinate transform, the divergence should be invariant.

represented by $[v]$ and $[\omega]$. Conversely, when $\nabla_p \cdot \mathbf{F} \neq 0$, the overturning circulation cannot balance the eddy fluxes, and an acceleration of the zonal flow must result.

More generally, when diabatic heating and friction are included, the divergence of the EP flux represents the net effect of eddies on the mean flow. If the EP flux is nonzero, eddies are causing an acceleration of the zonal-mean zonal wind. This acceleration either must be balanced by diabatic heating or friction, or it must cause a net change to the zonal wind.

7.3 The Eliassen-Palm flux

The EP flux shows up in a number of surprising ways when analysing the extratropical general circulation. In this section, we demonstrate the fundamental importance of EP fluxes to extratropical dynamics. We first explicitly show that the EP flux represents the net effect of eddies on the mean flow by putting the governing equations into a more illuminating form called the transformed Eulerian Mean (TEM). Next we show the connection between EP fluxes and the dynamics of potential vorticity. Finally, we show how EP fluxes are related to wave propagation in linearised dynamics. These tools will allow us to interpret the eddy-mean flow interactions in the midlatitude atmosphere, and they will provide a solid basis on which to reason about jet formation and jet maintenance in baroclinic fluids like the atmosphere.

7.3.1 the Transformed Eulerian Mean (TEM)

Let us define a streamfunction of the meridional overturning circulation χ by,

$$[v] = \frac{\partial \chi}{\partial p}, \tag{7.29a}$$

$$[\omega] = -\frac{\partial \chi}{\partial y}. \tag{7.29b}$$

The form of the streamfunction automatically implies the flow satisfies the zonal-mean continuity equation,

$$\frac{\partial [v]}{\partial y} + \frac{\partial [\omega]}{\partial p} = 0. \tag{7.30}$$

Expressing the zonal-mean thermodynamic equation in terms of the streamfunction, we have,

$$\frac{\partial [\theta]}{\partial t} - \sigma \left\{ \frac{\partial \chi}{\partial y} - \frac{\partial}{\partial y} \frac{[v^* \theta^*]}{\sigma} \right\} = \frac{[Q]}{c_p \pi}. \tag{7.31}$$

where we have taken the eddy forcing term to the left-hand side. We now make what appears to be a somewhat arbitrary definition of

$$\tilde{\chi} = \chi - \frac{[v^* \theta^*]}{\sigma}, \tag{7.32}$$

where $\tilde{\chi}$ is known as the "transformed" streamfunction. In analogy to (7.29), we may define "velocities" by,

$$[\tilde{v}] = \frac{\partial \tilde{\chi}}{\partial p}, \tag{7.33a}$$

$$[\tilde{\omega}] = -\frac{\partial \tilde{\chi}}{\partial y}. \tag{7.33b}$$

The overturning circulation defined by $(\tilde{v}, \tilde{\omega})$ is known as the *residual* overturning.

The benefit to the above definition is that it casts the thermodynamic equation in a particularly simple way,

$$\frac{\partial [\theta]}{\partial t} + \sigma [\tilde{\omega}] = \frac{[Q]}{c_p \pi}. \tag{7.34}$$

In particular, in steady state, the residual pressure velocity is balanced directly by diabatic heating, and there is no explicit eddy forcing[3]. The residual velocity is therefore upwards where the atmosphere is being heated, and downwards where the atmosphere is being cooled, as one would expect from a thermally direct circulation (see below).

The simplicity we have produced in the thermodynamic equation must be paid for in the zonal momentum equation. Applying our definition of $\tilde{\chi}$ to (7.15), we have,

$$\frac{\partial [u_g]}{\partial t} - f_0 [\tilde{v}] = [\mathcal{F}_x] - \frac{\partial [u_g^* v_g^*]}{\partial y} + f_0 \frac{\partial}{\partial p} \left(\frac{[v_g^* \theta^*]}{\sigma} \right). \tag{7.35}$$

Now the momentum equation has not one, but two explicit eddy terms. However, some progress has nevertheless been made, as we may express the eddy forcing in terms of the Eliassen-Palm flux divergence,

$$\frac{\partial [u_g]}{\partial t} - f_0 [\tilde{v}] = [\mathcal{F}_x] + \nabla_p \cdot \mathbf{F}_{EP}. \tag{7.36}$$

Equations (7.34) and (7.36) give the thermodynamic and momentum equations in the transformed Eulerian mean (TEM) form. This form is useful, because it confines the explicit eddy forcing terms into one equation. From these equations, it is clear that the effect of eddies on the zonal-mean flow is governed by the divergence of the EP flux.

The TEM equations also contain the meridional overturning circulation in its residual form. The residual overturning, while appearing rather arbitrary at first , has some very useful properties. In particular, the residual vertical velocity $\tilde{\omega}$ corresponds to the diabatically driven vertical motion. $\tilde{\omega}$ is upwards in regions that are diabatically heated and downwards

[3] Of course, eddies contribute to the zonal-mean heating, particularly when it comes to latent heat release in clouds.

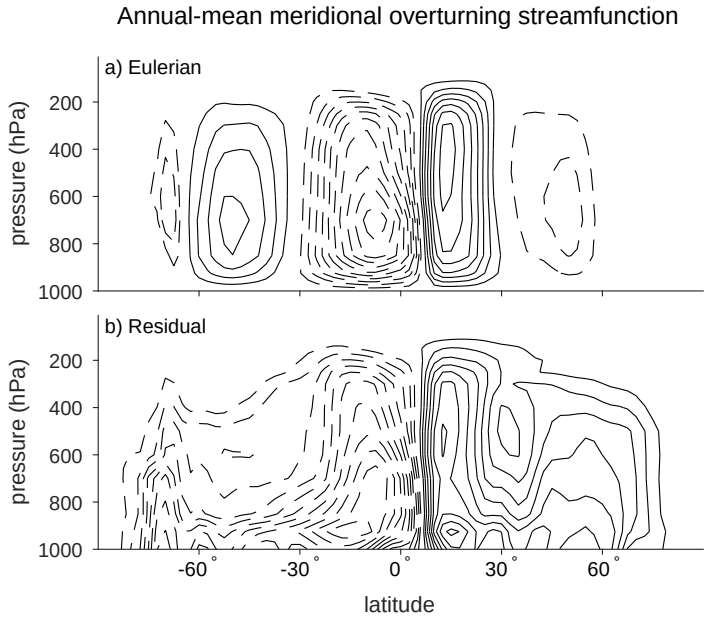

Figure 7.1: The annual-mean meridional overturning streamfunction calculated based on NCEP-DOE reanalysis data for the years 2008-2017. (a) Traditional Eulerian-mean streamfunction Ψ, and (b) residual streamfunction $\tilde{\Psi} = \Psi - 2\pi R_e \cos\phi \frac{\overline{[v^\dagger \theta^\dagger]}}{\partial_p [\overline{\theta}]}$. Solid lines represent clockwise motion and dashed lines represent anticlockwise motion and the contour interval is 10^{11} kg s^{-1}.

in regions that are diabatically cooled. As a result, the residual circulation is thermally direct, even at midlatitudes where the Eulerian mean circulation is thermally indirect (the Ferrel Cell). Indeed, the residual streamfunction is comprised of a single cell in each hemisphere, with rising motion in the tropics and descent in subpolar regions, although there is also a notable dip in the streamlines in the subtropics, roughly coinciding with the boundary between the Hadley and Ferrel Cells in the Eulerian perspective.

Away from the surface, the residual overturning also approximates the mass-weighted overturning circulation in isentropic coordinates (Held and Schneider, 1999), since vertical motions in isentropic coordinates correspond to diabatic heating or cooling. Near to the surface, however this similarity breaks down, and in fact the residual overturning does not close, with the residual streamfunction intersecting the surface. This is because the QG assumptions used to derive the TEM equations break down within the boundary layer.

According to (7.34) and (7.36) the EP flux divergence plays a central role in the dynamics of the zonal-mean flow in the extratropics. We shall see below that EP fluxes also have a close relationship to potential vorticity, and they are essential to understand when applying "PV thinking" to the general circulation.

7.3.2 Quasigeostrophic potential vorticity (QGPV)

The utility of viewing the extratropical circulation from the perspective of potential vorticity boils down to two properties of potential vorticity within the QG framework; the conservation principle and the invertibility principle. These two principles are derived briefly below, after which we discuss their physical interpretation.

Conservation of QGPV

Consider again the QG equation system. Taking $\partial_y(7.9) - \partial_x(7.8)$, we have,

$$\frac{\partial \zeta_g}{\partial t} + \mathbf{u_g} \cdot \nabla_p \zeta_g = -f_0 \left(\frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} \right) - \beta v_g + \frac{\partial \mathcal{F}_y}{\partial x} - \frac{\partial \mathcal{F}_x}{\partial y}, \tag{7.37}$$

where $\zeta_g = \partial_x v_g - \partial_y u_g$ is the geostrophic vorticity, and we have used that

$$\begin{aligned} \frac{\partial \mathbf{u_g}}{\partial y} \cdot \nabla_p u_g - \frac{\partial \mathbf{u_g}}{\partial x} \cdot \nabla_p v_g &= \frac{\partial u_g}{\partial y}\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y}\frac{\partial u_g}{\partial y} - \frac{\partial u_g}{\partial x}\frac{\partial v_g}{\partial x} - \frac{\partial v_g}{\partial x}\frac{\partial v_g}{\partial y} \\ &= \zeta_g \left(\frac{\partial u_g}{\partial x} + \frac{\partial v_g}{\partial y} \right) \\ &= 0 \end{aligned}$$

by the non-divergence of the geostrophic wind. Now, since $\partial_x u + \partial_y v = -\partial_p \omega$, we may write the QG vorticity equation,

$$\frac{\mathrm{D}_g}{\mathrm{D}t} (\zeta_g + f) = f_0 \frac{\partial \omega}{\partial p} + \frac{\partial \mathcal{F}_y}{\partial x} - \frac{\partial \mathcal{F}_x}{\partial y}, \tag{7.38}$$

where we have defined the geostrophic derivative,

$$\frac{\mathrm{D}_g}{\mathrm{D}t} = \frac{\partial}{\partial t} + \mathbf{u}_g \cdot \nabla_p \tag{7.39}$$

and used that $\frac{\mathrm{D}_g f}{\mathrm{D}t} = \beta v_g$.

