Preface

The 2014 Geophysical Fluid Dynamics Summer Study Program started on June 16th, with the topic of *Climate Physics and Dynamics*. The topic proved very timely and attracted an unprecedented number of applications from brilliant students. Professors Kerry Emanuel (MIT) and Geoff Vallis (Exeter) gave the principal lectures. They began with the simplest energy balance models and then included adjustment of the vertical profiles by convection (dry and moist). Kerry delved more deeply into convection and the processes found in "cloud-permitting" models, including island effects and the spontaneous formation of clusters surrounded by dry regions. Geoff discussed the larger-scale dynamics of the atmosphere and oceans, including the transports by eddies and the thermohaline circulation. This year, we webcast the principal lectures, with order 50 connections per day from viewers at many universities worldwide.

Antonello Provenzale, Raffaele Ferrari, and Glenn Flierl co-directed the summer and invited a large number of top researchers in the dynamics of climate, so that we had a broad critical view on the status of modelling and theory. As usual, the visitors presented new ideas and approaches to geophysical dynamics, and the visitors, staff, and fellows engaged in intense discussions on the porch and in the lecture room. The long-term staff members ensured that the fellows never lacked for guidance; many of them explored unfamiliar topics but were still able to develop new insights.

This summer's fellows were

- Jörn Callies, MIT/WHOI Joint Program in Oceanography
- Joseph Fitzgerald, Harvard University
- Shineng Hu, Yale University
- Alexis Kaminski, University of Cambridge
- Paige Martin, University of Michigan
- Daniel Mukiibi, University of Hamburg
- Ashley Payne, University of California, Irvine
- Erica Rosenblum, University of California, San Diego
- Geoff Stanley, University of Oxford
- Jim Thomas, New York University

In 2014, the Sears Public Lecture was delivered by Professor Cecilia Bitz, University of Washington, who discussed "The Future of Arctic and Antarctic Sea Ice." She noted that both the Arctic and Antarctic have experienced record sea ice coverage in the last decade. The Arctic has experienced record losses, especially in summer, while the Antarctic has had expanding sea ice. She considers how this is possible in an era of global climate change, and what we can expect in the future. Cecilia overcame problems with the microphone and engaged the audience both in answering question and in discussions at the reception and refreshments following the talk.

Anne Doucette, Julie Hildebrandt, and Janet Fields made sure that the administrative side of the program ran with admirable efficiency. Matt Barton ensured the webcast was high quality. We continue to be indebted to WHOI for their support and to the Academic Programs Office, who once more (with the cooperation of the weather!) provided a perfect atmosphere.

Table of Contents

2014 Geophysical Fluid Dynamics Participants

STAFF AND VISITORS

Abernathey, Ryan	Columbia University
Armour, Kyle	Massachusetts Institute of Technology
Barbante, Carlo	University of Venice
Bitz, Cecilia	University of Washington
Bordoni, Simona	California Institute of Technology
Caulfield, Colm-cille	Cambridge University
Cessi, Paola	University of California, San Diego
Chini, Greg	University of New Hampshire
Corti, Susanna	Inst. of Atmospheric Sciences & Climate, CNR
Cvitanovic, Predrag	Georgia Institute of Technology
Doering, Charlie	University of Michigan
Du, Hailiang	University of Chicago
Emanuel, Kerry	Massachusetts Institute of Technology
Ferrari, Raffaele	Massachusetts Institute of Technology
Flierl, Glenn	Massachusetts Institute of Technology
Fox-Kemper, Baylor	Brown University
Gentine, Pierre	Columbia University
Goluskin, David	Columbia University
Helfrich, Karl	Woods Hole Oceanographic Institution
Hussein, Aluie	Johns Hopkins University
Huybers, Peter	Harvard University
Ingersoll, Andy	California Institute of Technology
Jansen, Malte	Geophysical Fluid Dynamics Laboratory, Princeton
Johnson, Ted	University College London
Keller, Joseph	Stanford University
Kerswell, Richard	University of Bristol
Kuang, Zhiming	Harvard University
Lebovitz, Norman	University of Chicago
Levy, Marina	Laboratoire d'Ocanographie et du Climat
Mahadevan, Amala	Woods Hole Oceanographic Institution
Marshall, John	Massachusetts Institute of Technology
Meacham, Stephen	National Science Foundation
Morrison, Adele	Princeton University
Morrison, Phil	University of Texas, Austin
Muller, Caroline	Lab. d'Hydrodynamique de l'cole Polytechnique
Pedlosky, Joe	Woods Hole Oceanographic Institution
Provenzale, Antonello	Inst. of Atmospheric Sciences & Climate, CNR
Roe, Gerard	University of Washington
Schneider, Tapio	ETH Zurich
Shaw, Tiffany	Columbia University

Smith , Lenny	London School of Economics
Sobel, Adam	Columbia University
Spall, Mike	Woods Hole Oceanographic Institution
Spiegel, Edward	Columbia University
Straneo, Fiamma	Woods Hole Oceanographic Institution
Talley, Lynne	Scripps Research Institute
Taylor, John	University of Cambridge
Thiffeault, Jean-Luc	University of Wisconsin
Thomas, Jim	New York University
Thompson, Andy	California Institute of Technology
Timmermans, Mary-Louise	Yale University
Tziperman, Eli	Harvard University
Vallis, Geoff	University of Exeter
Veronis, George	Yale University
Von Hardenberg, Jost	ISAC-CNR
von Herzen, Brian	The Climate Foundation
Weiss, Jeff	University of Colorado
White, Brian	University of North Carolina
Wingate, Beth	University of Exeter
Wunsch, Carl	Harvard University
Yang, Da	California Institute of Technology
Yecko, Philip	Montclair State
Young, Bill	University of California, San Diego

GUEST STUDENTS

Bebieva, Yana	Yale University
Benavides, Santiago Jose	University of Texas, Austin
Miles, Christopher	University of Michigan
Shebley, Nicole	Yale University
Souza, Andre	University of Michigan
Wang, Lei	University of Chicago
Zhao, Mengnan	Yale University

2014 Principal Lecturers

Geoff Vallis and Kerry Emanuel

2014 Geophysical Fluid Dynamics Summer School Participants

1st Row (L-R): Jörn Callies, Jim Thomas, Geoff Stanley, Daniel Mukiibi, Erica Rosenblum, Geoff Vallis, Paige Martin, Ashley Payne, Kerry Emanuel, Alexis Kaminski, Joe Fitzgerald, Shineng Hu, Junyi Chai, Amala Mahadevan

2nd Row (L-R): Stephen Meacham (standing), Charles Doering, David Goluskin, Santiago Benavides, Andy Ingersoll, Phil Morrison, Joseph Keller, George Veronis, Adele Morrison, Simona Bordoni, Jeff Weiss, Rich Kerswell, Norm Lebovitz, Brian von Herzen, Mary-Louise Timmermans

3rd Row (L-R): Greg Chini, Lei Wang, Christopher Miles, Andre Souza, Ted Johnson, Martin Jucker, Lenny Smith, Da Yang, Glenn Flierl, Antonello Provenzale, Raffaele Ferrari, Joe Pedlosky, Beth Wingate, Bradley Hubbard-Nelson, Karl Helfrich

Missing from photo: Ryan Abernathey, Hussein Aluie, Kyle Armour, Carlo Barbante, Yana Bebieva, Cecilia Bitz, Colm-cille Caulfield, Paola Cessi, Susanna Corti, Predrag Cvitanovic, Hailiang Du, Baylor Fox-Kemper, Pierre Gentine, Peter Huybers, Malte Jansen, Zhiming Kuang, John Marshall, Caroline Muller, Gerard Roe, Hesam Salehipour, Tapio Schneider, Tiffany Shaw, Nicole Shebley, Adam Sobel, Mike Spall, Ed Spiegel, Fiamma Straneo, Lynne Talley, John Taylor, Jean-Luc Thiffeault, Andy Thompson, Eli Tziperman, Jost von Hardenberg, Brian White, Phil Yecko, Bill Young, Mengnan Zhao

Schedule

Principal Lectures

Monday, June 16: Elements of Radiative Transfer
Kerry Emanuel
Tuesday, June 17: Energy Balance Models and Simple Radiative Convective Equilibria (RCE)
Geoff Vallis
Wednesday, June 18: Convective Heat Transfer 1
Kerry Emanuel
Thursday, June 19: RCE and Tropopause Height
Geoff Vallis
Friday, June 20: Convective Heat Transfer II
Kerry Emanuel
Monday, June 23: Water and Radiation in the Climate System
Kerry Emanuel
Tuesday, June 24: Stability of RCE and Rotating RCE
Kerry Emanuel
Wednesday, June 25: Horizontal Convection
Geoff Vallis
Thursday, June 26: The Wind-driven Deep Ocean Circulation
Geoff Vallis
Friday, June 27: Surface Winds
Geoff Vallis

Seminars

Monday, June 30: Climate-biosphere Interaction across Multiple Scales
Antonello Provenzale, Institute of Atmospheric Sciences and Climate, CNR
Planetary Climates: Some Advantages of Sparse Data
Andy Ingersoll, California Institute of Technology
Tuesday, July 1: Mixing Stuff Up
Glenn Flierl, Massachusetts Institute of Technology
Wednesday, July 2: On the Relationship between GFD and Atmosphere/Ocean Models in the Era of Exascale Computing
Beth Wingate, University of Exeter, UK
Probabilistic Plume Model to Represent Dry, Shallow and Deep Convection
Pierre Gentine, Columbia University
Thursday, July 3: Baroclinic Equilibration with Topography
Ryan Abernathey, Columbia University
Finite-amplitude Wavepackets in Internal Waves in a Rotating Frame
Ted Johnson, University College London
Monday, July 7: Parameters, Probabilities and Insight(s)
Lenny Smith, London School of Economics and Political Science
On the Reliability of Multi-Year Forecasts of Climate

Susanna Corti, Institute of Atmospheric Sciences and Climate of the Italian National Research Council

Tuesday, July 8: Metriplecticism: Relaxation Paradigms for Computation and Derivation

Phil Morrison, University of Texas at Austin

Wednesday, July 9: Forcing and Feedback from Months to Millions of Years

Peter Huybers, Harvard University

Regional Climate Predictability from Regional Climate Feedbacks

Gerard Roe, University of Washington

Thursday, July 10: Low Clouds and Their Response to Climate Change: A GFD Problem Waiting to be Solved

Tapio Schneider, Swiss Federal Institute of Technology, Zurich

Mixing Mechanisms and Dynamics of the Deep Arctic Ocean

Mary-Louise Timmermans, Yale University

Friday, July 11: A Competition between Turbulence and Baroclinic Instability in the Upper Ocean

John Taylor, University of Cambridge

Retreating Greenland Glaciers and Warming Ocean: Dynamics at the Ice/Ocean Interface

Fiamma Straneo, Woods Hole Oceanographic Institution

Monday, July 14: Why are all my Power Cords Entangled? A Lecture on Random Entanglement

Jean-Luc Thiffeault, University of Wisconsin

Robotic GFD: Collaborative Tracking and Control in Time-Dependent and Stochastic Flows

Philip Yecko, Cooper Union New York

Tuesday, July 15: Climate Response to Increased CO2 - Effects of Explicit Representation of Atmospheric Convection and Connections with Past Warm and Hot Climates

Eli Tziperman, Harvard University

Wednesday, July 16: The Nonlinear, Downstream Development of Unstable Baroclinic Waves. Chaos and Scale.

Joe Pedlosky, Woods Hole Oceanographic Institution

Thursday, July 17: Dynamics of the Madden-Julian Oscillation

Adam Sobel, Columbia University

Friday, July 18: The Ocean's Role in Transient Climate Change

Kyle Armour, Massachusetts Institute of Technology

Monday, July 21: Wind Waves in the Coupled Climate System

Baylor Fox-Kemper, Brown University

Tuesday, July 22: The Northern Hemisphere Summertime Circulation: Mechanisms and Future Changes

Tiffany Shaw, Columbia University

Wednesday, July 23: Response of Antarctic Climate to Ozone-hole Forcing

John Marshall, EAPS MIT Cambridge

Thursday, July 24: Convectively Coupled Tropical Transients

Zhiming Kuang, Harvard University

Friday, July 25: Energetics of Semi-enclosed Basins with Two-layer Flows at the Strait

Paola Cessi, University of California, San Diego

Monday, July 28: Maintenance, Formation and Drift of Zonal Jets
Bill Young, University of California, San Diego
Tuesday, July 29: Ocean Heat Uptake and the Putative 'Hiatus'
Carl Wunsch, Massachusetts Institute of Technology
Wednesday, July 30: The Geometry of Transition in Shear Flows
Norman Lebovitz, University of Chicago
Thursday, July 31: Too Much of a Good Thing? Destabilisation and Mixing Induced by Stratification in Shear Flows
Colm-cille Caulfield, Cambridge University
Friday, August 1: A GFD Study of the Last Glacial Maximum Climate
Raf Ferrari, MIT
Monday, August 4: Geometry of Turbulence, or How to Slice Baroclinic Instability
Predrag Cvitanovic, Georgia Institute of Technology
Tuesday, August 5: Ice Cores as Archives of Past Climate and Atmospheric Composition
Carlo Barbante, University of Venice
Wednesday, August 6: Superrotation in Planetary Atmospheres
Jonathan Mitchell, UCLA
Thursday, August 7: The Coolest Part of the Ocean
Cecilia Bitz, University of Washington
Friday, August 8: On the Self-aggregation of Convection in the Tropical Atmosphere
Caroline Muller, Laboratoire dHydrodynamique de l'cole Polytechnique & CNRS/Ecole Polytechnique

Fellows' talks

Monday, August 18:The Role of Mixed Layer Instabilities in Submesoscale Turbulence
Jörn Callies, MIT/WHOI Joint Program in Oceanography
Understanding Eddy Saturation in the Southern Ocean using Mean Field Theory
Joseph Fitzgerald, Harvard University
A Study of Heat Transport and the Runaway Greenhouse Effect using an Idealized Model
Paige Martin, University of Michigan
Tuesday, August 19: The Role of the Lapse Rate Feedback in Arctic Amplification
Ashley Payne, University of California, Irvine
The Most Minimal Seed for the Onset of Shear Turbulence
Geoff Stanley, University of Oxford
Wednesday, August 20: Models for Tropopause Height and Radiative-Convective Equilibrium
Shineng Hu, Yale University
Vortex Filament dynamics in Two Dimensions
Jim Thomas, New York University
Thermobaric Effects on Double-diffusive Staircases
Erica Rosenblum, University of California, San Diego
Thursday, August 21: An Experimental Investigation of the Rossby Two-slit Problem
Alexis Kaminski, University of Cambridge
A Numerical Study of the Downstream Development of Baroclinic Instability
Daniel Mukiibi, University of Hamburg

Lecture 1: Elements of Radiative Transfer

Kerry Emanuel; notes by Jörn Callies and Joseph Fitzgerald

June 16, 2014

1 Preamble

As opposed to expectations, there is not going to be much GFD in this lecture. There will be no equations of motion. The reason is that to understand climate, it is necessary to merge understanding of the large-scale motion with the thermodynamics and specifically with radiative transfer. Most problems that have not been solved in climate science are connected to the interaction of radiative transfer with water substances.

This lecture will therefore review how radiation passing through the atmosphere interacts with its constituents, most importantly with water in its different phases. Much of the fundamentals will only be alluded to, because of the time constraint.

One example of where the interaction of radiation with water substances in the atmosphere is crucial is the Madden–Julian oscillation. This wave that moves around the globe in the tropics is not yet well understood—it is one of the phenomena for which there are more theories than theoreticians. As climate scientists are taking radiation and its interaction with water more seriously, this problem will likely be solved in the near future.

Another one of these problems is tropical cyclogenesis. This also is likely to be understood in the near future.

2 Useful facts and numbers

This is a loose compilation of useful facts and numbers that are fundamental controls of Earth's climate. Much of the detail that could take up entire courses will be skipped over here.

The sun has a luminosity of about $L_0 = 3.9 \times 10^{26}\,\mathrm{J\,s^{-1}}$, which corresponds to a radiative flux of $6.4 \times 10^7\,\mathrm{W\,m^{-2}}$ at the top of the photosphere. Sunlight is received at the Earth that orbits the sun at a mean distance of about $d = 1.5 \times 10^{11}\,\mathrm{m}$, so that the flux density is $S_0 = L_0/4\pi d^2 = 1370\,\mathrm{W\,m^{-2}}$. The solar flux exhibits temporal variability, especially associated with the 11-year solar cycle. The amplitude of these variations is of the order of a tenth of a percent of the mean value, and is treated as negligible in this lecture series. However, variations in the solar flux must be taken into consideration when studying climate variability and trends on decadal and multi-decadal time scales.

The incoming solar radiation must be balanced by outgoing terrestrial radiation. A blackbody of temperature T emits a spectral flux, integrated over a hemisphere, of σT^4, where $\sigma = 5.67 \times 10^{-8}\,\mathrm{W\,m^{-2}\,K^4}$ is the Stefan–Boltzmann constant. According to this

Stefan–Boltzmann law, the emitted radiative flux depends strongly on the body's temperature. The sun's emission temperature as inferred from the Stefan–Boltzmann law at the top of the photosphere is about 6000 K.

The amount of solar flux absorbed by the Earth is equal to $S_0(1-a_p)\pi r_p^2$, where a_p is the planetary albedo and r_p is Earth's radius. The planetary albedo is the fraction of the solar flux that is reflected back to space. This is due to the combined effect of scattering by clouds, by reflection off ice and other surfaces, and scattering by air molecules and aerosols. In the present climate, Earth's albedo is about $a_p = 0.3$.

Because the total area of Earth is $4\pi r_p^2$, the absorbed solar flux per unit area is $S_0(1-a_p)/4$. If the atmosphere were completely transparent and the Earth emitted as a blackbody of uniform temperature T_e, then

$$\frac{S_0}{4}(1-a_p) = \sigma T_e^4,$$

Using the numbers cited above, this gives $T_e = 255\,\mathrm{K} = -18\,^\circ\mathrm{C}$. This is much colder than what is observed, which is due to the fact that the atmosphere is not transparent in the infrared. In the more general case, the relation above can be taken as the definition of the effective emission temperature.

The atmosphere is not completely transparent to solar radiation. Due to the presence of water vapor (H_2O), clouds, ozone (O_3) in the stratosphere, and carbon dioxide (CO_2), some solar radiation is absorbed as it passes through the atmosphere. Terrestrial radiation is absorbed even more strongly, even though it barely interacts with the atmosphere's major constituents, nitrogen (N_2) and oxygen (O_2). The primary absorbers of terrestrial radiation are water in its three phases, CO_2, methane (CH_4), and nitrous oxide.

The distribution of radiative fluxes across frequencies for a blackbody in thermodynamic equilibrium at temperature T is given by Planck's law:

$$B_\nu(T) = \frac{2h\nu^3}{c^2\left(e^{\frac{h\nu}{kT}} - 1\right)},$$

where ν is the frequency, h is the Planck constant, k is the Boltzmann constant, and c is the speed of light. The atmosphere is in thermodynamic equilibrium except at very high altitudes, where the gas density is so low that the atmosphere is very nearly transparent. Planck's law tells us that the distribution of radiative fluxes across frequencies, or equivalently wavelengths, depends on its temperature T. Integrating this expression over all frequencies and all angles in a hemisphere gives the Stefan–Boltzmann law:

$$\pi \int_0^\infty B_\nu(T)\,\mathrm{d}\nu = \sigma T^4,$$

where the Stefan–Boltzmann constant can be written in terms of more fundamental physical constants, $\sigma = 2\pi^5 k^4/15c^2h^3$.

3 Absorption in the atmosphere

How does radiation interact with a gas? The basis of our understanding for how photons get absorbed by gas molecules is quantum physics, which will only be alluded to here.

An isolated atom can absorb photons only at discrete frequencies, those that correspond by $E_\nu = h\nu$ to differences in the available discrete energy levels of the atom's electron shell. Molecules have additional energy levels associated with additional degrees of freedom. These additional energy levels allow absorption of photons at a larger number of frequencies (Fig. 1).

For N_2 and O_2, only a few extra degrees of freedom are available. Compared to a single atom, they have an extra vibrational mode, but they remain relatively weak absorbers.

The more complex the molecule becomes, the more degrees of freedom it has, and thus the more absorption lines are present. Carbon monoxide (CO), for example, is asymmetric and thus has a permanent dipole, which gives it rotational modes of interaction with photons. More complex molecules allow additional rotational and vibrational degrees of freedom—and combinations of different modes of excitation. Photodissociation or ionization also leads to absorption.

The discrete energy levels of the molecules give rise to absorption lines at discrete frequencies. But if the absorption were only present at truly discrete frequencies, no absorption at all would be possible, as the continuous spectrum of solar radiation contains only an infinitesimal amount of energy in an infinitesimal frequency window. In reality, the atmosphere absorbs a finite amount of energy due to the broadening of absorption lines by two effects (Fig. 2). First, because of the movement of the molecules in a gas, the absorption frequencies get Doppler shifted slightly. The molecules move in random directions, so averaged over all molecules, absorption can occur over a distribution around the nominal absorption frequency. Second, the collision between molecules also broadens the absorption line. This is called pressure broadening, because the collision rate depends on pressure.

Elements of the absorption properties of gases have been known since the 19th century. John Tyndall measured the absorption capabilities of the various atmospheric gases in the infrared, concluding that water vapor was the most powerful absorber.

Water is a strong absorber because of its bent triatomic structure and its permanent dipole. It has a purely rotational absorption band and bands due to rotation, vibration, and translation. Ozone, while also absorbing in the infrared, is mostly important because it is involved in photodissociation in the stratosphere that absorbs ultra-violet (UV) light emitted by the sun. CO_2 has no permanent dipole, so it is not as strong an absorber as H_2O, but a dipole is present when the vibrational mode is excited, so CO_2 is an important absorber in the infrared as well. Other strongly absorbing constituents are nitrogen dioxide (N_2O) and CH_4.

It is interesting to note that the main absorbers H_2O and CO_2 only make up a small fraction of the total mass of the atmosphere (Fig. 3). The bulk of the atmospheric molecules, N_2 and O_2, are radiatively inert. The most abundant strong absorber is H_2O. While the CO_2 concentration is to leading order uniform, the H_2O concentration is highly variable. To leading order, the H_2O concentration is controlled by temperature, through the Clausius–Clapeyron relation, but deviations from this are possible when the air is sub-saturated.

The average lifetime of a water molecule in the atmosphere is on the order of two weeks. This is very short compared to that of CO_2, which decays on two exponential time scales, one on the order of decades that is controlled by upper ocean uptake, the other on the order of millennia that is controlled by carbon cycling into the deep ocean. Carbon is also taken out of the atmosphere by weathering of the rocky land surface, but that process operates

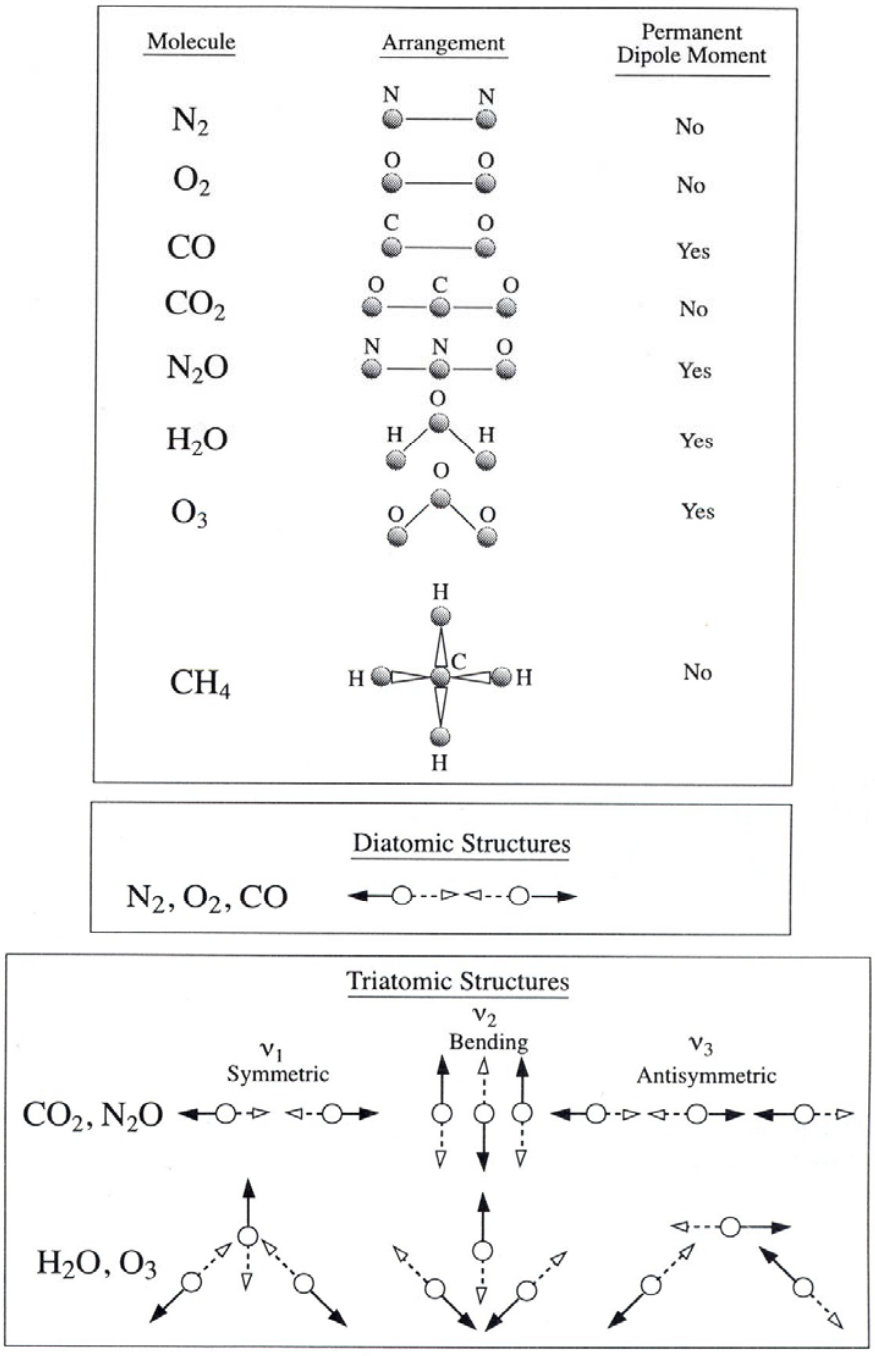


Figure 1: Illustration of the structure of molecules and their modes of interaction with radiation by dipole moments, vibrational modes, and rotational modes

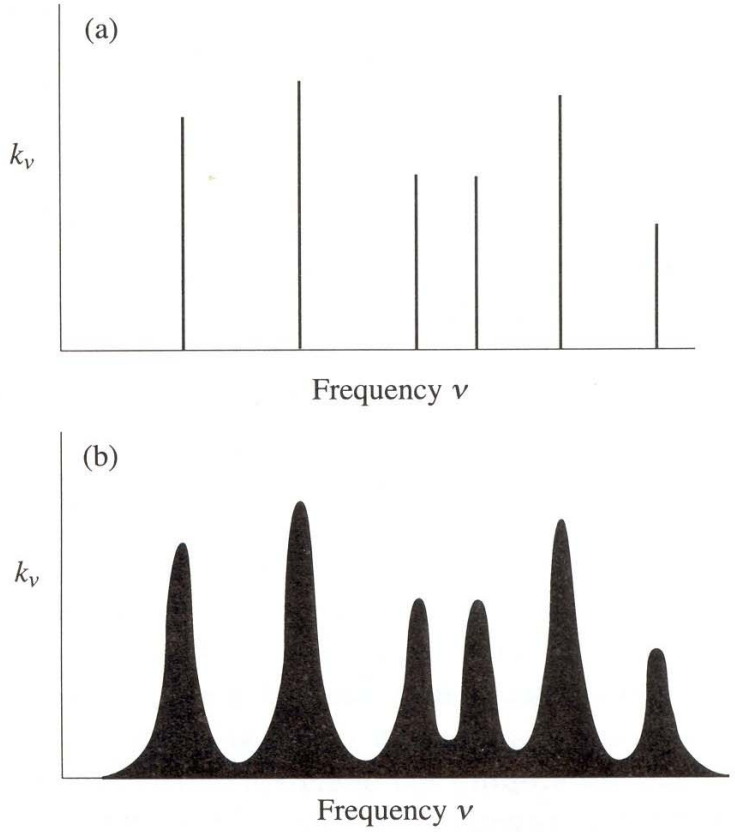


Figure 2: Illustration of the broadening of spectral lines (a) to continuous absorption bands (b) by Doppler and pressure effects

on geological time scales. The long lifetime of CO_2 in the atmosphere allows it to be well mixed. The CO_2 concentration is currently about 0.04 %.

Long-lived gases tend to be well-mixed. Gases with a shorter lifetime can be more highly concentrated near sources (Fig. 4). Ozone, for example, is enhanced in the stratosphere, where it is created by photodissociation. Carbon monoxide is mostly present in the troposphere, where its anthropogenic sources are. Water is also mostly present in the troposphere, but it is highly variable in space and time. It also undergoes phase transitions.

Water vapor is the most interesting of the trace gases, owing to its absorption properties and its highly variable concentration. The H_2O concentration is a function of the climate state, and the climate state is dependent on the distribution of H_2O. In fact, there are strong feedbacks, which have often been ignored. As alluded to in the preamble, recognizing the interplay of radiation and dynamics, which may lead to a self-regulation of the climate system, may allow progress in long-standing problems in atmospheric and climate science.

The CO_2 concentration has been rising due to anthropogenic emissions (Fig. 5). Direct measurements of CO_2 concentrations at Mauna Loa show an accelerating increase since 1959 to currently about 400ppm. Ice cores that contain air bubbles from pre-industrial times tell

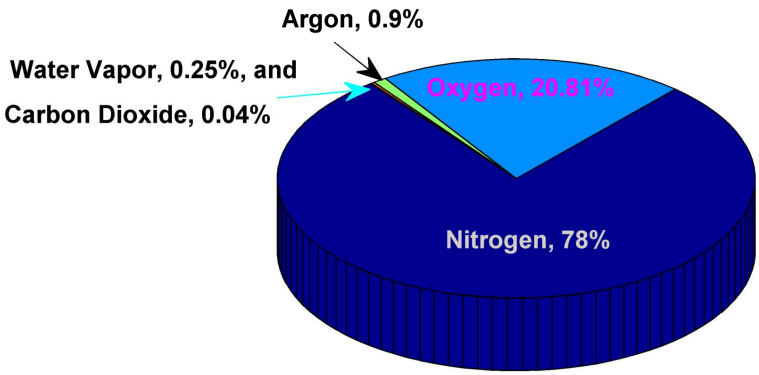


Figure 3: Atmospheric constituents by mass, showing that the strongly absorbing gases occur in small concentrations

us that the background, interglacial CO_2 concentration was about 280ppm. Superposed on this increase is a small seasonal cycle due to enhanced uptake of CO_2 by the deciduous forest in boreal summer that has no equivalent in the austral summer, because there is less land mass in the southern hemisphere.

Fig. 6 shows the Planck functions for a solar temperature and for typical terrestrial temperatures. The sun mostly emits in the visible range of the spectrum, while terrestrial emissions are mostly in the infrared range. This separation between solar and terrestrial radiation simplifies our conceptual understanding, as illustrated later in the lecture.

Also shown in this figure are the attenuations by absorption and scattering. The UV light is mostly absorbed by ozone in the stratosphere. The atmosphere is mostly transparent to visible light: some 90 % passes through. The absorption bands of H_2O and CO_2 tend to be in the infrared range. The transmission of terrestrial infrared radiation is small, except for a window around 10 μm.

It should be noted that this figure is a cartoonish illustration of the immensely complex absorption properties of the atmospheric gases. Very accurate numerical codes exist that compute the absorption line by line. These are computationally too expensive to be put in a climate model, so one resorts to simplified band models that are tested against the line-by-line computations. These radiation codes are quite reliable; most of the uncertainty in radiative transfer comes from scattering, especially multiple scattering in clouds.

4 Radiative equilibrium

Radiative equilibrium is the state of the atmosphere in which in every sample of gas, the amount of absorbed radiation equals the amount of emitted radiation. In calculating the radiative equilibrium, we make the assumption that the atmosphere is motionless, that no advective heat transfer occurs. When a system is out of radiative equilibrium, radiative transfer processes will relax it back toward the equilibrium. This relaxation will in general

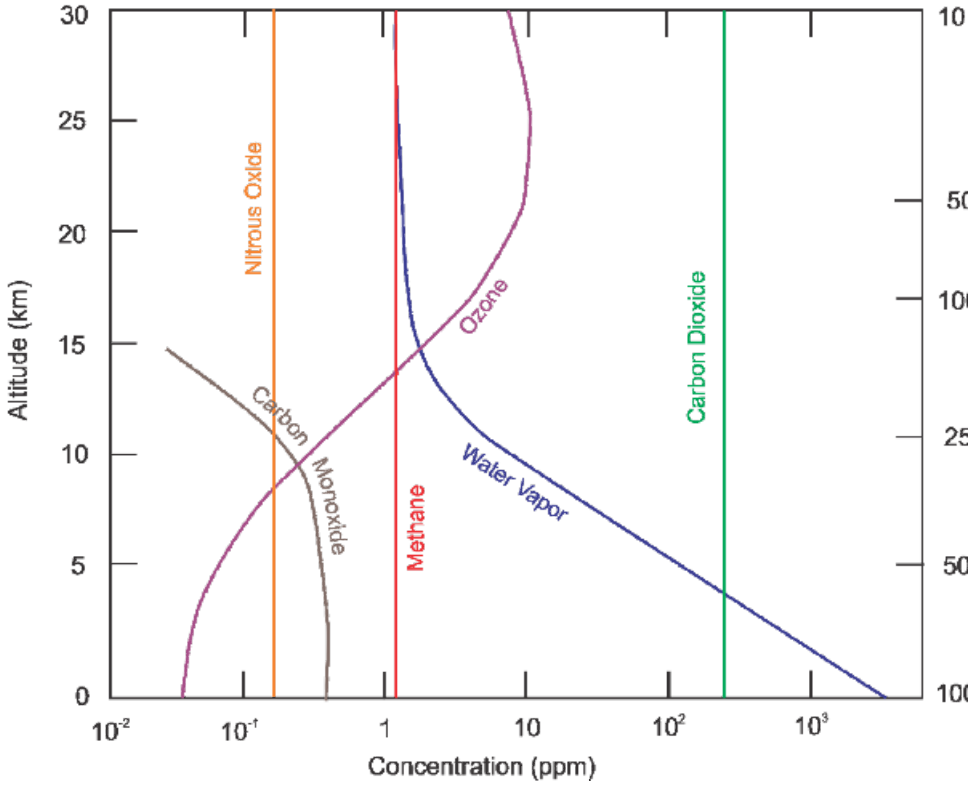


Figure 4: Vertical structure of concentrations of atmospheric constituents as inferred from climatology and specified for the band model of radiative equilibrium

be more complicated than simple Newtonian relaxation, due to the nonlocality of radiative transfer. For example, a temperature perturbation at one level will have a radiative effect on the temperature of other levels, which would not occur in a Newtonian cooling model.

As the simplest model of radiative equilibrium, consider a one-layer atmosphere (Fig. 7). The model consists of a single layer of atmosphere at temperature T_a and a surface at temperature T_s. The atmosphere is assumed to be completely transparent to solar radiation and completely opaque to terrestrial radiation. The solar flux received is then $S_0(1-a_p)/4 = \sigma T_e^4$, which should be taken as a definition of the emission temperature. The surface emits black-body radiation according to the Stefan–Boltzmann law, σT_s^4. The atmosphere similarly emits as a black body, σT_a^4, both upward and downward. Remember that the Stefan–Boltzmann law gives the radiative flux integrated over a hemisphere.

In radiative equilibrium, the incoming solar flux must balance the outgoing terrestrial flux, which trivially gives the atmospheric temperature $T_a = T_e$. At the surface, the balance is between the sum of the incoming flux from the sun and the atmosphere and the emitted black-body radiation:

$$\sigma T_e^4 + \sigma T_a^4 = \sigma T_s^4.$$

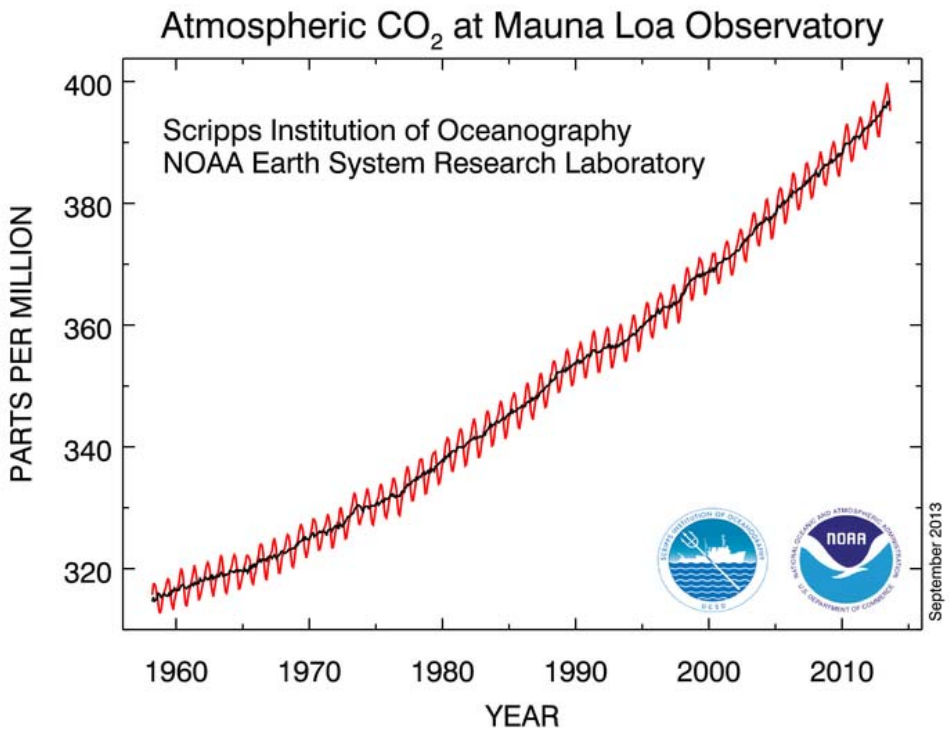


Figure 5: CO_2 concentration as measured at Mauna Loa Observatory (red) and running mean (black)

With $T_a = T_e$, this gives the surface temperature $T_s = 2^{1/4}T_e = 303\,\mathrm{K}$. The surface temperature is higher than the emission temperature. This is the greenhouse effect: the surface receives radiation from both the sun and the atmosphere. The resulting surface temperature is higher than the observed mean surface temperature on Earth, but not too far off, considering how simple a model this is. More importantly, the simple model captures the essence of the problem, allowing conceptual understanding. It can be considered the simplest climate model.

Note that in this one-layer model, the surface receives the same amount of flux from the sun as it does from the atmosphere. In reality, it turns out that the surface receives roughly twice as much radiation from the atmosphere as it does from the sun. This provides a puzzle, as our model surface receives less radiation from the atmosphere than the true earth surface, and yet our model surface temperature still comes out too hot. We will address this issue later in the lectures.

The model can be modified by adding extra layers. The case of two atmospheric layers that are completely opaque is depicted in Fig. 8. The extra layer increases the surface temperature to $T_s = 3^{1/4}T_e$. Generalizing to n layers results in a surface temperature $T_s = (n+1)^{1/4}T_e$. Every extra layer increases the surface temperature. It should be noted that the real atmosphere is not the continuous limit of this layered model, because as layers become thinner, they become less opaque, so in a continuous limit, the emissivity of the layers must go to zero.