Rearranging the thermodynamic equation (7.12), we may write the pressure velocity as,

$$\omega = \frac{1}{\sigma} \left\{ \frac{Q}{c_p \pi} - \frac{\mathrm{D}_g \theta}{\mathrm{D}t} \right\} \tag{7.40}$$

Multiplying by f_0 and taking the vertical derivative, we may write,

$$f_0 \frac{\partial \omega}{\partial p} = -\frac{\mathrm{D}_g}{\mathrm{D}t} \left\{ \frac{\partial}{\partial p} \left(\frac{f_0 \theta}{\sigma} \right) \right\} + \frac{\partial}{\partial p} \left(\frac{f_0 Q}{c_p \pi \sigma} \right), \tag{7.41}$$

where we have used that,

$$\frac{\partial \mathbf{u_g}}{\partial p} \cdot \nabla_p \theta = 0$$

as a result of the thermal wind relation.

Substituting (7.41) into (7.38), we may write a conservation equation for the quasigeostrophic potential vorticity, q,

$$\frac{\mathrm{D}_g q}{\mathrm{D}t} = \frac{\partial}{\partial p}\left(\frac{f_0 Q}{c_p \pi \sigma}\right) + \frac{\partial \mathcal{F}_y}{\partial x} - \frac{\partial \mathcal{F}_x}{\partial y}, \tag{7.42}$$

where q is defined,

$$q = \zeta_g + f + \frac{\partial}{\partial p}\left(\frac{f_0 \theta}{\sigma}\right). \tag{7.43}$$

Note that, for frictionless adiabatic flow, the QGPV is conserved following the geostrophic flow,

$$\frac{\mathrm{D}_g q}{\mathrm{D}t} = 0. \tag{7.44}$$

The above equation expresses the conservation principle for QGPV. Note that the advecting velocities in this equation are the geostrophic winds; QGPV is conserved following the geostrophic flow, not the total flow.

Invertibility of QGPV

To make the conservation of QGPV a useful principle, we must connect the QGPV to the variables we care about, the winds and potential temperature. To do so, note that we can express the geostrophic winds in terms of the geopotential,

$$u_g = \frac{1}{f_0}\frac{\partial \Phi}{\partial y} \tag{7.45}$$

$$v_g = -\frac{1}{f_0}\frac{\partial \Phi}{\partial x}, \tag{7.46}$$

and the vorticity is therefore,

$$\zeta_g = \frac{1}{f_0}\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)\Phi. \tag{7.47}$$

Furthermore, the potential temperature may be related to the geopotential through hydrostatic balance (7.19). Rearranging (7.19) we have,

$$\frac{1}{\mathcal{S}}\frac{\partial \Phi}{\partial p} = \frac{\theta}{\sigma}. \tag{7.48}$$

where $\mathcal{S} = -\sigma R_d \pi / p$. Substituting (7.47) and (7.48) into the definition of the QGPV (7.43), we have,

$$q = f + \frac{1}{f_0}\left(\frac{\partial^2 \Phi}{\partial x^2} + \frac{\partial^2 \Phi}{\partial y^2}\right) + \frac{\partial}{\partial p}\left(\frac{f_0}{\mathcal{S}}\frac{\partial \Phi}{\partial p}\right). \tag{7.49}$$

Equation (7.49) expresses the QGPV as the solution to an elliptic partial differential equation in the geopotential Φ. Furthermore, using (7.45), (7.46) and (7.48), the winds and potential temperature may be deduced directly from the geopotential. Thus, a given distribution of QGPV, combined with appropriate boundary conditions[4] may be "inverted" by solving (7.49), and all the variables of interest may be derived.

Fig. 7.2 shows an example of such an inversion for a Gaussian anomaly in the PGPV (a PV "ball"). The cyclonic QGPV anomaly produces an anomaly in the geopotential height which is associated with a cyclonic vortex in the geostrophic wind, and a dipole structure in the potential temperature. Note that the geopotential anomaly extends over a wider region than the PV anomaly. Indeed, the equation relating the QGPV to the geopotential is similar to the equation relating an electric charge to the electric field in electrostatics. This analogy is useful when thinking about the induced circulations associated with isolated anomalies of PV (see e.g., Hoskins et al., 1985).

PV thinking

The combination of the QGPV conservation principle and its invertibility makes the QGPV a powerful tool for analysing extratropical circulations. In particular, the QGPV tells us all we need to know about the flow; given the QGPV distribution, the invertibility principle tells us that all other quantities of the flow may be calculated, and its evolution my be determined by the conservation principle, allowing one to determine the future properties of the flow.

Since the complete description of a QG flow can be encapsulated by the QGPV, this has led to the widespread adoption of what is known as "PV thinking" when conceptualising extratropical circulations. This view focuses on the PV distribution and its evolution. For example, rather than analysing heat advection and momentum advection separately to determine causal mechanisms influencing the evolution of temperature and winds, one would analyse the PV and its evolution. The invertibility principle then allows this information to be translated into information about the winds and potential temperature distribution.

A QGPV-focussed view automatically accounts for the effects of the induced ageostrophic overturning circulation that accompanies changes in temperature and wind. Furthermore,

[4] In particular, potential temperature anomalies at boundaries may be interpreted as δ functions of PV; see Vallis (2017), section 5.4 for details.

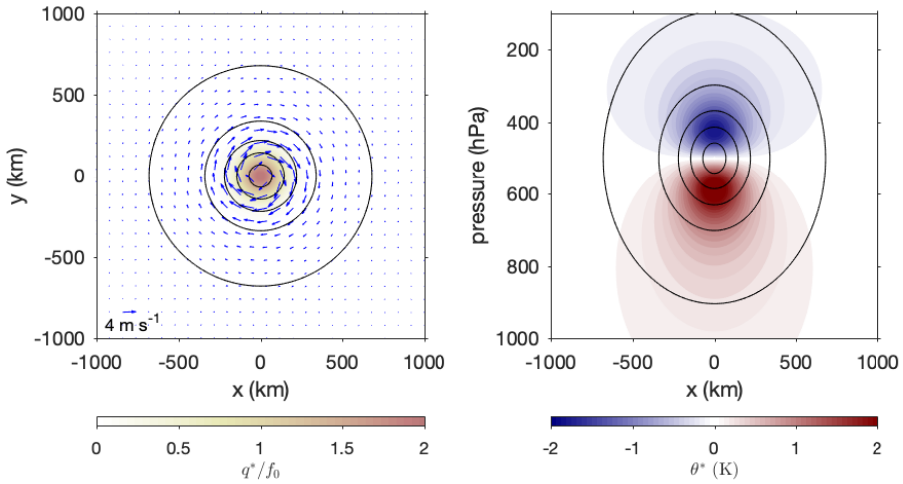

Figure 7.2: (left) Horizontal and (right) vertical cross sections of the geopotential height anomalies (black lines; contour interval 4 m) induced by a Gaussian anomaly of cyclonic QGPV with a maximum magnitude of $2f_0$, a horizontal standard deviation of 100 km, and a vertical standard deviation of 60 hPa. In the left panel, arrows give the induced geostrophic winds and colours give the PV anomaly, while in the right panel, colours give the potential temperature anomaly. Calculations are performed assuming $\mathcal{S} = 3 \times 10^{-6}$ m^4 s^2 kg^{-2}.

as we shall see below, "PV thinking" highlights the importance of the EP flux divergence through its influence on the PV distribution.

Evolution of the zonal-mean QGPV

Let us consider now how the zonal-mean QGPV evolves in time. For simplicity, we will start with frictionless, adiabatic flow, for which the QGPV is conserved following the geostrophic flow. Since the geostrophic wind is non divergent, we may write (7.44) in flux form as,

$$\frac{\partial q}{\partial t} + \frac{\partial u_g q}{\partial x} + \frac{\partial v_g q}{\partial y} = 0. \tag{7.50}$$

Taking the zonal mean, we have,

$$\frac{\partial [q]}{\partial t} = -\frac{\partial [v_g^* q^*]}{\partial y}, \tag{7.51}$$

where the mean meridional flux is zero since $[v_g] = 0$. This equation states that the zonal-mean QGPV distribution is forced (for adiabatic and frictionless flow) by the meridional eddy QGPV flux. But what exactly is the form of this flux? Using the definition of the

QGPV (7.43), we have,

$$[v_g^* q^*] = \left[v_g^* \frac{\partial v_g^*}{\partial x}\right] - \left[v_g^* \frac{\partial u_g^*}{\partial y}\right] + f_0 \left[v^* \frac{\partial}{\partial p}\left(\frac{\theta^*}{\sigma}\right)\right]. \tag{7.52}$$

Since the geostrophic wind is non divergent, we may subtract $[u^*(\partial_x u_g^* + \partial_y v_g^*)] = 0$ from the above equation to give, with some rearrangement,

$$[v_g^* q^*] = \left[\frac{1}{2}\frac{\partial}{\partial x}\left(v_g^* - u_g^*\right)\right] - \left[\frac{\partial}{\partial y}\left(u_g^* v_g^*\right)\right] + f_0 \left[v^* \frac{\partial}{\partial p}\left(\frac{\theta^*}{\sigma}\right)\right]. \tag{7.53}$$

The first term on the right-hand side is an exact derivative and is therefore zero in the zonal mean. Furtherore, we may use the product rule on the last term on the right-hand side to write,

$$[v_g^* q^*] = -\frac{\partial\left[u_g^* v_g^*\right]}{\partial y} + f_0 \frac{\partial}{\partial p}\left(\frac{[v^*\theta^*]}{\sigma}\right) - \left[\left(\frac{\theta}{\sigma}\right)\frac{\partial v^*}{\partial p}\right]. \tag{7.54}$$

Now, by thermal wind, we have that,

$$f_0 \frac{\partial v_g}{\partial p} = -\frac{R_d \pi}{p}\frac{\partial \theta}{\partial x}, \tag{7.55}$$

implying that the last term on the right-hand side of (7.54) is an exact derivative and equal to zero in the zonal mean. We therefore have that the eddy QGPV flux may be written,

$$\begin{aligned}[v_g^* q^*] &= -\frac{\partial\left[u_g^* v_g^*\right]}{\partial y} + f_0 \frac{\partial}{\partial p}\left(\frac{[v^*\theta^*]}{\sigma}\right) \\ &= \nabla_p \cdot \mathbf{F}_{EP}.\end{aligned} \tag{7.56}$$

The meridional flux of QGPV is given by the Elliasen-Palm flux divergence! This shows that our transformation of the equations into their TEM form above is far from arbitrary, but rather it is another aspect of so called "PV thinking".

7.3.3 EP fluxes and Rossby waves

So far we have shown that the EP flux divergence represents the net effect of eddies on the mean zonal wind and potential temperature field, and we have shown that the EP flux divergence is equal to the meridional flux of QGPV. We now show how EP fluxes are related to wave propagation.

We begin by deriving an equation for a quantity known as the "wave activity". For the QG system, we define the wave activity as,

$$\mathcal{A} = \frac{1}{2}\frac{\overline{[q^{*2}]}}{\partial [q]/\partial y}, \tag{7.57}$$

where, as in previous sections, the square brackets represent zonal mean and the asterisk represents a deviation from this mean. The wave activity is given by the eddy enstrophy q^{*2} divided by the mean QGPV gradient. Care must be taken in cases where the QGPV gradient changes sign, although we will not consider such situations here.

To derive a wave activity equation, we consider again the equation for the QGPV (7.42), which may be written in flux form,

$$\frac{\partial q}{\partial t} = -\frac{\partial u_g q}{\partial x} - \frac{\partial v_g q}{\partial y} + D \tag{7.58}$$

where $D = \partial_p(f_0 Q/c_p \pi\sigma) + \partial_y \mathcal{F}_x - \partial_x \mathcal{F}_y$ represents the effects of non-conservative terms associated with dissipation and diabatic heating. Taking the zonal mean, we have,

$$\frac{\partial [q]}{\partial t} = -\frac{\partial [v_g q]}{\partial y} + [D]. \tag{7.59}$$

Subtracting the above equation from (7.58) gives

$$\frac{\partial q^*}{\partial t} = -\frac{\partial u_g^*[q]}{\partial x} - \frac{\partial [u_g] q^*}{\partial x} - \frac{\partial v_g^*[q]}{\partial y} + D^* \tag{7.60}$$

where we have used that $[v_g] = 0$. Using the non-divergence of the geostrophic wind, we may write,

$$\frac{\partial q^*}{\partial t} = -v_g^* \frac{\partial [q]}{\partial y} - \frac{\partial [u_g] q^*}{\partial x} + D^*. \tag{7.61}$$

Multiplying by q^* and taking the time and zonal average, we may derive an equation for the enstrophy,

$$\frac{\partial}{\partial t}\left(\frac{1}{2}[q^{*2}]\right) = -[v_g^* q^*]\frac{\partial [q]}{\partial y} + [D^* q^*]. \tag{7.62}$$

Finally, dividing by $\partial_y [q]$, gives the governing equation for the wave activity,

$$\frac{\partial \mathcal{A}}{\partial t} = -[v^* q^*] + \mathcal{D}, \tag{7.63}$$

where $\mathcal{D} = [D^* q^*]/\partial_y [q]$. Since the QGPV flux is equal to the divergence of the EP flux, we have,

$$\frac{\partial \mathcal{A}}{\partial t} = -\nabla \cdot \mathbf{F}_{EP} + \mathcal{D}. \tag{7.64}$$

For conservative flow ($\mathcal{D} = 0$), regions where the wave activity is growing correspond to EP flux convergence, and regions where the wave activity is decaying correspond to EP flux divergence (provided the QGPV gradient remains positive).

Equation (7.64) is a rather general equation (we have made no small amplitude assumption for instance), and it shows the connection between the wave activity to the EP flux

divergence. However, it does not allow a clear interpretation in terms of wave dynamics. For that, we must consider small amplitude disturbances and linearise the QGPV equation.

As in section 6.3.1, we linearise the QGPV equation about a base state with uniform zonal flow $[u]$. Using (7.49) to express the QGPV in terms of the geopotential, we may write conservation of QGPV as,

$$\left(\frac{\partial}{\partial t}+[u]\frac{\partial}{\partial x}\right)\left\{\frac{\partial^2\Phi^*}{\partial x^2}+\frac{\partial^2\Phi^*}{\partial y^2}+\frac{\partial}{\partial p}\left(\frac{f_0^2}{\mathcal{S}}\frac{\partial\Phi^*}{\partial p}\right)\right\}+\beta\frac{\partial\Phi^*}{\partial x}=0. \tag{7.65}$$

where Φ^* is the perturbation geopotential. Assuming a wave-like solution of the form,

$$\Phi^*(x,y,p,t)=A\cos\left\{i\left(kx+ly+mp-\omega t\right)\right\} \tag{7.66}$$

we recover the dispersion relation for baroclinic Rossby waves,

$$\omega=[\overline{u}]k-\frac{\beta k}{\kappa^2} \tag{7.67}$$

where $\kappa^2=k^2+l^2+m^2\frac{f_0^2}{\mathcal{S}}$. Now, consider the group velocity of such waves. We have, for the meridional and pressure directions,

$$\begin{aligned}\mathbf{c_g}&=\frac{\partial\omega}{\partial l}\mathbf{j}+\frac{\partial\omega}{\partial m}\mathbf{k}\\&=\frac{2\beta kl}{\kappa^4}\mathbf{j}+\frac{2\beta km}{\kappa^4}\frac{f_0^2}{\mathcal{S}}\mathbf{k}.\end{aligned} \tag{7.68}$$

Recall that the group velocity $\mathbf{c_g}$ represents the direction in which the wave activity associated with a given frequency ω and wavenumbers k and l travels. We may also use geostrophic and hydrostatic balance to recover the perturbation winds and geopotential,

$$\begin{aligned}u_g^*&=\frac{1}{f_0}\frac{\partial\Phi^*}{\partial y}=-\frac{Al}{f_0}\sin\left(kx+ly+mp-\omega t\right)\\v_g^*&=-\frac{1}{f_0}\frac{\partial\Phi^*}{\partial x}=\frac{Ak}{f_0}\sin\left(kx+ly+mp-\omega t\right)\\\theta^*&=-\frac{1}{\mathcal{S}}\frac{\partial\Phi^*}{\partial p}=\frac{Am}{\mathcal{S}}\sin\left(kx+ly+mp-\omega t\right).\end{aligned}$$

Using the above equations, we have that,

$$[u_g^* v_g^*] = \frac{A^2 kl}{f_0^2},$$
$$[v_g^* \theta^*] = \frac{A^2 km}{\mathcal{S}}$$
$$[q^{*2}] = \frac{A^2 \kappa^4}{f_0^2}$$

Furthermore, since we have assumed the mean zonal wind to be uniform, $\partial_y[\overline{q}] = \beta$.

Comparing the above equations to (7.68), we note that,

$$\mathcal{A}\mathbf{c_g} = \mathbf{F}_{EP}. \tag{7.69}$$

This equation suggests a fundamental relationship between the EP flux and the group velocity of baroclinic Rossby waves. In particular, this means that the EP flux is always parallel to the group velocity; the EP flux tells us about how waves propagate.

Using (7.69), we may express the wave activity conservation relation as,

$$\frac{\partial \mathcal{A}}{\partial t} = -\nabla \cdot (\mathbf{c}_g \mathcal{A}) + \mathcal{D}. \tag{7.70}$$

The above equation makes clear why we have referred to $\mathcal{A}$ as the wave activity. The flux of $\mathcal{A}$ occurs via the group velocity for Rossby waves.

For a steady, propagating wave in adiabatic frictionless flow, we have that $\nabla \cdot (\mathbf{c}_g \mathcal{A}) = 0$, implying that the EP flux is also zero, and that the wave does not affect the zonal-mean flow. On the other hand, in a region that is a source of wave activity, we expect $\nabla \cdot (\mathbf{c}_g \mathcal{A}) > 0$, implying an acceleration of the zonal-mean zonal wind by (7.36), and in a region of wave dissipation (by friction for example) we expect $\nabla \cdot (\mathbf{c}_g \mathcal{A}) < 0$, implying a deceleration of the mean flow.

7.3.4 Summary and application to the atmosphere

In this section we have demonstrated the fundamental nature of the EP flux and its connection to midlatitude eddy dynamics. In particular, we have shown that,

1. The EP flux divergence represents the net effect of eddies on the mean zonal flow.
2. The EP flux divergence is equal to the meridional flux of QGPV.
3. The EP flux represents the flux of wave activity in the atmosphere, and it is parallel to the group velocity for Rossby waves.

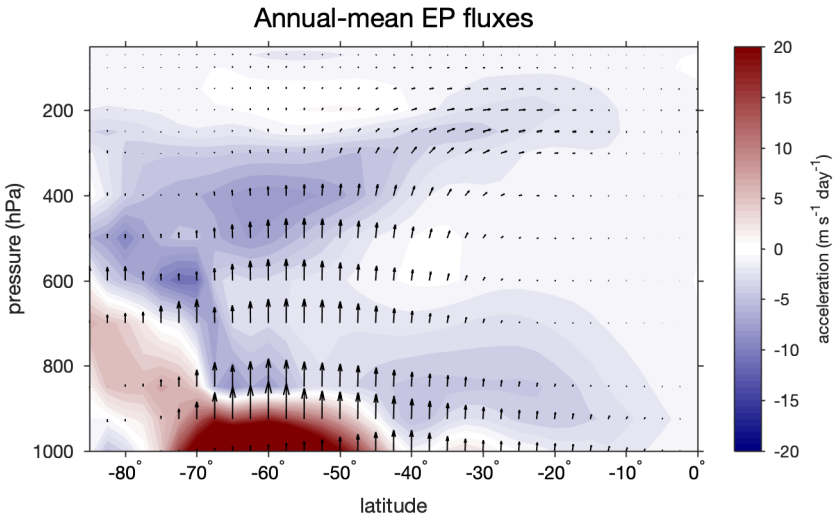

Figure 7.3: Annual-mean Eliassen-Palm flux [(7.71); vectors] and the acceleration of the zonal-mean flow by eddies [(7.72); coloours] calculated based on NCEP-DOE reanalysis for the years 2008-2017 (see footnote 5).