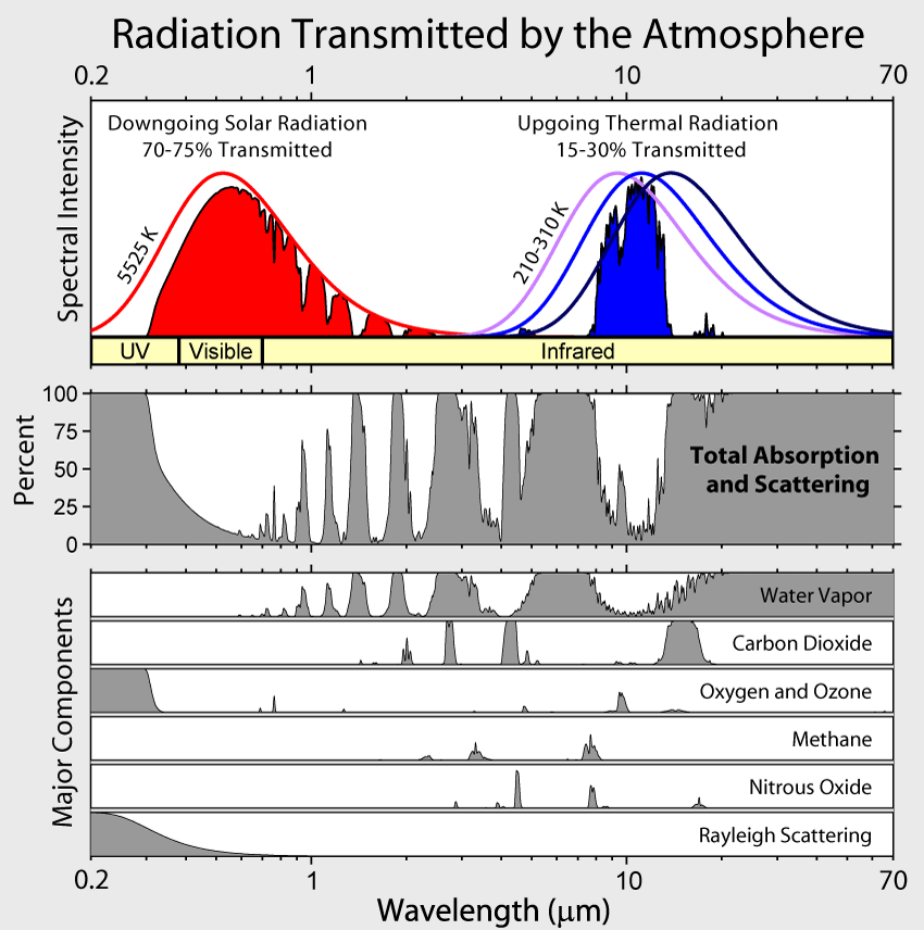


Figure 6: Planck curves of solar and terrestrial radiation as well as scattering and absorption by the atmospheric constituents

Instead of the continuous limit, let us consider two wrinkles on the two-layer model to illustrate two important aspects of radiative equilibrium.

First, consider a thin layer of gas at temperature T_t just above the layer 2. Let its emissivity be ε_t and let that emissivity tend to zero. This layer then does not affect any of the other layers, so we can consider the radiative balance of this layer independently. From

$$\varepsilon_t \sigma T_2^4 = 2\varepsilon_t \sigma T_t^4$$

we find $T_t = 2^{-1/4} T_2 = 2^{-1/4} T_e$. This illustrates that the radiative equilibrium temperature of an atmospheric layer can be lower than the emission temperature T_e.

Second, consider adding such a layer just above the surface. Let its temperature be T_a and its emissivity ε_a tend to zero. The balance of this layer is then

$$\varepsilon_a \sigma T_s^4 + \varepsilon_a \sigma T_1^4 = 2\varepsilon_a \sigma T_a^4,$$

so $T_a^4 = (T_s^4 + T_1^4)/2$. This layer therefore does not have the same temperature as the surface. This result is independent of ε_a, so long as it is sufficiently small, and illustrates that a

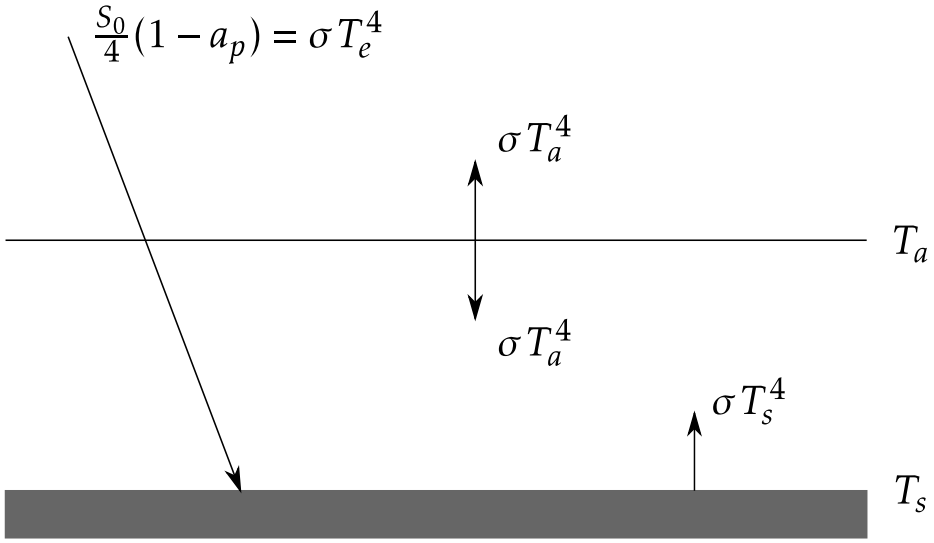


Figure 7: Radiative fluxes in a model with an atmosphere that consists of a single layer that is completely transparent to solar radiation and completely opaque to terrestrial radiation

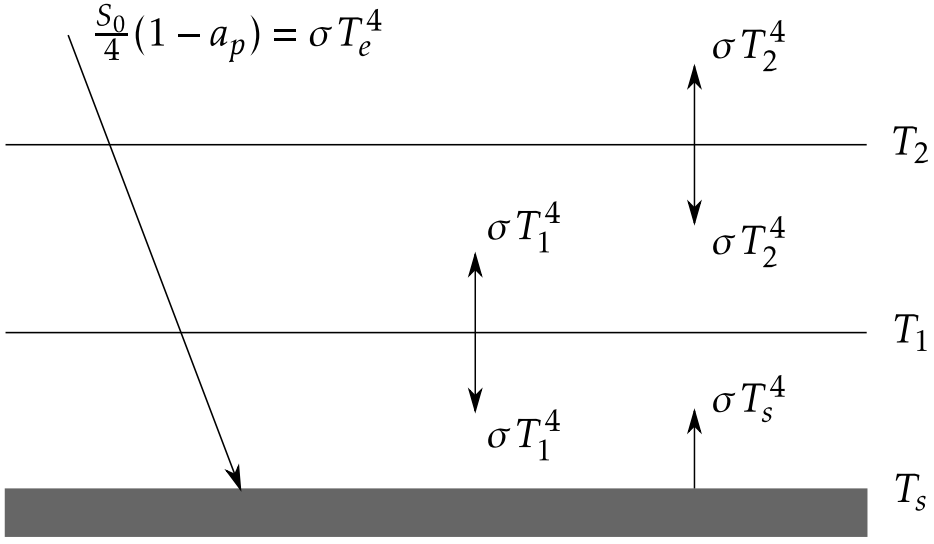


Figure 8: Radiative fluxes in a model with an atmosphere that consists of two layers that are completely transparent to solar radiation and completely opaque to terrestrial radiation

discontinuous emissivity entails a discontinuity in temperature. In radiative equilibrium, the surface atmospheric temperature is generally different from the temperature of the surface. Radiation drives the system into thermodynamic disequilibrium, which in reality is counteracted by heat diffusion or fluid motion. But this effect can be experienced in a desert during the day, where the sand is typically much hotter than the air at 2 m above ground.

We now discuss the radiative equilibrium in a much more realistic model. This model consists of a large number of layers for which the emissivity is calculated for wavelength bands from the concentrations of the atmospheric constituents. The concentrations are specified, with the exception of water vapor, for which a temperature-dependent relative humidity is prescribed.

There are a number of interesting features of the resulting radiative equilibrium (Fig. 9). There is a discontinuity of temperature at the surface—the surface temperature is different from the temperature of the atmosphere at the surface. Furthermore, the troposphere is very cold, much colder than is observed. Part of the reason is that there is a positive water vapor feedback: as the atmosphere gets cold, the amount of water vapor decreases for a fixed relative humidity, so the greenhouse effect is diminished. But the ultimate reason why the tropospheric temperature is unrealistic is that we disregard convection, which will be discussed in the next part.

The obtained temperature of the stratosphere is roughly consistent with the observed stratospheric temperatures, which indicates that the stratosphere is in a state not very far away from radiative equilibrium. There is an increase in temperature in the stratosphere, which is due to the absorption of UV light by ozone in the upper stratosphere. If we increased the CO_2 concentration, the stratosphere would cool. This counterintuitive effect can be explained by realizing that the stratospheric gas becomes a better emitter for a higher CO_2 concentration, so its temperature must be reduced to have the emitted radiation match the absorbed solar radiation. To understand this effect fully, a radiative-equilibrium model with two partially transparent layers would presumably be useful. In such a model, an increase in CO_2 concentration could be modeled by in increase in emissivity of the upper layer.

5 Convection

The reason for the unrealistically cold troposphere with a very steep temperature gradient in the lower troposphere is the omission of convection. The time scale of the radiative relaxation to equilibrium is on the order of tens of days. But convection operates on a few hours, much faster than radiation. Convection drives the system toward a state of convective neutrality. The separation of time scales means that where convection occurs, the system is very near this state of convective neutrality. The state resulting from the combination of radiation and convection is called the radiative–convective equilibrium. It should be noted, however, that such an equilibrium is a state of statistical equilibrium, which is highly turbulent in nature. We begin by calculating what this equilibrium state is if phase changes of water are neglected. Later, we will add such moist processes.

Historically, radiative–convective equilibrium was first analyzed in the 1960s. Since then, climate scientists have rapidly moved on to three-dimensional modeling, building

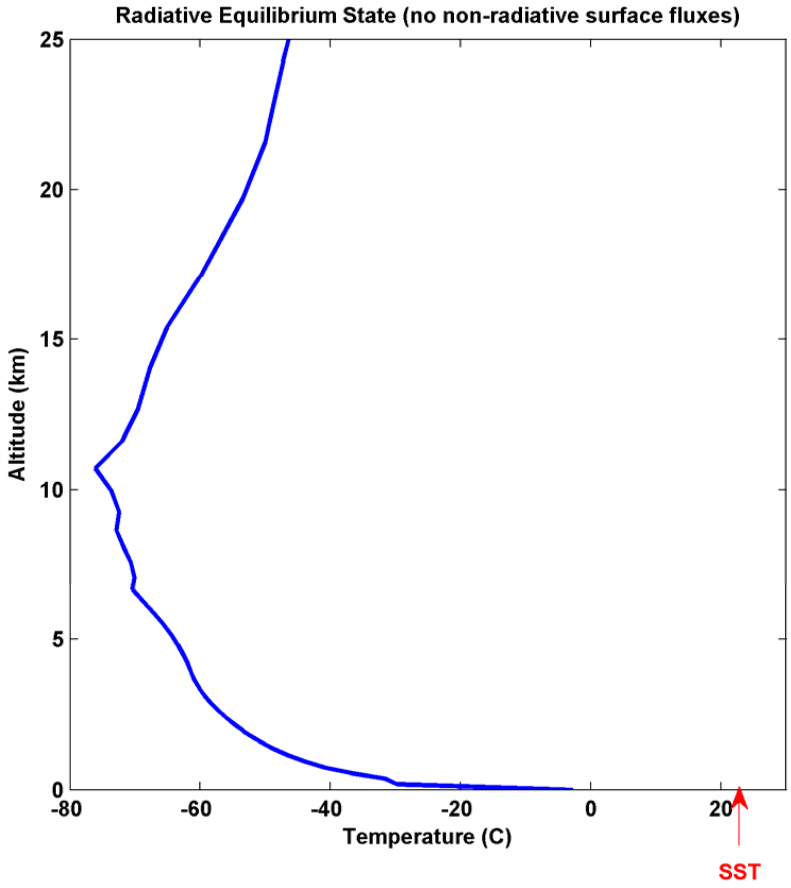


Figure 9: Radiative equilibrium of band model with concentrations of greenhouse gases specified and relative humidity prescribed

ever more complex models. Moving on so quickly might have been a historical mistake as much fundamental insight can still be gained from this very idealized radiative–convective equilibrium.

Let's consider the force balance of a cube of fluid of density ρ_b under the effect of gravity g (Fig. 10). The forces in the vertical on the box are then the sum of the box's weight and the pressure forces on the horizontal faces:

$$-\rho_b g \Delta x \Delta y \Delta z + p(z) \Delta x \Delta y - p(z + \Delta z) \Delta x \Delta y = 0.$$

Dividing through by the volume of the box and taking the limit $\Delta z \to 0$ gives

$$\frac{\partial p}{\partial z} = -\rho_b g,$$

which is called the hydrostatic balance.

Allowing a vertical acceleration modifies the balance to

$$-\rho_b g \Delta x \Delta y \Delta z + p(z) \Delta x \Delta y - p(z + \Delta z) \Delta x \Delta y = \rho_b \frac{\mathrm{d}w}{\mathrm{d}t} \Delta x \Delta y \Delta z,$$

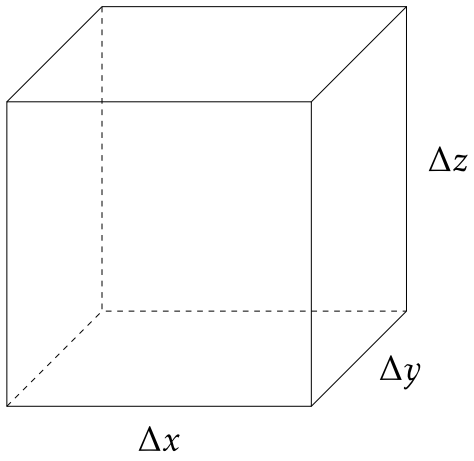


Figure 10: Box of air for which the vertical force balance is considered

where w is the vertical velocity. The vertical acceleration is then

$$\frac{\mathrm{d}w}{\mathrm{d}t} = -\frac{1}{\rho_b}\frac{\partial p}{\partial z} - g.$$

If the environment is in hydrostatic balance and the pressure is the same in the box and the environment, then

$$\frac{\mathrm{d}w}{\mathrm{d}t} = g\left(\frac{\rho_e}{\rho_b} - 1\right) = g\frac{\rho_e - \rho_b}{\rho_b} = B,$$

which is called buoyancy. The environmental density was here denoted by ρ_e.

Next we will assess whether a vertical profile of temperature is stable by displacing a parcel and calculating whether it gets pushed back to its original position, in which case we call it stable, or whether it gets accelerated away from its original position, in which case we call it unstable. This is not a trivial calculation, because the parcel's density changes as it is displaced, because its pressure changes. Throughout, we will neglect viscous stresses.

Lecture 2: Energy Balance and the Troposphere

Geoff Vallis; notes by Shineng Hu and Alexis Kaminski

June 17

The philosophy throughout these lectures is that in order to understand a complex system we must have a description of it at multiple levels, from a back-of-the-envelope calculation through idealized numerical models to a comprehensive simulation with all the bells and whistles. Because this is Walsh we will always try to include a back-of-the-envelope calculation and go from there, but other approaches are possible.

1 What are we trying to explain?

A schematic of the overall structure of the atmosphere and ocean is given in Fig. 1, with some pictures of the real atmosphere from observations given in Fig. 2 and Fig. 3. In Fig. 1 we sketch the troposphere, where temperature decreases with height, and the stratosphere, where temperature increases with height, and the dividing tropopause which is fairly high over the tropics (15km) and lower over polar regions (8km). We might immediately ask, what determines this structure? What determines the height of tropopause? Why is it about 10 km , and not 100 km or 1 km ? And what determines the width of the tropics where the tropopause is high? And so on.

Turning to the ocean, we have, again very schematically, warm water in the upper ocean and cold water below. The layer between them, where temperature varies very fast vertically, is called the thermocline. Sometimes we make an analogy between the thermocline and the tropopause, but actually the thermocline is more like the whole troposphere because they are both characterized by large vertical temperature gradients and relatively fast dynamics. Questions for oceanographers include what determines this structure of the ocean? What is the nature of the circulation that maintains it? More specifically, what determines the depth of the thermocline?

These are the *kinds* of questions we will consider in these lectures. We'll try to answer some of them, but not all. The philosophy throughout is that in order to understand a complex system we must have a description of the system at multiple levels, from a back-of-the-envelope calculation through idealized numerical models to a comprehensive simulation with all the bells and whistles. Because this is Walsh we will always try to include a back-of-the-envelope calculation and go from there, but other approaches are possible.

Our goals in this lecture are fairly fundamental:

- Understand at an elementary level what determines the surface temperature of Earth.
- Understand the need for a troposphere, and what determines its thickness.

To answer that we begin with a tutorial on radiation.

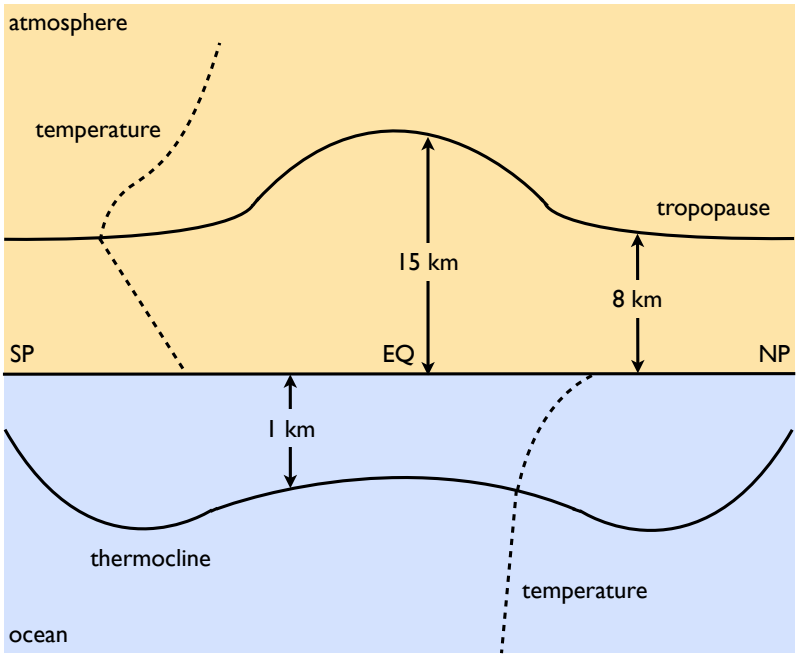


Figure 1: A schematic of thermal structures of ocean and atmosphere. The solid lines mark the tropopause and the base of the thermocline, and the near-vertical dashed lines are representative profiles of temperature.

2 Radiative Balance

2.1 The very basics

All macroscopic bodies except those at absolute zero (are there any?) emit thermal radiation. The black body emission per unit wavelength or per unit frequency are given by Planck's function which is, for the two cases respectively,

$$B_\lambda(T) = \frac{2\pi hc^2}{\lambda^5}\frac{1}{\exp(hc/\lambda k_B T) - 1}, \qquad B_\nu(T) = \frac{2h\nu^3}{c^2}\frac{1}{\exp(h\nu/k_B T) - 1} \tag{1}$$

where c is the speed of light, h is the Planck constant and k_B is the Boltzmann constant. Conventions for frequency and wavelength are such that $c = \omega/k = \omega\lambda/2\pi = \nu\lambda$. Integrating either of the above expressions over wavelength or frequency, respectively, gives the Stefan–Boltzmann law

$$B(T) = \sigma T^4, \tag{2}$$

where σ is Stefan's constant,

$$\sigma = \frac{2\pi^5 k_B^4}{15h^3c^2} = 5.6704 \times 10^{-8}\ \text{W m}^{-2}\,\text{K}^{-4}. \tag{3}$$

The maximum of Planck's function occurs at a wavelength $\lambda_m = b/T$ where $b = 2.898 \times 10^{-3}$ m K . This is Wien's displacement law, and it means that the higher the temperature

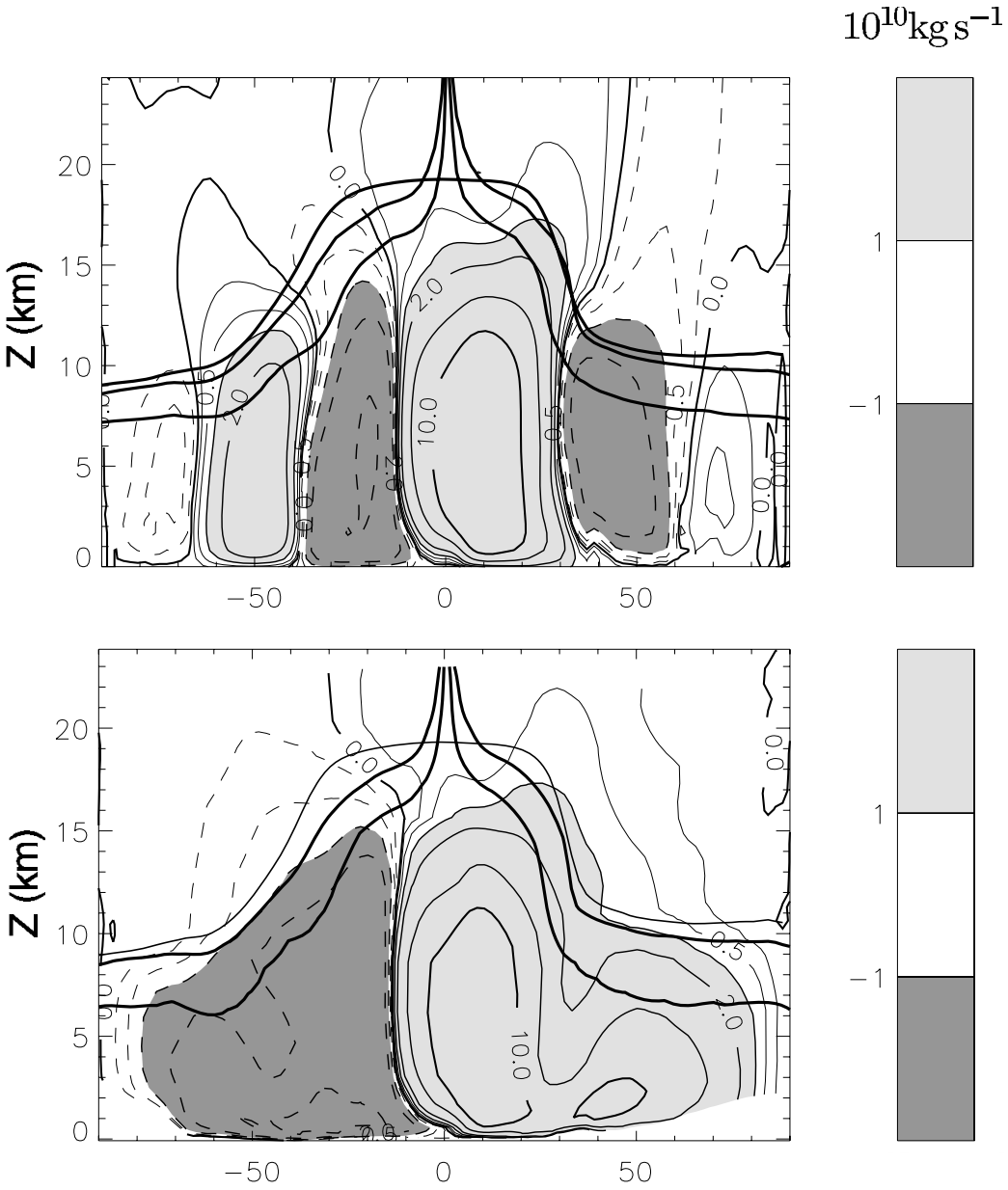


Figure 2: Overturning circulation of the atmosphere during a Northern Hemisphere winter. The contours and shading indicate an overturning streamfunction, rising just south of the equator. The top plot shows a conventional Eulerian average and the bottom plot is a residual circulation. Three measures of the tropopause are indicated with the more nearly horizontal solid lines.

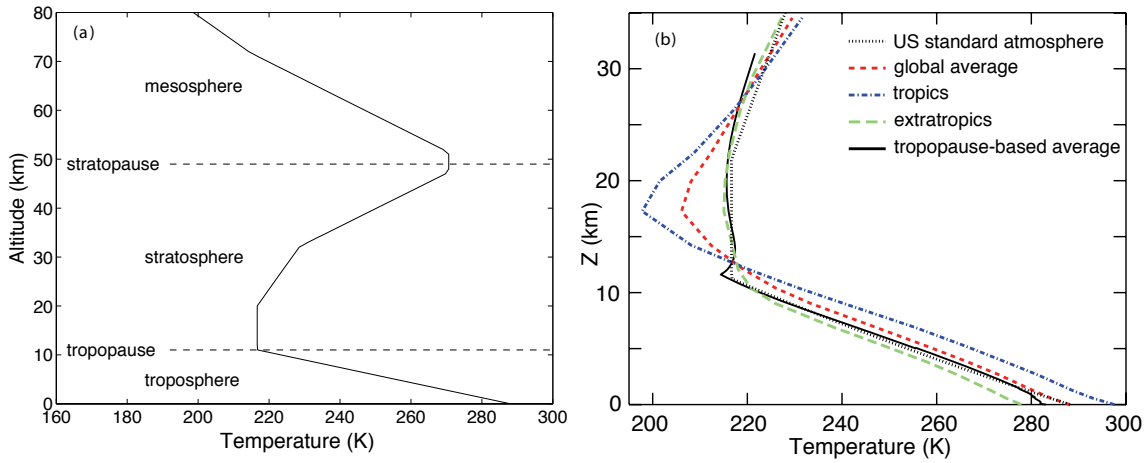


Figure 3: Temperature profiles in the atmosphere. On the left is the 'US standard atmosphere' and on the right are some observed profiles.

the shorter the wavelength at which emission predominantly occurs. For the Sun, at $T \approx 6000\,\mathrm{K}$, $\lambda_m = 5 \times 10^{-7}$ m , which is in the visible range; solar radiation is also sometimes called shortwave. For Earth, at $T = 280\,\mathrm{K}$, $\lambda_m = 1 \times 10^{-5}$ m , which is in the so-called infra-red, sometimes called longwave. The radiation itself is in units of W m $^{-2}$, and so is a flux of energy. The radiation reaching Earth from the Sun has an intensity of $S^* = 1366\,\mathrm{W\ m^{-2}}$, varying by about $1\,\mathrm{W\ m^{-2}}$ over the 11-year sunspot cycle.

2.2 Earth's global energy budget

The simplest model that gives the temperature of the Earth is to suppose that the incoming solar radiation is balanced by an outgoing flux of infra-red radiation at a single temperature so that

$$S_0(1-\alpha) = \sigma T^4, \tag{4}$$

where $S_0 = S^*/4 = 342\,\mathrm{W\ m^{-2}}$ and α is the Earth's albedo, the fraction of solar radiation reflected and measurements show that $\alpha \approx 0.3$. The resulting temperature, T_e is variously called the effective emitting temperature, the radiation temperature or the bolometric temperature. Plugging in numbers we find

$$T_e = \left(\frac{342 \times 0.7}{5.67 \times 10^{-8}} \right)^{1/4} = 255\,\mathrm{K}\,. \tag{5}$$

The actual surface temperature on Earth averages $288\,\mathrm{K}$. If you think $255\,\mathrm{K}$ is a good estimate of $288\,\mathrm{K}$, you are at heart a planetary scientist. If you think it is a bad estimate, you are a climate scientist or a meteorologist.

[Needed: table of emitting temperature and actual surface temperature for all planetary bodies in the solar system.]

A simple feedback we can put into such a model is the ice-albedo feedback, whereby we

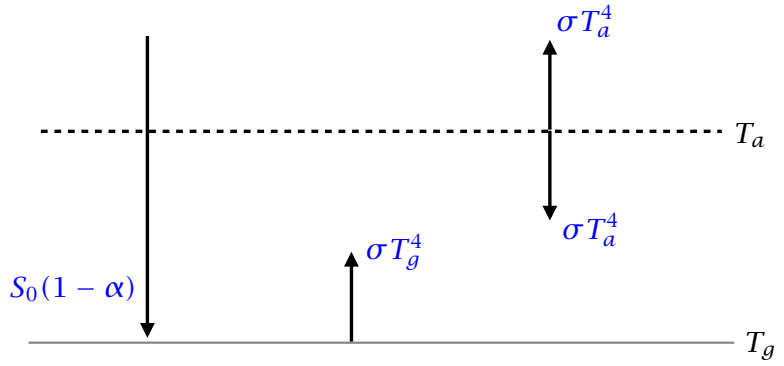


Figure 4: A simple EBM

suppose that α is a function of temperature. For example, we might suppose that

$$\alpha = \begin{cases} 0.3 & \text{for } T > T_0 \\ 0.8 & \text{for } T < T_0 \end{cases} \tag{6}$$

2.3 Effects of the atmosphere

The clear-sky atmosphere is largely transparent to solar radiation, but not to infra-red radiation. Most (but not all) of the solar radiation impinging on the atmosphere that is not reflected by clouds is thus absorbed at the Earth's surface, whereas most of the infra-red radiation emitted at the Earth's surface is absorbed by the atmosphere.

Given this, the next simplest model is to suppose there is an absorbing atmosphere above the surface, as illustrated in Fig. 4. If it is in equilibrium then the energy balance equations are:

$$\text{Top:} \qquad S_0(1-\alpha) = \sigma T_a^4, \tag{7}$$

$$\text{Surface:} \qquad S_0(1-\alpha) + \sigma T_a^4 = \sigma T_g^4. \tag{8}$$

(From these we also see the atmospheric balance, $2\sigma T_a^4 = \sigma t_g^4$.) The solution is

$$T_a = \left(\frac{(1-\alpha)S_0}{\sigma}\right)^{1/4}, \qquad T_g = 2^{1/4} T_a. \tag{9}$$

So that $T_a = 255\,\mathrm{K}$ (as it has to be) and $T_g = 303\,\mathrm{K}$, This is now too warm. One solution is to suppose the atmosphere has a finite emissivity, ϵ_a (which is less than one). This is getting ad hoc, but it will allow us to illustrate a nice effect. Thus,

$$\text{Top:} \qquad S_0(1-\alpha) = \epsilon_a \sigma T_a^4 + (1-\epsilon_a)\sigma T_g^4, \tag{10}$$

$$\text{Surface:} \qquad S_0(1-\alpha) + \epsilon_a \sigma T_a^4 = \sigma T_g^4 + F. \tag{11}$$

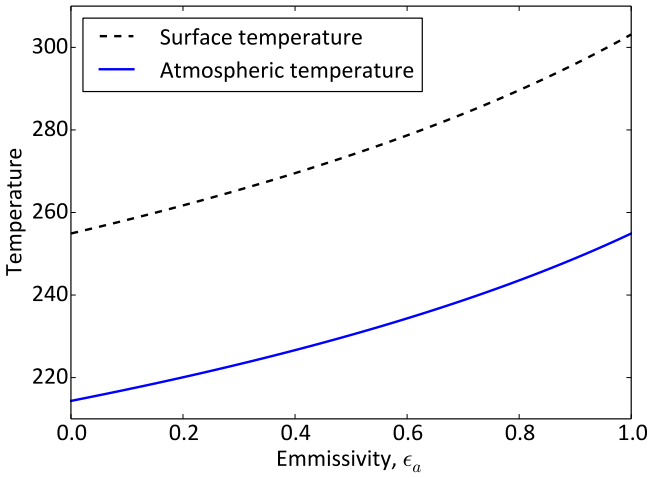


Figure 5: Temperature as a function of emissivity in the EBM

where we also introduce a flux F from surface to atmosphere. The solution for the surface temperature is

$$\sigma T_g^4 = \frac{S_0(1-\alpha) - F/2}{1 - \epsilon_a/2} \tag{12}$$

which, for $F = 0$ and $\epsilon = 0.77$, gives $T_g = 288\,\mathrm{K}$. The surface temperature obviously increases with ϵ_a as expected (Fig. 5).

3 Water Vapour Feedback

3.1 Saturation vapour pressure

The two main greenhouse gases are water vapour and carbon dioxide. Carbon dioxide is well mixed and is not volatile (it does not condense at Earthy temperatures). Its value is determined by geological and anthropogenic processes, and we can suppose its value to be specifiable. Water vapour levels are determined by the relative humidity of the atmosphere and, above all else, by the temperature through the Clausius–Clapeyron relation. This states that the saturation vapour pressure of water, e_s, or indeed of nearly any condensing material, varies as

$$\frac{\mathrm{d}e_s}{\mathrm{d}T} = \frac{L}{T(\rho_g^{-1} - \rho_c^{-1})} \approx \frac{L}{R_w T^2} e_s, \tag{13a,b}$$

where the second expression follows if $\rho_g \ll \rho_c$ (the density of the gas phase is much less than that of the condensed phase) and using the ideal gas law. The parameter L is the 'latent heat of condensation' and R_w is the gas constant for the gas in question, which for us is water. If L is constant (not a quantitatively good assumption, but good enough for now) we get

$$e_s(T) = e_{s0} \exp\left[\frac{L_s}{R_s}\left(\frac{1}{T_0} - \frac{1}{T}\right)\right]. \tag{14}$$

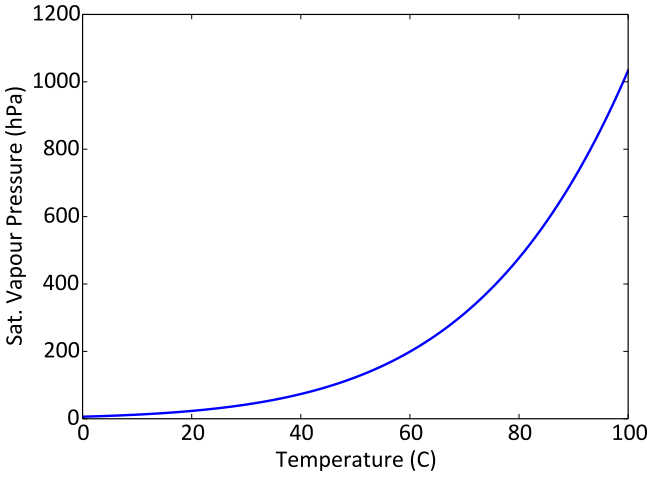


Figure 6: Saturation vapour pressure as a function of temperature

Evidently, saturation vapour pressure is a strongly increasing function of temperature. A liquid will boil when the temperature is sufficiently high that the saturation vapour pressure equals the ambient pressure, and for water at sea-level this occurs at 100° C . A good, semi-empirical approximation for saturation water vapour pressure is the Tetens–Bolton formula,

$$e_s = 6.112 \exp\left(\frac{17.67 * T_c}{T_c + 243.3}\right) \tag{15}$$

where T_c is temperature in Celsius and pressure is in hecto-Pascals (the same as millibars). This is actually a better approximation than (14) because it includes the variation of L with T. In any case, the main point is that water vapour content in the atmosphere increases fairly rapidly with temperature, at about 7% K $^{-1}$ (Fig. 6).

3.2 Radiative feedback and runaway greenhouse

Returning now to the EBM of the previous section, if we differentiate (12) we obtain

$$\frac{4\,\mathrm{d}T_g}{T_g} = \frac{\mathrm{d}\epsilon_a}{2-\epsilon_a}. \tag{16}$$

Now, ϵ_a may vary both because we add CO_2 (which we will denote as c) and because water vapour content may changes, so we write

$$\mathrm{d}\epsilon_a = \mathcal{A}\,\mathrm{d}c + \mathcal{B}\,\mathrm{d}e_s \tag{17}$$

where $\mathcal{A}$ and $\mathcal{B}$ are quantities that reflect the radiative properties of CO_2 and water vapour. If the main reason water vapour changes is because of the change in saturation vapour pressure with temperature then, using (13b) and (17), (16) becomes

$$(8-4\epsilon_a)\frac{\mathrm{d}T_g}{T_g} = \mathcal{A}\,\mathrm{d}c + \mathcal{B}\,\mathrm{d}e_s = \mathcal{A},\mathrm{d}c + \frac{\mathcal{B}L}{R_w T^2} e_s\,\mathrm{d}T_g. \tag{18}$$

or

$$\left(\frac{8-4\epsilon_a}{T_g} - \frac{\mathcal{B}Le_s}{R_w T_g^2}\right) \mathrm{d}T_g = \mathcal{A}\,\mathrm{d}c \tag{19}$$

Note that changes in atmospheric temperature is proportional to changes in surface temperature. Thus

$$\frac{\mathrm{d}T_g}{\mathrm{d}c} = \frac{\mathcal{A}T_g}{8-4\epsilon_a}\left(\frac{1}{1-y}\right) \quad \text{where} \quad y = \frac{\mathcal{B}e_s L}{R_w T_g(8-4\epsilon_a)}. \tag{20}$$

This is a rather interesting equation. It is not to be believed at a quantitative level, but it is perhaps the simplest model that captures in a physically plausible way the greenhouse-gas effects of both water vapour and CO_2 . The following is apparent:

- Adding carbon dioxide to the atmosphere causes temperature to go up (because $4\epsilon_a < 8$), providing $y < 1$, so the model delineates between forcing and feedback.
- The feedback is captured by the terms involving y, and it can be larger than the direct effect depending on the size of $\mathcal{B}$.
- As $y \to 1$ the feedback becomes very large, and this is called the *runaway greenhouse effect*. As the temperature increases the water vapour content increases, temperature further increases and so on.
- There is no a priori reason why y should be less than unity. For example, it will be large if the temperature is high, and so if e_s is high. It seems then that that T_g will *decrease* as c increases!

The last item seems totally unphysical, and to see what is going on we need to construct an explicit model of the greenhouse effect with water vapour feedback. We will do that soon but it will be easier if we must look in a bit more detail about radiation.

4 Radiative Transfer in a Grey Atmosphere

4.1 Assumptions

Radiative intensity, I is the radiative flux per solid angle and when dealing with radiation in three-dimensional problems we have to deal with directionality. We also have to deal with the dependence of absorption on wavelength, and with scattering. In dealing with radiation in the Earth's atmosphere we will make a number of main simplifications.

1. We can have completely separate treatments of solar and infra-red radiation.
2. Much of the time we can assume there is no solar absorption in the atmosphere. This is not quantitatively true but if it were the case, most of the atmosphere would be about the same.
3. We integrate over solid angles in the upward pointing hemisphere and again in the downward pointing atmosphere, so that we have two streams of radiation.

4. We'll integrate over wavelength in the infra-red and assume that a single emissivity suffices.

5. There is no scattering of infra-red radiation.

4.2 Equations of radiative transfer

Consider a monochromatic beam of radiation passing through a gas, and suppose for a moment the gas does not emit any radiation but only absorbs it. For a thin layer of gas the change in intensity of the beam is then

$$\mathrm{d}I = -I\mathrm{d}\tau \tag{21}$$

where τ is the *optical depth*. The equation may be regarded as a definition of optical depth — it is the fraction of the incoming radiation absorbed — with the difficulty then arising in relating it to the physical properties of the gas. Eq. (21) can be formally integrated to give $I = I_0 \exp(-\tau)$, where the factor $T = \exp(-\tau)$ is the *transmittance* of the layer. The optical depth of two layers is the sum of their optical depths and the total transmittance is the product of the two transmittances.

The optical depth of a gas is related both to the amount of gas and to its properties, and for a thin layer of gas of thickness $\mathrm{d}s$ we can write

$$\mathrm{d}\tau = k_A \rho \mathrm{d}s \tag{22}$$

where k_A is the mass absorption coefficient. In general the optical depth will depend on the wavelength but we shall assume it does not; that is, the atmosphere is *grey*. In the atmosphere if the pressure is hydrostatic then, in the vertical direction, $\mathrm{d}\tau = k_A \rho \, \mathrm{d}z = k_A \mathrm{d}p$ so that

$$\tau(p_1, p_2) = k_A(p_1 - p_2) \tag{23}$$

In fact the mass absorption coefficient increases with pressure so that in the atmosphere a somewhat better approximation is to write

$$\tau \approx \tau_r \frac{(p_1 - p_2)(p_1 + p_2)/2}{p_r^2} \tag{24}$$

where p_r is a reference pressure and τ_r is a reference optical depth, a function of the properties of the gas in question.

The slab of gas will also emit radiation, so taking this into account (21) becomes

$$\mathrm{d}I = (B - I)\mathrm{d}\tau \tag{25}$$

This is known as the Schwarzschild equation and it applies at each wavelength, but if we assume τ is not a function of wavelength and we integrate over all wavelengths then $B = \sigma T^4$. (You can either take this to be obvious or do a bit of algebra involving integrations over solid angles to convince yourself, or consult a radiation book like Goody or Petty or Pierrehumbert.) In terrestrial applications we assume that (25) applies in the infra-red, and do a separate calculation for solar radiation.

Now, in the atmosphere under two-stream approximation in the atmosphere we have upward, U, and downward, D, radiation and we write

$$-\frac{\mathrm{d}U}{\mathrm{d}\tau} = B - U, \qquad \frac{\mathrm{d}D}{\mathrm{d}\tau} = B - D. \tag{26a,b}$$

The *convention* we have chosen here is that τ increases downwards. This is convenient for atmospheric applications, for then we have $\tau = 0$ at the top of the atmosphere, but it is not mandated. We could choose it the other way and flip the signs of the right-hand sides and no physical result depends on this choice, or on the origin of τ. We will use these equations for the infra-red radiation and in what follows assume that solar radiation is all absorbed at the surface.

4.3 Solutions

Formal Solution

Consider the generic equation for radiation travelling in the direction of increasing τ or decreasing τ, B and U respectively

$$\frac{\mathrm{d}D}{\mathrm{d}\tau} = B - D, \qquad \frac{\mathrm{d}U}{\mathrm{d}\tau} = U - B. \tag{27}$$

Multiplying by the integrating factors $\exp(\tau)$ and $\exp(-\tau)$gives

$$\frac{\mathrm{d}}{\mathrm{d}\tau}(D\mathrm{e}^{\tau}) = B\mathrm{e}^{\tau}, \qquad \frac{\mathrm{d}}{\mathrm{d}\tau}(U\mathrm{e}^{-\tau}) = -B\mathrm{e}^{-\tau} \tag{28}$$

Integrating between $\tau = 0$ and τ' we obtain

$$D(\tau')\mathrm{e}^{\tau'} - D(0) = \int_0^{\tau'} B(\tau)\mathrm{e}^{\tau}\,\mathrm{d}\tau, \qquad U(\tau')\mathrm{e}^{-\tau'} - U(0) = -\int_0^{\tau'} B(\tau)\mathrm{e}^{-\tau}\,\mathrm{d}\tau \tag{29}$$

or

$$D(\tau') = \mathrm{e}^{-\tau'}\left[D(0) - \int_0^{\tau'} B(\tau)\mathrm{e}^{\tau}\,\mathrm{d}\tau\right], \qquad U(0) = U(\tau')\mathrm{e}^{-\tau'} + \int_0^{\tau'} B(\tau)\mathrm{e}^{-\tau}\,\mathrm{d}\tau \tag{30}$$

The first term in each solution is the attenuation of incoming radiation and the second is the cumulative emission. There are other ways to write the solution, but in general the solution of radiative problems can be written only in the form of integrals. Nevertheless, in some important special cases we can get a local solution as below.

Radiative equilibrium in planetary atmospheres

Consider an atmosphere with net incoming solar radiation S_{net} and suppose the planet is in radiative equilibrium with the incoming solar balanced by outgoing infra-red. The radiative transfer equations are thus to be solved with the boundary conditions that

$$D = 0,\ U = U_t \quad \text{at} \quad \tau = 0, \tag{31}$$

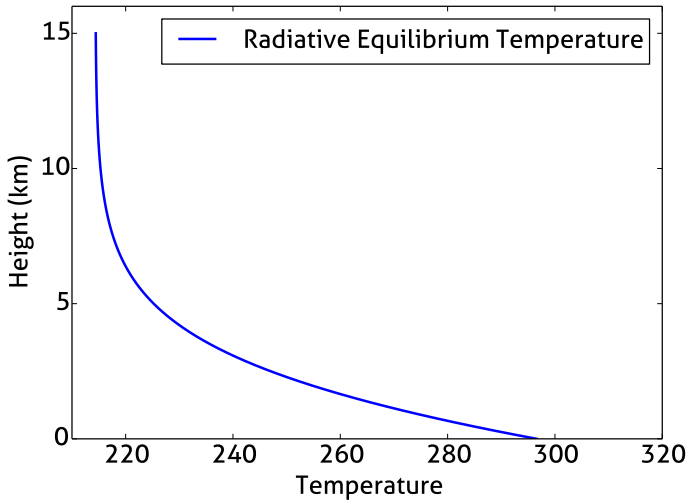


Figure 7: Radiative equilibrium temperature (solid curve) calculated using (36), with an optical depth of $\tau_0 = 8/3$, $H_a = 2\,\mathrm{km}$ and a net incoming solar radiation of $239\,\mathrm{W\ m^{-2}}$.

where $U_t = S_{\mathrm{net}}$ is the net outgoing long-wave radiation (OLR) at the top of the atmosphere. There are still too many variables as we don't know B, but we can obtain a *radiative equilibrium* solution if we assume there is no longwave heating in the column. The heating is proportional to the divergence of the net flux, so that if this is presumed zero then $\partial(U-D)/\partial z = 0$ so that

$$\frac{\partial (U-D)}{\partial \tau} = 0. \tag{32}$$

Let us rewrite (26) as

$$\frac{\partial}{\partial \tau}(U-D) = U + D - 2B, \tag{33a}$$

$$\frac{\partial}{\partial \tau}(U+D) = U - D. \tag{33b}$$

A solution of these equations that satisfies the boundary conditions is

$$D = \frac{\tau}{2} U_t, \qquad U = \left(1 + \frac{\tau}{2}\right) U_t, \qquad B = \left(\frac{1+\tau}{2}\right) U_t. \tag{34}$$

where U_t is the outgoing longwave radiation at the top of the atmosphere. The only thing remaining is to relate τ to z, and a simple recipe that is similar to (23) is to suppose that τ has an exponential profile.

$$\tau(z) = \tau_0 \exp(-z/H_a) \tag{35}$$

where typical values are $\tau_0 \approx 4$ and $H_a \approx 2\,\mathrm{km}$. The temperature then goes like

$$T^4 = U_t \left(\frac{1 + \tau_0 e^{-z/H_a}}{2\sigma} \right), \tag{36}$$

as illustrated in Fig. 7. Note the following aspects of the solution.

1. Temperature increases rapidly with height near the ground.

2. The upper atmosphere is nearly isothermal.

3. The temperature at the top of the atmosphere, T_t is given by

$$\sigma T_t^4 = \frac{U_t}{2} \tag{37}$$

   Thus, if we define the emitting temperature, T_e, to be such that $\sigma T_e^4 = U_t$, then $T_t = T_e/2^{1/4}$. Note also $B_t/U_t = 1/2$.

In fact, the temperature gradient near the ground varies so rapidly it is likely to be convectively unstable, which we come to in the next lecture. Also, note that we do not need to impose a temperature boundary condition at the ground; in fact there is no ground in this problem! — but what happens if we add one? That is, suppose that we declare that there is a black surface at some height, say $z = 0$, and we require that the atmosphere remain in radiative equilibrium. What temperature does that surface have to be?

From (34) the upward irradiance and temperature at any height z are related by

$$U = \left(\frac{2+\tau}{1+\tau}\right)\sigma T^4. \tag{38}$$

At $z = 0$ the surface will have to supply upwards radiation equal to that given by (38), and therefore its temperature, T_g is given by

$$\sigma T_g^4 = \left(\frac{2+\tau}{1+\tau}\right)\sigma T_s^4, \tag{39}$$

where T_s is the temperature of the fluid adjacent to the ground (the 'surface temperature'). That is, $T_g > T_s$ and there is a temperature discontinuity at the ground. Sometimes in very still conditions a very rapid change of temperature near the ground can in fact be observed, but usually the presence of conduction and convection will ensure that T_g and T_s are equal.

In the limit in which $\tau = 0$ in the upper atmosphere (let us prematurely call this the 'stratosphere') then we see that

$$D = 0, \qquad U = U_t, \qquad B = \frac{U_t}{2}. \tag{40}$$

That is, the atmosphere is isothermal, there is no downwelling irradiance and the upward flux is constant. The stratospheric temperature, T_{st} and the emitting temperature are related by

$$T_{st} = \frac{T_e}{2^{1/4}}. \tag{41}$$

Summary Points

To sum up, what have we found?

1. If we suppose the atmosphere is grey, and we know how optical depth varies with height, then if the atmosphere is in longwave radiative equilibrium we can construct an explicit solution for the temperature as a function of height.

2. The temperature will typically decrease very rapidly in height away from the surface. So much so it is likely to be convectively unstable, as we discuss in the next lecture.

3. The radiative equilibrium temperature does not care or know whether a surface (i.e., the ground) is present. If a surface *is* present, and we require that radiate equilibrium still hold, the temperature of the ground must be higher than the temperature of the air adjacent to it. This is because the ground must supply the same amount of radiation as would be supplied by an infinite layer of air below that level. Thus, there is a temperature discontinuity at the ground, which in reality would normally be wiped out by convection.

5 An explicit model of the Runaway Greenhouse Effect

We now come back to the greenhouse effect and construct an explicit model of runaway greenhouse. (The term 'runaway greenhouse' was coined by Ingersoll (1969).) Suppose the atmosphere is in radiative equilibrium. From the derivations above, we can relate the surface and ground temperatures to the incoming solar radiation through the relation

$$T_s^4 = \frac{T_e^4}{2}(1+\tau_0), \qquad T_g^4 = T_e^4(1+\frac{\tau_0}{2}) \tag{42}$$

Thus, if $\tau_0 = 1.254$ then, for $T_e = 255\,\mathrm{K}$ we find $T_g = 288\,\mathrm{K}$ and $T_s = 262\,\mathrm{K}$. The assumption of radiative equilibrium and the ensuing temperature discontinuity are unrealistic but the model will illustrate and important point. We'll construct a more realistic model in the next lecture.

Suppose that we let τ_0 be a function of temperature, increasing with the saturation vapor pressure at the surface. Thus, let

$$\tau_0 = A + Be_s(T_g) \tag{43}$$

where A and B are semi-empirical constants, and e_s is the saturation vapor pressure as given by the solution of the Clausius–Clapeyron equation, (15). We will tune their values such that $T_g = 288$ when $T_e = 255$, and with some experience of hindsight we set the ratio $A/B = 8$, whence we obtain $A = 1.12$ and $B = 0.14$. The reason for such a seemingly high ratio is that a grey model is too prone to give a runaway greenhouse because of its lack of windows in the infra-red. Thus, in reality, even as temperature and water vapor content increase some infra-red radiation *can* escape from the surface.

Putting the above together, the ground temperature is solution of

$$T_g^4 = T_e^4\left(1+\frac{1}{2}[A+Be_s(T_g)]\right). \tag{44}$$

This algebraic equation is quite nonlinear and must be solved numerically but a few points are apparent.

1. For any given T_e we can obtain a graphic solution by plotting T_g and $T_e^4(1+\tau_0/2)^{1/4}$ and seeing where the two curves intersect. For a range of values of T_e we will obtain two solutions, as illustrated in Fig. 8. However, if T_e is too high there will be no intersection of the curves because the value of $T_e^4(1+\tau_0/2)$ will always be larger than T_g.

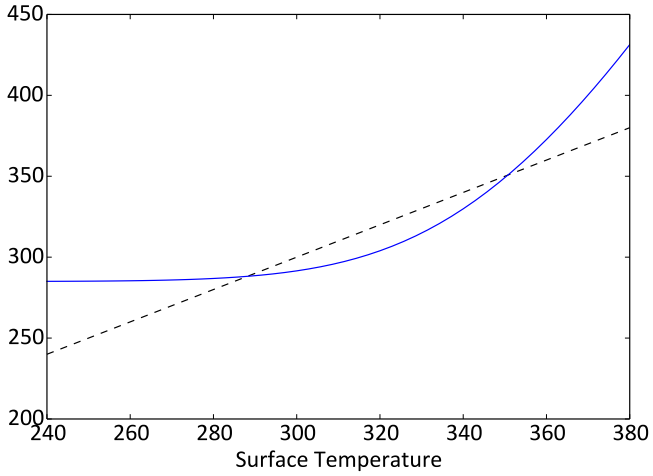


Figure 8: Graphical solution to the energy balance model (44) with $T_e = 255\,\mathrm{K}$. The dashed curve is T_g and the solid curve plots values of $T_e^4(1+\tau_0(T_g)/2)^{1/4}$, with τ_0 given by (43). Solutions occur at $T_g \approx 288\,\mathrm{K}$ and $T_g = 350\,\mathrm{K}$.

2. If T_e increases and Be_s is much smaller than A, then a solution is found by increasing T_g.

3. If T_e increases and Be_s is suitably large then we can imagine that a solution will be found with a *lower* value of T_g.

Numerical solutions, found iteratively, are illustrated in Fig. 9, and as expected there are two branches to the solution. [A much more detailed discussion with many extensions is to be found in the report by P. Martin in this volume.] For the parameters plotted, there is no solution if $T_e > 269\,\mathrm{K}$. That is to say, if a planet obeying the model above were in an orbit such that $T_e > 269\,\mathrm{K}$ then infra-red radiation would not be able to escape from the surface, and the surface temperature would keep on rising. All the water on the planet surface would boil, and eventually the water vapor would escape to space. Such a scenario may have occurred on Venus in the past.

5.1 Stability of solutions

The upper branch of the solution plotted in Fig. 9 runs counter to our intuition, in that temperature decreases as emitting temperature increases. The situation arises because the greenhouse effect is so strong, so that an increase in emitting temperature can lead to a decrease in surface temperature if the greenhouse effect also falls considerably. However, this solution is *unstable* as we now show.

We add a time dependence to the energy balance model and write

$$C\frac{\mathrm{d}T_g}{\mathrm{d}t} = \sigma T_e^4 - \sigma\frac{T_g^4}{1+\tau_0/2} \tag{45}$$

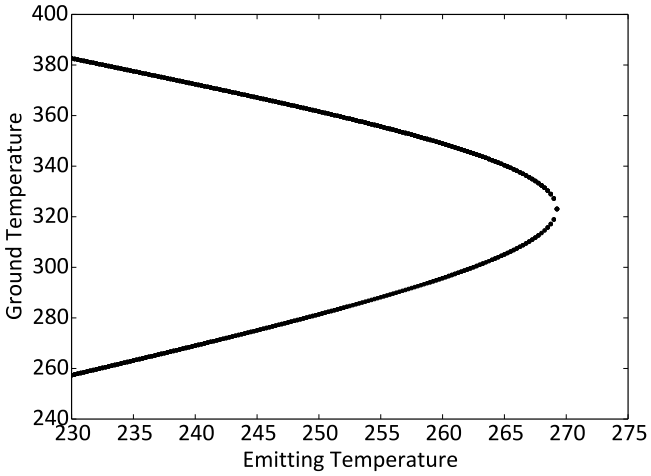


Figure 9: Solutions the energy balance model (44) obtained numerically. Plotted are values of T_g as a function of The dashed curve is T_g and the solid curve plots values of $T_e^4(1 + \tau_0(T_g)/2)^{1/4}$, with τ_0 given by (43) .

We perturb the system about an equilibrium point and so obtain

$$C\frac{\mathrm{d}T_g'}{\mathrm{d}t} = \frac{-4\sigma T_g^3 T_g'}{1+\tau_0} + \frac{\sigma T_g^4 \tau_0'}{(1+\tau_0)^2}, \tag{46a}$$

$$= \left(\frac{T_g}{1+\tau_0/2}\frac{\mathrm{d}\tau_0}{\mathrm{d}T_g} - 4\right) T_g' \tag{46b}$$

Thus, the solution will be stable or unstable according as whether the term in brackets is negative or positive, respectively.

A tiny bit of algebra will reveal that the ratio of the two terms in brackets in (46b) precisely the same as the ratio of the gradients of the solid curve and the dashed curve at the intersection points in Fig. 8. Thus, the solution at the higher temperature (about 350 K in the graph) is unstable, because the gradient of the blue curve is greater than the gradient of the dashed curve. Similarly, the solution at the lower temperature (288 K) is stable. All of the solutions on the upper branch on Fig. 9 are therefore unstable.

References

Ingersoll, A. P., 1969. The runaway greenhouse: A history of water on Venus. *J. Atmos. Sci.*, **26**, 6, 1191–1198.

Lecture 3: Convective Heat Transfer I

Kerry Emanuel; notes by Paige Martin and Daniel Mukiibi

June 18

1 Introduction

In the first lecture, we discussed radiative transfer in the climate system. Here, we will delve into the convective heat transfer within the climate system, and how this leads to the radiative-dry convection equilibrium as an extension of the radiative equilibrium that we saw in Lecture 1.

2 The Buoyancy Equation

From the first lecture, we have that the buoyancy, B, of a fluid particle is given by the formula (stemming from Archimedes' Principle)

$$B = g\frac{\rho_e - \rho_b}{\rho_b}, \tag{1}$$

where subscripts e and b denote that the variable pertains to the environment and the sample fluid parcel, respectively. For convenience, we make two variable substitutions. First, we switch from using density to specific volume α using the relation $\alpha = 1/\rho$, which yields

$$B = g\frac{\alpha_b - \alpha_e}{\alpha_e} = g\frac{\delta\alpha}{\alpha_e}. \tag{2}$$

Specific volume, however, is not conserved under adiabatic compression, but the entropy s is conserved. Hence comes our second substitution from α to s. To make this substitution, we go through a few steps as follows:

$$\delta\alpha = (\tfrac{\partial\alpha}{\partial p})_s\delta_p + (\tfrac{\partial\alpha}{\partial s})_p\delta_s.$$

Setting the first term on the right side to zero, we are left with

$$\delta\alpha = (\tfrac{\partial\alpha}{\partial s})_p\delta_s.$$

We can then employ a Maxwell Relation

$$(\frac{\delta\alpha}{\delta s})_p = (\frac{\delta T}{\delta p})_s. \tag{3}$$

(The Maxwell Relations are a set of equations stating the relation between derivatives of thermodynamic variables. For a more in-depth discussion and derivation of the Maxwell Relations, refer to a standard Thermodynamics textbook.)

The term on the right can be plugged into the above equation for buoyancy (2):

$$B = \frac{g}{\alpha}(\frac{\delta T}{\delta p})_s \delta s.$$

Assuming a hydrostatic environment, we can replace p with vertical height z according to $\alpha dp = -gdz$, giving the final equation

$$B = -(\frac{\delta T}{\delta z})_s \delta s = \Gamma \delta s, \tag{4}$$

where Γ is the adiabatic lapse rate.

3 Stability

One of the primary questions in this lecture pertains to the instability of the atmosphere to convection. We can consider a parcel of air that we displace upwards. If the buoyancy of the parcel is also upwards, then the parcel is unstable and will continue to accelerate upwards. If, however, the buoyancy is downwards, then the parcel will accelerate back toward its original position. Relating this now to entropy, we see that if entropy decreases with height, then a particle displaced upwards will have a higher entropy than its surrounding and will thus be unstable. If entropy is constant in height, then the atmosphere is neutrally stable.

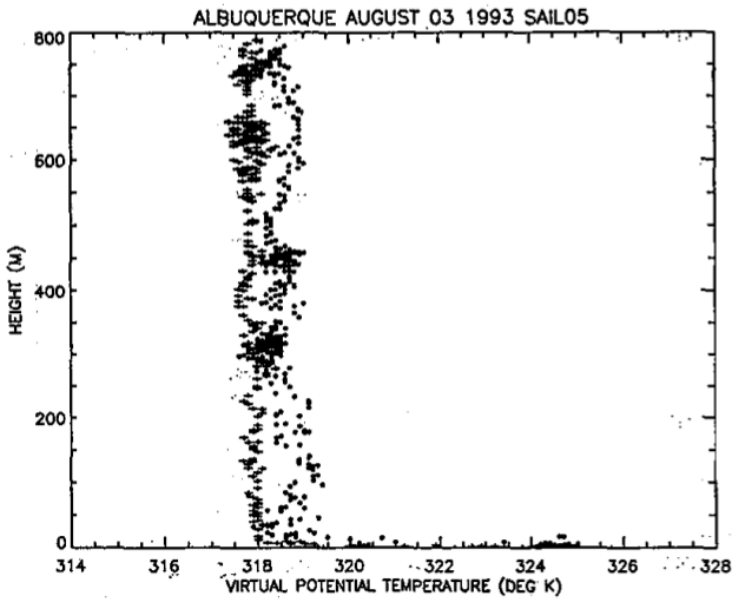


Figure 1: The profile of virtual potential temperature, which can be considered to be entropy for our purposes.

Figure 1 shows measurements taken from a model airplane of the entropy (the x-axis shows the virtual potential temperature, which behaves essentially like entropy) in the

atmosphere up to around 800m. It is clear that the entropy profile is nearly constant, with the exception of a few meters above the surface. This constant entropy layer implies the existence of a convecting layer, which can account for the mismatch in temperatures between the ground and the layer of air just above the ground in the radiative equilibrium model. Convection acts on a small enough time scale that it is able to destabilize the atmosphere, yielding a constant entropy layer.

4 The Prandtl Problem

(For a detailed description of the Prandtl problem discussed below, see Prandtl L 1925 Z. Angew. Math. Mech. 5 136.)

We consider now the Prandtl Problem, as shown in Figure 2. There is a rough bottom at constant temperature, which we set to be z=0. The fluid, which is subject to gravity, is cooled such that the vertical integral of cooling over the whole depth is constant. In mathematical terms, this can be written

$$\int_0^\infty \dot{Q} dz = const. \tag{5}$$

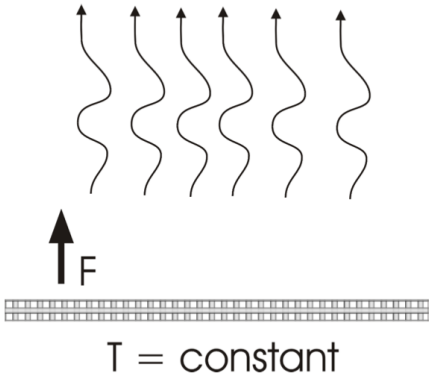


Figure 2: Schematic of the Prandl problem.

There is a flux of heat from the lower boundary, which is the convective (buoyancy) flux F. F should be constant, and can be written as an average defined by

$$F = \overline{w'B'}, \tag{6}$$

where w' is the perturbation vertical velocity and B' the perturbation buoyancy.

To find a scale for the velocity in terms of F and z, we perform dimensional analysis. The following are the relevant dimensions in the problem:

$$F \sim length^2/time^3$$
$$z \sim length$$

The desired scaling for $q \sim length/time$ can be achieved by

$$q \sim (Fz)^{\frac{1}{3}}. \tag{7}$$

Similarly, we can find a scaling for the buoyancy (which has units of length/time^2) in terms of F and z:

$$B' \sim (\frac{F^2}{z})^{\frac{1}{3}}. \tag{8}$$

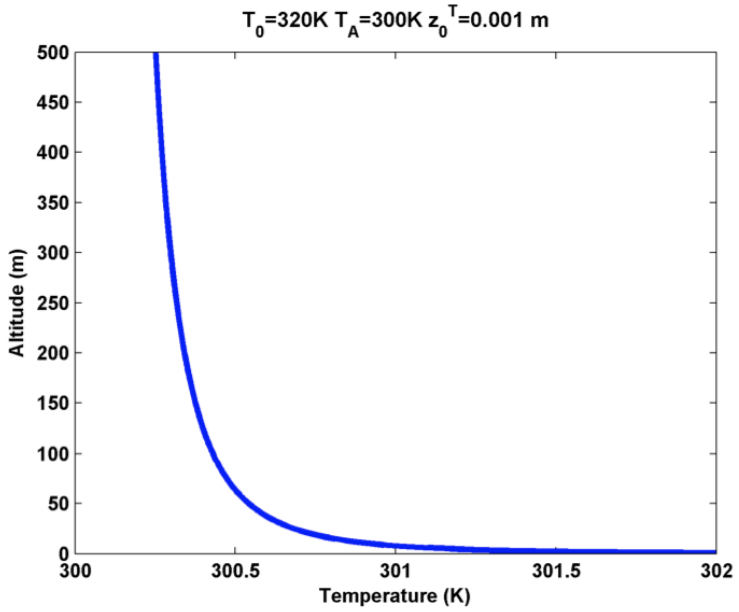


Figure 3: The vertical temperature profile.

We can then express the buoyancy B in terms of the average buoyancy B_0 and the perturbation buoyancy B':

$$B \sim B_0 - cF^{\frac{2}{3}}[(z_0^T)^{\frac{-1}{3}} - z^{\frac{-1}{3}}], \tag{9}$$

where we define z_0^T to be the thermal roughness scale. This dependence is shown in figure 3, where temperature is plotted on the x-axis in lieu of buoyancy, but they are simply related by the coefficient of thermal expansion and thus have qualitatively similar behavior.

The radiative-dry convective equilibrium is shown in figure 4, with the pure radiative equilibrium plotted for comparison. From the graph, a few observations can be made. First, the temperature in the radiative-convective equilibrium in the troposphere and near the surface is significantly larger than in the purely radiative state. Why do we see this temperature increase in the troposphere and surface? This is because we assumed a constant relative humidity (a big assumption!), which implies more water vapor, hence a greater greenhouse effect, and thus a warmer temperature. However, there remains an important question - why is this graph so different than what we observe in the real world? This will be answered in the next section!

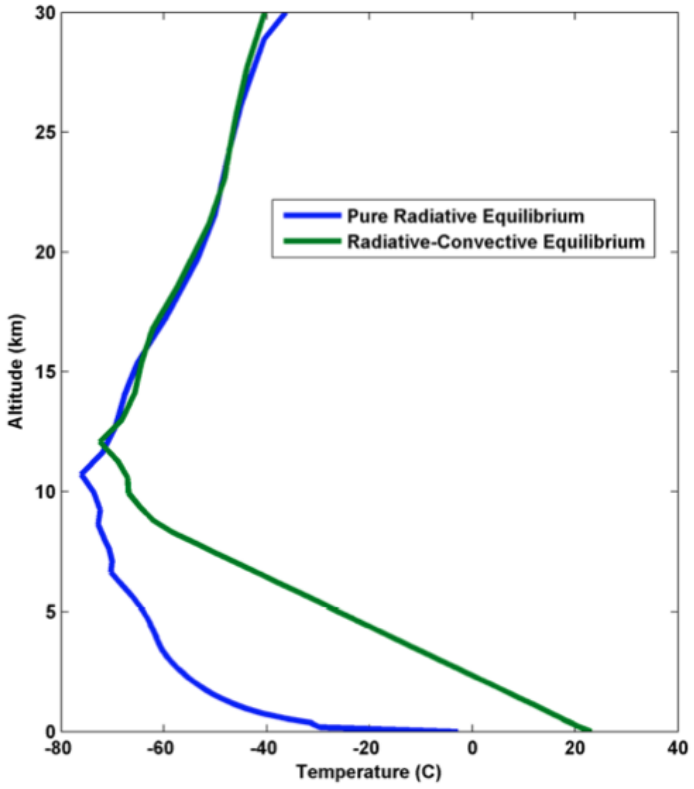


Figure 4: Full calculation of radiative-dry convective equilibrium.

5 Radiative - Dry Convective Model

In figure 5 is shown a simple model, very much like the one discussed in the first lecture. There are two layers, both opaque, but we are now assuming that, in addition to radiative fluxes, there are convective fluxes, denoted by F. By assumption (i.e. because we are forcing convective neutrality), we write

$$T_1 = T_2 + \Delta T$$
$$T_S = T_1 + \Delta T = T_2 + 2\Delta T.$$

We can still write $\sigma T_e^4 = \sigma T_2^4$, yielding the temperature of the second layer $T_2 = T_e$. This dictates the temperatures of the other layers:

$$T_1 = T_e + \Delta T$$
$$T_S = T_e + \Delta T.$$

At the surface we can write the equation

$$\sigma T_e^4 + \sigma T_1^4 = \sigma T_S^4 + F_S.$$

This can be rearranged to solve for the surface convective flux:

$$F_S = \sigma(-T_S^4 + T_e^4 + T_1^4) = \sigma T_e^4[1 + (1+x)^4 - (1+2x)^4], \tag{10}$$

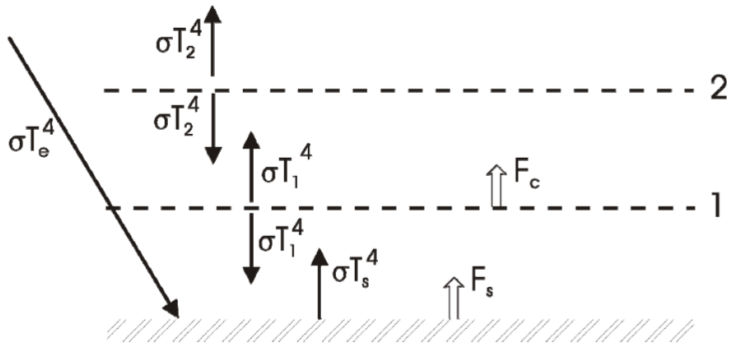


Figure 5: A simple radiative-convective energy balance model.

where x is defined to be $x = \frac{\Delta T}{T_e}$.

Similarly, the equation for the second layer can be written

$$2\sigma T_e^4 = \sigma T_1^4 + F_C,$$

and then solving for the convective flux

$$F_C = \sigma T_e^4[2 - (1 + x)^4]. \tag{11}$$

The convective fluxes are needed to maintain a constant lapse rate. This leads into the topic introduced below, and discussed in detail in Lecture 5 about moist convection.

6 Introduction to Moist Convection

Moist convection is important for a number of reasons, and it will shape the equilibrium curve into something that looks more similar to the real-world picture.

Water in the atmosphere is responsible for a significant amount of heating due to phase changes (discussed below), and is considered one of the most important greenhouse gasses. Moist convection also plays a key role in stratiform cloudiness, and thus the planet's albedo and long-wave trapping.

6.1 Water Variables

The following are the variables which will be used in subsequent lectures.

$q = \frac{M_{water}}{M_{air}}$	specific humidity
e	vapor pressure (partial pressure of water vapor)
$e^* = 6.112hPa\ e^{\frac{17.67(T-273)}{T-30}}$	saturation vapor pressure
$H = \frac{e}{e*}$	relative humidity
$q*$	saturation specific humidity

The ideal gas law will be used, where R^* and $\bar{m}$ are the universal gas constant and the molecular weights of the constituents, respectively:

$$p = \rho \frac{R^* T}{\bar{m}},$$

which can be rewritten in terms of the vapor pressure

$$e = \rho_v \frac{R^* T}{\bar{m}_v},$$

with $\bar{m}_v$ referring to the mass of the vapor particles. The specific humidity is the ratio of water vapor density to total density:

$$q = \frac{\rho_v}{\rho} = \frac{m_v}{\bar{m}} \frac{e}{p},$$

and thus the saturation specific humidity is expressed as

$$q^* = \frac{m_v}{\bar{m}} \frac{e^*}{p}. \tag{12}$$

In Lecture 5, which continues the discussion of convective heat transfer, we will use these quantities and equations defined above to study moisture in the atmosphere, which will lead us to the Radiative-Moist Convective Equilibrium (RCE).

Lecture 4: Radiative-Convective Equilibrium and Tropopause Height

Geoff Vallis; notes by Erica Rosenblum and Ashley Payne

June 19

We now consider what the effect of convection might be on all the concepts and solutions found in lecture 2. Because our interest in mainly in the large scale structure of the atmosphere we will take a somewhat simplistic view of convection and suppose that it acts to restore an unstable lapse to something that is neutrally stable, that lapse rate being given by either the dry adiabatic or moist adiabatic lapse rate. Readers interested in finding out more about convection and radiative-convective equilibrium should consult Kerry Emanuel's lecture notes.

1 Radiative-convective equilibrium

In lecture 2 we found that in radiative equilibrium the temperature falls off very rapidly with height in the lower atmosphere, so much so that it is likely to be convectively unstable. We imagine the atmosphere will convect and that the lapse rate will adjust until it is stable, as in Fig. 1, up to some height H_T. Sometimes, either instead of or in addition to, heat may be transported upwards by large-scale motion such as baroclinic waves. In either case, let us suppose that the dynamics acts such as to produced constant lapse rate up to some height H_T, which we will later associate with the tropopause. We wish to obtain an expression for that height. That is, we seek a solution for which

$$z \le H_T : \qquad T = T_s - \Gamma z \tag{1a}$$

$$z > H_T : \qquad \text{Radiative equilibrium, satisfying (lec.2:26) and (lec.2:32)} \tag{1b}$$

Further, since we are imposing a convective heat flux, we can suppose that at the surface the temperature is continuous, so that the ground temperature is such that $\sigma T_g^4 = U(z = 0)$.

To obtain a solution we might just think of adjusting the lapse rate in (lec2:fig.7) so that there is no net heating, and this may indeed be what convection does on a short timescale. However, an overall radiative balance is not necessarily then achieved, so that the system will then evolve further. The variable in this equation is H_T, and this can be adjusted until (lec.2:36) is satisfied, with the outgoing radiation The solution of these equations requires an iterative approach and the algorithm is as follows.

1. First solve the radiative transfer equations for radiative equilibrium.
2. Make a guess for the height of the tropopause, and hence obtain the temperature all the way down to the ground.

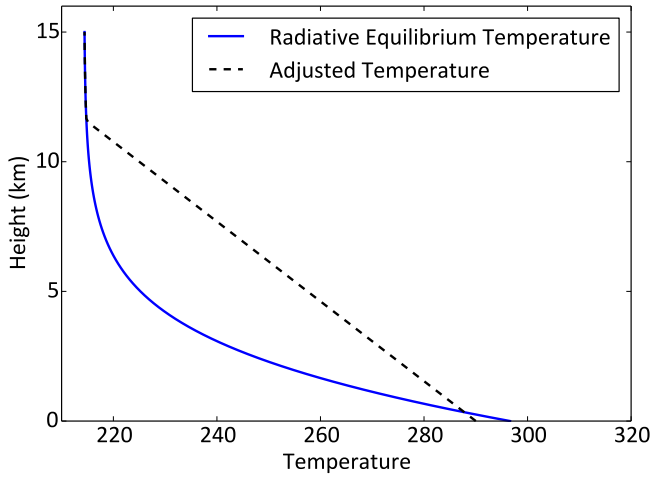


Figure 1: Radiative equilibrium temperature (solid curve) calculated using (lec.2:36), with an optical depth of $\tau_0 = 8/3$, $H_a = 2\,\mathrm{km}$ and a net incoming solar radiation of $239\,\mathrm{W\ m^{-2}}$. The dashed line shows a schematic adjusted temperature with a lapse rate of $6.5\,\mathrm{K\ km^{-1}}$ up to a tropopause (at about 11 km here) and a radiative equilibrium temperature in the stratosphere.

3. Integrate the radiative transfer equations down from the top. The outgoing radiative balance is achieved this way, but there is no balance at the surface if temperature is continuous. That is, $\sigma T_g^4 \neq U(z=0)$.

4. Change the height of the tropopause, find another solution, and iterate until the surface radiative balance is achieved.

An alternative is to specify the surface temperature and integrate the radiative transfer equations up along a given lapse rate from the bottom to a certain height, beyond which we suppose that radiative equilibrium holds. This procedure will not give the correct outgoing radiation, so the procedure must again be iterated.

1.1 Global Warming

Without actually solving the RCE equations we can make an important deduction as to what happens to the height of the tropopause under global warming, that is what happens when additional carbon dioxide is added to the atmosphere. If the atmosphere stays in radiative balance (which it will in the long term) then the outgoing radiation remains the same. If the stratosphere has a small optical depth then its temperature stays the same from (lec.2:40). Therefore *the temperature of the tropopause must stay the same!* However, the height of the emitting temperature must increase, because the emissivity of the lower atmosphere increases, and the photons that reach space come, on average, from a higher level in the troposphere. And, as a consequence, the troposphere warms as illustrated in Fig. 2. But if the temperature of the tropopause is to stay the same then its height must increase, and a simple calculation tells us by how much.

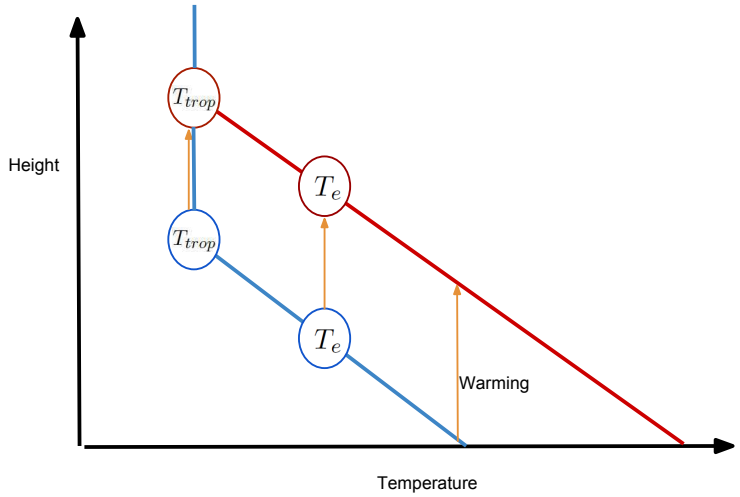


Figure 2: Schematic of temperatures before (blue line) and after an increase in optical depth of the atmosphere, such as happens in global warming. The troposphere warms but the emitting temperature stays the same. Hence the tropopause temperature stays the same and the height of the tropopause increases.

If the lapse rate stays the same then the tropopause height will increase by an amount ΔH_T given by

$$\Delta H_t = \frac{\Delta T_s}{\Gamma} \tag{2}$$

where ΔT_s is the change in surface temperature. If we allow the lapse rate to change also, then

$$\Delta H_t = \frac{\Delta T_s}{\Gamma} - H_t \frac{\Delta \Gamma}{\Gamma} \tag{3}$$

or

$$\frac{\partial H_t}{\partial T_s} = \frac{1}{\Gamma} - \frac{H_T}{\Gamma} \frac{\partial \Gamma}{\partial T_s}. \tag{4}$$

If we suppose that Γ is the moist adiabatic lapse rate then we can calculate this expression analytically, and some results are shown in Fig. 3, where the lapse rate is assumed constant with height and a function of surface temperature. It is interesting that the increase in tropopause height is quite significant – about 400 m per degree – and that both the direct temperature effect and the lapse rate effect are important (at least in regions where the lapse rate is moist adiabatic). An increase in tropospheric height is one of the most robust results we have concerning changes of the structure of the atmosphere under global warming, as discussed more in Vallis *et al.* (2014).

2 The Height of the Tropopause

We now provide an approximate, analytic, expression for the height of the tropopause.[1] We assume the following.

[1] This section is joint work with Pablo Zurita-Gotor.