The above properties suggest that the EP flux is an ideal quantity for use in visualising the effects of eddies on the mean flow in the atmosphere. Fig. 7.3 shows a visualisation of the EP flux and its divergence for the southern hemisphere following Edmon et al. (1980) and defined in spherical coordinates rather than the cartesian coordinates used here[5]. The dominant pattern of $\nabla \cdot \mathbf{F}_{EP}$ is the strong divergence of the EP flux near the surface at midlatitudes and to a lesser extent polar latitudes, with a broad region of EP flux convergence aloft. The EP flux vectors themselves are almost vertical at low levels in the midlatitudes, bending equatorward in the upper atmosphere.

Since the EP flux represents the flux of wave activity, we may also interpret the EP flux vectors as the mean motion of Rossby wavepackets. The midlatitude lower troposphere is a source of Rossby waves (associated with baroclinic instability); this source of wave activity

[5]The EP flux may be defined in spherical coordinates as

$$\mathbf{F}_{EP}^{\text{spherical}} = -[\overline{u^\dagger v^\dagger}] \cos\phi\, \mathbf{j} + f \cos\phi \frac{[\overline{v^\dagger \theta^\dagger}]}{\partial_p [\overline{\theta}]}\, \mathbf{k}, \tag{7.71}$$

where $\mathbf{j}$ and $\mathbf{k}$ are the unit vectors in the meridional (latitude) and vertical (pressure) directions. The net acceleration of the zonal-mean flow by eddies is then given by

$$\frac{1}{\cos\phi} \nabla_p \cdot \mathbf{F}_{EP}^{\text{spherical}} = -\frac{1}{R_e \cos^2\phi} \frac{\partial}{\partial \phi} \left\{ [\overline{u^\dagger v^\dagger}] \cos^2\phi \right\} + f \frac{\partial}{\partial p} \left\{ \frac{[\overline{v^\dagger \theta^\dagger}]}{\partial_p [\overline{\theta}]} \right\}. \tag{7.72}$$

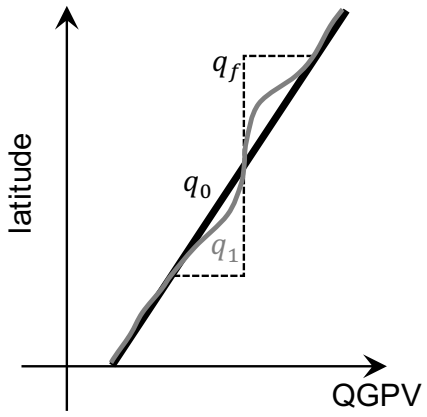

Figure 7.4: Effect of eddy mixing on the QGPV distribution. A fluid with an initially uniform PV distribution (q_0, black solid) undergoes some mixing, which locally reduces the QGPV gradient, but produces an increase in the QGPV gradient both to the north and south (q_1, grey). This change in the mean state has the effect of promoting mixing in the region with low QGPV gradient and suppressing it elsewhere. Further mixing therefore enhances the anistropy in the QGPV distribution, producing a mixed region with close to zero QGPV gradient, and regions with very high QGOV gradients on either side (q_f, dashed).

is associated with an acceleration of the flow. The waves then propagate upwards, into the upper troposphere, where some of them decay (decelerating the upper tropospheric flow). Other waves continue propagating and are deflected meridionally (but primarily equtoraward), where they eventually reach their critical latitude and decay, decelerating the flow at subtropical and subpolar latitudes.

7.4 Jets in a baroclinic atmosphere

Let us now think a little more generally about jet formation in baroclinic atmospheres. We have shown above that there is a fundamental connection between the meridional QGPV flux and the EP flux divergence. In particular, we may rewrite the TEM form of the zonal-mean momentum equation (7.36) using the QGPV flux,

$$\frac{\partial [u_g]}{\partial t} = [v^* q^*] + f_0 [\tilde{v}] + [\mathcal{F}_x]. \tag{7.73}$$

The explicit eddy forcing is by the QGPV flux, while the effect of the mean meridional circulation is through the residual meridional velocity $[\tilde{v}]$. Since the residual overturning balances the mean diabatic heating in the atmosphere, it is zero in adiabatic flow. Under

adiabatic, frictionless flow, we therefore have that,

$$\frac{\partial [u_g]}{\partial t} = [v^* q^*], \tag{7.74}$$

the zonal-mean acceleration is controlled by the eddy QGPV flux.

The utility of thinking about the flow from the perspective of QGPV fluxes is that the QGPV is a conserved variable, and therefore its flux must be (on average) down gradient. This provides a strong constraint on the QGPV gradient, which, as we have seen in the previous chapter, need not apply to momentum. This provides us with a natural and simple method for parameterising the effects of eddies on the mean flow by assuming that the action of eddies is to diffuse QGPV down its zonal mean gradient. Such diffusion, if it were to occur with a uniform diffusivity, would tend to smooth out QGPV gradients, thereby smoothing the zonal wind distribution. How then, can we explain the formation of jets in quasigeostrophic flows from the perspective of the QGPV?

Consider a flow that initially has a uniform gradient of QGPV. This could be a flow at rest; the gradient of QGPV would then equal β the gradient of the Coriolis parameter. Suppose an eddy forms in this flow and produces some irreversible mixing within some region. Since QGPV is conserved, we would expect such mixing to result in a local reduction of the QGPV gradient. Far away from the eddy, however, we would expect the QGPV to be unchanged. As shown in Fig. 7.4, this requires the QGPV gradient to be *increased* in the surrounding regions. The action of the irreversible mixing by the eddy has therefore created alternate regions of increased and decreased QGPV gradient.

Now consider how this new mean state affects the propagation of eddies. A key quantity for Rossby wave propagation is the Doppler-shifted phase speed,

$$[\overline{u}] - c = \frac{\partial_y [\overline{q}]}{\kappa}. \tag{7.75}$$

Where $[\overline{u}] - c$ is large, wave propagation is supported, and mixing is suppressed, while where $[\overline{u}] - c$ is small, wave breaking occurs, and mixing is enhanced. Since the Doppler-shifted phase speed is proportional to the QGPV gradient, this implies that mixing is enhanced in regions where the QGPV gradient is already low. This produced a feedback mechanism that produces regions of strong mixing, and uniform QGPV, separated by jets, associated with step changes in QGPV, that act as mixing barriers.

Jets may be produced by the action of QGPV diffusion with a diffusivity that is nonuniform and depends on the flow. In fully developed QG turbulence, this can lead to multiple jets separated by regions of enhanced mixing known as a "PV staircase" for the shape of its QGPV profile as a function of meridional distance. What sets the separation distance of these jets? The answer to this question is beyond the scope of these notes, but it depends

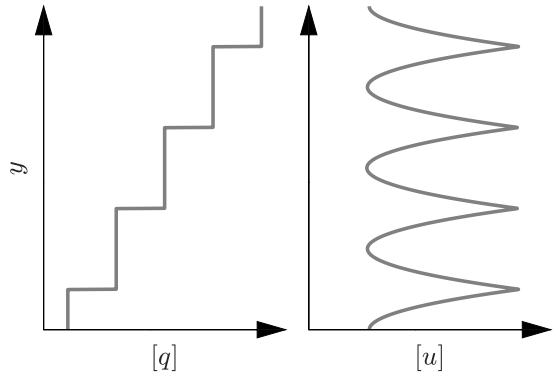

Figure 7.5: Example of a PV staircase, with the PV increasing stepwise with latitude, and the associated distribution of zonal wind exhibiting multiple jets. In certain parameter regimes, the spacing of the PV stairs is related to the Rhines scale, but in other regimes, the size of the steps may also be influenced by the role of friction in arresting the upscale energy cascade.

on the rotation rate, the size of the planet, and the role played by surface friction in the flow (see e.g., Vallis, 2017, ch 9). A key length scale involved, is the Rhines scale, defined,

$$L_\beta = \sqrt{\frac{U}{\beta}}, \tag{7.76}$$

where U is a characteristic velocity, and under certain conditions the Rhines scale sets the jet spacing (Rhines, 1975). Note that this length scale depends on the variation of f with latitude, and not its absolute value.

Chapter 8

The Hadley cell redux: eddy influences and the seasonal cycle

In the preceding chapters we have presented two contrasting views of the atmospheric general circulation. In chapters 6 and 7 we have described how eddies drive the extratropical overturning circulation through their transports of momentum and heat. On the other hand, chapter 4 presented a theory for the tropical circulation based on axisymmetric dynamics, where eddies are entirely neglected. How can we bridge these two views of the factors that drive meridional overturning circulation? In this chapter we consider this question, and we consider how to incorporate the influence of eddies on the Hadley Cell. This discussion will lead us to consider how the Hadley Cell varies seasonally, and how the role of eddies changes between the equinoctial circulation and the solsticial circulation that is dominated by the monsoons.

8.1 A non-axisymmetric Hadley Cell

8.1.1 Angular-momentum balance of the subtropics

We begin by considering the angular-momentum budget of the Hadley Cell. In chapter 4, we limited ourselves to an axisymmetric atmosphere, in which angular momentum is conserved in the absence of friction. This led us to the steady-state inviscid equation for angular-momentum (4.1), which may be written in pressure coordinates as,

$$\mathbf{u} \cdot \nabla_p M = 0 \tag{8.1}$$

where $\mathbf{u} = (u, v, \omega)$ and $M = R_e \cos\phi(u + R_e \cos\phi)$ is the angular momentum. This equation states that angular momentum is conserved along streamlines.