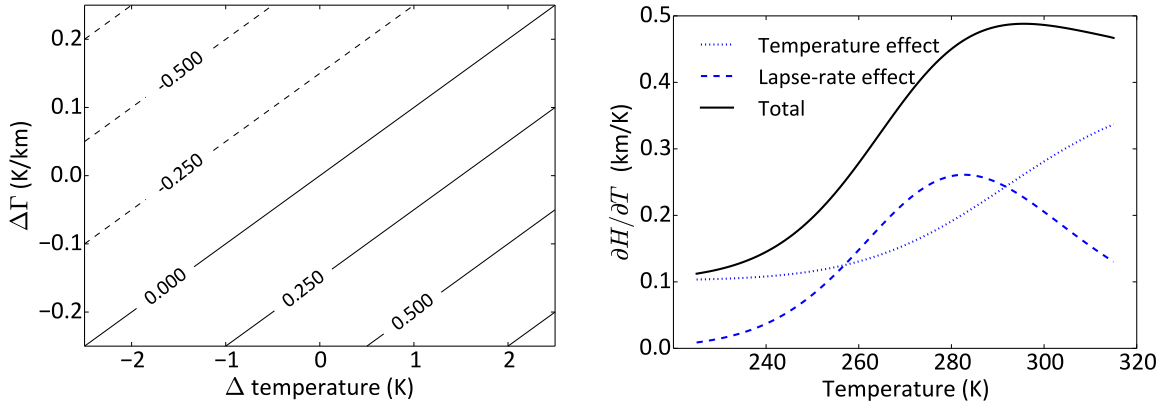


Figure 3: (a) Contours of change in tropopause height (km) as a function of temperature change and lapse rate change, calculated using (3). (b) Rate of change of tropopause height with temperature ($\partial H_t/\partial T$) as a function of temperature, calculated using (4).

1. Single column (so a one-dimensional calculation).
2. Grey atmosphere with an optical thickness that decays exponentially with height.
3. A specified lapse rate to some height H_T, beyond which there is radiative equilibrium.
4. An optically thin atmosphere in the upper troposphere and stratosphere.
5. An overall radiative balance. So the outgoing IR radiation is specified (equal to net incoming solar).
6. No surface temperature discontinuity. So ground temperature equals surface air temperature ($T_g = T_s$), and the upwards radiation at $z = 0$ is given by σT_g^4.

2.1 Algorithm

To find an exact solution the equations must be iterated, and an algorithm for that is as follows.

1. First numerically integrate (lec.2:33) to obtain a radiative equilibrium solution.
2. Guess a height for the tropopause and thus obtain a temperature at all levels below that, including the ground, using the given lapse rate.
3. Calculate the radiative fluxes by integration of (lectire 2:33) down from the top. The upwards radiation at the ground will in general not equal σT_g^4.
4. Adjust the height of the tropopause and repeat step (2) and (3).
5. Iterate the calculation until a surface balance is achieved.

An alternative procedure is to guess a surface temperature and integrate the equations up, assuming a constant lapse rate up to a height H_T, with radiative equilibrium beyond. When this is done the temperature at H_T will not be the correct one, and outgoing radiation will not equal to the incoming radiation, and again we have to iterate.

2.2 Analytic approximation

The analytic approach involves obtaining an analytic expression for the outgoing radiation for a given temperature profile along the lines of (lec.2:30). The OLR so obtained will be a function of the height of the tropopause, and by making the expression equal to the incoming solar radiation we obtain an expression for the tropopause height. Instead of actually using (lec.2:30) it is easier to solve the equations approximately *ab initio.* We make one other approximation, that the value of B/U varies linearly from the tropopause (where its value is 0.5) to its value at the surface (where $B/U = 1$). Thus,

$$\frac{B}{U} = 1 - \frac{z}{2H_T}. \tag{5}$$

Numerical calculations suggest this is a decent approximation (can it be improved upon?). Rewrite (lec.2:26a) as

$$\frac{\mathrm{d}\log U}{\mathrm{d}\tau} = 1 - \frac{B}{U} = \frac{z}{2H_T}. \tag{6}$$

Using $\tau(z) = \tau_s \exp(-z/H_a)$ we obtain

$$\frac{\mathrm{d}\log U}{\mathrm{d}z} = -\frac{z}{2H_T H_a}\tau_s \exp(-z/H_a). \tag{7}$$

We can integrate this expression by parts to obtain a value of the upwelling radiation at the tropopause $U(H_T)$, namely

$$\log\left(\frac{U(H_T)}{U(0)}\right) = -\frac{\tau_s}{2H_T}\int_0^{H_T} \exp(-z/H_a)\,\mathrm{d}z \approx -\frac{\tau_s H_a}{2H_T}. \tag{8}$$

for $H_T \gg H_a$. This is an expression for the outgoing longwave radiation, and we see that the only variable in the equation is H_T – note that the upwelling radiation at the surface is given by the surface temperature, which is a function of the tropopause temperature, H_T and the lapse rate, Γ.

To obtain a closed form for the tropopause height assume that the stratosphere is optically thin and note that $U(H_T) = U(H_T) = 2\sigma T_T^4$ and $U(0) = \sigma T_g^4 = \sigma T_s^4$. Furthermore, T_T and T_s are related by $T_T = T_s - \Gamma H_T$. The left-hand side of (8) then becomes

$$\begin{aligned}\log\left(\frac{2\sigma T_T^4}{\sigma T_s^4}\right) &= \log 2 + 4\log\frac{T_T}{T_s} = \log 2 + 4\log\left(\frac{T_T}{T_T + \Gamma H_T}\right) \\ &\approx \log 2 - \frac{4\Gamma H_T}{T_T}.\end{aligned} \tag{9}$$

Using (9), (8) becomes

$$\log 2 - \frac{4\Gamma H_T}{T_s} = -\frac{\tau_s H_a}{2H_T} \tag{10}$$

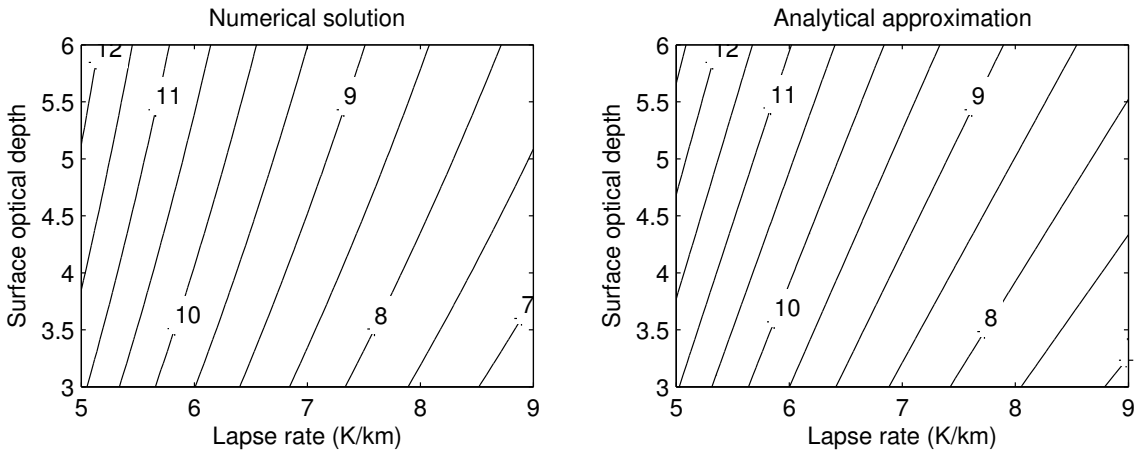


Figure 4: Analytic approximation and numerical calculation for tropopause height.

or

$$8\Gamma H_T^2 - \mathrm{C}\, H_T T_T - \tau_s H_a T_T = 0. \tag{11}$$

where $\mathrm{C} = 2\log 2 \approx 1.38$. The solution of this equation is

$$H_T = \frac{1}{16\Gamma}\left(\mathrm{C}T_T + \sqrt{\mathrm{C}^2 T_T^2 + 32\Gamma\tau_s H_a T_T}\right). \tag{12}$$

For Earth's atmosphere, $H_a \approx 2\,\mathrm{km}$, $\tau_s \approx 8/3$ and $\Gamma \approx 6.5\,\mathrm{K\ km^{-1}}$. All three terms in the quadratic are then approximately the same size and $H_T = 10.3\,\mathrm{km}$, which is in fact reasonably close to the exact numerical solution (obtained iteratively) of the radiative-convective equations (Fig. 4).

The numerical approximation of the logarithm in (9) can be improved by using T_m instead of T_T, where T_m is the temperature half way between the surface and the tropopause. However, we still want to have T_T as a parameter in the quadratic for H_T (because T_T is given if the OLR is known). Thus, we have to do some more algebraic fiddling and the upshot is that we get a quadratic similar to (11) but with different coefficients. [Student exercise. See also Vallis et al (2014) for another way to proceed.]

Once we have the tropopause height we can obtain an expression for the temperature everywhere in the troposphere, and the surface. We could then perform a calculation similar to that of section (lec.2:3.2) and obtain an analytic expression for how the surface temperature increases with carbon dioxide content, and the conditions for a runaway greenhouse effect. [With extensions this could be a student project.]

Optically thick and thin limits

The above approach allows us to be precise about what it means for an atmosphere to be optically thin or thick. Using (12) and approximating $C^2 = 2$ we easily find that the optically thick limit arises when

$$\tau_s H_a \gg \frac{T_T}{16\Gamma} \quad \text{whence} \quad H_T \approx \sqrt{\frac{T_T \tau_s H_a}{8\Gamma}} \tag{13}$$

The optically thin case has

$$\tau_s H_a \ll \frac{T_T}{16\Gamma} \quad \text{whence} \quad H_T \approx \frac{1.38 T_T}{8\Gamma}. \tag{14}$$

With parameters appropriate for Earth's atmosphere both of the above limits give estimates in the range 5–10 km , and note that they are additive effects. What is the interpretation of these expressions? Do they work on other planets? What is the role of lateral heat transport?

A number of these issues have been taken up by Shineng Hu in his summer project, and the interested reader is referred to his report for more details.

3 Lateral Transport

The actual tropopause height is determined by a combination of lateral heat transport and the RCE state above. Suppose we think of the energy balance of a column of air. If there is no horizontal divergence of heat flux into the column then we have the same situation as before, and the tropopause height is determined by the RCE argument. This in fact will be the situation somewhere in mid-latitudes where the horizontal heat flux is a maximum. At this particular location we have the RCE problem above.

Elsewhere there is a flux of heat into or out of the column and this will affect the tropopause height. As far as the column is concerned these effects are similar to changes in the outgoing longwave radiation, and so it would change the tropopause height in the same way that changing the outgoing longwave radiation would. That is, suppose you have separately solved the problem on how heat is transferred horizontally as well as the RCE problem above. In that case you know the outgoing radiation at any particular latitude and you know the lapse rate and so you can determine the tropopause height. In fact, this effect probably will not make a big difference to the tropopause height because the sensitivity of tropopause height on tropopause temperature is fairly weak. Thus, if the lapse rate is fixed independently of the horizontal dynamics the height of the tropopause will only be affected by these dynamics to a limited degree.

However, there are other ways that the dynamics affects things, and one is in the determination of the lapse rate itself. This may occur in either the tropics or the extratropics, where the mechanisms will be slightly different. In the tropics the prime determinant of the lapse rate is moist convection, and a simple possibility is that the lapse rate is given by the moist adiabatic lapse rate. However, this is itself a fairly strong function of temperature, so that the tropopause height becomes a function of temperature mainly through the effect that horizontal transfer has on the moist adiabatic lapse rate, not temperature itself.

Another argument (Held?) is that the moist static energy at the tropopause is almost the same as what it is at the surface, whence

$$N^2 \approx \frac{L q_s g}{c_p T_s H_T} \tag{15}$$

where H_T is, again, tropopause height and q_s and T_s are the surface values of water vapour and temperature. Note that H_T decreases with increasing stability in this expression,

whereas it increases with stability (i.e., decreases with lapse rate) in the RCE expression (12). Only one value of the tropopause height is consistent with both.

In midlatitudes the situation is complicated because the heat transport is effected by baroclinic instability and this has a characteristic height. One possibility is to try to adjust things so that height of the tropopause is consistent both with baroclinic instability and with the radiative constraint. There are a number of possibilities that we might wish to consider (good student projects!).

Suppose that baroclinic instability is like the Eady problem. In this case the instability goes from the surface to the tropopause; there is no additional height scale.

But suppose that we have a β-plane. In this case the vertical scale is given by the 'Charney height' which, in a Boussinesq system, is

$$h = \frac{f^2 \Lambda}{\beta N^2} \tag{16}$$

where Λ is the vertical shear of the zonal wind, $\Lambda = \partial U/\partial z$. We might argue that the height, h, in (16) must be the same as that given by the radiative constraint, and this gives us a theory of the stratification of the atmosphere. Just as in the tropics, h decreases with increasing stability so there is only one solution. If we suppose that $h = H_T$, where H_T is the height of the tropopause given by the radiative constraint, then using the thermal wind equation, $f\Lambda = -\partial b/\partial y$ and with $N^2 = \partial b/\partial z$ we find

$$H_T = -\frac{f \partial b/\partial y}{\beta \partial b/\partial z} \tag{17}$$

or

$$s = \frac{f}{\beta H_T} \sim \frac{a}{H_T}, \tag{18}$$

where s is the slope of the isopycnals and a is the Earth's radius. This means that an isopycnal roughly goes from the surface at the equatorward edge of the midlatitudes to the tropopause at the pole. This hypothesis and its friends have generated quite a lot of controversy in the community....

4 Appendix: Dry and Wet Lapse Rates

4.1 A dry ideal gas

The negative of the rate of change of the temperature in the vertical is known as the *temperature lapse rate*, or often just the lapse rate, and the lapse rate corresponding to $\partial\theta/\partial z = 0$ is called the *dry adiabatic lapse rate* and denoted Γ_d. Using $\theta = T(p_0/p)^{R/c_p}$ and $\partial p/\partial z = -\rho g$ we find that the lapse rate and the potential temperature lapse rate are related by

$$\frac{\partial T}{\partial z} = \frac{T}{\theta}\frac{\partial \theta}{\partial z} - \frac{g}{c_p}, \tag{19}$$

so that the dry adiabatic lapse rate is given by

$$\Gamma_d = \frac{g}{c_p}. \tag{20}$$

The conditions for static stability are thus:

$$\begin{array}{lll} \text{stability :} & \dfrac{\partial\widetilde{\theta}}{\partial z} > 0; \qquad \text{or} & -\dfrac{\partial\widetilde{T}}{\partial z} < \Gamma_d \\ \text{instability :} & \dfrac{\partial\widetilde{\theta}}{\partial z} < 0; \qquad \text{or} & -\dfrac{\partial\widetilde{T}}{\partial z} > \Gamma_d \end{array}, \tag{21a,b}$$

where a tilde indicates that the values are those of the environment. The atmosphere is, in fact, generally stable by this criterion: the observed lapse rate, corresponding to an observed buoyancy frequency of about $10^{-2}\,\mathrm{s}^{-1}$, is often about $7\,\mathrm{K\ km^{-1}}$, whereas a dry adiabatic lapse rate is about $10\,\mathrm{K\ km^{-1}}$. Why the discrepancy? One reason, particularly important in the tropics, is that the atmosphere contains water vapour.

4.2 Saturated lapse rate

The amount of water vapour that can be contained in a given volume is an increasing function of temperature (with the presence or otherwise of dry air in that volume being largely irrelevant). Thus, if a parcel of water vapour is cooled, it will eventually become saturated and water vapour will condense into liquid water. A measure of the amount of water vapour in a unit volume is its partial pressure, and the partial pressure of water vapour at saturation, e_s, is given by the Clausius–Clapeyron equation,

$$\frac{\mathrm{d}e_s}{\mathrm{d}T} = \frac{L_c e_s}{R_v T^2}, \tag{22}$$

where L_c is the latent heat of condensation or vapourization (per unit mass) and R_v is the gas constant for water vapour. If a parcel rises adiabatically it will cool, and at some height (known as the 'lifting condensation level', a function of its initial temperature and humidity only) the parcel will become saturated and any further ascent will cause the water vapour to condense. The ensuing condensational heating causes the temperature and buoyancy of the parcel to increase; the parcel thus rises further, causing more water vapour to condense, and so on, and the consequence of this is that an environmental profile that is stable if the air is dry may be unstable if saturated. Let us now derive an expression for the lapse rate of a saturated parcel that is ascending adiabatically apart from the effects of condensation.

Let w denote the mass of water vapour per unit mass of dry air, the mixing ratio, and let w_s be the saturation mixing ratio. ($w_s = \alpha e_s/(p - e_s) \approx \alpha_w e_s/p$ where $\alpha_w = 0.622$, the ratio of the mass of a water molecule to one of dry air.) The diabatic heating associated with condensation is then given by

$$Q_{cond} = -L_c \frac{\mathrm{D}w_s}{\mathrm{D}t}, \tag{23}$$

so that the thermodynamic equation is

$$c_p \frac{\mathrm{D}\ln\theta}{\mathrm{D}t} = -\frac{L_c}{T}\frac{\mathrm{D}w_s}{\mathrm{D}t}, \tag{24}$$

or, in terms of p and and T

$$c_p \frac{\mathrm{D}\ln T}{\mathrm{D}t} - R\frac{\mathrm{D}\ln P}{\mathrm{D}t} = -\frac{L_c}{T}\frac{\mathrm{D}w_s}{\mathrm{D}t}. \tag{25}$$

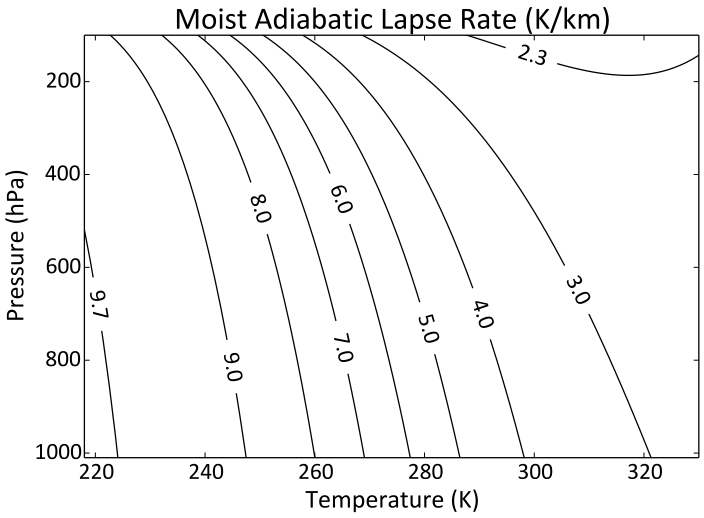


Figure 5: The saturated adiabatic lapse rate as a function of temperature and pressure when water (H_2O) is the condensate.

If these material derivatives are due to the parcel ascent then

$$\frac{\mathrm{d}\ln T}{\mathrm{d}z} - \frac{R}{c_p}\frac{\mathrm{d}\ln p}{\mathrm{d}z} = -\frac{L_c}{Tc_p}\frac{\mathrm{d}w_s}{\mathrm{d}z}, \tag{26}$$

and using the hydrostatic relationship and the fact that w_s is a function of T and p we obtain

$$\frac{\mathrm{d}T}{\mathrm{d}z} + \frac{g}{c_p} = -\frac{L_c}{c_p}\left[\left(\frac{\partial w_s}{\partial T}\right)_p \frac{\mathrm{d}T}{\mathrm{d}z} - \left(\frac{\partial w_s}{\partial p}\right)_T \rho g\right]. \tag{27}$$

Solving for $\mathrm{d}T/\mathrm{d}z$, the lapse rate, Γ_s, of an ascending saturated parcel is given by

$$\Gamma_s = -\frac{\mathrm{d}T}{\mathrm{d}z} = \frac{g}{c_p}\frac{1-\rho L_c(\partial w_s/\partial p)_T}{1+(L_c/c_p)(\partial w_s/\partial T)_p} \approx \frac{g}{c_p}\frac{1+L_c w_s/(R_d T)}{1+L_c^2 w_s/(c_p R_v T^2)}. \tag{28}$$

where the last near equality follows with use of the Clausius–Clapeyron relation. The quantity R_d is the gas constant for dry air and R_v is the gas constant for water vapor, and $R_v = R_d/\alpha_w$. The quantity Γ_s is variously called the *pseudoadiabatic* or *moist adiabatic* or *saturated adiabatic* lapse rate, and it is plotted in Fig. 5.

Because g/c_p is the dry adiabatic lapse rate Γ_d, $\Gamma_s < \Gamma_d$, and values of Γ_s are typically around $6\,\mathrm{K\ km^{-1}}$ in the lower atmosphere; however, $\mathrm{d}w_s/\mathrm{d}T$ is an increasing function of T so that Γ_s decreases with increasing temperature and can be as low as $3.5\,\mathrm{K\ km^{-1}}$. For a saturated parcel, the stability conditions analogous to (21) are

$$\text{stability :} \qquad -\frac{\partial \widetilde{T}}{\partial z} < \Gamma_s, \tag{29a}$$

$$\text{instability :} \qquad -\frac{\partial \widetilde{T}}{\partial z} > \Gamma_s. \tag{29b}$$

where $\widetilde{T}$ is the environmental temperature. The observed environmental profile in convecting situations is often a combination of the dry adiabatic and moist adiabatic profiles: an unsaturated parcel that is is unstable by the dry criterion will rise and cool following a dry adiabat, Γ_d, until it becomes saturated at the lifting condensation level, above which it will rise following a saturation adiabat, Γ_s. Such convection will proceed until the atmospheric column is stable and, especially in low latitudes, the lapse rate of the atmosphere is largely determined by such convective processes.

References

Vallis, G. K., Zurita-Gotor, P., Cairns, C. & Kidston, J., 2014. Response of the large-scale structure of the atmosphere to global warming. *Quart. J. Roy. Meteor. Soc.*. 10.1002/qj.2456.

Lecture 5: Convective Heat Transfer II

Kerry Emanuel; notes by Geoff J. Stanley and Jim Thomas

June 20

1 Introduction

In this lecture we will study classical Köhler theory and the effects of aerosols and their size distribution to explain raindrop formation. We will then study the stability of a moist atmosphere (building on the stability of a dry atmosphere in Lecture 3). We then conclude with a discussion of radiative moist convective equilibrium.

2 Raindrop Formation

How does rain actually form in Earth's atmosphere? The answer is quite complex, and there is a long history of scientists who have attempted to explain the phenomenon. In this lecture we will examine the process and along the way note its historical development.

We must begin with a phase diagram of water, which has been obtained through both theory and laboratory experiment. Figure 1 shows the phase of water as a function of pressure p and temperature T. When pressure and/or temperature change so as to cross a line on the diagram, water changes from one phase to another, as the theory goes. However, phase transitions in Earth's atmosphere, being catalysed by a distribution of aerosols, are more complex than in the laboratory.

First let us recall our definitions. Let ρ be the density of air and ρ_v the density of water vapour in the air. Let T be the air temperature. The gas constant for water vapor is $R_v = R^*/m_v = 462$ J kg^{-1} K^{-1}, with $R^* = 8.3144621$ J mol^{-1} K^{-1} the universal gas constant and $m_v = 18$ g mol^{-1} the molar mass of water. The latent heat of condensation of water is $L_v = 2.5 \times 10^6$ J / kg. The specific humidity q, which is the actual mass concentration of water vapor in the air, the vapor pressure e, and the saturation vapor pressure e^* are given by

$$q = \frac{\rho_v}{\rho},$$

$$e = \frac{\rho_v R^* T}{m_v},$$

$$e^* = e_0 \exp\left(\frac{L_v}{R_v}({T_0}^{-1} - T^{-1})\right). \tag{1}$$

For temperatures well below the critical temperature, the Clausius-Clapeyron relation takes the above exponential form and determines e^* as a function of T, with e_0 and T_0 empirical

constants. This is the curve between the Triple Point and the Critical Point in the phase diagram (Figure 1). The above formulas treat L_v as constant, but it is, in fact, a weak function of temperature; correspondingly, there are more accurate semi-empirical formulas for e^*, but the above will suffice here. An air parcel holding water vapor becomes saturated when $e > e^*$, i.e. when it crosses this curve. This can occur by increasing its vapor pressure (by increasing q such as through sea spray, or by increasing the total pressure) or by decreasing T so as to decrease e^*. From the ideal gas law $p = \rho R^* T/\bar{m}$, with $\bar{m}$ the mean molecular weight of air (which changes very little between dry and saturated air, owing to the overwhelming concentration of N_2 and O_2) we may write

$$e = \rho_v \frac{R^* T}{m_v} = qp\frac{\bar{m}}{m_v}.$$

A subsaturated air parcel that ascends adiabatically (no heat content change) and reversibly (no change in q) will experience a lower p as well as T, such that both e and e^* decrease. While e decreases nearly perfectly exponentially with height (through $p \approx p_0 \exp(-z/H)$ with H the scale height and p_0 the surface pressure), e^* decreases as the exponential of a term $(-T^{-1})$ that itself decreases with height: $\partial_z \left(-T^{-1}\right) = \Gamma_D T^{-2}$ with Γ_d the dry lapse rate. Thus e^* decreases with height faster than e, so condensation routinely occurs by lifting a moist air parcel.

Cloud physicists realized long ago the difficulty of rain droplet formation. Forming a water droplet without a condensation nucleus, purely by condensation (homogeneous nucleation) is possible but requires supersaturation of the air by 10% or more, or equivalently (by Clausius-Clapeyron) supercooling of the air down to about $-40°$C. This supersaturation is required to overcome the tendency of small droplets to evaporate: water molecules on the edge of a small droplet, which has a small radius of curvature, are more exposed and separated from their neighbouring molecules; hence they experience weaker cohesive forces (surface tension), and will evaporate more easily. This is the Kelvin effect: the amount of supersaturation required for a droplet to be in condensation-evaporation equilibrium is larger for smaller droplets.

However, observations rarely show air in Earth's atmosphere to be supersaturated by more than about 1%. There had to be another process at work. Assistance comes from condensation nuclei (aerosols) which, by decreasing the number of liquid H_2O molecules at the interface (the edge of a droplet, in this case), reduce the tendency for evaporation. For a flat interface, this is known as the Raoult effect. In Earth's atmosphere, there are plenty of aerosols that catalyse this phase transition very well, so that it truly is a good approximation to assume that water vapor condenses when crossing this curve. However, simply condensing water vapor is not sufficient to form rain; droplets must grow in size until their terminal speed exceeds the updraft speed.

In a seminal paper, Köhler [1] transformed Raoult's law for a spherical interface (imitating a water droplet) and simultaneously added the Kelvin effect. Very small droplets—those for which the Kelvin effect inhibits condensational growth—that form on a condensation nuclei have a high particulate to water ratio. So by Raoult's effect the environment can be *sub*saturated and yet the droplet can maintain condensation-evaporation equilibrium. However, condensation will not grow such a droplet: if it grows the required vapor pressure for equilibrium is raised (Raoult's effect) and so, without adjusting the ambient vapor pres-

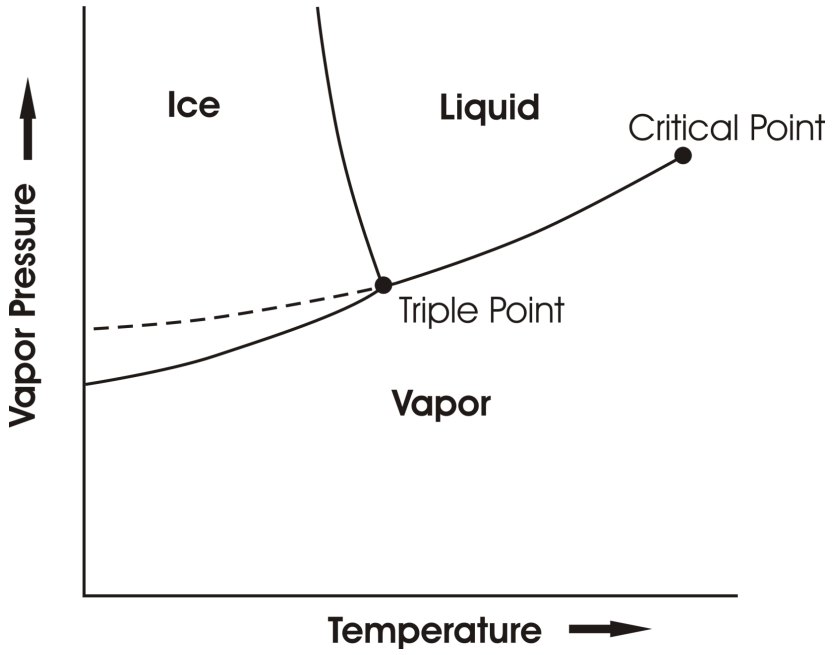


Figure 1: The phases of water dependent on temperature and pressure. At the triple point ($T = 0°$C and $p = 6.112$ hPa) all three phases can coexist in equilibrium. Beyond the critical point ($T > 647$ K and $p > 218$ atm, so not relevant for Earth's atmosphere), the liquid and gaseous phases become indistinguishable. Atmospheric aerosols that catalyse the condensation of water vapor to liquid droplets are plentiful in the atmosphere, so transitioning from vapor to liquid is just a matter of lowering the temperature or raising the vapor pressure to cross the phase transition line. However, these aerosols are not as well suited for catalysing deposition (vapor to ice). The dashed line indicates that in Earth's atmosphere, significant supersaturation or supercooling is required to form ice crystals from vapor.

sure, the droplet will evaporate back to its original size: a stable equilibrium (corresponding to positive slopes in Figure 2). When the ambient vapor pressure is above a critical supersaturation level, as Köhler's theory goes, Kelvin's effect begins to dominate: the edge of a growing water droplet becomes flatter and the vapor pressure required for equilibrium falls. If the ambient vapor pressure remains constant, an even larger vapor pressure differential drives more condensation onto the droplet: an unstable equilibrium (corresponding to negative slopes in Figure 2). The critical supersaturation, predicted by Köhler's theory for a given aerosol number density, is well below 1%; concerning observations, this theory of droplet growth was plausible.

Once the critical supersaturation has been passed and a water droplet has formed, one might expect that, since the saturation vapor pressure falls as the water droplet grows (from the Kelvin effect), that it will grow and grow without end—or until so large that it falls as rain. However, in the atmosphere CCN [cloud condenstation nuclei] are generally abundant, and are all competing for available water vapor, so none grow particularly large. Pure condensational growth does not produce water droplets large enough to fall as rain,

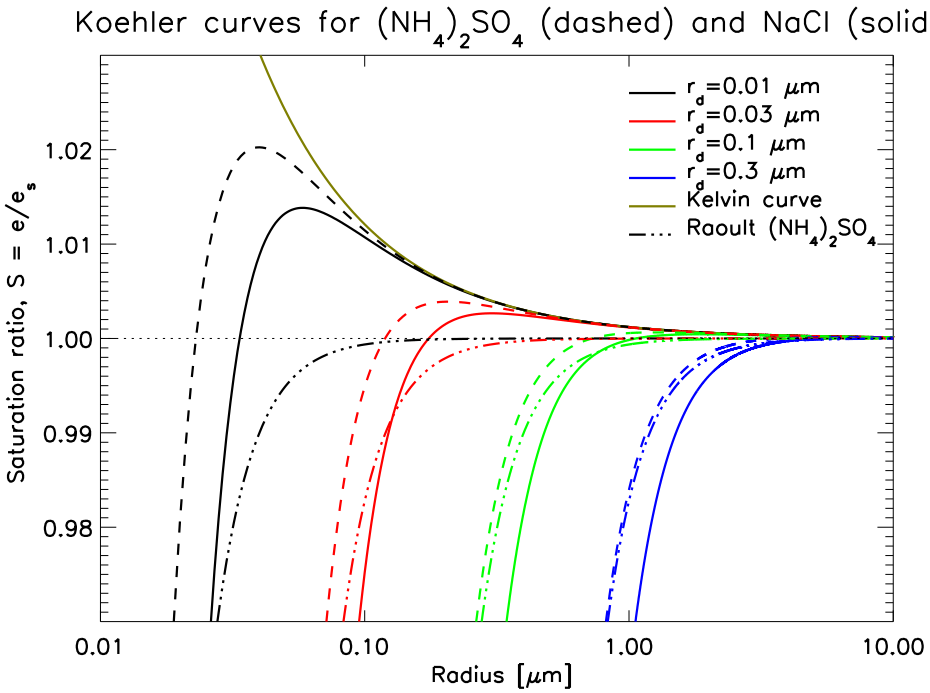


Figure 2: vapor pressure required for a droplets of different radii to be in condensation-evaporation equilibrium, as described by Köhler theory. The Kelvin effect (solid beige curve) increases the required vapor pressure with decreasing droplet size, due to an increasing radius of curvature. The Raoult effect (dash-dot curves) reduces the required vapor pressure with decreasing droplet size, due to higher aerosol concentration. Köhler theory (solid curves) combines these two effects, and accounts for both the type of aerosol (dashed vs. solid) and the size of the aerosol (colours). Figure from Ulrike Lohmann's lectures (IACETH); available at `http://www.iac.ethz.ch/edu/courses/bachelor/vertiefung/atmospheric_physics/Slides_2012/koehler.pdf`

and yet another process is required for rain formation.

Surely the collection of smaller droplets by larger droplets through collisions must be of some importance. However, not only does condensational growth produce droplets too small to rain out, those droplets are also roughly all the same size. That is, the distribution of droplet sizes narrows with time. Equal size droplets have the same buoyancy and drag forces and so will accelerate at the same rate (recall Stokes' flow past a sphere). All moving at similar velocities, then, they will not have much opportunity for collision. Turbulent air flows can force collisions of like-sized droplets; there is a significant body of literature on this topic but will not be pursued here. For the more laminar case, the efficiency of the collisional process for raindrop formation is a very strong function of not just the number density of droplets, but also of the droplet size distribution. A narrow distribution results in rare collisions. A variety of condensation nuclei with different sizes could initiate droplets with a distribution of sizes, but aerosols in continental regions are predominantly dust without much size variance. Again, another mechanism was needed to explain rain formation, and

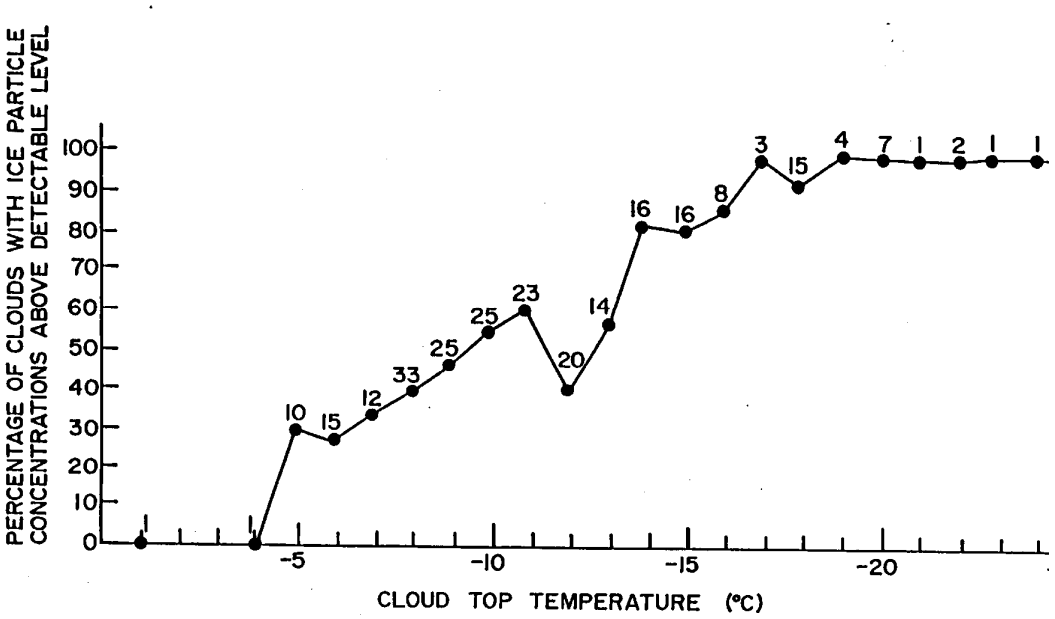


Figure 3: An observational study showing the fraction of clouds with ice in them as a function of cloud top temperature, being generally the coldest temperature in the cloud. The numbers indicate the sample number of clouds for each temperature. Above 0°C no ice exists, and below (for this study) −19°C all clouds contain ice.

it came from Bergeron in 1933.

There are two points to discuss before describing Bergeron's mechanism. First, we've said that atmospheric aerosols are well suited to catalyse water droplet formation; but those same aerosols are not so efficient for ice formation. Homogeneous nucleation of ice crystals from liquid water (i.e. in a perfectly clean atmosphere) also requires very low temperatures, around −40°C. Thus for ice formation, it is safe to assume water crystallizes below −40°C, and that none crystallizes above 0°C. But there is some complicated function (of temperature, and the aerosol population) that determines the fraction that crystallizes between these temperatures. Figure 3 shows this function from one study; the function will vary from study to study and is not meant to be taken as universal. The non-monotonicity found in this particular study indicates the potential difficulty in determining a more universal function. The relevant point for this discussion, however, is that this preference for liquid droplet formation over ice crystal formation leads to a large population of droplets but a small population of ice crystals in mixed phase clouds.

Second, deriving the Clausius-Clapeyron relation involved a thermodynamic study of the phase transition to water vapor from liquid water *or* from ice. Now, the latent heat of fusion for water is about one eighth of its latent heat of condensation[1], and so the latent heat

[1]The latent heat of fusion is often neglected in climate models, but is certainly non-negligible in detailed

of deposition (ice formation), being the sum of the latent heat of fusion and condensation, is larger than the latent heat for condensation alone. These different values for L go into (1), and this shows that the saturation vapor pressure over ice is always lower than that over liquid water. Thus for a given ambient vapor pressure between the two saturation vapor pressures, liquid water droplets evaporate and ice crystals grow by deposition.

Eleven years before presenting his theory Bergeron noted that a stratus cloud deck could extend down only to the canopy, not to the ground, when temperatures were below freezing. He proposed that a small ice crystal population could rapidly form large crystals—large enough to fall, and possibly melt back to rain—by scavenging water from a large population of water droplets. Given that temperatures below 0°C exist even in the tropics at high altitudes, this process had the potential to explain both mid-latitude and tropical rainfall.

Indeed, the Bergeron process does explain much rain formation, but not all. In the 1950s cloud physicists convened for a conference in Puerto Rico, where they observed intense rainfall from a cloud whose top was certainly warmer than 0°C. Indeed, in this case of a maritime climate, salt particles from the sea are plentiful in the atmosphere, giving a wide distribution of condensation nuclei sizes, and rain formation can occur via condensational growth and collision collection alone without any cold cloud microphysics.

We have seen how aerosols, and in particular their size distribution, play a critical role in rain drop formation, and thus in determining whether it will rain or not. We have also seen how cold cloud microphysics involving ice crystals can produce rain when the aerosol distribution does not particularly support rain drop formation. We have considered the condensational and collisional processes of rain droplet growth. It is important that we understand these fundamental processes of rain formation if we are to accurately predict the weather, as well as the climate.

3 Stability

In Lecture 3 we derived the entropy of a dry air parcel and considered adiabatic vertical displacements of this air parcel. Such a process is reversible so its entropy is conserved, and we found that the environment was stable (unstable) if the entropy increased (decreased) with height. Determining a stability criterion for moist air, however, turns out to be much more difficult. The difficulty arises because the density of moist air depends not just on its entropy and pressure, but also on the water content and the fraction of that water content that is vapor versus liquid.

We can define the entropy of moist air as a function of temperature, pressure, and specific humidity. This is well defined because, as discussed in the previous section, water vapor supersaturates very little. The moist air entropy s is that for dry air plus two additional terms (the last two, below):

$$s = c_p \ln \frac{T}{T_1} - R_d \ln \frac{p}{p_1} + \frac{L_v q}{T} - q R_v \ln \mathcal{H}. \tag{2}$$

Here $\mathcal{H} \equiv e/e^*$ is the relative humidity, c_p is the specific heat capacity for air at constant pressure, R_d (R_v) is the specific gas constant for dry air (water vapor), L_v is the latent heat

studies of cloud physics.

of condensation, and T_1 and p_1 are arbitrary constants (note that entropy is only relevant up to an additive constant). Let us also define the saturation entropy, which has $\mathcal{H} = 1$ so $\ln \mathcal{H} = 0$, giving

$$s^* = c_p \ln \frac{T}{T_1} - R_d \ln \frac{p}{p_1} + \frac{L_v q^*}{T}. \tag{3}$$

In addition to adding latter two terms, water content changes s by modifying the molar mass of the air mixture: hence the specific gas constant used should technically be a function of q. Similarly, the heat capacity of air is modified by water vapor, and even more strongly by condensed water: c_p should be a function of q. These modifications should be accounted for in numerical weather prediction models, but since these changes are more of a numerical nature than conceptual, we shall treat R_d and c_p as constants here. We have also chosen to neglect any ice processes, because of the difficulty in knowing at what temperature ice formation occurs (recall Figure 3). These, however, are the more minor difficulties.

The major difficulty in defining the stability of moist air comes from the additional dependence of density on the water content. The specific volume α is now a function of three variables: entropy and pressure, as well as a modification due to water. This makes it impossible to construct a 2D thermodynamic diagram (with axes of pressure and entropy) that classifies the stability of an atmospheric column. That is, there is no simple stability criterion based on $\partial s/\partial p$, as there was for dry air. An extra dimension, the water content, would be needed.

However, it is possible to define a modified entropy, s', that, to good approximation, solves this problem. Let us define $q_t = (\rho_v + \rho_l)/\rho$ to be the total mass fraction of water content in air, whether from water vapor (ρ_v) or liquid droplets (ρ_l). This q_t is conserved in a reversible vertical displacement, even under condensation or evaporation. Thus an arbitrary function of q_t is also conserved in a reversible process. It is possible to choose this function carefully such that, when added to s to define s', we get the specific volume as a function of just two variables: $\alpha \approx \alpha(p, s^{*\prime})$.

With this caveat, we may now consider the 2D thermodynamic diagram shown in Figure 4. Consider an adiabatic vertical displacement of an air parcel originating from the ground with (modified) entropy s'_0. If this is reversible, then s and s' are conserved. We need to know about the specific volume α of our parcel, and of the environment, to determine stability, and by the argument above we can get this from p and s'. Suppose that we know, from other thermodynamic information not shown in the figure, where the parcel will condense. This is the Lifting Condensation Level (LCL), above which the air parcel is saturated. Comparing its (modified) entropy (the rightmost vertical line) to the (modified) saturation entropy of the environment (dashed line) will tell us, through $\alpha(p, s^{*\prime})$, its buoyancy. Where $s'_0 > s^{*\prime}$ the air parcel will be positively buoyant: an unstable situation. Where $s'_0 < s^{*\prime}$ it is negatively buoyant and stable.

However, the moist atmosphere can exhibit meta-stability: it can be stable to small vertical displacements but unstable to large vertical displacements. An air parcel lifted from the ground to anywhere below the bottom intersection of the vertical s'_0 line and the $s^{*\prime}$ curve, called the Level of Free Convection (LFC), will be negatively buoyant and hence stable. But lift that same air parcel even slightly above the LFC and it will be positively buoyant, hence unstable. It will rise until it reaches the top intersection of the vertical s'_0 line and the $s^{*\prime}$ curve, called the Level of Neutral Buoyancy (LNB). There is no analogue

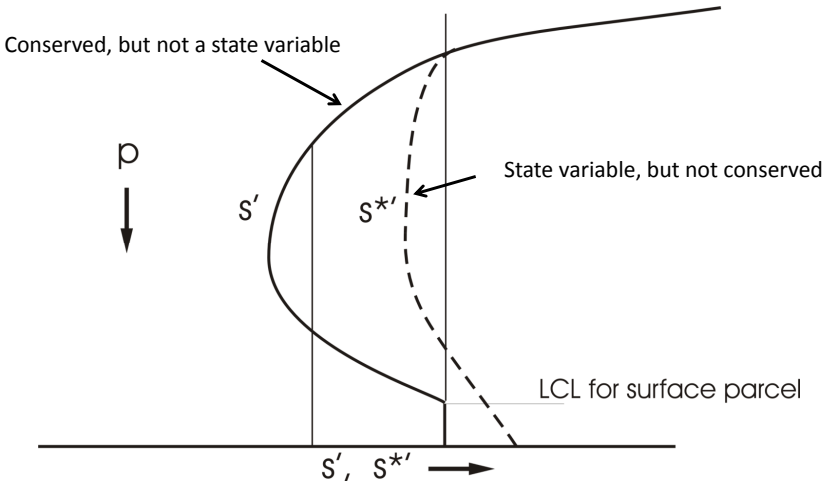


Figure 4: A schematic illustration of moist meta-stability. A typical environmental profile for s' is shown (thick solid line). In the lower and subsaturated air, turbulent mixing or convection has brought the environment to the dry adiabatic lapse rate, equivalent to a constant s'. Above the Lifting Condensation Level (LCL) water vapor condenses, q decreases in (2) and s' decreases. Higher up the decreasing pressure begins to dominate and s' begins to increase with decreasing p. Near the top of the diagram s' increases even more rapidly with decreasing p due to increasing T in the stratosphere. For full details, see the text.

of this in a homogeneous (e.g. dry) atmosphere.

Basic thermodynamics tells us that the heat added during a cyclic process is $\oint \mathrm{d}Q = \oint T\ \mathrm{d}s$. Thus if we transformed Figure 4 to have T as its vertical coordinate, the area between s'_0 and $s^{*\prime}$ curves is a measure of energy. Specifically, the area where $s'_0 > s^{*\prime}$ (above the LFC and below the LNB) is the the amount of potential energy available for convective release, termed the Convective Available Potential Energy (CAPE). Similarly, the area where $s'_0 < s^{*\prime}$ (below the LFC) is the amount of energy required for the air parcel to overcome the stable stratification and rise above the LFC. This is termed the Convective Inhibition (CIN).

Large-scale radiative cooling above the LFC lowers the entropy, thereby building up a reservoir of CAPE. Meanwhile CIN can be built up by large-scale subsidence and/or surface cooling (e.g. at night). CIN can be thought of as a potential barrier, and it is crucial to the meta-stability exhibited by moist air: without it, convection would immediately consume any CAPE, as it does in a dry atmosphere. During the day, surface heating raises the surface entropy which can reduce the CIN, or localized processes can add enough energy to overcome the CIN for a particular air parcel. When this occurs, air rises above the LFC and begins to tap the (potentially large) reservoir of CAPE. This can create extremely powerful thunderstorms, through a somewhat rare phenomenon called supercell convection, as well as tornadoes and hailstorms. The meteorological effects of CAPE, however, are beyond the scope of this lecture.

In the meteorological community there are two schools of thought. The first holds that CAPE is built up and released repeatedly during large events. The second holds that moist

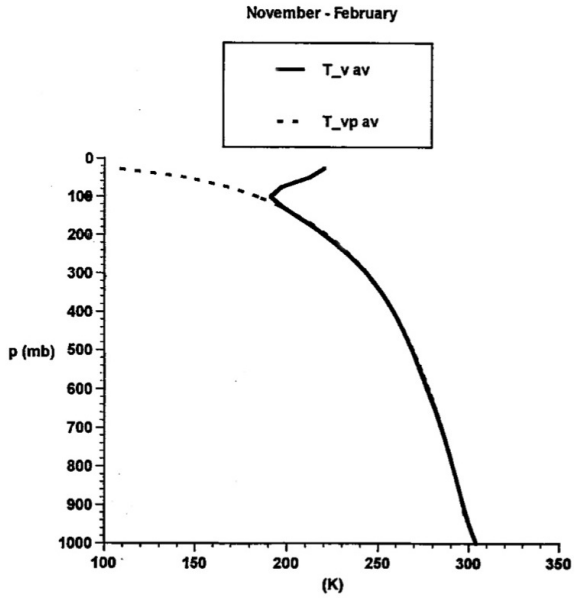


Figure 5: A typical tropical sounding. The solid line is a result of averaged measurements and the dashed line is a moist adiabat. The region where the slope changes sign is the tropopause.

convection is a process in quasi-statistical equilibrium (like dry convection), except in a few cases over continents.

Note that if we lift a different air parcel (i.e. from elsewhere in the column than the ground), then its LCL will be at a different pressure level, and its stability properties will also differ. Thus to fully characterize the stability of a moist water column, one must take air parcels from all pressure levels and subject them to all possible vertical displacements, large or small, which we shall explore in the next section.

4 Radiative moist convective equilibrium

The moist adiabatic lapse rate (s^*) is a function of pressure and temperature, though in the stratosphere where water vapor is scarce it is quite near to the (constant) dry adiabatic lapse rate. This is evident from a typical sounding data shown in Figure 5.

Figure 6 depicts the pressure variation experienced by fluid parcels that are raised from a range of starting pressure values. As seen in the figure, a parcel lifted from 900 hPa never becomes positively buoyant, because it has much less moisture to start with. For a parcel starting around 940 hPa it is striking that neutral buoyancy appears throughout the whole troposphere.

The process of forming a cumulus cloud is sketched in Figure 7. Air begins to rise and starts to form a cumulus cloud. Due to appreciable coalescence time, the precipitation formation occurs slowly as the air mass rises. A part of the precipitation falls outside the envelope during the mature stage. Re-evaporation can happen in this time, leading to

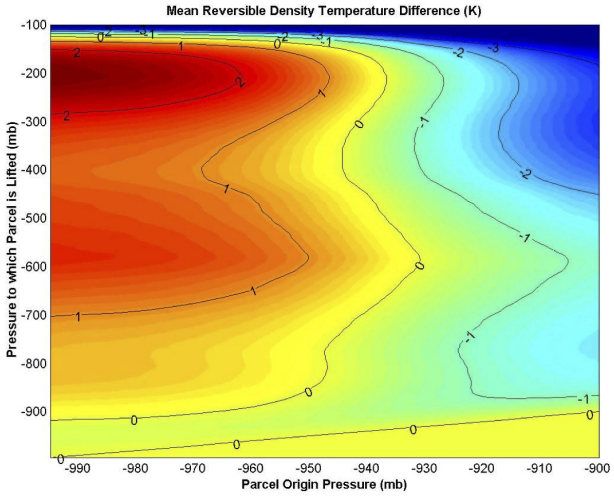


Figure 6: A plot showing the stability of air parcels being elevated. Red color indicates positive buoyancy and blue indicates negative. Note that a sample starting around 900 mb is positively buoyant, neutrally buoyant if it starts around 940 mb and negatively buoyant around 900 mb

chilling of surrounding air. In the final stage, a density current appears in the atmosphere, leading to 'raining' of low entropy air which in the process blocks the rise of high entropy air from the lower atmosphere. This leads to the destruction of the plume that formed, its life cycle being of the order of an hour.

Consider now a moist RCE situation where homogeneous conditions are set in the horizontal plane, i.e. constant ocean surface temperature, constant solar radiation along the horizontal and so on. On letting moist convection occur in such a system, interesting things happen. A cartoon representation of this is depicted in Figure 8, where a cross section is indicated. In spite of a spectrum of clouds appearing in the real world, only two are indicated here for simplicity. Air ascends through the deep clouds, precipitating in the process as rain. The air falling after reaching the top of the cloud is depleted of condensed water and is under saturated. The descent is therefore dry and hence the air is clear, as is seen in the figure between the clouds. The asymmetry in ascend and descending parts is interesting. There are strong updrafts in localized regions, which are widely spaced. Larger areas of clear air appear surrounding them, where the air descends at speeds roughly two orders of magnitude less than the ascending speeds. While the ascend can be on the order of a few m/s while the descends are usually of the order of a few cm/s.

Temperature variations with altitude are indicated in Figure 9 which were obtained from different models, holding relative humidity fixed. The calculations were carried out for pure radiative, dry radiative-convective and moist radiative-convective cases. Note that the height of the tropopause increases with these models. The RCE model fits well as a zeroth order approximation for a real data plotted at a typical station.

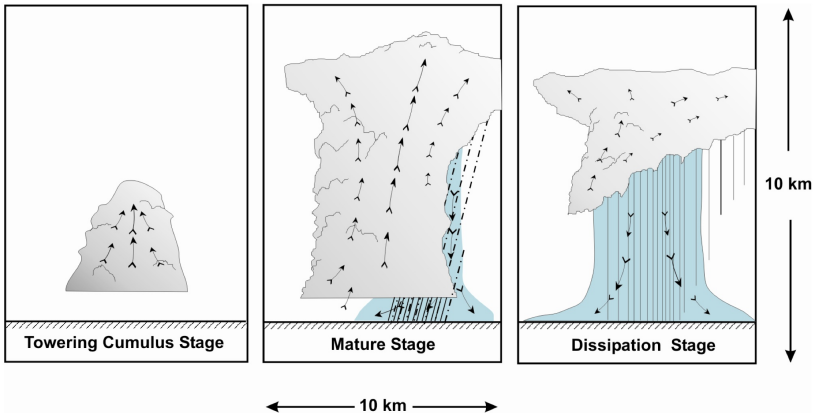


Figure 7: The formation of a cumulus cloud and its dissipation

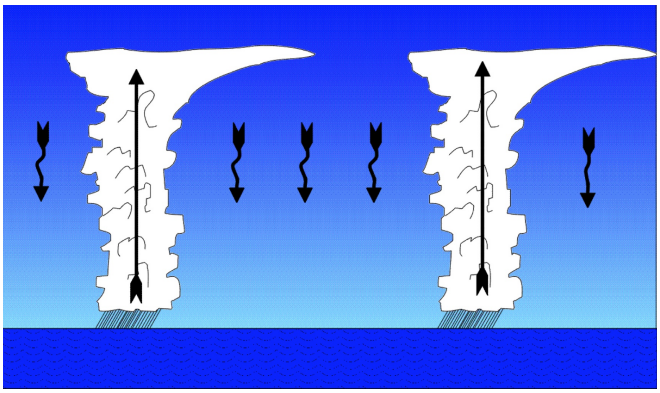


Figure 8: Precipitating widely spaced clouds

References

[1] Köhler, H., 1936. *The nucleus in and the growth of hygroscopic droplets.* Trans. Faraday Soc., 32, 11521161.

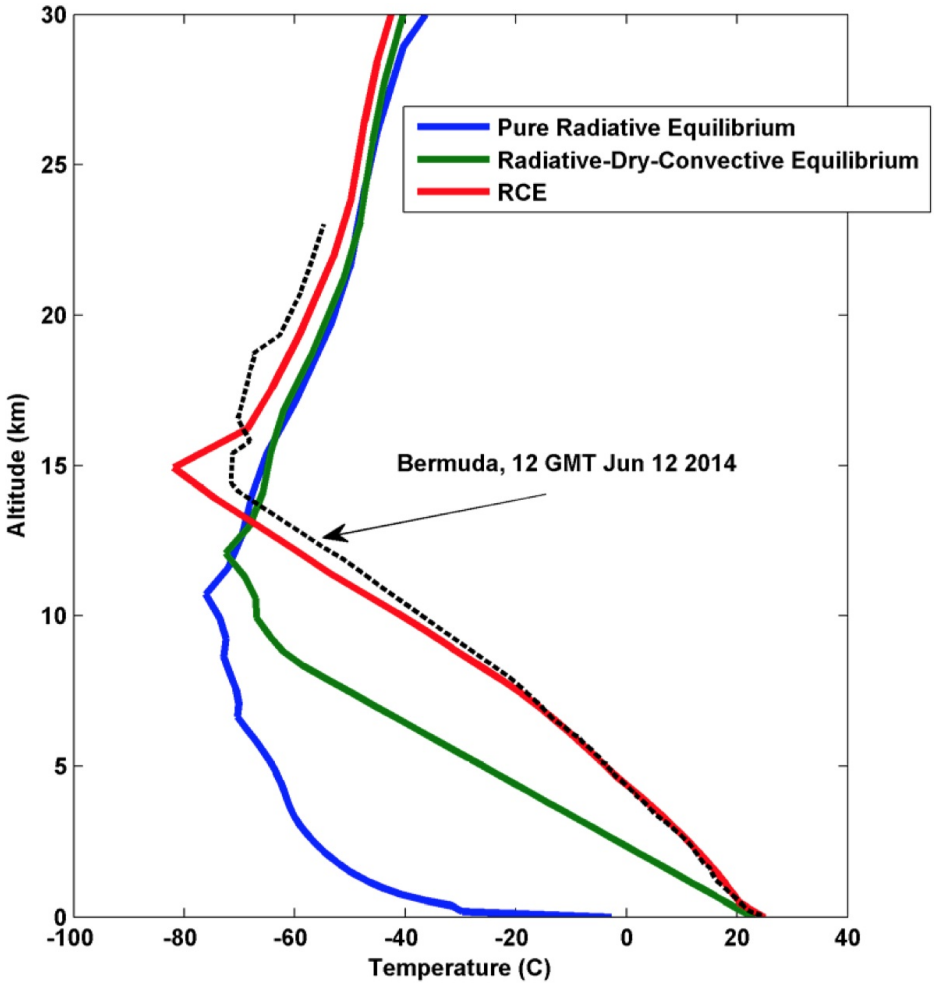


Figure 9: A comparison between different models predicting temperature as a function of altitude.

Lecture 6: Water and Radiation in the Climate System

Kerry Emanuel; notes by Joe Fitzgerald and Jörn Callies

June 23

1 Introduction

Over the next two lectures we will progress toward problems involving the interaction of water substances and radiation that lie at the frontier of climate research. At the end of the previous lecture, we left off with a discussion of moist radiative-convective equilibrium (RCE). Using a single column model with many levels in the vertical, we solved the band-integrated radiative transfer equations coupled to a representation of moist convection and obtained a temperature profile $T(z)$ that was an essentially realistic representation of the temperature profile of the tropical atmosphere. We note, however, that obtaining a realistic temperature profile is not a stringent test of our model's moist convection scheme: the observed $T(z)$ profile is relatively easily reproduced, because it follows the moist adiabat in the bulk of the troposphere.

Moist RCE is a good starting point for our understanding of the tropical atmosphere's leading-order structure, and examining our RCE solution can lead us to several physically important observations. By comparing the temperature profile resulting from RCE to that which results from radiative equilibrium, we see that the troposphere is cooling radiatively. Additionally, the RCE temperature profile indicates that the temperature at the tropical tropopause is approximately $-80°$C, a value which may remain fixed even as the climate changes. However, our RCE model also failed to reproduce some features of the real atmosphere, especially those due to the large-scale circulation, which was only crudely represented by an increased albedo. In the tropical upper troposphere/lower stratosphere, the observed temperature profile is shifted somewhat upward compared with our RCE calculation. This is due to the Brewer-Dobson circulation, a meridional overturning circulation in the tropical stratosphere which features upwelling through the tropical tropopause. This circulation can produce temperature deviations near the tropopause relative to our RCE solution because the timescale of radiation is not fast compared with that of the Brewer-Dobson upwelling.

2 The Earth's Energy Balance from Observations

We now move on to a reckoning of the Earth's energy balance based on observations. To build intuition, we consider what happens to solar energy as it arrives at Earth. Although the real atmosphere is continuous in the vertical, for simplicity we consider the energy balance by separating the atmosphere's vertical structure into three boxes: the troposphere,

the stratosphere, and the surface (which we take to include land as well as ocean). As a normalization, we consider the passage of 100 units of solar radiation (which could be thought of as individual photons) through the atmosphere. Figure 1 shows a schematic of how this solar radiation is passed between our three boxes.

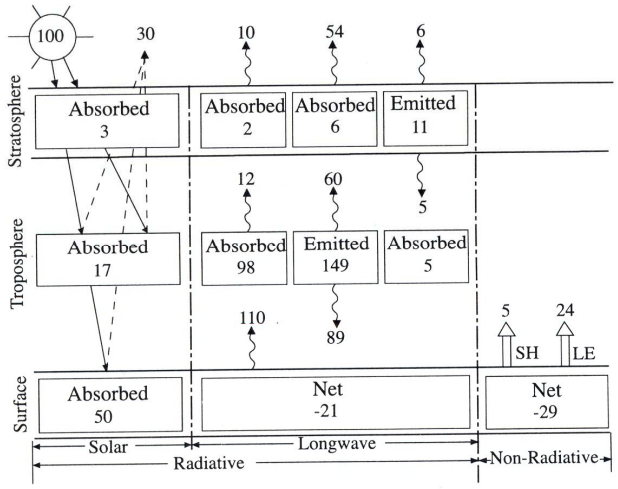


Figure 1: A schematic of the Earth's energy balance. The atmosphere is separated into surface, troposphere, and stratosphere boxes, and we consider the passage of 100 units of solar radiation through the system. The values in each box indicate how much energy is passed from box to box by various transfer processes.

In this approximate picture of the energy balance, 30 units of shortwave (SW) radiation are reflected back to space. The largest fraction of this reflection is due to clouds, and the next largest fraction is due to ice on the surface. After reflection, 70 units remain to be absorbed by the atmosphere and surface. Three of these units are absorbed in the stratosphere. This absorption filters harmful incoming UV photons out of the incoming SW radiation field. The troposphere absorbs 17 of the remaining units, with water vapor and clouds being the dominant absorbers. The 50 remaining units of SW are absorbed by the surface. These SW contributions to the energy balance are indicated on the left-hand side of Fig. 1.

The longwave (LW) contributions to the energy balance are shown in the centre portion of Fig. 1. The surface emits 110 units of LW radiation toward space. Eighty-nine of these units are absorbed and re-emitted back toward the surface by the troposphere. Note that this downwelling LW flux is greater than the 50 units the surface receives directly from the sun. The troposphere also emits 60 units upward toward space. The asymmetry between the 89 units of downward emission and 60 units of upward emission results from our bulk averaging over the troposphere and from the variations of temperature with height.

As discussed in previous lectures, the lower atmosphere is in thermodynamic disequilibrium with the surface. This thermal discontinuity drives a turbulent flux of enthalpy from the surface to the atmosphere with both sensible and latent heat components. These non-radiative surface fluxes are indicated on the right-hand side of Fig. 1.

The net contributions to the energy budget from solar SW radiation, LW radiation, and

the non-radiative surface fluxes are shown at the bottom of Fig. 1. Note that these values indicate that the primary cooling mechanism in the tropics is not radiation: the net effect of outgoing LW radiation in the tropics is small when the 89 units of longwave returning to the surface from the atmosphere are taken into account. The main balance in the tropics is instead between incoming solar radiation and evaporation. This implies that the tropical surface is much cooler than it would otherwise be if water could not evaporate, assuming that water vapor was still present in the atmosphere.

3 Contributions of Various Absorbers to the Temperature Profile

We now consider the effects of various greenhouse gases (GHGs) on the tropical temperature profile. As a simple conceptual method to assess the effect of each gas, we re-run the RCE column model (which includes variable clouds interacting with radiation) with GHGs removed one at a time from the column and compare the resulting temperature profile with our original RCE solution. The result of this procedure is shown in Fig. 2.

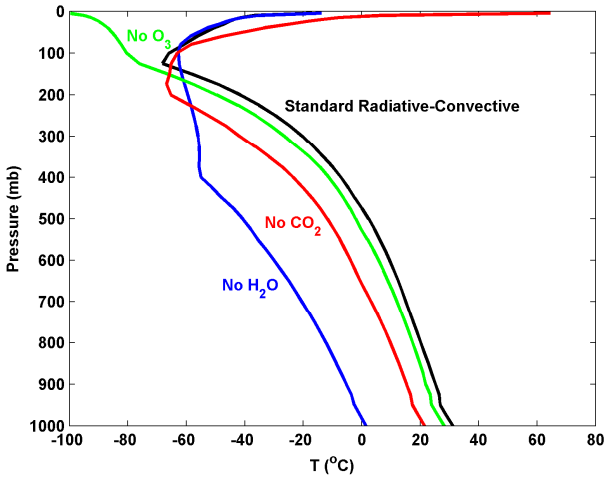


Figure 2: Temperature profiles $T(z)$ in RCE computed using a single-column model. Greenhouse gases are removed one at a time from the model and the equilibrium profile is computed to demonstate the approximate effect that each gas has on $T(z)$.

In the absence of CO_2 (red curve), the tropospheric temperature profile remains on a moist adiabat but is approximately 10°C colder than the original RCE profile. The stratospheric temperature, on the other hand, increases substantially. This stratospheric warming occurs because O_3 continues to absorb UV radiation but CO_2 is no longer available to emit infrared (IR) radiation toward space. The stratospheric thus warms to radiate additional energy and retain energy balance.

Removal of O_3 from the column model (green curve) also reduces the tropospheric temperature slightly, as O_3 has a weak greenhouse effect in the troposphere. In the stratosphere, removing O_3 results in a significant cooling as absorption of incoming UV no longer occurs.

Removal of water vapor from the column model (blue curve) results in a much colder troposphere alongside significant changes in the vertical structure of $T(z)$ in the upper troposphere. Some of the change in vertical structure may be artificial, as water vapor interacts chemically with other GHGs and we do not include these effects in our model.

The stratosphere is compressed into the uppermost portion of Fig. 2 due to the use of pressure coordinates. To show the effects of various gases on the stratospheric temperature more clearly, we plot the RCE profiles again in Fig. 3 using log-pressure coordinates. Comparison of the profiles computed in the absence of O_3 (green curve) and CO_2 (red curve) with the standard RCE profile (black curve) demonstrates that O_3 and CO_2 have strong and opposite-sign effects on stratospheric temperatures. As the first order radiative balance of the stratosphere is between UV absorption by ozone and IR emission by CO_2, removal of ozone cools the stratosphere while removal of CO_2 warms the stratosphere. The RCE calculation without O_3 also indicates that UV absorption in the stratosphere is not required to produce a tropopause: a transition in the shape of the temperature profile from moist adiabatic to a stable profile in radiative balance still occurs near the same level as in the standard RCE solution, although this transition is less sharp when O_3 is absent.

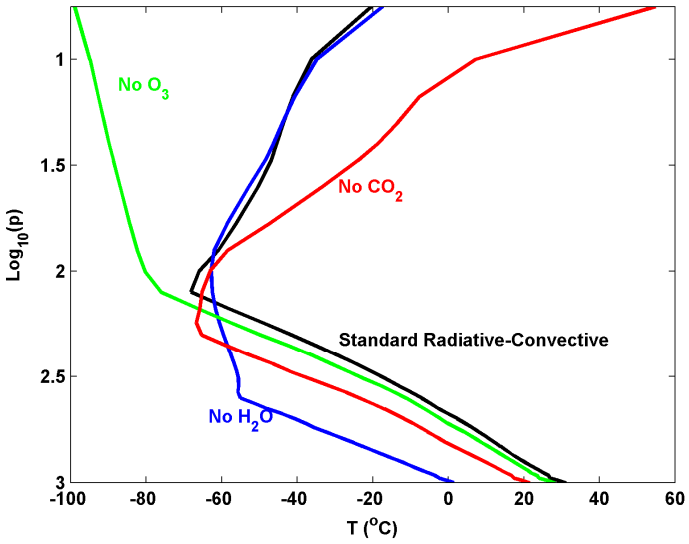


Figure 3: As in Fig. 2, but with log-pressure on the vertical axis to show the stratosphere more clearly.

4 Timescale of Approach to Radiative-Convective Equilibrium

4.1 Observations from a Single-Column Model

We now turn our attention to a basic physical issue: If $T(z)$ is perturbed away from its RCE solution, on what timescale does $T(z)$ return to the RCE state? Figure 4 shows the first ten days of the approach to equilibrium in our RCE column model. The time-

dependent calculations assume a fixed relative humidity profile, taken from the original RCE solution. The initial condition is an isothermal atmosphere, indicated by the Day 0 curve in Fig. 4. After ten days, $T(z)$ is still changing significantly with time. Further time evolution is shown in Fig. 5, which shows $T(z)$ in 15-day increments beginning from the state of the atmosphere at Day 10. After 25 days the tropospheric temperature has essentially returned to its RCE state, but the stratosphere takes much longer to return to equilibrium. Temperature variations in the stratosphere do not become small until nearly Day 150. Note that this long stratospheric return timescale is comparable to the change of the seasons, which implies that there is a lag in the stratospheric response to the seasonal cycle. As we will discuss later in this lecture, the coupling between the atmosphere and the surface (taken to be an ocean surface in these calculations; on land one would have to specify the water availability) is critical to these results. We also note that, in the stratosphere, $T(z)$ features small-scale vertical structures that persist for a long time during the return to equilibrium. These are likely due to small-scale structure in the specified relative humidity profile and the long timescale associated with water vapor in the stratosphere.

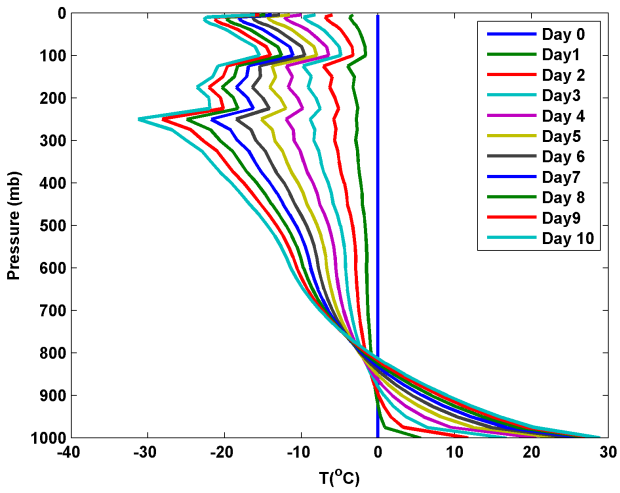


Figure 4: Daily snapshots of the temperature profile from a single-column model during evolution toward RCE. The initial condition is an isothermal atmosphere, and relative humidity is specified based on the RCE state.

4.2 A Simple Model of Relaxation to RCE

To better understand the controls on the timescale of the atmosphere's return to RCE, we now examine a simple analytical model recently presented in Cronin, T.W. and K. A. Emanuel, 2013: The climate time scale in the approach to radiative-convective equilibrium. *J. Adv. Model. Earth Sys.*, **5**, doi:10.1002/jame.20049. The goal is to strip this problem down to its essence. As we will see, this model reveals an important counter-intuitive result.

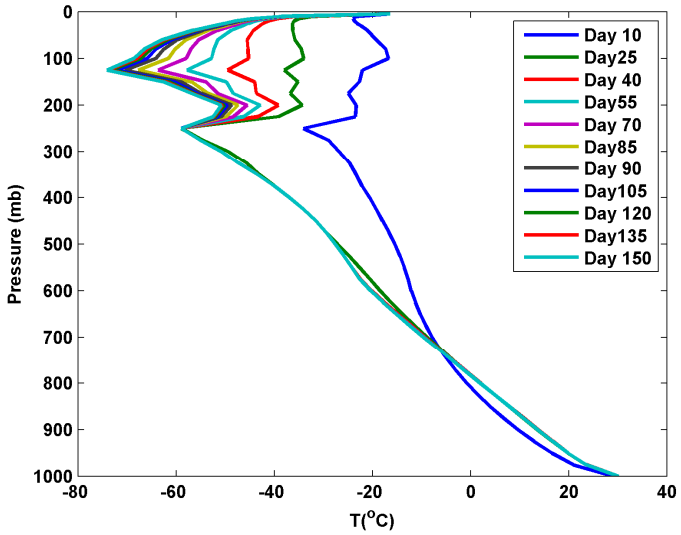


Figure 5: As in Fig. 4, but with snapshots every 15 days to show the slow evolution of the stratosphere toward equilibrium.

4.2.1 Temperature and Moist Static Energy Profiles in the Analytical Model

Figure 6 shows a schematic of our model setup. We consider an atmospheric column coupled to an ocean surface with mixed layer depth Δz_{ML}. The temperature profile is assumed to lie on a moist adiabat between the top of the boundary layer and the tropopause. This is equivalent to the statement that the saturation moist static energy (MSE) is constant with height up to the tropopause, where

$$h = c_p T + L_v q + g z \tag{1}$$

is the MSE and

$$h^\star = c_p T + L_v q^\star + g z \tag{2}$$

is the saturation MSE. Here c_p is the heat capacity of air at constant pressure, T is the temperature, L_v is the latent heat of vaporization of water, q is the specific humidity, $q^\star$ is the saturation specific humidity, g is the gravitational acceleration, and z is the height. MSE and saturation MSE are useful quantities as they are conserved during adiabatic and saturated adiabatic motions of air parcels, respectively.

To understand the equivalence of the two conditions 1) that the temperature profile lies on a moist adiabat (i.e., the column is neutral to moist convection) and 2) that the saturation MSE is constant with height, we consider the following argument. Suppose we adiabatically lift a parcel from some level z_1 to a higher level z_2. The parcel initially has the saturation MSE of the environmental profile at z_1, so that $h^\star_{\mathrm{parcel}} = h^\star(z_1)$. The parcel retains this value of saturation MSE during its adiabatic ascent and arrives at level z_2 with $h^\star_{\mathrm{parcel}} = h^\star(z_1)$. The environmental air at z_2 has saturation MSE $h^\star_{\mathrm{environment}} = h^\star(z_2)$. The difference between the saturation MSE of the parcel and that of the environment is then given by

$$h^\star_{\mathrm{parcel}} - h^\star_{\mathrm{environment}} = c_p(T_{\mathrm{parcel}} - T_{\mathrm{environment}}) + L_v(q^\star_{\mathrm{parcel}} - q^\star_{\mathrm{environment}}) \tag{3}$$

or equivalently,

$$h^\star(z_1) - h^\star(z_2) = c_p(T_{\text{parcel}} - T(z_2)) + L_v(q^\star_{\text{parcel}} - q^\star(z_2)). \tag{4}$$

No contribution to this saturation MSE difference arises from the gz term as we compare the parcel to the environment at the same altitude. To assess stability we are interested in the sign of the temperature difference appearing on the RHS of this equation: if $T_{\text{parcel}} > T_{\text{environment}}$, our lifted parcel will continue to rise and the column is unstable. The sign of this temperature difference can be inferred directly from the sign of the saturation MSE difference because the saturation specific humidity $q^\star$ is an increasing function of temperature at a given pressure. That is, if $T_{\text{parcel}} > T_{\text{environment}}$, then $q^\star_{\text{parcel}} > q^\star_{\text{environment}}$. This means that the two terms on the RHS must have the same sign. So, if the LHS is positive, both terms on the RHS must be positive, and similarly if the LHS is negative. The condition for convective neutrality is then that the saturation MSE is constant with height and equal to the actual MSE of the subcloud layer, so that both terms on the RHS are zero and adiabatically lifted parcels are neutrally buoyant.

As discussed in previous lectures, the ocean surface is in thermodynamic disequilibrium with the lower atmosphere. This thermodynamic disequilibrium is primarily due to a jump in moisture properties: the temperature at the bottom of the atmospheric boundary layer is close to that of the sea surface. However, boundary layer air is unsaturated, with a jump to saturation occurring very close to the sea surface. This leads to a discontinuity in h at the surface. (Note, however, that $h^\star$, whose profile is not shown in the boundary layer in Fig. 6, is nearly continuous at the surface due to the near agreement of atmospheric and surface temperatures.) As the boundary layer air is unsaturated, water vapor does not condense until parcels are lifted sufficiently high. The level at which condensation first occurs is referred to as the lifted condensation level, and below this level, the temperature profile assumes the dry adiabatic lapse rate. The MSE is constant with height in this dry adiabatic region: $c_pT + gz$ is constant with height due to the adiabatic lapse rate and q is constant with height due to homogenization of q in the vertical by dry convective turbulence. (In reality, q is only approximately constant with height and actually exhibits a slight decrease as we move up through the boundary layer. MSE therefore decreases in correspondence with the q variations.) Above the dry adiabatic region, the atmospheric temperature profile follows a moist adiabat up the the tropopause. Above the tropopause, radiative equilibrium leads to an increase in h with height, indicating stability to convection.

The "sickle" shape of the MSE profile $h(z)$ in Fig. 6 also warrants explanation. As mentioned above, h is constant in the boundary layer as water does not change phase and q is approximately constant with height. Between the top of the boundary layer and the tropopause, $h(z)$ exhibits a minimum as a function of height, even though h is conserved under adiabatic displacements. To understand the shape of $h(z)$, we note that moist convection occurs in a small fractional area of the atmosphere, and that in the remainder of the atmosphere we have large-scale subsidence balanced by radiative cooling. This radiative cooling acts to decrease the MSE of the descending air, causing h to decrease as we move down through the column starting from the tropopause. As we continue to descend, the $h(z)$ profile begins to increase again before we reach the dry adiabatic region in which h is constant. This increase of MSE is due to the action of shallow convection which mixes high-MSE boundary layer air into the lower levels of the free troposphere.

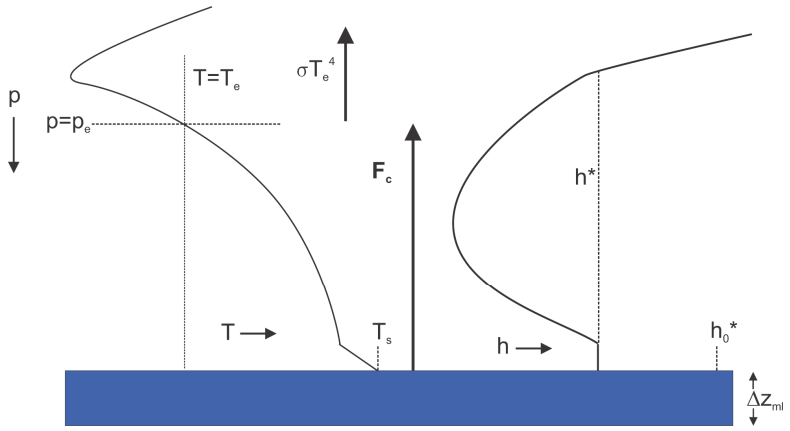


Figure 6: Schematic temperature $T(z)$, moist static energy $h(z)$, and saturation moist static energy $h^\star(z)$ profiles for our analytical model.

4.2.2 The Effect of Greenhouse Gases in the Analytical Model

Greenhouse gases are represented in the analytical model through their effect on the pressure level p_e at which the temperature on the moist adiabat equals the emission temperature $T_e = 255$ K. The effect of changing GHG concentrations in this model is shown schematically in Fig. 7. Increasing GHG concentrations causes the emission level to move upward (p_e to decrease). The moist adiabat that the tropospheric temperature profile follows is determined by the emission level p_e and the emission temperature T_e. Increasing the emission level then shifts the temperature profile onto a warmer moist adiabat. Alternatively, we can think of the moist adiabat as being determined by its value of saturation MSE, which in turn depends on T_e and p_e: $h^\star = h^\star(T_e, p_e)$.

Comparison of the adiabatic temperature profiles before and after the addition of GHGs in Fig. 7 demonstrates that temperature change at the original emission level ΔT_a is larger than the temperature change ΔT_s at the surface. Differential warming with height is a result of the curvature of the moist adiabat. For parameter values appropriate for today's atmosphere, this effect implies that a temperature increase of 3° at the emission level maps to a temperature increase of only 1° at the surface. This points to two competing climate feedbacks associated with water vapor: 1) The positive water vapor feedback due to increasing water vapor concentrations in warmer climates. 2) The ratio of surface warming to emission level warming decreases as the climate warms due to changes in the curvature of the moist adiabat. This second effect, known as the lapse rate feedback, can be thought of as buffering of surface temperature changes by evaporation.

It is important to note, however, that it is the wet bulb temperature, rather than the absolute temperature, that is most important to human comfort. The wet bulb temperature is defined as the temperature that an air parcel would have if it were cooled to saturation. This distinction is important, since even though future surface warming will be buffered by water vapor, humans will still feel the increased surface humidity. The biological upper bound is 35°C wet bulb temperature, past which human beings cannot survive. In the present climate, a maximum of 33°C is reached in the tropics, so it is likely that the

biological limit will be exceeded in global warming. Air conditioning can alleviate that, but will also likely increase GHG emissions even more.