Under more general conditions, where we allow for zonal asymmetries, (8.1) is no longer valid. Instead, we must consider the eddy fluxes of angular momentum as we did in chapter 5. For flows that are not axisymmetric, the corresponding steady-state equation governing the zonal- and time-mean angular momentum may be written,

$$[\overline{\mathbf{u}}] \cdot \nabla [\overline{M}] = -\nabla_p \cdot \left([\overline{\mathbf{u}^\dagger M^\dagger}] \right). \tag{8.2}$$

Using the definition of the angular momentum, this may be rearranged to give

$$\frac{[\overline{v}]}{R_e} \frac{\partial [\overline{M}]}{\partial \phi} + [\overline{\omega}] \frac{\partial [\overline{M}]}{\partial p} = -\frac{1}{\cos\phi} \frac{\partial [\overline{u^\dagger v^\dagger}] \cos^2\phi}{\partial \phi} - R_e \cos\phi \frac{\partial [\overline{u^\dagger \omega^\dagger}]}{\partial p}. \tag{8.3}$$

Noting that $\partial_\phi [\overline{M}] = -R_e^2 \cos\phi (f + [\overline{\zeta}])$, where ζ is the relative vorticity, we may write this as

$$[\overline{v}](f + [\overline{\zeta}]) + [\overline{\omega}] \frac{\partial [\overline{u}]}{\partial p} = -S, \tag{8.4}$$

where, as in chapter 5, we have denoted the eddy-momentum flux convergence by

$$S = -\frac{1}{R_e \cos^2\phi} \frac{\partial [\overline{u^\dagger v^\dagger}] \cos^2\phi}{\partial \phi} - \frac{\partial [\overline{u^\dagger \omega^\dagger}]}{\partial p}. \tag{8.5}$$

The above equation was derived in chapter 5 as (5.17).

In chapter 5, we were focussed on the extratropical circulation, and we therefore were justified in assuming that the flow was in the quasigeostrophic (QG) regime. In this regime, vertical advection of angular momentum may be neglected, and the relative vorticity may be assumed to be small in comparison to the planetary vorticity f. The equation above therefore provides a direct relationship between the meridional flow $[\overline{v}]$ and the eddy fluxes in the QG regime. In the present case, we are interested in the tropical circulation, and we therefore cannot make the QG approximation. Nevertheless, we can still gain some insight from the balance (8.4).

Consider evaluating (8.4) in the upper troposphere at the latitude of the Hadley Cell streamfunction maximum ϕ_m (Fig. 8.1). At this latitude, streamlines of the flow are quasi-horizontal[1], and therefore we may neglect the term involving ω on the left-hand side. Integrating with mass weighting from the level of the streamfunction maximum to the tropopause, we may write,

$$(1 - \mathrm{Ro}) \int_{p_t}^{p_m} [\overline{v}] \frac{dp}{g} = -\int_{p_t}^{p_m} \frac{S}{f} \frac{dp}{g}, \tag{8.6}$$

[1]While the arrows in Fig. 8.1 are not horizontal at the latitude ϕ_m, our approximation is still justified provided $f[\overline{v}] >> [\overline{\omega}]\partial_p[\overline{u}]$.

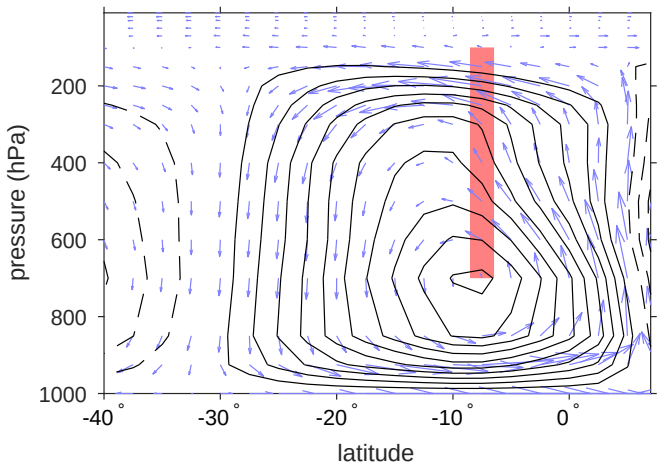

Figure 8.1: Annual-mean streamfunction [solid (anticlockwise) and dashed (clockwise) black lines; contour interval 1×10^9 kg s^{-1}] and meridional overturning mass flux (arrows) estimated based on the NCP-DOE reanalysis. Red line gives the integration path used to construct (8.8).

where we have defined the bulk Rossby number,

$$\mathrm{Ro} = -\frac{1}{f}\frac{\int_{p_t}^{p_m}[\overline{v}][\overline{\zeta}]\,\frac{dp}{g}}{\int_{p_t}^{p_m}[\overline{v}]\,\frac{dp}{g}}. \tag{8.7}$$

That is, the bulk Rossby number is defined as the negative of the ratio of the mass-flux-weighted average relative vorticity to the planetary vorticity. Using the definition of the streamfunction (2.63), we may relate the integral of $[\overline{v}]$ to the maximum in the streamfunction $\Psi_{\max}$, so that we may write,

$$(1-\mathrm{Ro})\Psi_{\max} = -\frac{\langle S\rangle}{f}, \tag{8.8}$$

where we have defined $\langle S\rangle = 2\pi R_e \cos\phi \int_{p_t}^{p_t} S\,\frac{dp}{g}$.

Equation (8.8) provides a relationship between the Hadley Cell strength (as measured by its maximum mass flux) and a quantity proportional to the upper-tropospheric eddy-momentum flux convergence at the latitude of the streamfunction maximum. This relationship is mediated by the non-dimensional number Ro, which we have identified as the bulk Rossby number.

8.1.2 Linear and angular-momentum conserving limits

To understand (8.8) it is helpful to consider two limits. In the first limit, we assume the relative vorticity is much smaller than the planetary vorticity in the subtropical atmosphere,

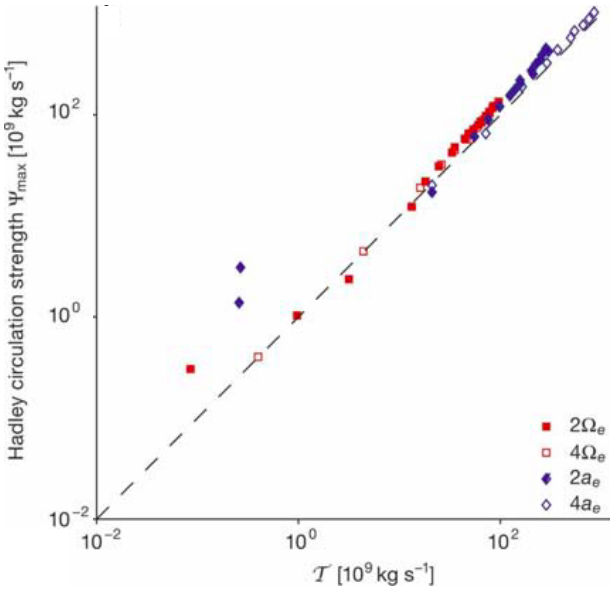

Figure 8.2: Hadley Cell strength $\Psi_{\max}$ plotted against a measure of the subtropical eddy momentum flux divergence $\mathcal{T}$ for idealised simulations with different planetary rotation rates Ω_e and planetary radii a_e (colours). Simulations were conducted with an idealised dry GCM as detailed in Walker and Schneider (2006). Figure adapted from Fig. 6b of Walker and Schneider (2006).

and as a result Ro$<<1$. In this regime, the strength of the Hadley Cell is directly related to the eddy momentum flux divergence in the upper tropsophere. Any change in the strength of the Hadley Cell must therefore occur via a change in these eddies; the Hadley Cell does not respond directly to thermal driving! The low bulk Rossby number limit is the limit in which QG scaling is valid, and it is the relevant limit for the extratropics. As we showed in chapter 5, the strength of the Ferrel cell is directly tied to the eddy driving.

The second limit to consider is one in which the eddy fluxes are small and angular momentum is quasi-conserved along streamlines. In this limit, regions of quasi-horizontal flow must have that $\partial_\phi[\overline{M}] \approx 0$. This implies that,

$$[\bar{\zeta}] \approx -f, \tag{8.9}$$

and the Rossby number $Ro \approx 1$. In this regime, the left and right hand sides of (8.8) both approach zero, and the equation tells us nothing about the strength of the Hadley Cell. This is the relevant limit for the Held-Hou model of the Hadley Cell, and in this limit, the circulation strength is directly controlled by energetic constraints; the Hadley Cell can respond to thermal driving directly.

8.1.3 What is the bulk Rossby number for the Hadley Cell?

The above discussion suggests that a key question for understanding the Hadley Cell on Earth is the extent to which it may be described in one or the other of the above limits. Walker and Schneider (2006) provided some hints at the answer to this question. The authors conducted a suite of simulations of the equinoctial circulation with an idealised dry general circulation model forced by Newtonian relaxation to an imposed "radiative equilibrium" state. Starting from an Earth-like control simulation, the planetary rotation rate, planetary radius, radiative equilibrium temperature profile, and convective lapse rate were varied over a broad range. Fig. 8.2 summarises their results. It plots, for a number of simulations with varying planetary rotation rates and planetary radii, the strength of the simulated Hadley Cell $\Psi_{\max}$ against a measure of the strength of the subtropical eddy momentum flux divergence $\mathcal{T} = -\langle S \rangle / f$ evaluated at the latitude of the Hadley Cell streamfunction maximum. Almost all simulations lie close to the one-to-one line, indicating that Ro$<<$ 1. Those simulations that do not lie close to the one-to-one line have relatively weak Hadley Cells. This implies that, over a broad range of parameters, the equinoctial Hadley Cell remains close to the linear, low Rossby number regime where eddy-momentum fluxes play a role in determining the Hadley Cell strength.

Walker and Schneider (2006) further showed that the response of the Hadley Cell strength to changes in planetary parameters did not agree with the theoretical predictions based on the Held-Hou model. Instead, the Hadley Cell in their simulations behaved in a manner consistent with the notion that the Hadley Cell strength is controlled by extratropical processes, and in particular by eddies propagating from the midlatitudes and breaking in the subtropics thereby producing a flux of angular momentum from the subtropics to the midlatitudes.

Consistent with the idealised discussed above, observational estimates suggest that, in the annual mean, the ratio $[\overline{\zeta}]/f \lesssim 0.4$ in the subtropical upper tropsophere, implying that Earth's Hadley Cell is relatively close to the linear regime (Fig. 8.3). Indeed, for the annual mean, contours of angular momentum, which for a resting fluid are vertical, are only moderately perturbed by the circulation, and streamlines cross angular-momentum contours regularly in the descending branch. Recall that in the Held-Hou model, streamlines are exactly parallel to angular momentum contours; the observed annual-mean Hadley Cell is rather far from this limit!