The diurnal cycle of incoming solar radiation has important effects on tropospheric temperature which we exclude here by taking a single average value for the incoming solar radiation. In making such an approximation we introduce an error, as the average temperature resulting from the real diurnal cycle is not equal to the temperature that would result from constant insolation at the diurnal average value. This rectification of the diurnal cycle arises from the nonlinear dependence of the surface enthalpy fluxes on the surface temperature disequilibrium as well as from additional effects due to the radiative effects of clouds.

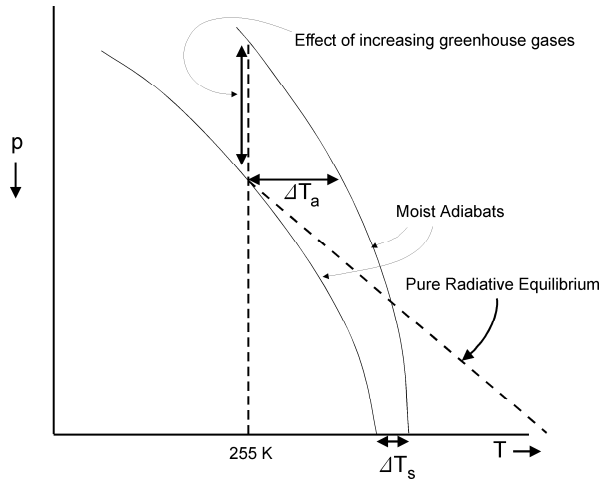


Figure 7: Schematic demonstrating the effect of adding greenhouse gases in our analytical model. The addition of GHGs moves the emission level upward and shifts the temperature onto a warmer moist adiabat.

4.2.3 Energy Balance in the Analytical Model

We now formulate the energy balance equations for our analytical model and study their linear stability to determine the timescale of relaxation to RCE. Based on the approximate surface energy balance inferred from observations in Section 2, we assume that the net LW flux at the surface is zero, and that that the steady-state energy balance in the troposphere is between the bulk turbulent flux of enthalpy F_c through the boundary layer (primarily due to evaporation) and IR emission to space σT_e^4 (see Fig. 6):

$$F_c = \rho_s C_k |V_s| (h_0^\star - h^\star) = \sigma T_e^4. \tag{5}$$

Here ρ_s is the density of air near the surface, C_k is the enthalpy exchange coefficient, V_s is the surface velocity provided by large-scale dynamics, σ is the Stefan-Boltzmann constant, $h^\star$ is the vertically-constant value of saturation MSE in the free troposphere and T_e is the emission temperature corresponding to the solar forcing. Equation (5) allows us to write the saturation MSE of the ocean surface $h_0^\star$ in terms of the free tropospheric $h^\star$ and the

solar forcing:

$$h_0^\star = h^\star + \frac{\sigma T_e^4}{\rho_s C_k |V_s|}. \tag{6}$$

We will now examine what governs the time scale of relaxation to equilibrium if the surface temperature T_s and tropospheric saturation MSE $h^\star$ are infinitesimally perturbed. (Note that our perturbations to $h^\star$ are taken to be constant with height, i.e., we perturb which moist adiabat the temperature profile follows. Non-constant perturbations would be adjusted to constancy on the fast convective timescale). We assume in our calculation that T_e, p_e and V_s remain constant. The linear dynamical equations for the evolution of the saturation MSE perturbation $h^{\star\prime}$ and surface temperature perturbation T_s' are then

$$\frac{\Delta p}{g}\frac{\mathrm{d}h^{\star\prime}}{\mathrm{d}t} = -\frac{\partial F_{\mathrm{rad}}}{\partial h^\star}h^{\star\prime} + F_c' \tag{7}$$

$$c_\ell \rho_\ell \Delta z_{\mathrm{ML}}\frac{\mathrm{d}T_s'}{\mathrm{d}t} = -F_c'. \tag{8}$$

Here Δp is the pressure thickness of the troposphere, $\Delta p = p_{\mathrm{surface}} - p_{\mathrm{tropopause}}$, ρ_ℓ is the density of liquid water, c_ℓ is the heat capacity of liquid water, Δz_{ML} is the mixed layer depth, and F_{rad} is the net upwelling IR radiation at the top of the atmosphere. In equation (7), the $\Delta p/g$ factor results from integrating the density of air through the troposphere to calculate the tropospheric heat capacity using hydrostatic balance. Equation (7) is the linearization of the tropospheric MSE balance, while equation (8) is the linearization of the surface energy balance.

In their current form the linear evolution equations are not closed as we do not know $\partial F_{\mathrm{rad}}/\partial h^\star$ or F_c'. Using the chain rule, the relationship (6) between $h^\star$ and $h_0^\star$, and the Clausius-Clapeyron relation we write

$$\frac{\partial F_{\mathrm{rad}}}{\partial h^\star} = \frac{\partial F_{\mathrm{rad}}}{\partial T_s}\frac{\partial T_s}{\partial h^\star} = \frac{B}{1 + \frac{L_v^2 q_s^\star}{c_p R_v T_s^2}}, \tag{9}$$

in which we have defined $B \equiv \partial F_{\mathrm{rad}}/\partial T_s$ to be the climate sensitivity parameter. Here R_v is the water vapor gas constant, and subscript s indicates surface values so that $q_s^\star$ is the surface saturation specific humidity. Note that T_s refers to the basic state value of the surface temperature and not its linear deviation. The quantity B depends on the structure of GHGs and other factors, but is for simplicity here treated as a constant, taken from the single column model. The linear variation of the convective flux F_c' is given by

$$F_c' = \rho_s C_k |\mathbf{V}_s| \left(h_0^{\star\prime} - h^{\star\prime}\right) \tag{10}$$

which completes the closure of (7,8).

To cast our simple model into its final form, we re-write our surface temperature perturbation T_s' in terms of the perturbation to the surface saturation MSE $h_0^{\star\prime}$ using the relation

$$T_s' = \frac{h_0^{\star\prime}/c_p}{1 + \frac{L_v^2 q_s^\star}{R_v c_p T_s^2}} \tag{11}$$

and nondimensionalize according to the following definitions:

$$A \equiv \rho_s C_k |V_s| c_p \qquad b \equiv 1 + \frac{L_v^2 q_s^\star}{R_v c_p T_s^2} \qquad \chi \equiv \frac{Ab}{B + Ab} \tag{12}$$

$$\tau_a \equiv \frac{c_p \Delta p b}{g(B + Ab)} \qquad \tau_o \equiv \frac{c_\ell \rho_\ell \Delta z_{\rm ML}}{Ab}. \tag{13}$$

After these manipulations we arrive at the final form of our analytical linear model:

$$\tau_a \frac{{\rm d}h^{\star\prime}}{{\rm d}t} = -h^{\star\prime} + \chi h_0^{\star\prime} \tag{14}$$

$$\tau_o \frac{{\rm d}h_0^{\star\prime}}{{\rm d}t} = h^{\star\prime} - h_0^{\star\prime}. \tag{15}$$

In these equations, τ_a is an atmospheric relaxation timescale and τ_o is an ocean surface timescale. The nondimensional parameter χ plays an important role in the linear dynamics and has value $\chi \sim 0.9$ for our current climate. We now use this two-dimensional linear system of ordinary differential equations to examine how ocean-atmosphere coupling affects the timescale of return to RCE.

Consider first the artificially restricted case in which the ocean surface temperature is held fixed, so that $T_s' = 0$ (and correspondingly $h_s' = 0$). The atmospheric temperature profile then returns to its original moist adiabat on the atmospheric timescale τ_a, which for our current climate has the value $\tau_a \sim 10$ days. In the correspondingly restricted case for the ocean in which the ocean surface temperature is perturbed holding the atmosphere fixed, the ocean surface returns to its equilibrium temperature on the ocean timescale $\tau_o \sim 100$ days.

We now consider the general solution for the decay timescale of perturbations away from RCE by computing the eigenvalues of the 2×2 matrix characterizing our linear dynamics. We find that the least damped eigenmode decays according to the timescale

$$\tau = \frac{2\tau_o}{1 + \frac{\tau_o}{\tau_a} - \sqrt{\left(1 + \frac{\tau_o}{\tau_a}\right)^2 - 4(1-\chi)\frac{\tau_o}{\tau_a}}}. \tag{16}$$

It is informative to investigate the behaviour of τ in two interesting limits: $\tau_o \gg \tau_a$, and $\tau_o \ll \tau_a$. In the case in which the oceanic timescale is much longer than the atmospheric timescale, we find that

$$\tau \sim \tau_a + \frac{\tau_o}{1 - \chi}. \tag{17}$$

This timescale is longer than either the atmospheric or oceanic timescales, and can be very large if χ is close to 1. In the case in which the atmospheric timescale is much longer than the ocean timescale, we obtain

$$\tau \sim \frac{\tau_a}{1 - \chi} \tag{18}$$

which is also larger than either τ_a or τ_o (as $\tau_a \gg \tau_o$), and can also grow large if χ is close to 1. This analysis indicates that the timescale for return to equilibrium for a coupled system is always longer than the atmospheric timescale. Even in the limit of a very thin mixed

layer ($\Delta z_{\rm ML} \to 0$), in which case τ_o becomes very small, the timescale for the return to equilibrium can still be much longer than the atmospheric timescale τ_a as a result of the factor $(1-\chi)$ appearing in the denominator due to the atmosphere–ocean coupling. This effect can be seen in general circulation models, which require much longer integrations to reach equilibrium when coupled to an ocean.

The linear model defined by equations (7) and (8) performs well when compared to a single-column model in spite of its apparent simplicity. Figure 8 shows a comparison of the return rate to equilibrium as calculated from the analytical model alongside best fit decay rates measured in a single-column model. The single-column model was used to measure the climate sensitivity parameter B which appears in the linearized dynamics. The degree of agreement between the simple model and column model is impressive both in the approximate numerical magnitude of the timescale as well as in the dependence of the timescale on the mixed layer and the surface temperature of the equilibrium state.

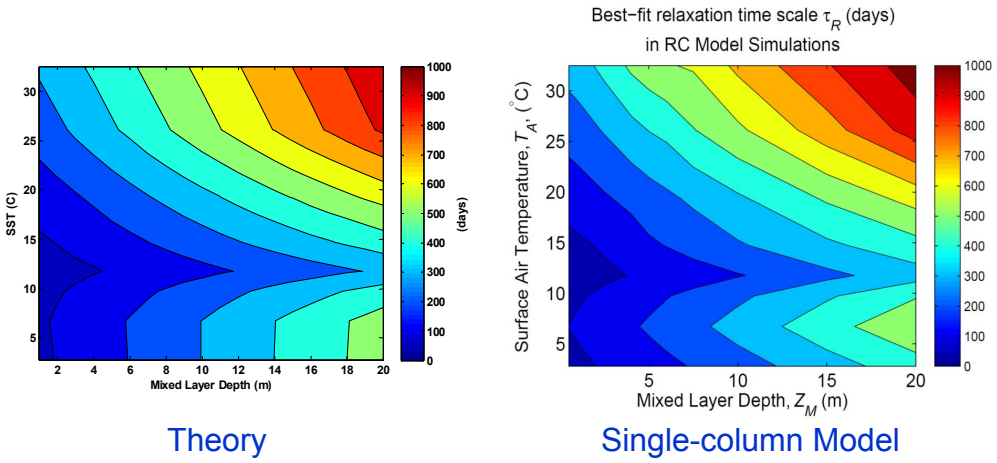


Figure 8: Comparison of the return timescale to RCE between the analytical linear model and a single-column model. The single-column model is used to parameterize the analytical model through the single climate sensitivity parameter B. The agreement between the simple theory and the numerical model is apparent.

5 Explicit Simulation of Radiative-Convective Equilibrium

We next turn our attention to a problem at the frontier of research in atmospheric science: convective aggregation. The numerical modeling and analysis in this section follows a recently published paper: Wing, A. A., and K. A. Emanuel (2014), Physical mechanisms controlling self-aggregation of convection in idealized numerical modeling simulations, *J. Adv. Model. Earth. Syst., 6, 59-74*, doi:10.1002/2013MS000269.

5.1 Model Observations of Convective Aggregation

The study of convective aggregation requires a more sophisticated modelling approach than we have taken so far in these lectures. We now move from simple dynamical systems models

and single-column models to idealized three-dimensional simulations of moist convection using the cloud-resolving model SAM (System for Atmospheric Modelling). The model geometry and forcing are shown in Fig. 9. The domain is doubly periodic in the horizontal with rigid boundaries at the top and bottom. The model is forced by fixed and spatially homogeneous sea surface temperatures (SSTs) in the range of 297–312 K at the lower boundary and a constant value of the solar forcing set to be 413.98 W/m^2. The domain size is 768 km $\times$ 768 km $\times$ 28 km, with a horizontal resolution of 3 km and 64 levels in the vertical. The initial conditions are specified according to a profile computed from the domain average of a previous model run of radiative–convective equilibrium in a smaller domain. The model resolution permits the marginal resolution of moist convection, and the model features a fully-interactive rapid radiative transfer model (RRTM), cloud microphysics, as well as surface fluxes.

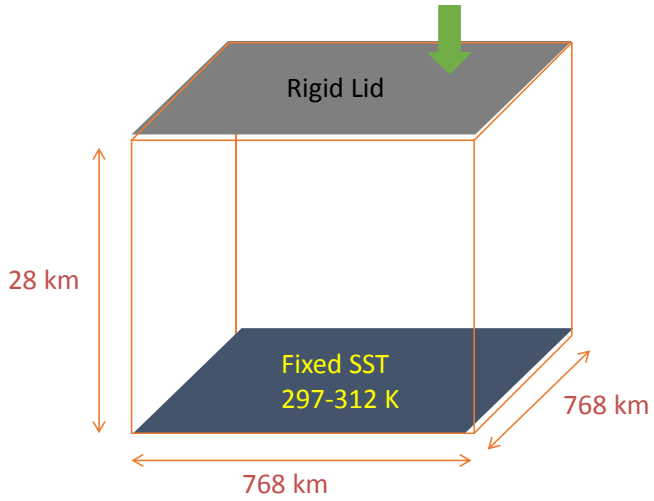


Figure 9: Schematic of the domain used in cloud-resolving modelling to study convective aggregation.

When the cloud-resolving model is run in this configuration for a sufficiently long time, the convection, which is initially statistically homogeneous and apparently random, spontaneously focuses into a cluster. This can be visualized in two dimensions by plotting the vertically-integrated water vapor (also referred to as the total precipitable water, or TPW) at various times throughout the model integration. Figure 10 (left) shows the time-averaged TPW over the first day of model integration as a function of (x, y) for a case in which the SST is set at 305K. The moisture field at each point in space is near its spatial average value. As the integration continues, TPW becomes spatially inhomogeneous. Figure 10 (right) shows the average TPW on Day 10. A relatively dry area is seen to form in the upper left corner of the domain. The temporal and spatial evolution of this dry patch, or "hole", can be seen in Figs. 11–14. By Day 90, the hole has expanded to cover most of the outer edge of the (periodic) domain, and the convective region (the complement of the hole) has become focused at the centre of the domain. We note that the boundary conditions and forcing for this problem are spatially homogeneous, so that the final position of the convecting region (the red region in Fig. 14) is random and dependent on the structure of

the noise in the initial conditions. We also note that the spatial structure of the convective cluster/hole system has a complex time evolution, sometimes forming horizontal bands as can be seen in Fig. 13. This phenomenon of hole formation followed by convective clustering is often referred to under the name "convective aggregation" and has been observed in a variety of cloud-resolving models. However, the evolution shown in Figs. 10–14 indicates that the convective cluster actually results from the initial formation of a small hole in the otherwise statistically homogeneous convection, and so perhaps the phenomenon could be more accurately called "convective self-annihilation".

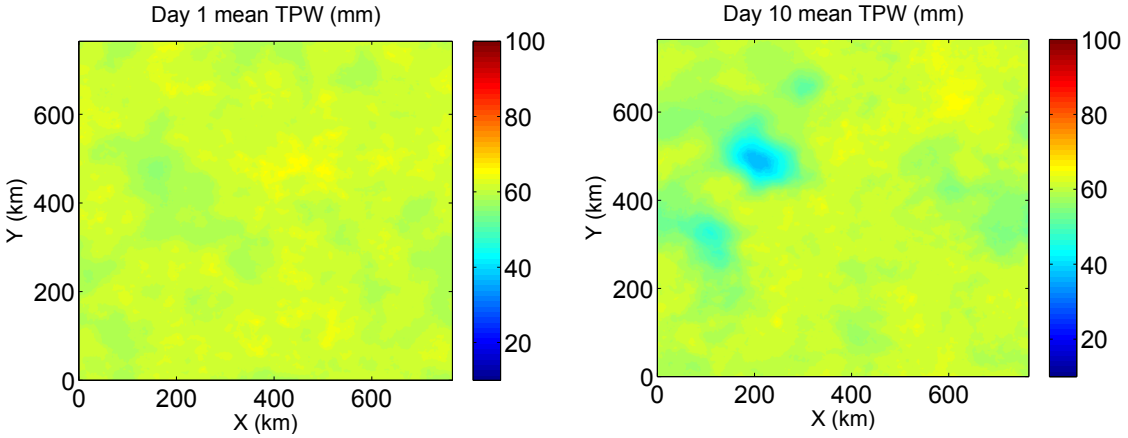


Figure 10: Average total precipitable water (TPW) over Day 1 (left) and Day 10 (right) during a cloud-resolving model simulation in the configuration shown in Fig. 9. The SST is set at 305K. Note the formation of a small hole in the convection on Day 10 in the upper left corner of the domain.

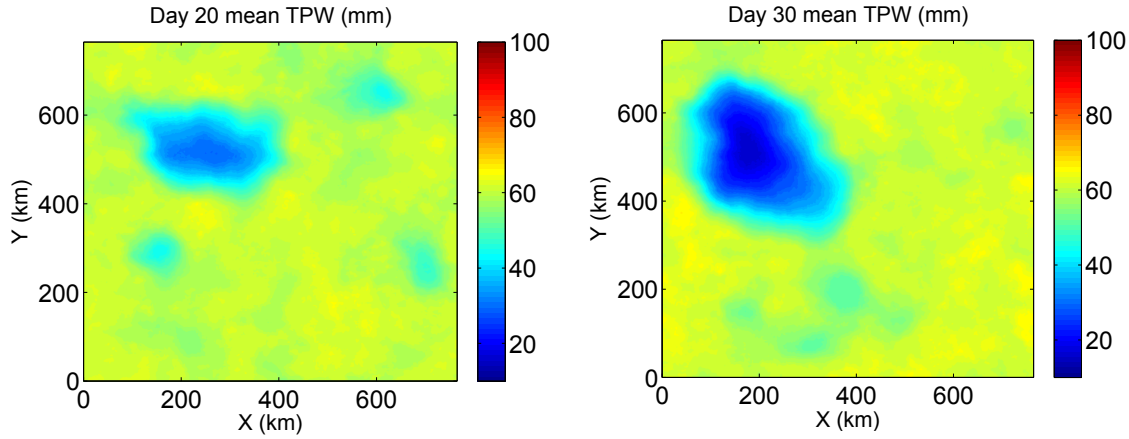


Figure 11: As in Fig. 10, but for Day 20 (left) and Day 30 (right).

Two physically interesting and important questions concerning convective aggregation are: 1) What controls the spatial scale of the final cluster? 2) How does convective ag-

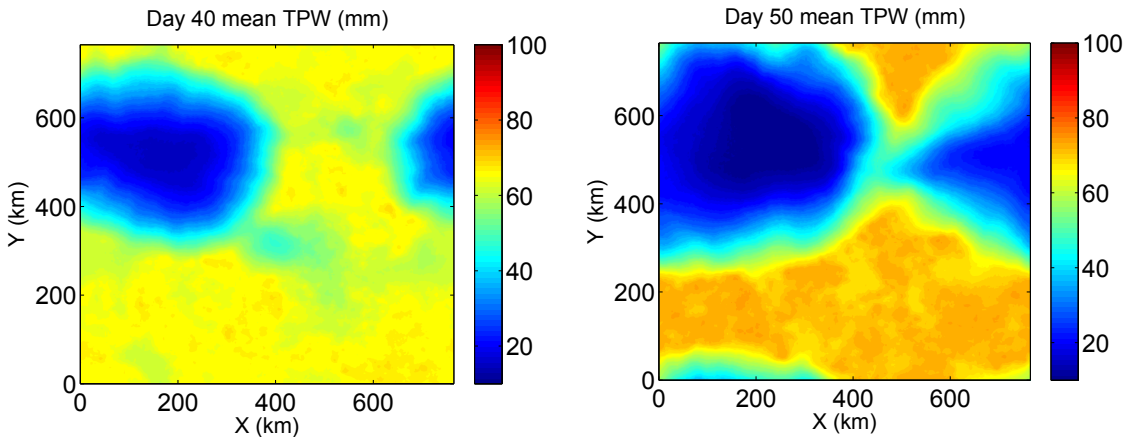


Figure 12: As in Fig. 10, but for Day 40 (left) and Day 50 (right).

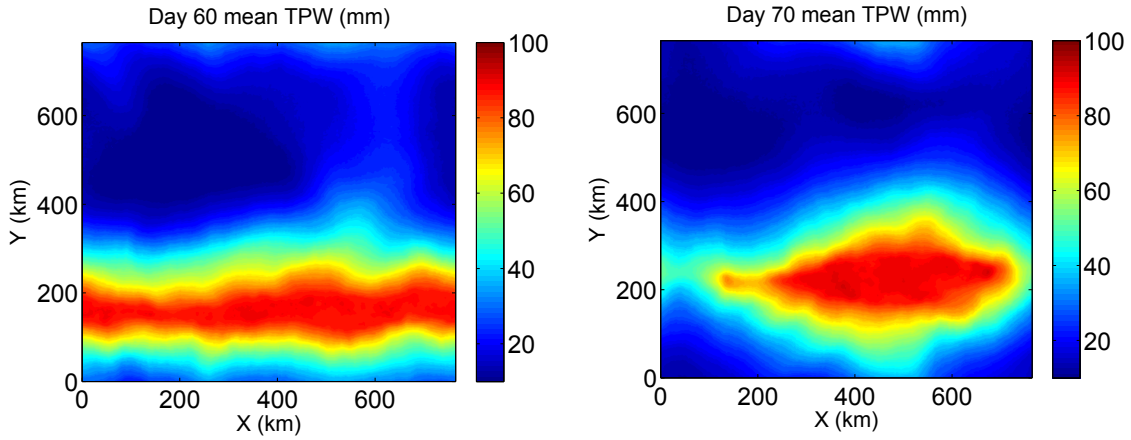


Figure 13: As in Fig. 10, but for Day 60 (left) and Day 70 (right).

gregation depend on the imposed SST? The answers to these questions are not completely understood. Toward addressing the first question, it is observed in nonrotating cloud-resolving model simulations that convective clustering always results in a final cluster state that takes on the scale of the domain. An example of this can be seen in Fig. 14, in which there is only one convectively active (bright red) region, with the rest of the domain being dry and quiescent. If simulations are run in a larger domain, the size of the final cluster increases as well. It has also been found that convective aggregation does not occur if the simulation is run in a sufficiently small domain. We will discuss a physical hypothesis for explaining this observation in the next lecture. We note before proceeding that it is relevant to ask whether our fixing of SSTs results in an effective forcing of the convective cluster once it has formed. Such an effect could arise because we do not allow the SST to adjust under the convective region, thereby excluding the shadowing effect the clouds would have on the surface. It would be important to test the effect of variable SSTs on cluster formation by coupling our cloud-resolving model to a slab ocean. However, such computations are very

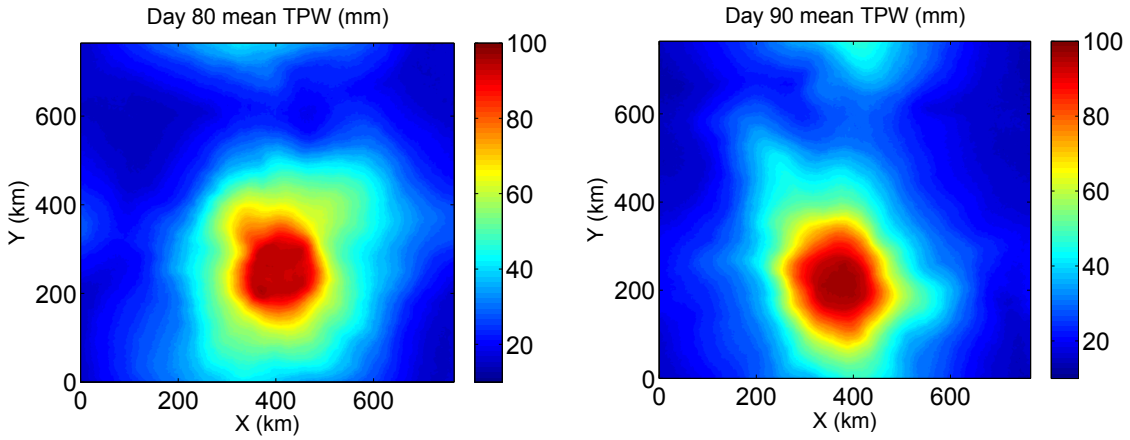


Figure 14: As in Fig. 10, but for Day 80 (left) and Day 90 (right).

expensive to run to equilibrium as the time scale for equilibration becomes very long due to the coupling, as discussed previously in this lecture. It is not currently known whether convective aggregation occurs when coupled to an interactive ocean surface.

To address the dependence of convective aggregation on SST, we examine the results of our cloud-resolving model simulations for different fixed, spatially-constant SST values. Figure 15 shows the time evolution of the outgoing LW radiation (OLR) for a variety of simulations with different SSTs. After an initial transient equilibration period, all integrations reach an approximate statistical equilibrium after a few days with OLR in the range of 250–270 W/m^2. After this period, some model runs show a significant increase in OLR, indicating that this initial statistical equilibrium is unstable. This increase in OLR is a signal that convective aggregation is occurring: LW flux to space increases when the convection aggregates as there is very little water vapor outside the cluster, and the cluster occupies a relatively small areal fraction of the domain. The SST values for each integration are shown in the panel on the RHS of Fig. 15. Inspection of the SST values reveals an apparent transition point in the large-scale behaviour of the convection as a function of SST. For the relatively low SST values 297 K, 298 K, 300 K (dark blue and black lines with OLR $\sim$ 255 W/m^2 after the transient), the OLR appears to remain statistically steady, and the convection remains statistically homogeneous. As the SST is increased beyond 300 K, the behaviour transitions to convective organization: model integrations for 301–307 K show increasing OLR associated with convective aggregation beginning around 40 days. These results suggest that 300 K is an SST threshold above which convection aggregates and below which convection remains statistically homogeneous. We note that the time required for the clustering to complete appears to depend on the initial noise in the system and varies between model runs, although the *rate* of increase of OLR appears to be determined by the boundary conditions and forcing.

A complication appears to arise as we continue to increase SST above 307 K. The integrations for 310 K, 311 K, and 312 K, shown in purple/red/dark orange, appear to have constant OLR $\sim$ 275 W/m^2, never forming a convective cluster. This appears to indicate a non-monotonic dependence of convective clustering on SST. However, this behaviour may

be artificial. An additional model run with an SST value of 310 K but with the size of the domain doubled is shown by the red curve that exhibits increasing OLR and that has the highest final value of OLR at day 100. Although the 310 K run with the standard domain size does not aggregate, a cluster does form in the enlarged domain. This suggests that the apparent lack of monotonicity may actually indicate that the critical domain size beyond which clustering is possible depends on temperature: we may need to use systematically larger model domains to observe clustering as we increase SSTs. As the calculations are expensive, however, it is not currently feasible to continually increase the domain size and test whether all higher SST cases aggregate given a sufficiently large domain.

This leaves us with the impression that the critical temperature above which aggregation happens is about 300 K or 27°C. This is suspiciously close to the tropical surface temperature in the current climate. This may be a coincidence, but we will speculate that it arises from a self-regulatory process in the next lecture.

Aggregation similarly happens in climate models, affecting dramatically the simulated climate. Whether clusters in these models form for the right reasons is an open question. Clusters are, of course, also observed in the real world, giving us confidence that aggregation plays an important role in the present climate.

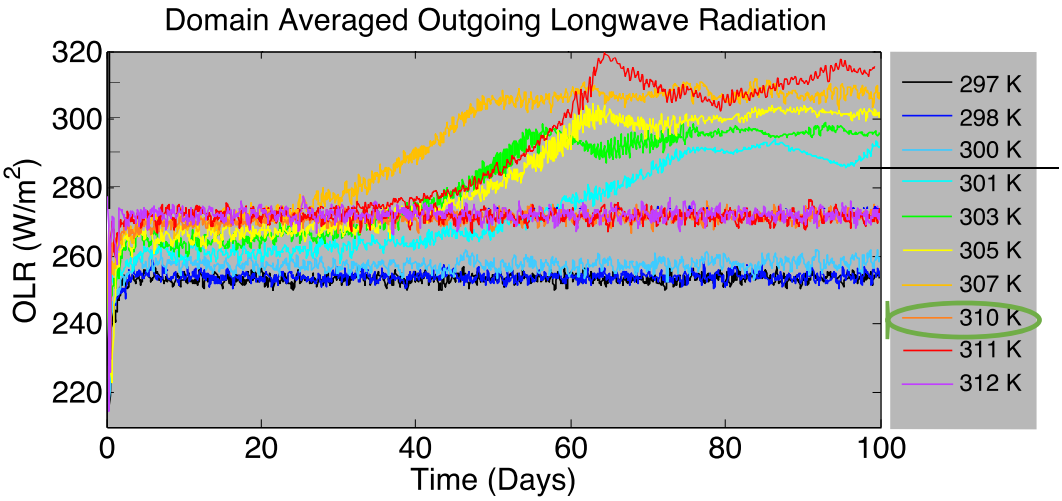


Figure 15: Time evolution of the outgoing longwave radiation (OLR) for various cloud-resolving model runs with different fixed SSTs (right). Increases in OLR after the initial equilibration period indicate the formation of a convective cluster.

5.2 Energy Budget during Aggregation

To better understand the feedback processes involved in convective aggregation, we now examine the energy budget throughout the domain and as a function of time. We will work with the mass-weighted column integral of the frozen moist static energy (FMSE), defined as

$$h = c_p T + gz + L_v q_v - L_f q_i \tag{19}$$

in which we now use the symbol q_v to indicate the water vapor mixing ratio, q_i is the ice mixing ratio, and L_f is the latent heat of fusion of water. We will denote the mass-weighted vertical integral operator using a caret: $\hat{h}$. The vertically integrated FMSE $\hat{h}$ is not changed by moist convection, which is a transport process, but can be only be changed by radiation, surface fluxes, and horizontal transport. We will consider the deviations of $\hat{h}$ from the horizontal mean, and denote these deviations by $\hat{h}'$. The evolution equation for $\hat{h}'$ is

$$\frac{\partial \hat{h}'}{\partial t} = \mathrm{LHF}' + \mathrm{SHF}' + \mathrm{NetSW}' + \mathrm{NetLW}' - \boldsymbol{\nabla}_h \cdot \widehat{\mathbf{u}h} \tag{20}$$

where LHF is the surface latent heat flux, SHF is the surface sensible heat flux, NetSW is the column SW radiative flux convergence, NetLW is the column LW radiative flux convergence, and $\boldsymbol{\nabla}_h \cdot \widehat{\mathbf{u}h}$ is the advective transport term, with $\boldsymbol{\nabla}_h$ denoting the gradient operator in the horizontal. Note that the advection term has zero mean as it is a flux divergence. The evolution equation for the variance of $\hat{h}'$ is obtained by multiplying eq. (20) by $\hat{h}'$:

$$\frac{1}{2}\frac{\partial}{\partial t}(\hat{h}')^2 = \hat{h}'\mathrm{LHF}' + \hat{h}'\mathrm{SHF}' + \hat{h}'\mathrm{NetSW}' + \hat{h}'\mathrm{NetLW}' - \hat{h}'\boldsymbol{\nabla}_h \cdot \widehat{\mathbf{u}h}. \tag{21}$$

The variance of FMSE is a useful quantity to examine because it increases with the degree of convective aggregation. The source and sink terms in eq. (21) are correlations between the mass-weighted column-integrated FMSE and the various diabatic terms as well as with the horizontal convergence of $\hat{h}$. When one of these correlations is positive, that term contributes to self-aggregation by either drying a relatively dry region or moistening an already moist region. By examining the sign and magnitude of these RHS terms, we will be able to identify the effect that each of these physical processes has on the formation of convective clusters. In the analysis, the LHF, SHF, NetSW, and NetLW terms are calculated from the model. The horizontal convergence term is then diagnosed as a residual.

To visualize the results of the analysis, each RHS term is calculated and then averaged over each non-overlapping 48 km × 48 km subdomain and also averaged over one day. For each RHS term, the average values for each subdomain are then ordered according to the subdomain's average value of the column relative humidity, from dry to moist. This ordering procedure allows us to visualize the effect of each term in the FMSE variance budget as it depends on both time and moisture. We also normalize each term by the instantaneous horizontal mean value $\{(\hat{h}')^2\}$, where curly braces indicate the horizontal average. This normalization prevents the early stages of the dynamics from being washed out due to the smallness of the FMSE variance as well as the smallness of the individual RHS budget terms. Figure 16 shows an example of such a visualization from the 305 K run, whose OLR can be seen as a function of time in Fig. 15. In Fig. 16 the x-axis gives the columns ranked by their moisture content, while the y-axis shows time evolution from 0 to 90 days. Colors show the value of the correlation between $\hat{h}'$ and the total diabatic term, that is, the contribution to the rate of change of $(\hat{h}')^2$ from

$$\hat{h}'\mathrm{LHF}' + \hat{h}'\mathrm{SHF}' + \hat{h}'\mathrm{NetSW}' + \hat{h}'\mathrm{NetLW}'. \tag{22}$$

The black line running vertically through Fig. 16 is the $\hat{h}' = 0$ contour: the line separating the relatively dry columns from the relatively moist columns. We note that this way of

presenting the x-axis does not track the actual values of the column moisture. Over time, the range of column relative humidities from the driest to the moistest may increase, so that the x-axis would expand over time if we tracked the absolute values of humidity rather than using the ranking system presented here.

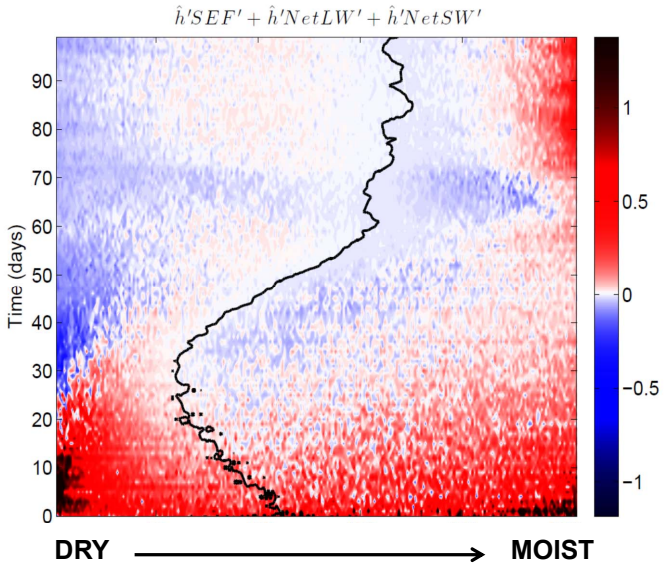


Figure 16: Contribution to the variance budget of mass-weighted column-integrated frozen moist static energy from the total diabatic term in the moisture-time plane. The x-axis orders air columns from driest to moistest, while time increases upward along the y-axis. Colors show the contribution to the variance tendency due to the sum of the diabatic terms on the RHS of eq. (21).

Figure 16 demonstrates that the total diabatic term on the RHS of eq. (21) acts to increase the FMSE variance in the first 20 days of the simulation, and that at early times the strongest correlations are observed in the dry regions. This is consistent with our observation that the convective clustering originates as a 'hole' seen to develop in the TPW field over time: increasing variance in dry regions corresponds to continued drying of the hole. These positive correlations then decrease in magnitude and shift into the moist regions, coinciding with the expansion of the hole into the convecting region.

Having examined the total effect of the diabatic terms on convective organization, we now consider them one at a time. Figure 17 visualizes the contribution to the rate of change of $(\hat{h}')^2$ from the column SW flux convergence NetSW. The SW flux convergence correlation is weak over most of domain for the majority of the integration with two important exceptions. Over the first 50 days in the driest regions, the SW effect is positive and tends to increase variance. This results from a feedback whereby dry columns ($\hat{h}' < 0$) absorb less solar radiation (NetSW$' < 0$) because water vapor is an important SW absorber. Negative solar absorption anomalies induce anomalous sinking, continuing to dry out the column. This positive feedback also acts in the moist regions once the convective cluster has formed. In the latter part of the integration (beyond 50 days or so) the moist regions have a strongly

positive SW effect because moist columns absorb more solar radiation leading to anomalous rising motion and moistening of the column. Both directions of this feedback act to enhance clustering and correspondingly increase the FMSE variance.

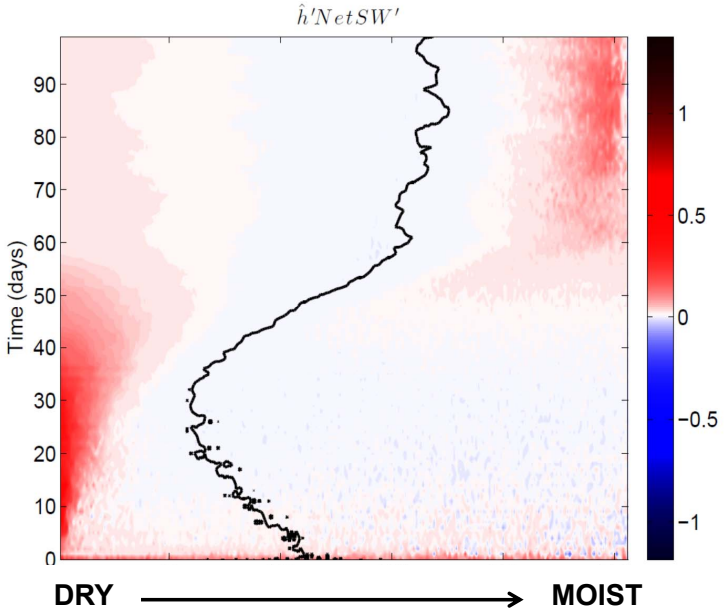


Figure 17: As in Fig. 16, but for the SW flux convergence.

The contribution to the FMSE variance from the LW term $\hat{h}'\text{NetLW}'$ is shown in Fig. 18. Early in the integration the LW contribution is essentially positive, but weakens after the first 20 days or so and is of varying sign in both relatively dry and relatively moist regions. However, once the convective cluster has been formed, the LW term is important for its maintenance. This can be seen in the top-right corner of Fig. 18, where the moistest columns (associated with the cluster) tend to have their FMSE variance increased by the LW effect. This feedback leads to hysteresis: once a cluster is formed, it is maintained by this effect, even if the SST is lowered below the critical value for the formation of clusters. The LW positive feedback is primarily due to the presence of high clouds that are opaque to LW radiation and have relatively low temperatures, reducing the LW emission to space. Anomalously low LW cooling manifests as a positive contribution in the anomaly energy budget $\text{NetLW}' > 0$, which has a positive correlation with $\hat{h}'$ in the moist columns within the convective cluster.