The above results suggest a strong influence of extratropical processes on the Hadley cell for annual-mean or equinoctial conditions. However, as shown in chapter 1, Earth's tropical circulation undergoes a large seasonal shift from an equinoctial regime, with two Hadley Cells of roughly equal strength, to a solsticial regime, in which strong monsoon circulations contribute to a dominant winter cell. For this solsticial circulation, the ratio $[\overline{\zeta}]/f$ reaches ~ 0.8 in regions of the Hadley Cell's upper branch, indicating a higher bulk Rossby number, and implying that the circulation is closer to the nonlinear, angular-momentum conserving

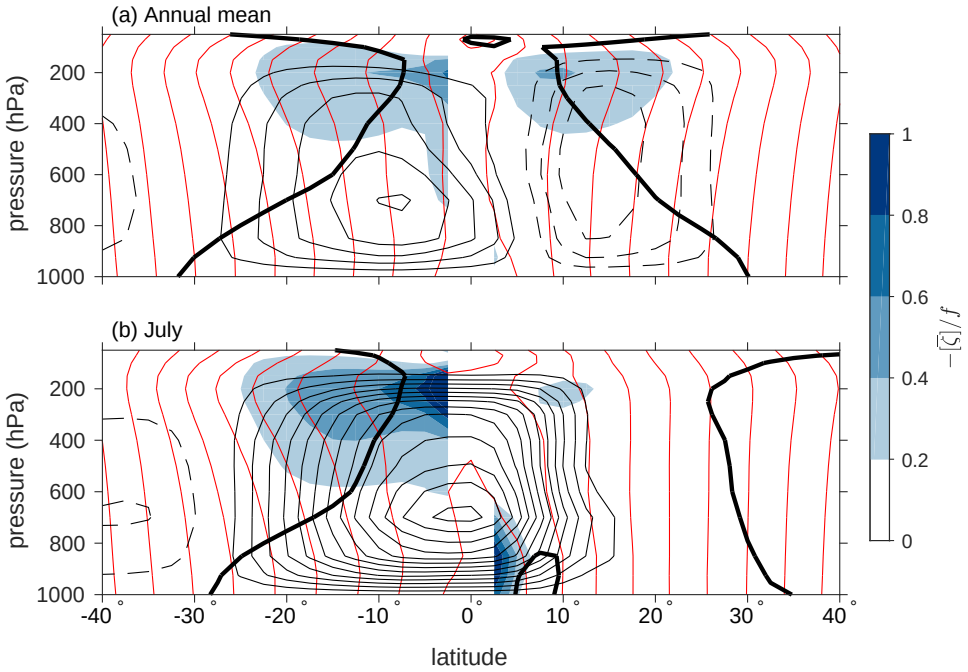

Figure 8.3: Streamfunction [thin black solid (anticlockwise) and dashed (clockwise) lines; contour interval 2×10^9 kg s^{-1}], angular-momentum contours (contoured for values of angular momentum equal to that of the Earth at latitudes of $0°, \pm 4°, \pm 8°, ...$), and ratio $-[\bar{\zeta}]/f$ (colours; latitudes equatorward of 5° masked out) according to the NCEP-DOE relanalysis for (a) the annual mean and (b) July. Thick black line gives the zero contour of the zonal- and time-mean zonal wind.

regime. Understanding the reasons for this seasonal variation in the character of the Hadley Cell, and the implications for the tropical general circulation, are the subject of the next section.

8.2 Monsoons and the solsticial Hadley Cell

8.2.1 Geometrical constraints on the Hadley Cell

To understand some of the differences between the equinoctial and solsticial Hadley Cells, it is useful to consider the geometry of angular momentum contours. Recall that, under our shallow fluid approximation, the planetary angular momentum is given by,

$$M_p = R_e^2 \cos^2(\phi). \tag{8.10}$$

Near to the equator, we may approximate this as,

$$M_p \approx R_e^2 (1 - \phi^2), \tag{8.11}$$

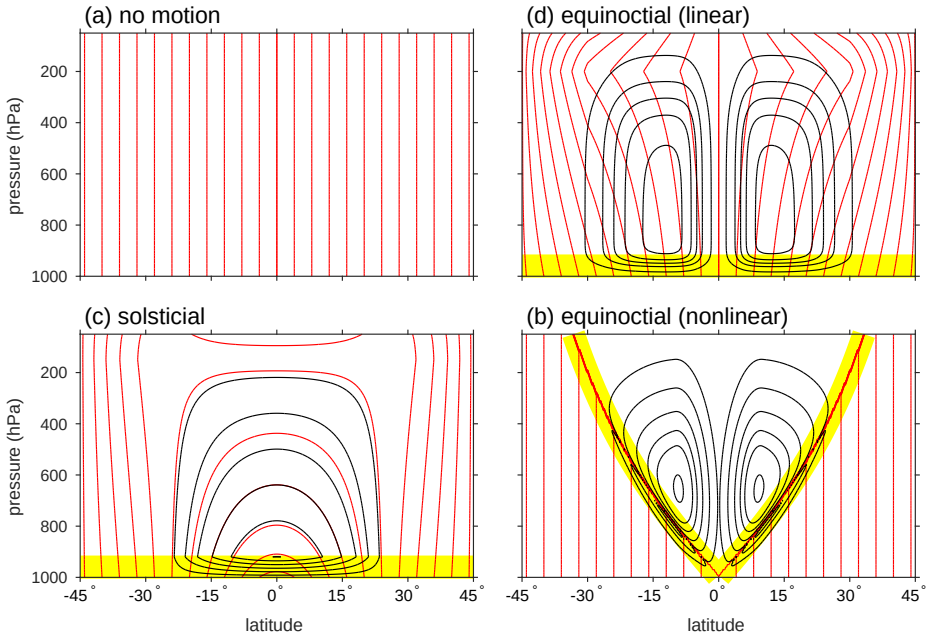

Figure 8.4: Contours of angular momentum and the streamfunction for idealised Hadley Circulations of different "shapes". (a) resting atmosphere, (b) angular-momentum conserving equinoctial Hadley Cell in the nonlinear Ro ≈ 1 regime (c) angular-momentum conserving solsticial Hadley Cell, and (d) eddy-influenced equinoctial Hadley Cell in the linear Ro $<< 1$ regime. Yellow shading represents boundary layers associated with the Earth's surface or with a discontinuity in the zonal wind.

indicating that the planetary angular momentum decreases roughly quadratically away from its maximum at the equator. Contours of angular momentum for a resting atmosphere are therefore vertical; Fig. 8.4a shows such contours at values of angular momentum corresponding to the planetary angular momentum at intervals of four degrees of latitude.

Now let us consider how these contours are perturbed by an equinoctial Hadley Cell. The observed equinoctial circulation includes a subtropical jet, which implies a poleward displacement of the M contours. However, observational estimates indicate that this displacement is relatively weak, and streamlines of the Hadley Cell cross angular momentum contours throughout the descending branch, as depicted in Fig. 8.4b. This is characteristic of a Hadley Cell in the linear regime.

In contrast, if we consider an angular-momentum conserving Hadley Cell, streamlines cannot cross angular momentum contours, except within boundary layers. As pointed out in section 4.2.2, this requires the descending branch of the Hadley Cell to be sloped so that

air following a streamline reaches the surface near the equator. As shown in Fig. 8.4d, this produces an angular momentum distribution with no contours in the region of the Hadley Cell, and a sharp discontinuity in angular momentum at its sloping edge [see Fang and Tung (1996) for detailed solutions of this "sloping" Hadley Cell]. For the equinoctial case, conservation of angular momentum along streamlines implies a distribution of angular momentum that is homogenised throughout the Cell.

Now let us consider the cross-equatorial cell in the solsticial case. Air rises from the surface at a latitude ϕ_I in the summer hemisphere with angular momentum $M_I = M_p(\phi_I)$, reaching the tropopause and flowing across the equator and descending on the winter side. Assuming angular-momentum is conserved, the angular momentum contour M_I also follows this trajectory. For the cross equatorial cell, this contour may intersect the surface at latitude $-\phi_I$, at which point the implied zonal wind is zero, while the associated streamline may be closed within the boundary layer (Fig 8.4d). Streamlines may be parallel to angular-momentum contours in the free troposphere without requiring either a homogenised angular-momentum distribution or an exotic Hadley Cell geometry.

The above considerations of how to close streamlines while ensuring they do not cross angular momentum contours outside of the boundary layer provides a hint as to why we might expect the solsticial, cross-equatorial cell to be more likely to reside close to the angular-momentum conserving regime. We now consider the solsticial cell more quantitiatively by generalising the Held and Hou (1980) model to the case in which the forcing maximises off the equator.

8.2.2 Extending Held & Hou to the solstice: the Lindzen & Hou model

The angular-momentum conserving axisymmetric solution of Held and Hou (1980) was extended to the off-equatorial case by Lindzen and Hou (1988). As in chapter 4, the model was originally phrased in terms of a Boussinesq fluid, but we will assume an atmosphere in convective quasi-equilibrium here. We generalise the RCE distribution of boundary-layer entropy (3.37) to the case in which the maximum occurs at a latitude ϕ_0,

$$s_b^{\mathrm{RCE}} = s_{b0}^{\mathrm{RCE}} - \delta s_b^{\mathrm{RCE}} (\sin\phi - \sin\phi_0)^2. \tag{8.12}$$

Note that this distribution of entropy has a non-zero gradient at the equator and therefore automatically violates Hide's theorem.