To discuss the effects of the surface fluxes LHF$'$ and SHF$'$ it is useful to decompose their fluctuations into components due to fluctuating winds and fluctuations in the air-sea disequilibrium in temperature and water vapor. We can write the surface enthalpy fluxes in terms of bulk formulae as

$$\text{LHF} = \rho c_E L_v U(q^\star_{Ts} - q_v) \tag{23}$$

$$\text{SHF} = \rho c_H c_p U(T_s - T_a) \tag{24}$$

where c_E is the latent heat exchange coefficient, c_H is the sensible heat exchange coefficient,

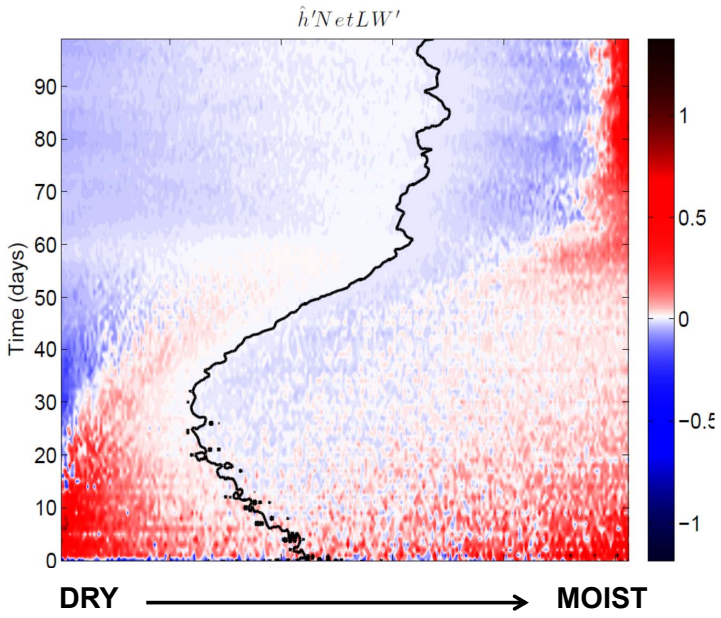


Figure 18: As in Fig. 16, but for the LW flux convergence.

U is the surface wind speed, T_s is the surface temperature, $q^\star_{Ts}$ is the saturation specific humidity at the surface, T_a is the atmospheric temperature at the lowest model level, and q_v is the specific humidity at the lowest model level. When computing the deviations from the horizontal mean of these enthalpy fluxes, we can isolate the total contribution from the fluctuations in surface wind speed:

$$\text{Surface flux wind feedback term} = \hat{h}'\rho L_v (c_E U)' \{\Delta q\} + \hat{h}'\rho c_p (c_H U)' \{\Delta T\}\,. \tag{25}$$

Here the Δ symbol denotes the difference between the surface and the lowest model level as in eqns. (23) and (24). Figure 19 shows the effect of the wind fluctuation term in the moisture–time plane. This term is essentially positive everywhere and also contributes strongly to the maintenance of the cluster (upper right corner). This effect arises due to a positive feedback. The formation of the convective cluster produces more thunderstorms leading to anomalously strong surface winds. These enhanced surface winds produce greater surface enthalpy fluxes, so that the wind contribution to SHF + LHF is positive, giving a positive correlation with $\hat{h}'$ in the moist columns.

We now separately consider the contribution to the FMSE variance budget of the surface enthalpy flux anomalies due to fluctuations in the air–sea disequilibrium. From the bulk formulae (23) and (24) we obtain

$$\text{Surface flux air–sea disequilibrium feedback term} = \hat{h}'\rho L_v \{c_E U\} \Delta q' + \hat{h}'\rho c_p \{c_H U\} \Delta T'. \tag{26}$$

Figure 20 shows the effect of this feedback term. The air–sea disequilibrium fluctuations act to reduce the tendency toward cluster formation everywhere in the domain and over all time periods. In the cluster region in particular, this surface flux acts to reduce the FMSE variance and weaken the cluster. This is due to a feedback. The presence of the convective

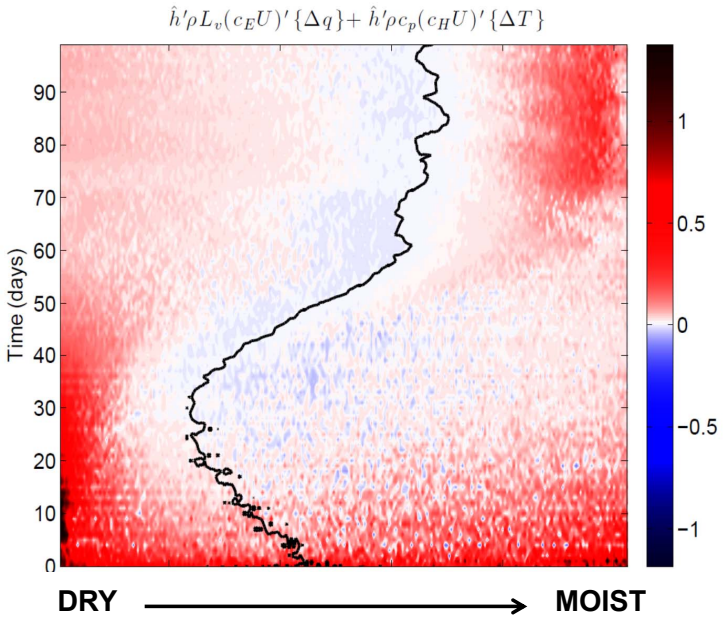


Figure 19: As in Fig. 16, but for the contribution to the total surface enthalpy flux from wind fluctuations.

cluster leads to moistening in the boundary layer. As a result, for a fixed surface wind speed, the boundary layer air has a specific humidity anomalously close to that of the ocean surface, leading to a negative anomalous surface enthalpy flux in a relatively moist region. This is a negative correlation and acts as a variance sink in the FMSE budget.

The combined effect of the surface flux feedback term is shown in Fig. 21. This includes the contributions discussed above due to U' and $(\Delta q', \Delta T')$ as well as additional contributions from $U'\Delta q'$ and $U'\Delta T'$. Comparison of the combined surface flux contribution with its components shown in Figs. 19 and 20 demonstrates that there is a large amount of cancellation in the FMSE variance budget between the mostly-positive effect of the wind fluctuations and the mostly-negative effect of the air–sea disequilibrium fluctuations. However, these cancellations are not complete. The net effect of the surface flux feedback is positive everywhere during the first 20 days of the integration, and is strongest in the dryest regions. The variations in strength and sign of this combined feedback term in the moisture–time plane demonstrate the effects of competing feedbacks controlling the development of the convective cluster.

For completeness, we show in Fig. 22 the contribution to the FMSE variance budget from horizontal divergence. This term is calculated as a residual from the rest of the budget, and is noisy and of varying sign in moisture–time space. This term can be computed more accurately by evaluating it directly, but the necessary quantities were not saved in the model simulation.

In the final lecture, we will try to understand the physics of aggregation. We will add rotation to our idealized model problem, taking us toward the dynamics of hurricanes and tropical cylones. The addition of rotation to the problem will provide an external scale for

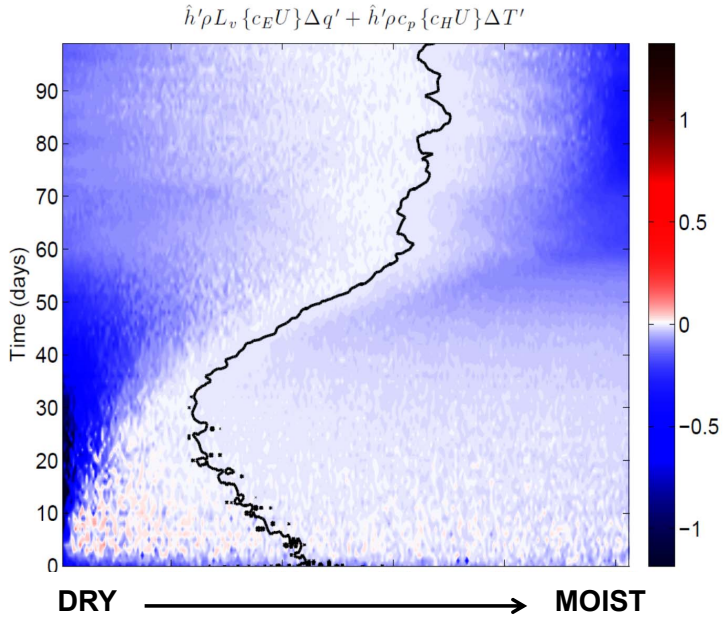


Figure 20: As in Fig. 16, but for the contribution to the total surface enthalpy flux from fluctuations in the air-sea thermodynamic disequilibrium.

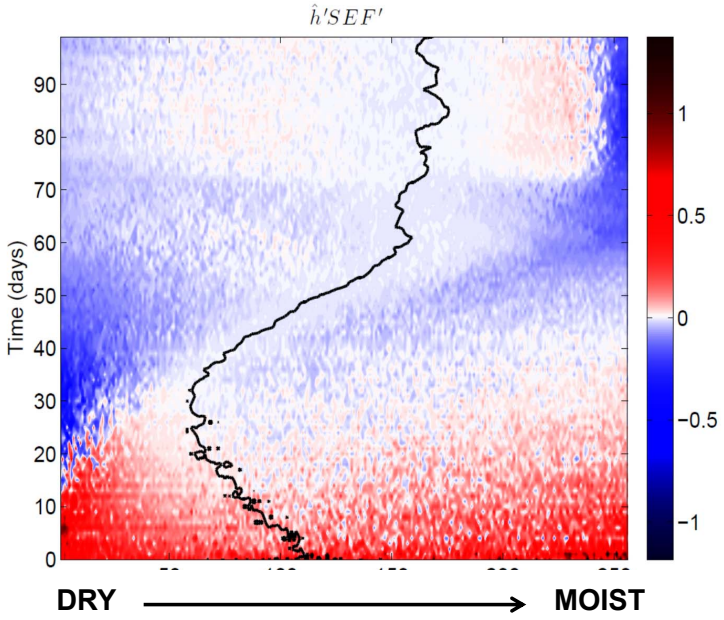


Figure 21: As in Fig. 16, but for the total surface enthalpy flux.

the problem, leading to scale-selection in the convective aggregation process. Using a linear stability analysis of the radiative–convective equilibrium, we will find that the physics when the hole is forming are different from the physics that sustain the cluster. The linear stability

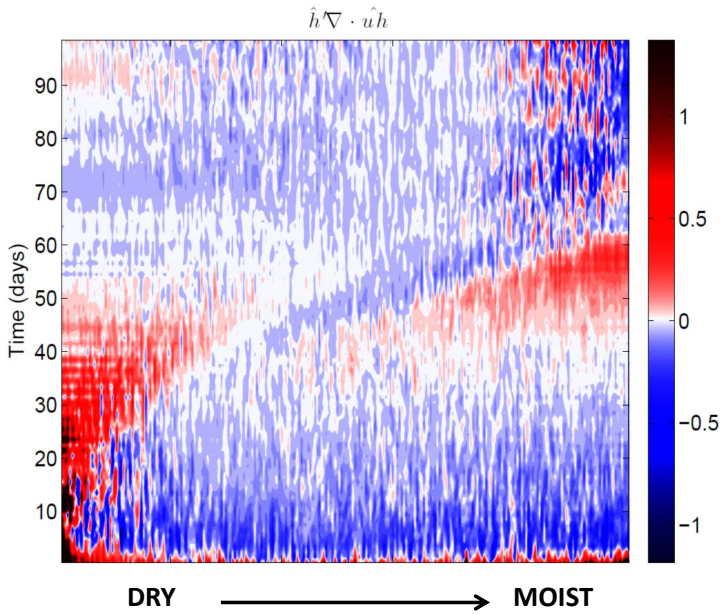


Figure 22: As in Fig. 16, but for the horizontal convergence. Note that this term is computed as a residual from the rest of the FMSE budget.

analysis will reveal that the initial instability depends on the temperature dependence of LW emissivity and help explain why aggregation only occurs at sufficiently large SSTs. In our analysis, we will make the weak temperature gradient approximation—in contrast to classic GFD problems, we thus parameterize the dynamics and focus our attention on the interaction of water substances and radiation. We will end with a speculative discussion of what we can infer from these idealized simulations for the real climate.

Lecture 7: Stability of RCE and Rotating RCE

Kerry Emanuel; notes by Alexis Kaminski and Shineng Hu

June 24

1 Stability of RCE

The radiative moist convective equilibrium (RCE) state we have been considering so far is evidently unstable in some conditions. By considering the linear stability of the RCE state (recalling that RCE is a statistical equilibrium), is it possible to explain the clustering seen previously in the cloud-resolving simulations? In order to proceed, some additional concepts must be introduced to allow for dynamics in RCE. In some sense, we will take the opposite approach to classical geophysical fluid dynamics in that the dynamics will be parameterized while the physics will be resolved.

1.1 Weak temperature gradient approximation

As the RCE state considered so far is non-rotating, there is no way to sustain lateral temperature gradients (e.g. $\mathrm{d}T/\mathrm{d}x$). As a result, the weak temperature gradient (WTG) approximation can be made, which assumes that the dynamics operate so as to keep temperature constant in the horizontal [Sobel and Bretherton, 2000].

We can suppose that we have an RCE state and add an SST anomaly, essentially making a small patch of ocean slightly warmer. In this case, the atmosphere wants to be warmer above the SST anomaly, but with the WTG approximation this is not allowed to happen. This results in large-scale ascent above the patch such that adiabatic heat balances the anomaly. In essence, the SST anomaly induces that vertical motion which is necessary to keep the atmosphere relatively cool. The vertical velocity also advects water vapour vertically, leading to a moistening of the column, an increase in clouds, and less sunlight reaching the surface.

The WTG approximation can be extended to a weakly rotating framework. In this case, if w is known from WTG, then the vorticity equation can be used to solve for the horizontal component of motion. This allows us to solve for atmospheric flow without solving the equations of motion, even in a single column model. It should be noted that the convection in the system does not allow the ascending air to reach saturation on the macro scale.

WTG may also be used to empirically determine the stability of the RCE state. A single column model can be run until equilibrium, at which point it is stopped and restarted in WTG mode with a small amount of noise added. If the RCE state is stable, at this point the noise dies away. However, for an unstable RCE state, the noise will lead to ascent or descent throughout the column.

The WTG approximation may also be derived asymptotically by assuming convection and gravity waves act on sufficiently fast timescales (in analogy with the quasigeostrophic approximation) that they are able to smooth out temperature perturbations quickly. Virtual temperature (i.e. temperature corrected for water vapour content) above the planetary boundary layer (PBL) is assumed to be invariant. The vertical velocity w is thus calculated in order to maintain constant virtual temperature, using the vertical motion to advect water vapour.

1.2 Single column model

The model used is the MIT Single-Column model. It employs Fouquart & Bonnel shortwave radiation, Morcrette longwave radiation, Emanuel-Zivkovic-Rothman convection, and the Bony-Emanuel cloud scheme. 25 hPa grid spacing is used in the troposphere, with higher resolution in the stratosphere. The model is run with fixed SST until the RCE state is reached, then re-initialized in WTG mode with fixed temperature at 850 hPa and above to simulate the PBL. Small perturbations in w are added to the initial condition.

When the SST is less than approximately 32°C, there is no drift from the RCE state, and the initial perturbations decay away. However, for higher SSTs, there is a migration toward states with either ascent (formation of a cluster) or descent (formation of a hole), depending on the random initial perturbations. These states correspond to multiple equilibria in 2-column models, observed by Nilsson and Emanuel [1999], Raymond and Zeng [2000] and others.

Figure 1 shows timeseries data, starting from the addition of noise, for the descending branch. Plotted are contours of specific humidity, $\omega = \mathrm{d}p/\mathrm{d}t$ (an analogue for vertical velocity in pressure coordinates, where $\omega \simeq -\rho g w$), radiative heating, and convective heating. Note that the height of the cutoff in data for ω, and the height at which convective heating changes sign, corresponds to the PBL at 850 hPa. Qualitatively, figure 1 shows that in the descending branch air is sinking as the atmosphere is drying out. Radiative cooling is observed low in the troposphere. The growth rate of the formation of the hole (in the case of the descending branch) is similar to that seen in the cloud-resolving model discussed earlier.

Figure 2 shows the effect of an instantaneous, vertically-uniform reduction of the specific humidity by 20% from the RCE state with two different SST values. In the stable case (SST=25°C), the reduced specific humidity leads to a negative shortwave radiation anomaly, as there is less water vapour present to absorb the incoming solar radiation. The longwave radiation anomaly is mostly positive, as less water vapour leads to less radiative cooling by infrared.

However, the unstable case (SST=40°C) exhibits markedly different behaviour. While the shortwave radiation anomaly is similar to that of the stable scenario (indicating reduced absorption of incoming solar radiation), the longwave radiation anomaly is an order of magnitude larger than that of the stable scenario. The longwave radiation anomaly is positive above approximately 750 mbar but negative below. This is a consequence of the nonlocal nature of radiation, with the lower layers receiving less infrared radiation from above. In this circumstance where RCE is very warm, the basic state has large quantities of water vapour and very high emissivity in the lower troposphere, though the emissivity

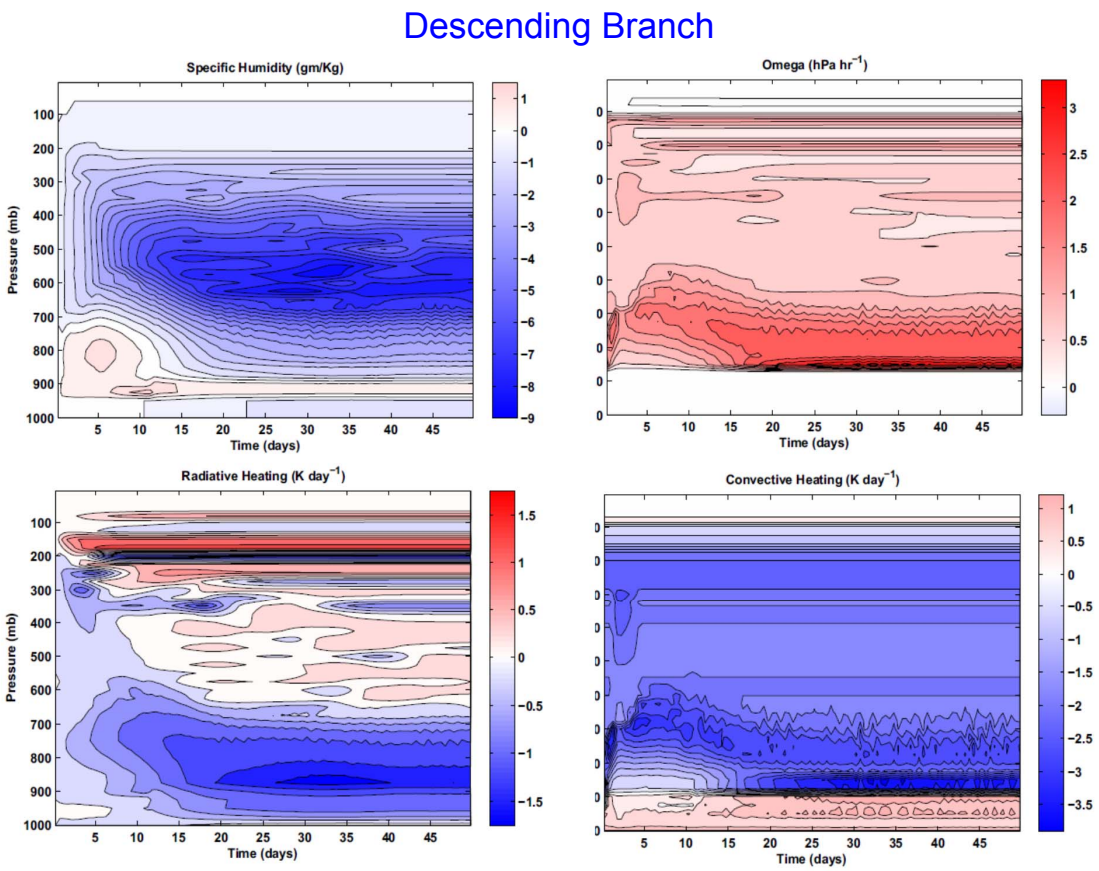


Figure 1: Timeseries contours of specific humidity in gm/kg (top left), ω in hPa hr^{-1} (top right), radiative heating in K day^{-1} (bottom left), and convective heating in K day^{-1} (bottom right) for the descending branch of the WTG single-column model results. From figure 3 of Emanuel et al. [2014].

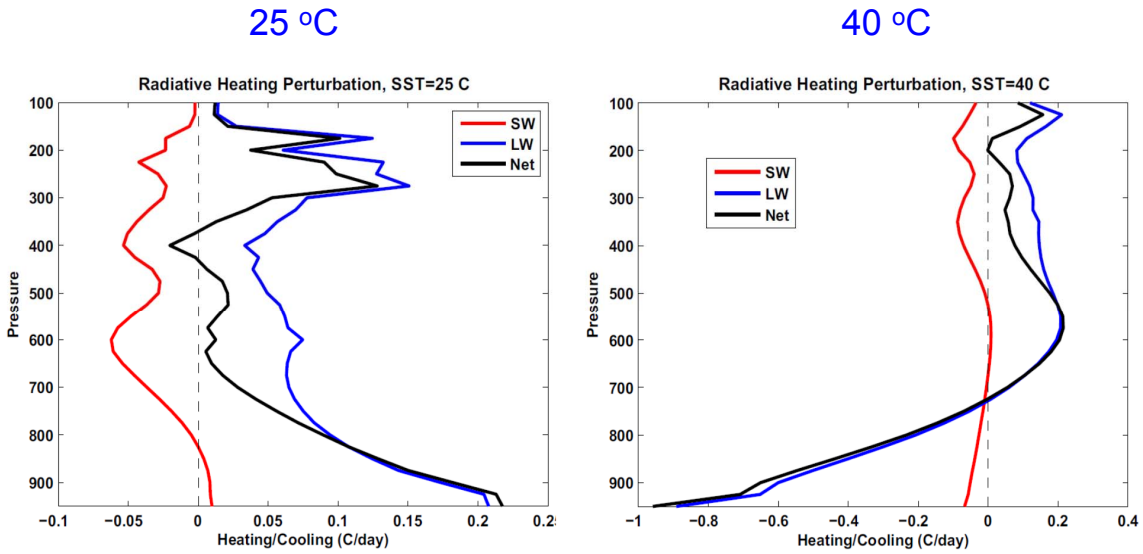


Figure 2: The response of shortwave (SW), longwave (LW), and net perturbation heating rates to an instantaneous reduction in specific humidity for (left) SST = 25° and (right) SST = 40°. From figure 5 of Emanuel et al. [2014].

is not as high in the higher troposphere. The basic state dries: warming occurs locally but emits less radiation to the lower troposphere, which in turn receives less radiation and cools.

Figure 3 shows the net radiation anomaly for the specific humidity perturbation used in figure 2 for a range of SST. A clear qualitative transition is observed between SST values of 30°C and 35°C, corresponding to states above and below the critical SST for stability of approximately 32°C.

The drying of the basic state has consequences for stability, as drying introduces downward motion which further dries in the unstable cases. Convection is also reduced, which dries the upper troposphere and lower stratosphere due to downward flux of water vapour. It should be noted that even if convection shut down entirely in the column, the troposphere would not be expected to dry out entirely as neighbouring columns would be in RCE, leading to horizontal motions bringing in water vapour.

1.3 Two-layer model

The mechanism described above requires at least two troposphere layers to be understood, and cannot be parameterized as Newtonian relaxation. As such, we consider a two-layer troposphere model, shown in figure 4, for which RCE can be calculated and the linear stability of the resulting state can be analyzed.

The critical SST of approximately 30°C can be thought of in terms of the Clausius-Clapeyron equation: the troposphere needs to be sufficiently opaque to longwave radiation, with optical depth increasing with increasing temperature.

The basic state of this system is moist RCE with a grey atmosphere. The temperatures T_1 and T_2 are fixed by the longwave infrared emissivities ε_1 and ε_2, respectively, and the

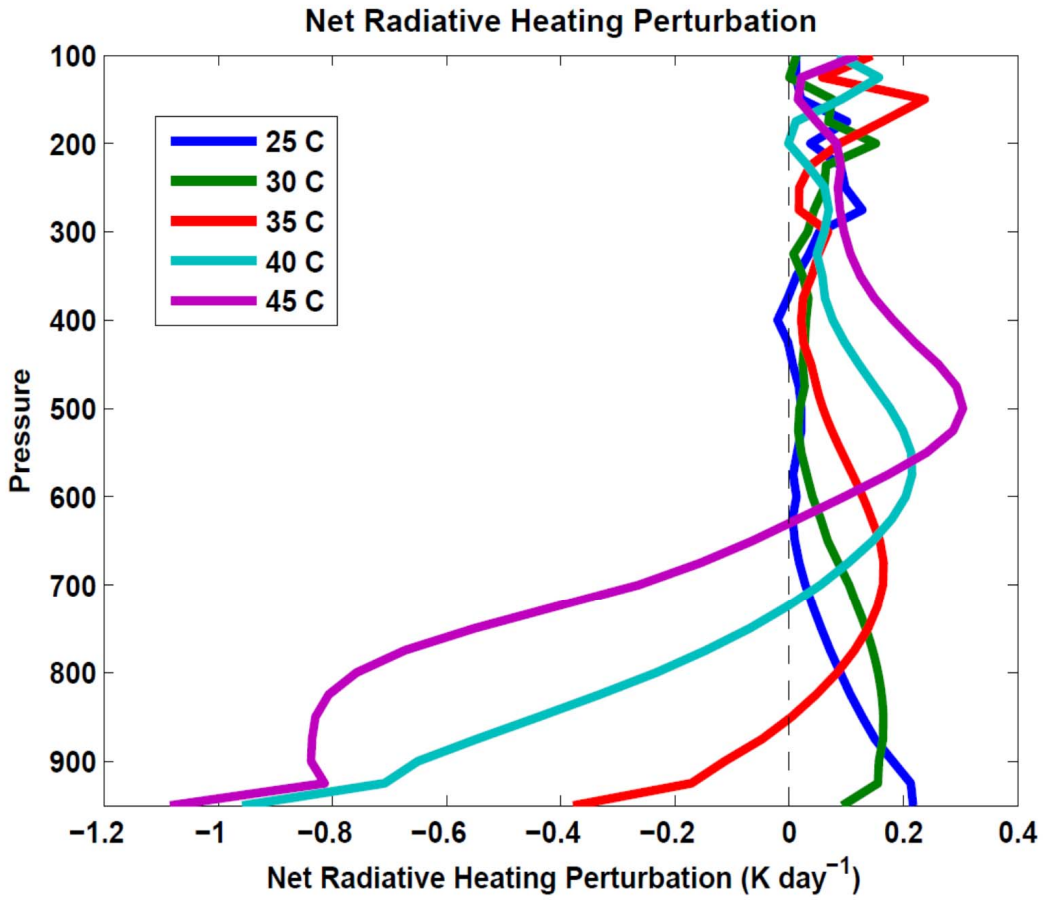


Figure 3: The response of perturbation net heating rates at several SST values to an instantaneous reduction in specific humidity. From figure 6 of Emanuel et al. [2014].

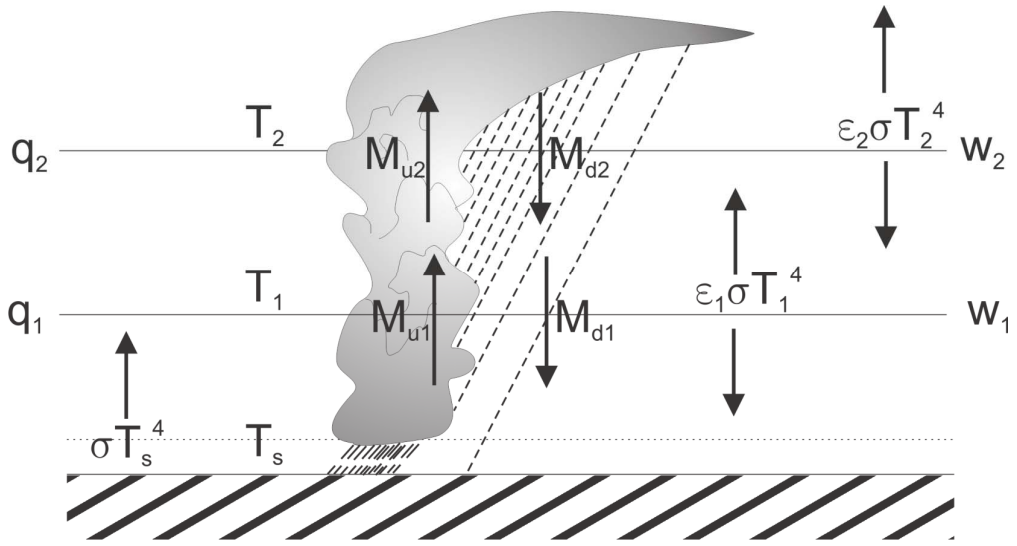


Figure 4: Schematic for the two-layer troposphere model for RCE stability. From figure 1 of Emanuel et al. [2014].

surface temperature is specified. The water vapour-dependent emissivities are controlled by q_1 and q_2, which are the only time-dependent quantities in the model. The boundary layer is held in quasi-equilibrium, with the convective mass fluxes M calculated such that the enthalpy of the boundary layer is held fixed in time, and the vertical velocities w are calculated from WTG [Emanuel et al., 2014].

As seen in the cloud-resolving model, the resulting structure is independent of any horizontal scale, and the scale of the self-aggregating patches appears to increase with the domain size. The physics do not explain the scale of the patches. One possible mechanism by which the size of the clusters may be set is related to the convective turbulence of the surroundings. Convective clouds have an order one aspect ratio and a well-defined horizontal length scale, which can be used to establish a scale for the turbulent diffusion. The vertical velocities associated with convective turbulence, though unsolved, are $\mathcal{O}(\mathrm{m/s})$, and the vertical scale of the clouds does not change much. If we let H be the scale of the clouds and v_c be the characteristic velocity, then the turbulent diffusivity scales like $v_c H$. The descent of air in a patch acts to dry the column; conversely, the turbulence is acting to remoisten the patch. For air descending in a patch of size L with velocity w, the ratio of drying to moistening can be scaled as $(wH)/(v_c L)$. So, for small holes, turbulent diffusion is relatively more influential, and making the domain large is equivalent to decreasing the effect of the turbulent diffusion on the formation of the holes. This may explain why larger domain sizes were needed to see patches in the cloud-resolving models for higher SSTs. This mechanism gives a lower bound of sorts on the required size of the patches, but does not provide an upper bound.

Based on this model, a criterion for instability is

$$\frac{\overline{\dot{Q}_1}}{\varepsilon_1}\frac{\partial \varepsilon_1}{\partial q_1} + (1-\varepsilon_p)\frac{\overline{\dot{Q}_2}}{\varepsilon_2}\frac{\partial \varepsilon_2}{\partial q_2} + \varepsilon_p \frac{S_2}{S_1}\frac{\sigma \varepsilon_1 T_2^4}{\rho_1}\frac{\partial \varepsilon_2}{\partial q_2} > 0\,, \tag{1}$$

where $\varepsilon_{1,2}$ are emissivities, $\dot{Q}$ is the radiative heating per unit mass, S_2/S_1 is the $\mathcal{O}(1)$ ratio of dry static stabilities, and ε_p is the precipitation efficiency with typical values $0.5 < \varepsilon_p < 1$. Convection produces downdrafts driven by precipitation; $\varepsilon_p = 0$ corresponds to all precipitation re-evaporating leading to no heating, while $\varepsilon_p = 0$ corresponds to no re-evaporation and no downdrafts.

In (1), the first term is negative but small when atmospheric moisture is low. The second term is negative but not necessarily small. It is only the third term which is positive, and instability only occurs if this term is sufficiently large. The resulting instability is radiatively driven due to the dependence of the emissivities $\varepsilon_{1,2}$ on water vapour and convection: RCE becomes linearly unstable if the infrared opacity of the lower troposphere is sufficiently large and if ε_p is large.

We can consider ordinary RCE, as seen in the top row of figure 5 and perturb locally with downward vertical velocity. For low SST, i.e. when the system is stable, the shortwave radiative heating is not strongly affected, while longwave radiative cooling is reduced throughout the column. The convective heating is also somewhat reduced. The net result is positive perturbation heating, leading to large scale ascent through WTG and negative feedback on the initial downward perturbation, as shown in the middle row of figure 5.

However, for high SST (shown in the bottom row of figure 5), the downward vertical velocity perturbation leads to strong negative perturbations in the shortwave heating. Longwave radiative cooling is reduced in the upper troposphere, rather than throughout the column, as the change in optical depth in the lower troposphere due to the velocity perturbation is very small. The longwave cooling of the lower troposphere is increased, and convective heating is decreased. The net result is thus negative perturbation heating which, as a consequence of WTG, induces large scale descent, and as such acts as a positive feedback on the initial downward perturbation.

When the resulting instability becomes nonlinear, cloud effects on radiation begin to take over. These effects are dominant once a cluster has formed, as the central dense overcast causes an intense anomaly in outgoing longwave radiation. However, it should be emphasized that cloud feedbacks are not important for instigating instability, but maintain the clusters when already formed; there is a strong hysteresis in the RCE system. Figure 6 shows the hypothesized subcritical bifurcation for the system, in which a "clustering metric" (vertical velocity w on the cluster scale) is used.

In preliminary attempts at a nonlinear two-layer model of RCE, updraft mass flux does show evidence of aggregation. However, it is unclear as to whether the aggregated state is a unique attractor of sorts for RCE.

Given the idealized nature of the two-layer model considered here, it is natural to ask how relevant the results obtained are. It can be seen that in a model of deep convective self-aggregation above uniform SST [Bretherton et al., 2005], aggregation dramatically dries the atmosphere (in the sense of the whole domain average), as seen in figure 7. As a result, the greenhouse effect is reduced and SST would be expected to drop, though disaggregation may not occur due to the hysteresis of the system. In considering several datasets (including

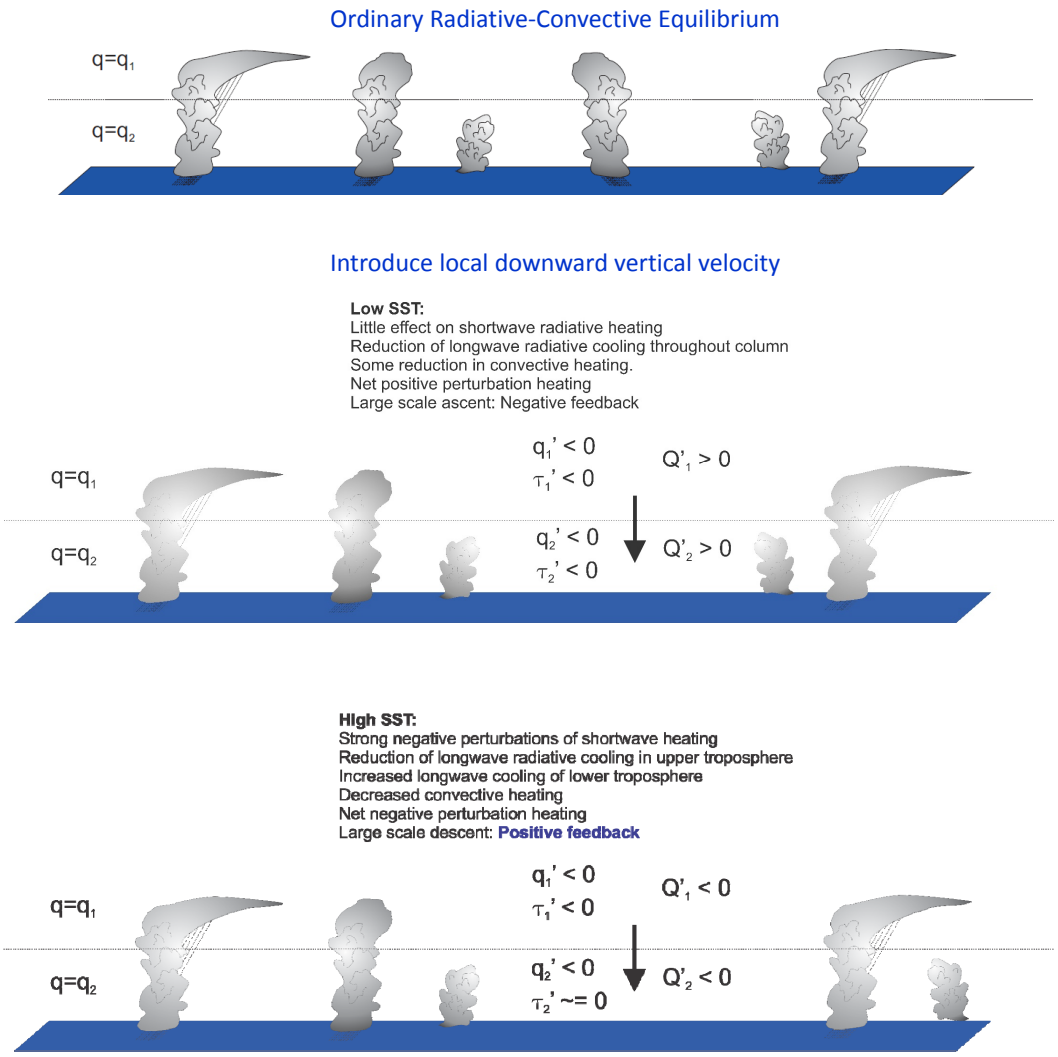


Figure 5: Response of the two-layer RCE model (top) to a local downward perturbation in vertical velocity for (middle) low SST and (bottom) high SST. (Original artwork by Kerry Emanuel.)

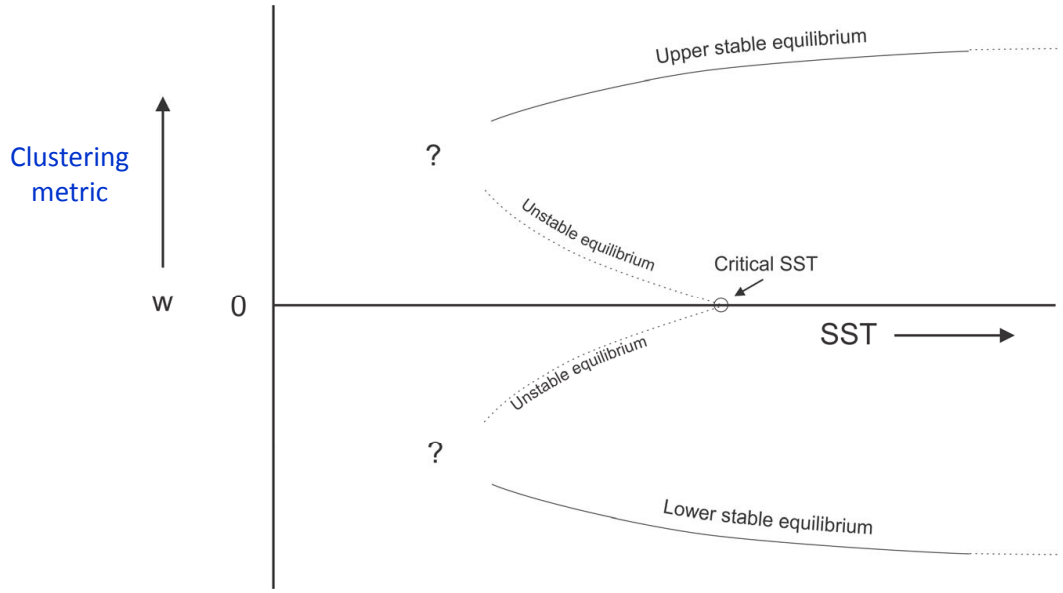


Figure 6: Hypothesized subcritical bifurcation. From figure 7 of Emanuel et al. [2014].

satellite observations and reanalysis data), Tobin et al. [2012] defined an empirical "Simple Convective Aggregation Index" (SCAI) based on the number of convective clusters and the distances between clusters. They found that the atmosphere was drier in locations where convection was more clustered.