Like the Held-Hou model, the Lindzen and Hou (1988) model assumes that air rises in the ITCZ to the upper troposphere, carrying with it the angular momentum of the surface, and then flows both north and south conserving its angular momentum. Denoting the ITCZ latitude ϕ_I (which in general will differ from the forcing maximum ϕ_0), this implies that the upper-tropospheric zonal velocity within the region of the Hadley Cells is given

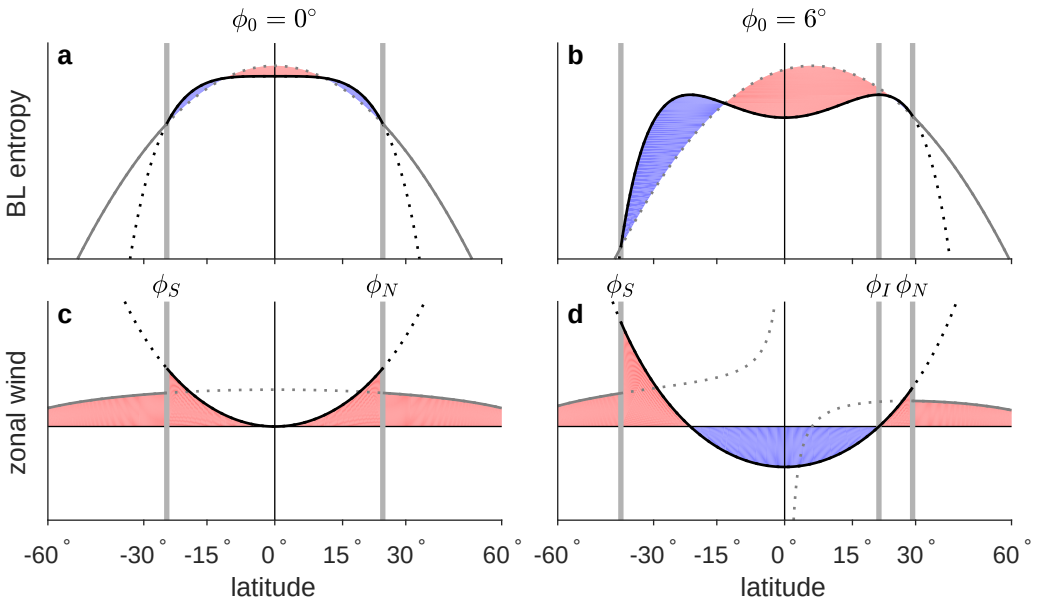

Figure 8.5: Boundary layer entropy (top) and zonal wind (bottom) for the Lindzen and Hou (1988) model of the Hadley Cell under (a,c) equinoctial conditons [$\phi_0 = 0°$; equivalent to Held and Hou (1980)] and (b,d) solsticial conditions ($\phi_0 = 6°$). Black line shows angular-momentum conserving solution, valid between the latitudes ϕ_S and ϕ_N, and gray line shows RCE solution, valid outside this region. Solutions are plotted using solid lines in regions where they are valid and using dotted lines otherwise. In top panels, colours show regions of radiative-convective heating (red) and cooling (blue), and in bottom panels colours show westerly (red) and easterly (blue) zonal winds. In right panels, ϕ_I marks the latitude of the boundary between the summer and winter Hadley Cells, corresponding to the maximum in boundary-layer entropy and the zero zonal wind line.

by,

$$u_t^{\mathrm{AMC}} = \Omega R_e \frac{(\cos^2\phi_I - \cos^2\phi)}{\cos\phi}. \tag{8.13}$$

Assuming that the surface winds are weak, we may use the thermal wind relation (3.35) to derive an expression for the boundary layer entropy given by,

$$s_b^{\mathrm{AMC}} = s_{b0} - \frac{\Omega^2 R_e^2}{2(T_b - T_t)} \frac{(\cos^2\phi_I - \cos^2(\phi))^2}{\cos^2\phi}. \tag{8.14}$$

The unknowns s_{b0} and ϕ_I are then constrained by assuming that $s_b^{\mathrm{AMC}} = s_b^{\mathrm{RCE}}$ at the northern and southern boundaries ϕ_N and ϕ_S, and by assuming each cell is energetically

closed, so that,

$$\int_{\phi_S}^{\phi_I} (s_b^{\mathrm{AMC}} - s_b^{\mathrm{RCE}}) \cos\phi \, d\phi = 0,$$
$$\int_{\phi_I}^{\phi_N} (s_b^{\mathrm{AMC}} - s_b^{\mathrm{RCE}}) \cos\phi \, d\phi = 0.$$

An example of such a solution is shown in Fig. 8.5 for $\phi_0 = 6^\circ$ along with the original Held-Hou solution recovered by setting $\phi_0 = 0$ (holding all other parameters the same). Some of the main features of the solution are,

- The requirement of angular-momentum conservation in the upper branch implies that both the zonal wind and boundary-layer entropy are symmetric about the equator within the region of the Hadley Cells. This implies a maximum in s_b within the winter hemisphere. Such a secondary maximum is not observed in the zonal mean solsticial temperature or boundary-layer entropy distributions, indicating that the flow is somewhat outside the fully nonlinear regime, despite the higher bulk Rossby number.
- The latitude of the ITCZ, ϕ_I, corresponding to the maximum in boundary-layer entropy is substantially poleward of the forcing maximum ϕ_0.
- Equatorward of ϕ_I, the tropopause level winds are easterly; in the ϕ_0 case, the tropopause level winds are everywhere westerly.
- The implied transport of heat by the cross equatorial cell is much larger than the equinoctial cells, as indicated by the size of the coloured regions in Fig. 8.5. Assuming the gross moist stability of the cells are the same, this implies that the cross-equatorial cell is much stronger than the equinoctial cells. This strengthening occurs nonlinearly, and requires only a shift of ϕ_0 only a few degrees from the equator.

Some of the above features of the Lindzen & Hou model appear to be relevant to Earth's solsticial circulation. For example, we have seen that the solsticial cross-equatorial Hadley Cell is much stronger than its equinoctial counterparts, consistent with the behaviour above. Furthermore, the easterly wind shear associated with temperatures that increase poleward may be seen in Earth's solsticial circulation as well as in Fig. 8.5. This latter feature may be of particularly importance for governing the seasonal cycle of the tropical overturning circulation for reasons we discuss below.

8.2.3 Eddy-mean flow interactions and monsoons

In the previous sections we have argued that Earth's Hadley Cell varies in its proximity to the nonlinear angular momentum conserving regime depending on the time of year. In

the equinoctial regime, the bulk Rossby number is relatively small, and eddies are thought to play a substantial role in influencing the strength of the Hadley Cell (Fig. 8.3). The cross-equatorial solsticial Hadley Cell, on the other hand, achieves higher values of the bulk Rossby number, and it has been argued to have some similarities with the angular-momentum conserving model of Lindzen and Hou (1988) [although it should be noted that evidence for an influence from midlatitude eddies has been found in the solsticial case as well (Caballero, 2007, 2008)]. In this section we briefly discuss some reasons why we might expect the Solsticial Cell to be closer to the nonlinear regime, and some of the implications of this regime transition for the tropical seasonal cycle.

To begin, we note that, the model of Lindzen and Hou (1988) produces easterly zonal winds equatorward of the cross-equatorial cell's rising branch at ϕ_I and westerly winds elsewhere (Fig. 8.5). This means that, for the equinoctial case, the winds are westerly everywhere, but when the rising branch is displaced sufficiently far from the equator, the equatorial region contains upper tropospheric easterlies. Also recall that Rossby waves can only propagate meridionally in regions in which the zonal phase speed c satisfies $[u] - c > 0$. Finally, note that the energy of disturbances in the atmosphere is dominated by phase speeds $c > 0$ (e.g., Fig. 6.6). As a result, Rossby wave propagation is prevented in regions of easterly winds, and this limits the influence of midlatitude eddies on the solsticial Hadley Cell.

The above argument was made based on the zonal wind distribution produced by the Lindzen and Hou (1988) model, but it may also be seen in estimates of the observed wind distribution (Fig. 8.3). For the annual mean, the upper-tropospheric zonal winds are westerly poleward of about 10° either side of the equator, leaving the centre of the equinoctial cells able to be penetrated by Rossby waves. During the solstice, however, the centre of the cross-equatorial cell is shielded from midlatitude eddies by easterly winds throughout the troposphere, thus explaining the cell's proximity to the nonlinear regime.

The transition of the tropical circulation from the equinoctial, eddy-influenced regime to a solsticial, angular-momentum conserving regime has been argued to be relevant to the rapid onset of monsoons on Earth. In particular, Bordoni and Schneider (2008) argued that the shielding of the Hadley Cell rising branch from midlatitude eddies by equatorial easterlies immediately before monsoon onset allows for the subsequent transition of the cell to the nonlinear regime (see also Schneider and O'Gorman, 2008). Once in this regime, the Hadley Cell strengthens rapidly as the maximum in forcing shifts further from the equator as was found in the Lindzen and Hou (1988) model. The eddy-mean flow and nonlinear feedbacks associated with this process allow for a rapid amplification of the cross-equatorial cell that in some respects resembles the onset of the Asian monsoon in the Boreal summer.

The above argument implies that an important aspect of monsoon onset may be produced simply by feedbacks associated with the zonal-mean flow and eddies. Bordoni and Schneider (2008) argues that this implies that land-sea contrasts are unnecessary for the development

of monsoons, rather, the important aspect is the existence of a surface with low thermal inertia. This is part of a greater body of work which has sought to replace the soewhat dated view of monsoons as large-scale sea breezes with the more modern view of the monsoon as the regional manifestation of the cross-equatorial solsticial Hadley Cell (e.g., Gadgil, 2018). This view highlights the constraints placed on monsoons through the angular-momentum budget, and the connection of monsoons to moist convection, over the simplified view of the monsoon as a response to temperature differences between land and ocean.

Bibliography

Bordoni, S. and T. Schneider, 2008: Monsoons as eddy-mediated regime transitions of the tropical overturning circulation. *Nat. Geosci.*, **1**, 515–519, doi: 10.1038/ngeo248.

Caballero, R., 2007: Role of eddies in the interannual variability of Hadley cell strength. *Geophys. Res. Lett.*, **34**, L22 705, doi: 10.1029/2007GL030971.

Caballero, R., 2008: Hadley cell bias in climate models linked to extratropical eddy stress. *Geophys. Res. Lett*, **35**, L18 709, doi: 10.1029/2008GL035084.

Edmon, H. J., B. J. Hoskins, and M. E. McIntyre, 1980: Eliassen-Palm cross sections for the troposphere. *J. Atmos. Sci.*, **37**, 2600 – 2616, doi: 10.1175/1520-0469(1980)037⟨2600:EPCSFT⟩2.0.CO;2.

Emanuel, K., 2007: Quasi-equilibrium dynamics of the tropical atmosphere. *The Global Circulation of the Atmosphere*, T. Schneider and A. H. Sobel, Eds., Princeton University Press, chap. 7, 186–218.

Emanuel, K., A. A. Wing, and E. M. Vincent, 2014: Radiative-convective instability. *J. Adv. Model. Earth Syst.*, **6**, 75–90, doi: 10.1002/2013MS000270.

Emanuel, K. A., 1994: *Atmospheric convection*. Oxford University Press, 580 pp.

Emanuel, K. A., J. D. Neelin, and C. S. Bretherton, 1994: On large-scale circulations in convecting atmospheres. *Q. J. R. Meteorol. Soc.*, **120**, 1111–1143, doi: 10.1002/qj.49712051902.

Fang, M. and K. K. Tung, 1996: A simple model of nonlinear Hadley circulation with an ITCZ: Analytic and numerical solutions. *J. Atmos. Sci.*, **53**, 1241–1261, doi: 10.1175/1520-0469(1996)053⟨1241:ASMONH⟩2.0.CO;2.

Gadgil, S., 2018: The monsoon system: Land–sea breeze or the itcz? *Journal of Earth System Science*, **127**, 1.

Hartmann, D. L., 1994: *Global physical climatology*. Academic Press, 411 pp.

Held, I. and T. Schneider, 1999: The surface branch of the zonally averaged mass transport circulation in the troposphere. *J. Atmos. Sci.*, **56**, 1688–1697.

Held, I. M., 2000: The general circulation of the atmosphere. *Proc. Geophysical Fluid Dynamics Program*, Woods Hole Oceanographic Institution, 1–54, URL https://www.whoi.edu/fileserver.do?id=21464&pt=10&p=17332.

Held, I. M., 2001: The partitioning of the poleward energy transport between the tropical ocean and atmospheres. *J. Atmos. Sci.*, **58**, 943–948, doi: 10.1175/1520-0469(2001)058⟨0943:TPOTPE⟩2.0.CO;2.

Held, I. M. and A. Y. Hou, 1980: Nonlinear axially symmetric circulations in a nearly inviscid atmosphere. *J. Atmos. Sci.*, **37**, 515–533, doi: 10.1175/1520-0469(1980)037⟨0515:NASCIA⟩2.0.CO;2.

Hide, R., 1969: Dynamics of the Atmospheres of the Major Planets with an Appendix on the Viscous Boundary Layer at the Rigid Bounding Surface of an Electrically-Conducting Rotating Fluid in the Presence of a Magnetic Field. *J. Atmos. Sci.*, **26**, 841–853, doi: 10.1175/1520-0469(1969)026⟨0841:DOTAOT⟩2.0.CO;2.

Holton, J., 2004: *An introduction to dynamic meteorology*. 4th ed., Academic Press, 535 pp.

Hoskins, B., 1975: The geostrophic momentum approximation and the semi-geostrophic equations. *J. Atmos. Sci.*, **32**, 233–242.

Hoskins, B. J., M. E. McIntyre, and A. W. Robertson, 1985: On the use and significance of isentropic potential vorticity maps. *Quarterly. Journal of the Royal Meteorological Society*, **111**, 877 – 946.

Houghton, J., 2002: *The Physics of Atmospheres*. Cambridge University Press.

Hurley, J. V. and W. R. Boos, 2013: Interannual variability of monsoon precipitation and local subcloud equivalent potential temperature. *J. Climate*, **26**, 9507–9527, doi: 10.1175/JCLI-D-12-00229.1.

Inoue, K. and L. E. Back, 2015: Gross moist stability assessment during TOGA COARE: various interpretations of gross moist stability. *J.Atmos. Sci.*, **72**, 4148–4166.

Inoue, K. and L. E. Back, 2017: Gross moist stability analysis: Assessment of satellite-based products in the gms plane. *J. Atmos. Sci.*, **74**, 1819–1837.

Jansen, M. and R. Ferrari, 2012: Macroturbulent equilibration in a thermally forced primitive equation system. *J. Atmos. Sci.*, **69**, 695–713.

Juckes, M., 2000: The static stability of the midlatitude troposphere: The relevance of moisture. *J. Atmos. Sci.*, **57**, 3050–3057.

Kanamitsu, M., W. Ebisuzaki, J. Woollen, S. Yang, J. Hnilo, M. Fiorino, and G. Potter, 2002: NCEP-DOE AMIP Reanalysis (R-2). *Bull. Amer. Meteor. Soc*, **83**, 1631–1643.

Lindzen, R. S. and A. Y. Hou, 1988: Hadley circulations for zonally averaged heating centered off the equator. *J. Atmos. Sci.*, **45**, 2416–2427, doi: 10.1175/1520-0469(1988)045⟨2416:HCFZAH⟩2.0.CO;2.

Lorenz, E., 1983: A history of prevailing ideas about the general circulation of the atmosphere. *Bulletin of the American Meteorological Society*, **64**, 730–769.

Manabe, S. and R. Strickler, 1964: Thermal equilibrium of the atmosphere with a convective adjustment. *J. Atmos. Sci.*, **21**, 361–385.

Neelin, J. D. and I. M. Held, 1987: Modeling tropical convergence based on the moist static energy budget. *Mon. Wea. Rev.*, **115**, 3–12, doi: 10.1175/1520-0493(1987)115⟨0003:MTCBOT⟩2.0.CO;2.

Nie, J., W. R. Boos, and Z. Kuang, 2010: Observational evaluation of a convective quasi-equilibrium view of monsoons. *J. Climate*, **23**, 4416–4428, doi: 10.1175/2010JCLI3505.1.

Osborn, T. J. and P. D. Jones, 2014: The CRUTEM4 land-surface air temperature data set: construction, previous versions and dissemination via Google Earth. *Earth System Science Data*, **6**, 61–68, doi: 10.5194/essd-6-61-2014.

Pauluis, O. and I. M. Held, 2002: Entropy budget of an atmosphere in radiative-convective equilibrium. Part I: Maximum work and frictional dissipation. *J. Atmos. Sci.*, **59**, 125–139.

Peixoto, J. P. and A. H. Oort, 1992: *Physics of Climate*. AIP Press, 520 pp.

Privé, N. C. and R. A. Plumb, 2007a: Monsoon dynamics with interactive forcing. Part I: Axisymmetric studies. *J. Atmos. Sci.*, **64**, 1417–1430, doi: 10.1175/JAS3916.1.

Privé, N. C. and R. A. Plumb, 2007b: Monsoon dynamics with interactive forcing. Part II: Impact of eddies and asymmetric geometries. *J. Atmos. Sci.*, **64**, 1431–1442, doi: 10.1175/JAS3917.1.

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

Raymond, D., 2000: The hadley circulation as a radiative–convective instability. *J. Atmos. Sci.*, **57**, 1286–1297.

Raymond, D., Ž. Fuchs, S. Gjorgjievska, and S. Sessions, 2015: Balanced dynamics and convection in the tropical troposphere. *J. Adv. Model. Earth Syst.*, **7**, 1093–1116, doi: 10.1002/2015MS000467.

Raymond, D. J., S. L. Sessions, A. H. Sobel, and Ž. Fuchs, 2009: The mechanics of gross moist stability. *J. Adv. Model. Earth Syst.*, **1**, doi: 10.3894/JAMES.2009.1.9.

Rhines, P. B., 1975: Waves and turbulence on a beta-plane. *Journal of Fluid Mechanics*, **69**, 417–443.

Salmon, R., 1998: *Lectures on Geophysical Fluid Dynamics*. Oxford University Press, 378 pp.

Satoh, M., M. Shiobara, and M. Takahashi, 1995: Hadley circulations and their rôles in the global angular momentum budget in two-and three-dimensional models. *Tellus*, **47A**, 548–560, doi: 10.1034/j.1600-0870.1995.00104.x.

Schneider, T. and P. A. O'Gorman, 2008: Moist convection and the thermal stratification of the extratropical troposphere. *J. Atmos. Sci.*, **65**, 3571–3583.

Singh, M. S., 2019: Limits on the extent of the solsticial hadley cell: The role of planetary rotation. *J. Atmos. Sci.*, **76**, 1989–2004.

Singh, M. S. and Z. Kuang, 2016: Exploring the role of eddy momentum fluxes in determining the characteristics of the equinoctial Hadley circulation: fixed-SST simulations. *J. Atmos. Sci.*, **73**, 2427–2444, doi: 10.1175/JAS-D-15-0212.1.

Singh, M. S. and P. A. O'Gorman, 2013: Influence of entrainment on the thermal stratification in simulations of radiative-convective equilibrium. *Geophys. Res. Lett.*, **40**, 4398–4403, doi: 10.1002/grl.50796.

Sobel, A. H., J. Nilsson, and L. M. Polvani, 2001: The weak temperature gradient approximation and balanced tropical moisture waves. *J. Atmos. Sci.*, **58**, 3650–3665, doi: 10.1175/1520-0469(2001)058⟨3650:TWTGAA⟩2.0.CO;2.

Starr, V., 1968: *Physics of negative viscosity phenomena*. McGraw-Hill Education.

Stone, P., 2005: General circulation of the Earth's atmosphere. *MIT Open Courseware*, Woods Hole Oceanographic Institution, URL https://ocw.mit.edu/courses/earth-atmospheric-and-planetary-sciences/12-812-general-circulation-of-the-earths-atmosphere-fall-2005/.

Vallis, G. K., 2017: *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. 2nd ed., Cambridge University Press, 946 pp.

Vallis, G. K., 2019: *Essentials of Atmospheric and Oceanic Dynamics*. Cambridge University Press, 366 pp.

Walker, C. C. and T. Schneider, 2006: Eddy influences on Hadley circulations: simulations with an idealized GCM. *J. Atmos. Sci.*, **63**, 3333–3350, doi: 10.1175/JAS3821.1.

Wallace, J. and P. Hobbs, 2006: *Atmospheric science: an introductory survey.* 2d ed., Academic Press.

Xie, P. and P. A. Arkin, 1997: Global precipitation: A 17-year monthly analysis based on gauge observations, satellite estimates, and numerical model outputs. *Bull. Amer. Met. Soc.*, **78**, 2539–2558, doi: 10.1175/1520-0477(1997)078⟨2539:GPAYMA⟩2.0.CO;2.

Xu, K.-M. and K. A. Emanuel, 1989: Is the tropical atmosphere conditionally unstable? *Mon. Wea. Rev.*, **117**, 1471–1479, doi: 10.1175/1520-0493(1989)117⟨1471:ITTACU⟩2.0.CO;2.