One conjecture for the self-organization of RCE, shown in figure 8, is as follows. In cases of high SST, convection self-aggregates. However, this causes the horizontally-averaged humidity to drop dramatically, which in turn leads to a decreased greenhouse effect and cooling of the system, causing disaggregation of the convection. Thus, we can hypothesize that the system wants to be near the phase transition to the aggregated state, i.e. with SST near the critical temperature for self-aggregation.

Based on this hypothesis, we can consider a situation of "self-organized criticality", first proposed by David Neelin for a different mechanism. This proposal says that the system should reside near the critical temperature for self-aggregating, thus regulating tropical SST. In addition, the convective cluster size should follow a power-law distribution.

Several questions remain which could be asked about the two-layer RCE model:

1. How good is the assumption of convective equilibrium?
   There is a substantial amount of debate regarding this point, and the answer is, at present, unknown. However, there is some suggestion that the assumption is more justified on the macro scale, i.e. for longer scales in space and time.

2. If self-aggregation occurs on large scales, is there a breakdown of the range over which RCE exists? Might this be related to the behaviour of the Madden-Julian Oscillation?

3. Is self-aggregation seen in climate models?

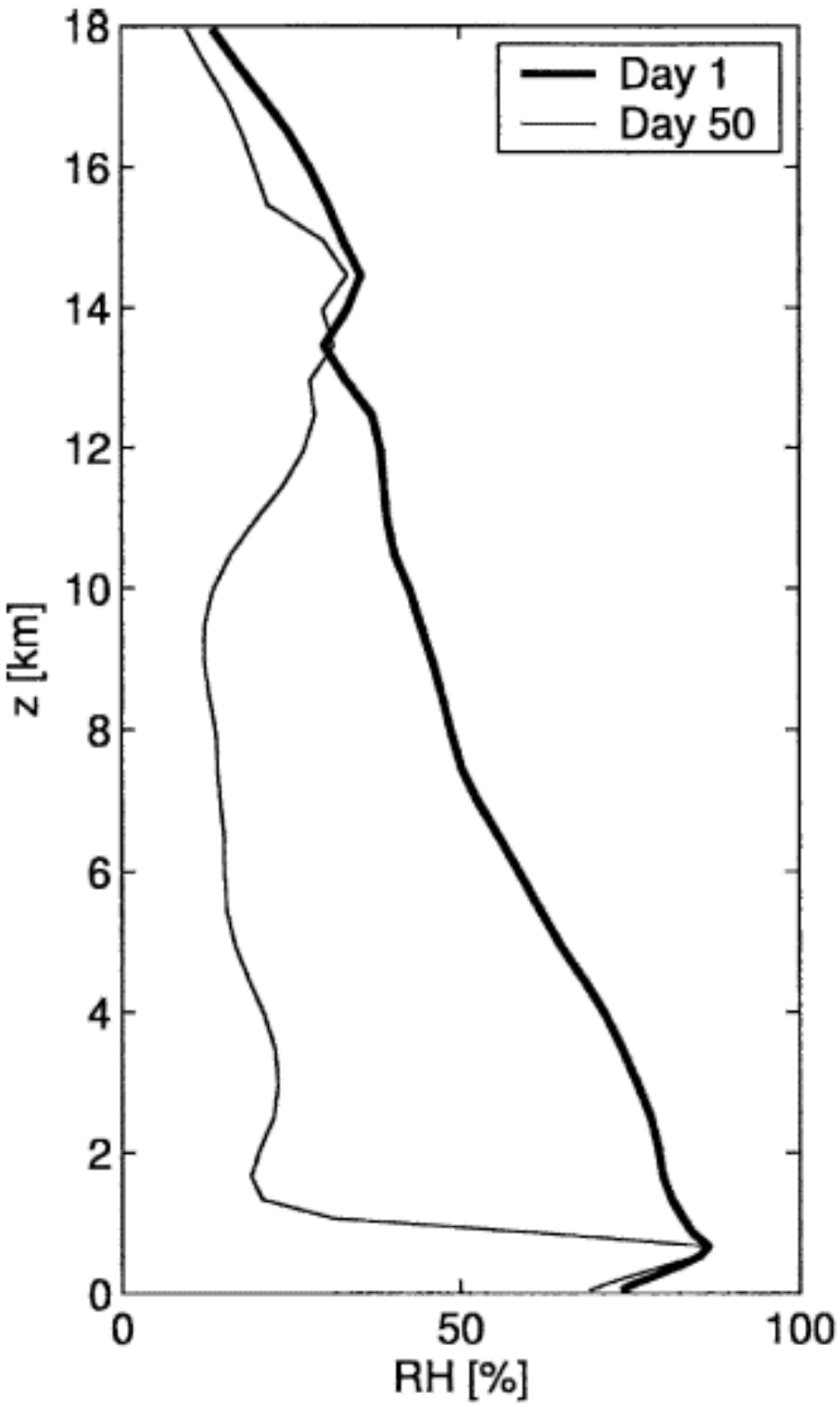


Figure 7: Horizontally averaged profile of relative humidity on day 1 and day 50 of a cloud-resolving model with self-aggregating convection. From figure 4 of Bretherton et al. [2005].

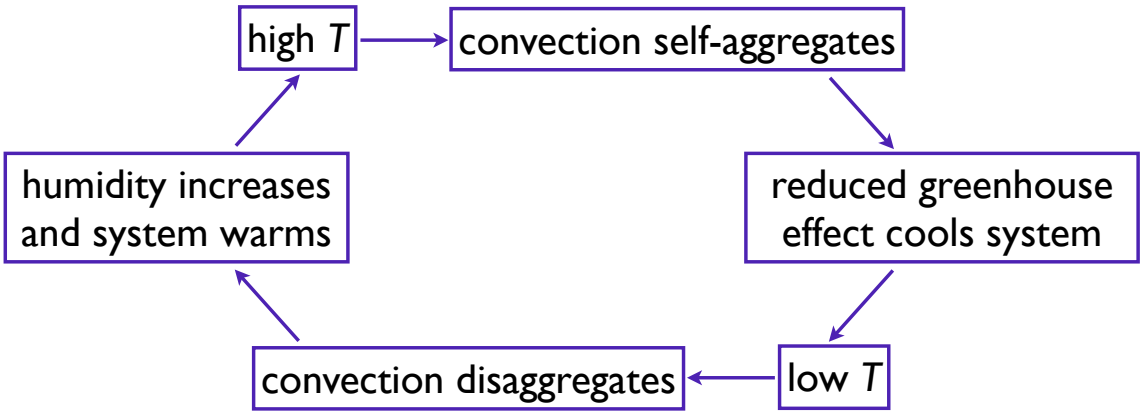


Figure 8: Hypothesis for self-organization of RCE.

Probably not, or not accurately, as there is no explicit coupling of convective downdrafts to surface fluxes, and so the necessary feedbacks are missing. However, the physics required to *maintain* clusters once initialized are there in global climate models – what is lacking are the physics need for spontaneous formation of clusters.

2 Hurricanes

In further considering the self-aggregation problem discussed above, it is natural to ask what the result would be if rotation were included. As figure 9 shows, for a 3000 km 3D box with vertically-integrated water vapour, the self-aggregated clusters appear to have become hurricanes. In figure 10, where the rotation rate has been increased, multiple smaller hurricanes are seen in the box, thereby indicating that the hurricanes have an associated size and spacing, unlike the non-rotating clusters which scaled with domain size. However, the transition temperature from the non-aggregated to aggregated states is not strongly affected by the addition of rotation.

The energy production associated with an ideal hurricane is shown in figure 11. The energetics are comparable to a Carnot cycle, in which maximum efficiency results from the particular cycle of isothermal expansion, adiabatic expansion, isothermal compression, and adiabatic compression. This is similar to what occurs in a hurricane, in which air rises due to excess enthalpy, expands, and descends. It should be noted that the last leg, however, is not adiabatic in a hurricane – while air cools radiatively, the environmental temperature profile is moist adiabatic and so the amount of cooling is equivalent to that of saturated air descending moist adiabatically. This is also shown for a real hurricane (Hurricane Inez) in figure 12. The maximum rate of energy production is

$$P = \frac{T_s - T_0}{T_s}\dot{Q} \tag{2}$$

where T_s and T_0 are the surface and outflow temperatures, respectively, and $\dot{Q}$ is the rate of heat input. The hurricane has a high-entropy core which takes heat out of the ocean. Based

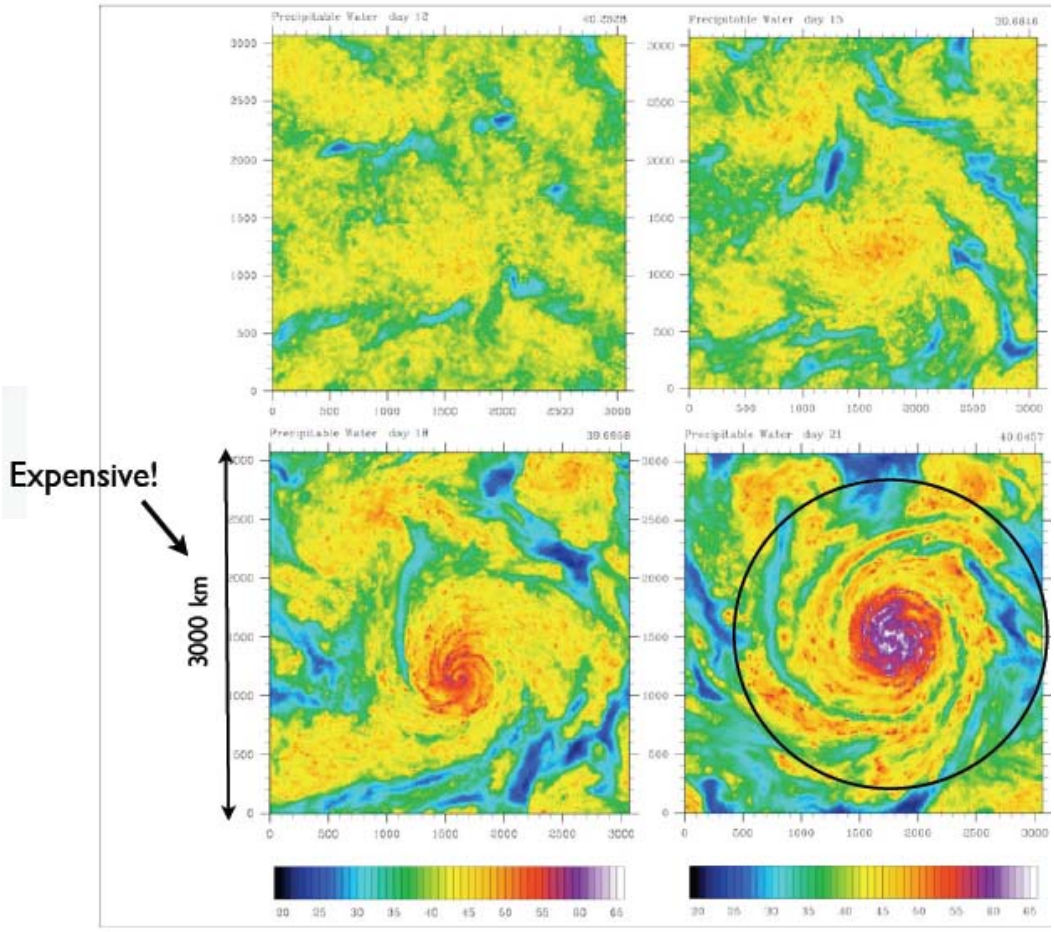


Figure 9: Several snapshots of precipitable water (days 12, 15, 18, and 21) for rotating RCE. The formation of a hurricane is indicated in the bottom right panel.

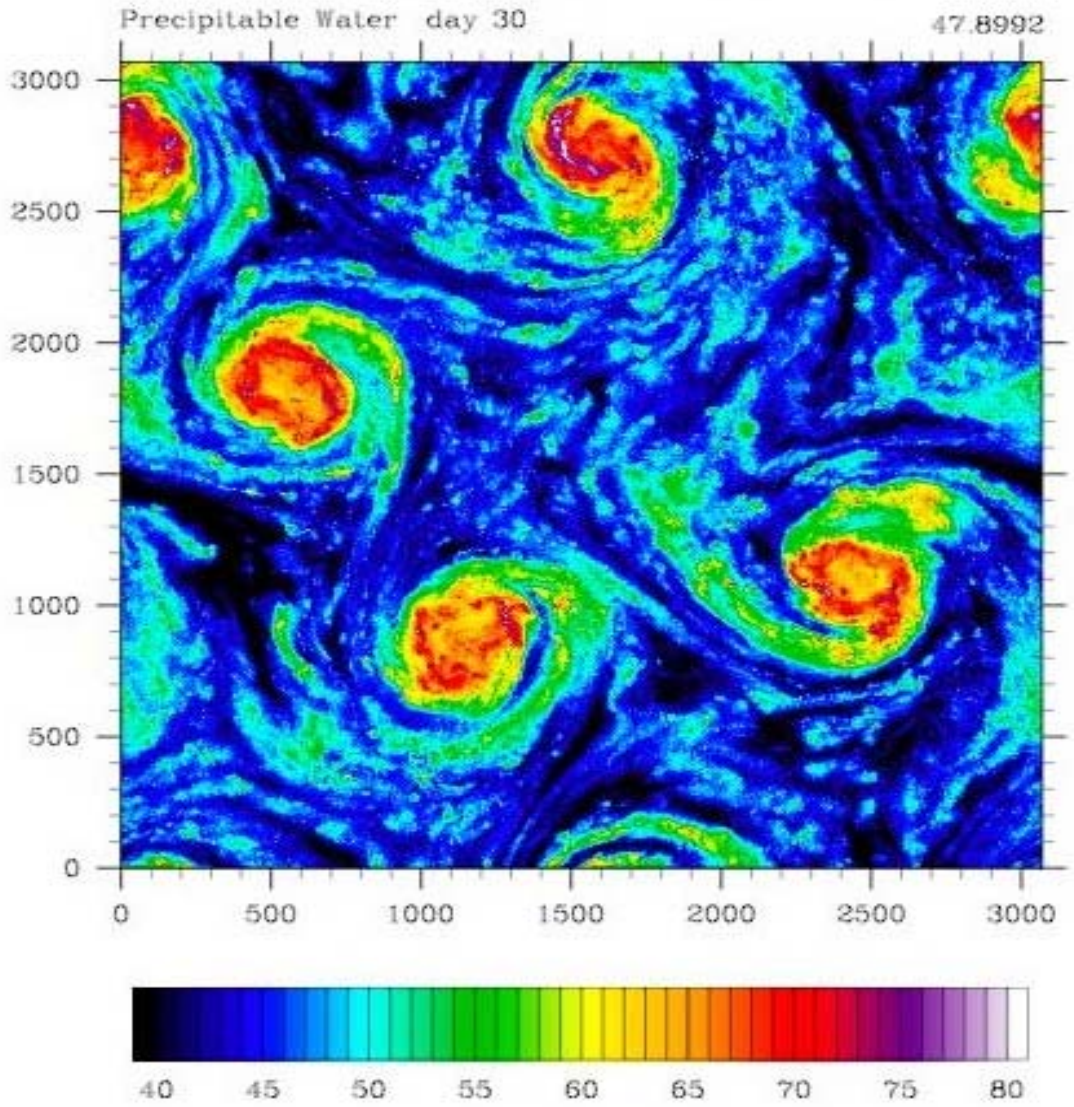


Figure 10: Snapshot of precipitable water (day 30) for rotating RCE. Here, the rotation rate f has been increased from that of figure 9.

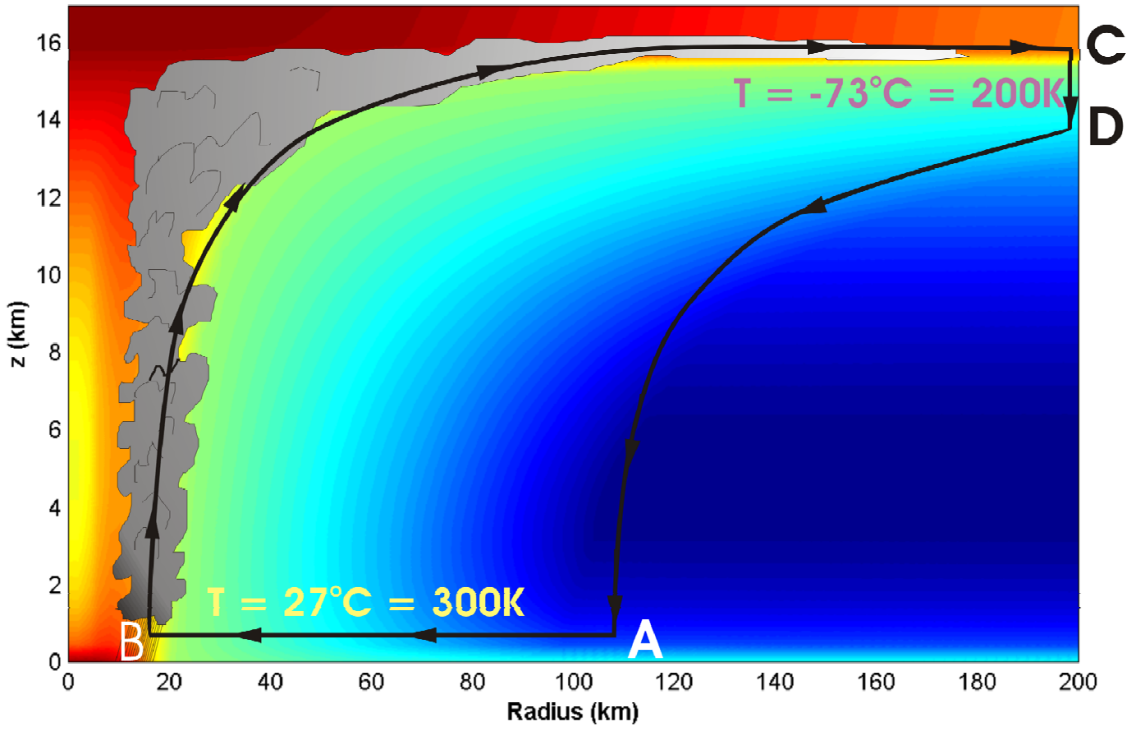


Figure 11: Energy production associated with an idealized hurricane. Red (blue) denotes regions of high (low) entropy.

on the energy cycle in figure 11, the maximum steady intensity that the storm can achieve, also known as the "potential intensity", can be computed and related to a maximum wind speed [Emanuel, 1999].

The theoretical maximum wind speed for the hurricanes can be derived as

$$|V_{pot}|^2 \simeq \frac{C_k}{C_D}\frac{T_s - T_0}{T_0}(k_0^* - k) \tag{3}$$

where C_k/C_D is the ratio of exchange coefficients of enthalpy and momentum and $k_0^* - k$ is the air-sea enthalpy disequilibrium. This air-sea enthalpy disequilibrium is the driver of hurricanes, and occurs due to the presence of greenhouse gases in the atmosphere. Using the maximum wind speed, V_{pot}, a scaling for rotating RCE can be derived.

Using the above quantities, as well as the modified thermodynamic efficiency $\varepsilon \equiv (T_s - T_0)/T_0$, angular velocity of the Earth's rotation Ω, saturation concentration of water vapour at the sea surface $q_s \sim \mathrm{e}^{T_s}$, and net upward radiative fluxes at the surface and the top of the atmosphere, F_s and F_{TOA} respectively, scalings for the hurricanes can be derived. These include the potential intensity, $V_p^3 \approx \varepsilon(F_{TOA} - F_s)/C_D$, the radius at which maximum winds occur , $r_m \sim V_{pot}/\Omega$, the distance between storm centres, $D \sim \sqrt{L_v q_s}/\Omega$ (alternately, the deformation radius), and the number density (i.e. number of storms per unit area), $n \sim \Omega^2/(L_v q_s)$.

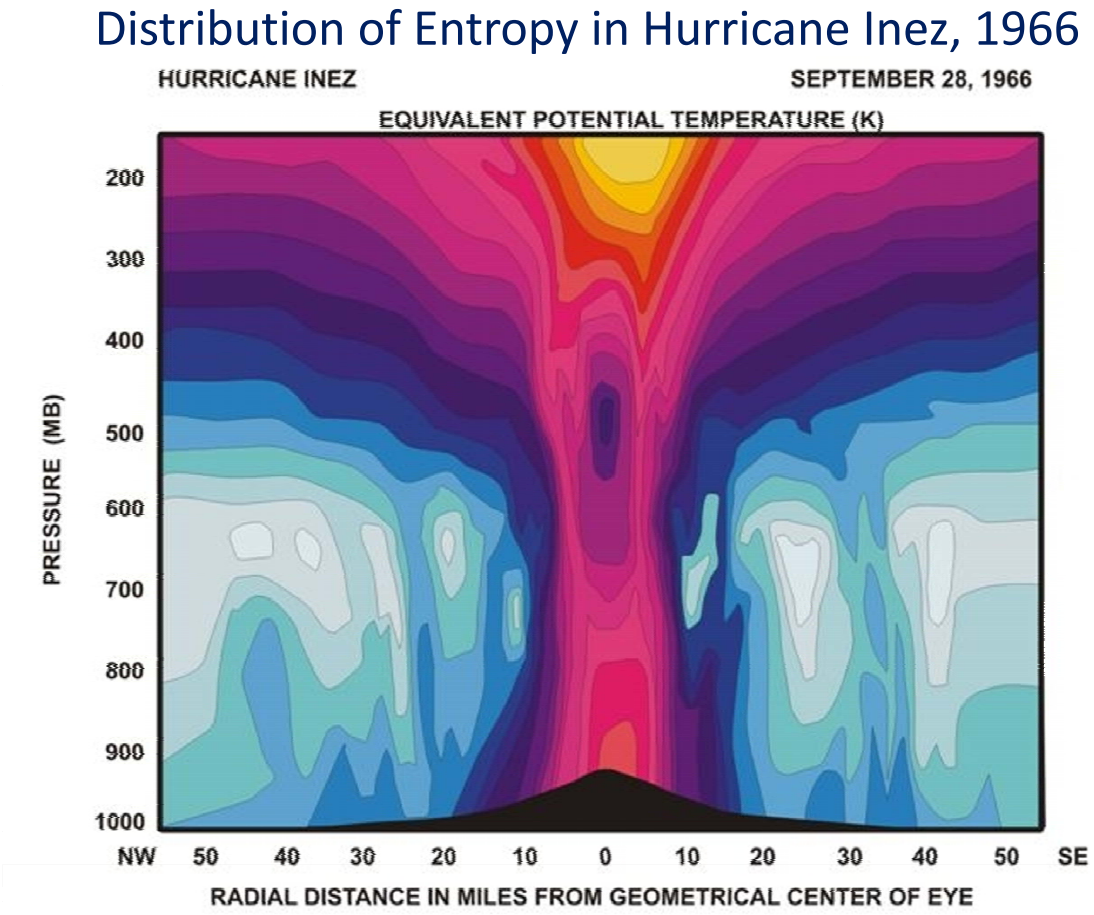


Figure 12: Equivalent potential temperature in Hurricane Inez. Graphic by Kerry Emanuel, based off of figure 6 of Hawkins and Imbembo [1976].

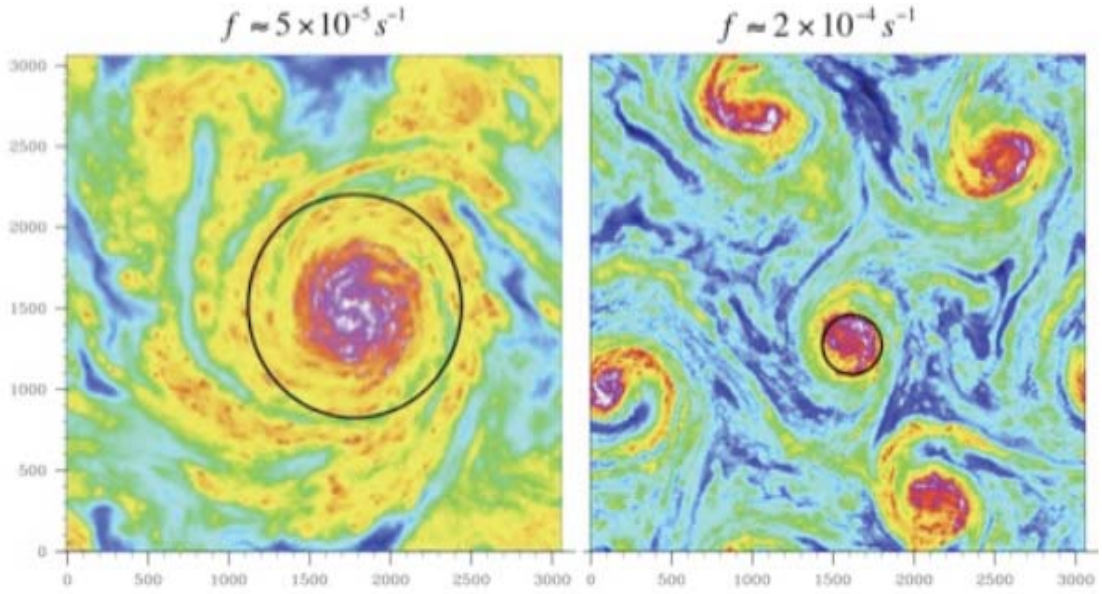


Figure 13: Snapshots of precipitable water in rotating RCE with $f \approx 5 \times 10^{-5}\,\mathrm{s}^{-1}$ (left) and $f \approx 2 \times 10^{-4}\,\mathrm{s}^{-1}$ (right). Approximate sizes from scaling analysis are indicated by the black circles. (From figure 1 of Khairoutdinov and Emanuel [2013]

Simulations agree well with this scaling. For example, figure 13 shows the result of increasing the rotation speed with the eye of storm deformation radius overlaid. The decreasing radii, increasing number density, and decreasing distance between storm centres all agree well with the predictions of the hurricane world scaling. Increasing the SST is seen to lead to fewer, but more intense, events, as shown in figure 14.

3 RCE of an Earth-like aquaplanet

The question remains of how to relate RCE with large-scale dynamics. To answer this, we can consider a hypothetical Earth-like planet without continents (i.e. an aquaplanet) or seasons, for which the only friction that acts on the atmosphere is at the surface itself. As the RCE state depends only on the latitude, we can imagine calculating the RCE state at each latitude of the planet (bearing in mind that this would lead to difficulties at the poles, for which this problem is not defined owing to the lack of radiation). While there are some rather substantial differences between Earth and the planet described here, we can compare between the real world and the hypothetical planet, asking "why do we not observe the same balance in reality?"

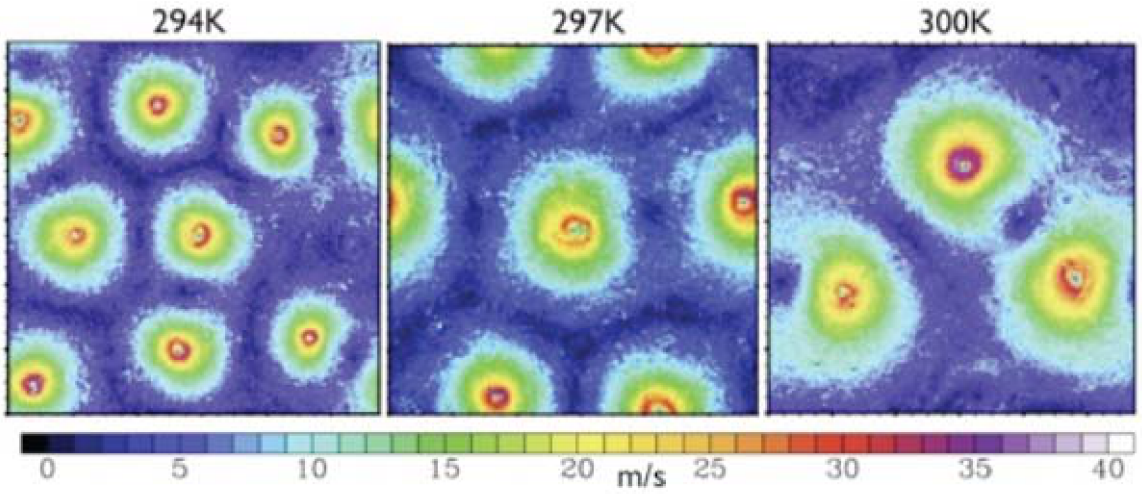


Figure 14: Snapshots of near-surface wind (m/s) in rotating RCE for three different values of SST: 294 K (left), 297 K (centre), and 300 K (right).

The primary advantage of the setup described here is that there is an *exact* nonlinear equilibrium solution for the atmospheric flow. This solution is characterized by

1. every column of the atmosphere, as well as the surface beneath it, is in RCE
2. the wind vanishes at the surface of the planet
3. horizontal pressure gradients are balanced by Coriolis accelerations
4. surface pressure is constant

The pressure above the surface decreases poleward. This decrease in pressure occurs more rapidly at higher altitudes. The pressure gradient is in geostrophic balance with a west wind, and the wind experiences a vertical shear due to the horizontal temperature gradient from incoming radiation, i.e. the system is in thermal wind balance, as shown in figure 15. The atmosphere-ocean heat flux from the equator to the poles means that there is a heat balance at the top of the atmosphere.

However, this setup does have the potential for some difficulty. For instance, it is possible that the planet does not have sufficient angular momentum to support the west-east wind required by the thermal wind balance. Also, it is possible that the equilibrium solution may be unstable. Instabilities such as baroclinic instability in mid- to high-latitudes may lead to the formation of eddies, which would in turn cause radiative imbalance, thus moving away from the thermal wind balance previously computed.

References

C. S. Bretherton, P. N. Blossey, and M. Khairoutdinov. An energy-balance analysis of deep convective self-aggregation above uniform SST. *J. Atmos. Sci.*, 62:4273–4292, 2005.

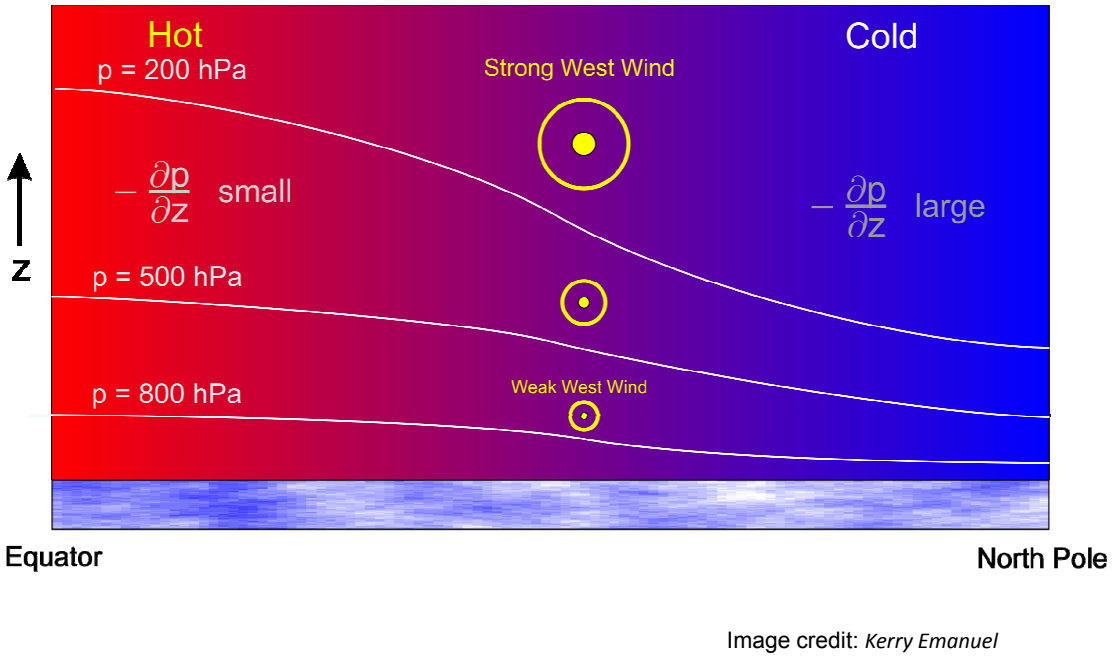


Figure 15: Thermal wind balance for an Earth-like planet without continents or seasons. Original artwork by Kerry Emanuel.

K. Emanuel, A. A. Wing, and E. M. Vincent. Radiative-convective instability. *J. Adv. Model. Earth Syst.*, 6:75–90, 2014.

K. A. Emanuel. Thermodynamic control of hurricane intensity. *Nature*, 401:665–669, 1999.

H. F. Hawkins and S. M. Imbembo. The structure of a small, intense hurricane – Inez 1966. *Mon. Weather Rev.*, 104:418–442, 1976.

M. Khairoutdinov and K. Emanuel. Rotating radiative-convective equilibrium simulated by a cloud-resolving model. *J. Adv. Model. Earth Syst.*, 5:816–825, 2013.

J. Nilsson and K. A. Emanuel. Equilibrium atmospheres of a two-column radiative-convective model. *Q. J. R. Meteorol. Soc.*, 125:2239–2264, 1999.

D. J. Raymond and X. Zeng. Instability and large-scale circulations in a two-column model of the tropical troposphere. *Q. J. R. Meteorol. Soc.*, 124:3117–3135, 2000.

A. H. Sobel and C. S. Bretherton. Modeling tropical precipitation in a single column. *J. Clim.*, 13:4378–4392, 2000.

I. Tobin, S. Bony, and R. Roca. Observational evidence for relationships between the degree of aggregation of deep convection, water vapor, surface fluxes, and radiation. *J. Clim.*, 25:6885–6904, 2012.

Lecture 8: Horizontal Convection

Geoff Vallis; notes by Daniel Mukiibi and Paige Martin

June 19

1 The Ocean: Constraints on a Thermal Circulation

We now turn our attention to the ocean. Our goal in this lecture and the next is to describe a theory for the deep circulation of the ocean, sometimes called the meridional overturning circulation (MOC) and occasionally the thermohaline circulation. We begin in this lecture by showing that there have to be winds or some other form of mechanical forcing in order to drive a substantial deep ocean circulation. The root effect goes back to Sandström, and although his rigour was suspect it seems his intuition was right.

2 Sandström's Effect

We first give an argument that is similar in spirit to the one that Sandström gave in his original papers (Sandström, 1908, 1916).

2.1 Maintaining a steady baroclinic circulation

The Boussinesq equations are

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\nabla\phi + b\mathbf{k} + \mathbf{F}, \qquad \frac{\mathrm{D}b}{\mathrm{D}t} = \dot{Q}, \qquad \div\mathbf{v} \ = 0, \tag{1a,b,c}$$

where F represents frictional terms and $\dot{Q} = J + \kappa\nabla^2 b$ (that is, the heating term here includes the effects of diffusion). The circulation, C, changes according to

$$\begin{aligned}\frac{\mathrm{D}C}{\mathrm{D}t} = \frac{\mathrm{D}}{\mathrm{D}t}\oint \mathbf{v} \ \cdot \mathrm{d}\mathbf{r} &= \oint \left(\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t}\cdot \mathrm{d}\mathbf{r} + \mathbf{v} \ \cdot \mathrm{d}\mathbf{v}\right) \\ &= \oint b\mathbf{k}\cdot\mathrm{d}\mathbf{r} + \oint \mathbf{F}\cdot\mathrm{d}\mathbf{r},\end{aligned} \tag{2}$$

(Note that rotation does no work.) Furthermore, we can write rate of change of circulation itself as

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \oint \left(\frac{\partial\mathbf{v}}{\partial t} + \mathbf{v} \ \cdot \nabla\mathbf{v}\right)\cdot\mathrm{d}\mathbf{r} = \oint \left(\frac{\partial\mathbf{v}}{\partial t} + \omega\times\mathbf{v}\right)\cdot\mathrm{d}\mathbf{r}. \tag{3}$$

Let us assume the flow is steady, so that $\partial\mathbf{v}\,/\partial t$ vanishes. Let us further choose the path of integration to be a streamline, which, since the flow is steady, is also a parcel trajectory.

The second term on the right-most expression of (3) then also vanishes and (2) becomes

$$\oint b\,\mathrm{d}z = -\oint \mathbf{F}\cdot\mathrm{d}\mathbf{r} = -\oint \frac{\mathbf{F}}{|\mathbf{v}\,|}\cdot\mathbf{v}\;\mathrm{d}r, \tag{4}$$

where the last equality follows because the path is everywhere parallel to the velocity. Let us now assume that the friction retards the flow, and that $\oint \mathbf{F}\cdot\mathbf{v}\,/|\mathbf{v}\,|\,\mathrm{d}r < 0$. (One form of friction that has this property is linear drag, $\mathbf{F} = -C\mathbf{v}$ where C is a constant. The property is similar to, but not the same as, the property that the friction dissipates kinetic energy over the circuit.) Making this assumption, if we integrate the term on the left-hand side by parts we obtain

$$\oint z\,\mathrm{d}b < 0. \tag{5}$$

Now, because the integration circuit in (6) is a fluid trajectory, the change in buoyancy $\mathrm{d}b$ is proportional to the heating of a fluid element as it travels the circuit $\mathrm{d}b = \dot{Q}\,\mathrm{d}t = \mathrm{d}Q$, where the heating includes diffusive effects.

$$\oint z\,\mathrm{d}Q < 0. \tag{6}$$

Thus, the inequality implies that the net heating must be negatively correlated with height: that is, *the heating must occur, on average, at a lower level than the cooling in order that a steady circulation may be maintained against the retarding effects of friction.*

A compressible fluid

A similar result can be obtained for a compressible fluid. We write the baroclinic circulation theorem as

$$\frac{\mathrm{D}C}{\mathrm{D}t} = \oint p\,\mathrm{d}\alpha + \oint \mathbf{F}\cdot\mathrm{d}\mathbf{r} = \oint T\,\mathrm{d}\eta + \oint \mathbf{F}\cdot\mathrm{d}\mathbf{r}, \tag{7}$$

where η is the specific entropy. Then, by precisely the same arguments as led to (6), we are led to the requirements that

$$\oint T\,\mathrm{d}\eta > 0 \qquad \text{or equivalently} \qquad \oint p\,\mathrm{d}\alpha > 0. \tag{8a,b}$$

Equation (8a) means that parcels must gain entropy at high temperatures and lose entropy at low temperatures; similarly, from (6b), a parcel must expand ($\mathrm{d}\alpha > 0$) at high pressures and contract at low pressures.

For an ideal gas we can put these statements into a form analogous to (6) by noting that $\mathrm{d}\eta = c_p(\mathrm{d}\theta/\theta)$, where θ is potential temperature, and using the definition of potential temperature for an ideal gas. With these we have

$$\oint T\,\mathrm{d}\eta = \oint c_p\frac{T}{\theta}\,\mathrm{d}\theta = \oint c_p\left(\frac{p}{p_R}\right)^{\kappa}\mathrm{d}\theta, \tag{9}$$

and (8a) becomes

$$\oint c_p\left(\frac{p}{p_R}\right)^{\kappa}\mathrm{d}\theta > 0. \tag{10}$$

Because the path of integration is a fluid trajectory, $\mathrm{d}\theta$ is proportional to the heating of a fluid element. Thus [and analogous to the Boussinesq result (6)], (10) implies that *the heating (the potential temperature increase) must occur at a higher pressure than the cooling in order that a steady circulation may be maintained against the retarding effects of friction.*

These results may be understood by noting that the heating must occur at a higher pressure than the cooling in order that work may be done, the work being necessary to convert potential energy into kinetic energy to maintain a circulation against friction. More informally, if the heating is below the cooling, then the heated fluid will expand and become buoyant and rise, and a steady circulation between heat source and heat sink can readily be imagined. On the other hand, if the heating is above the cooling there is no obvious pathway between source and sink.

2.2 A rigorous result

Following Paparella & Young (2002), we now show more rigorously that, if the diffusivity is small, the circulation is in a certain sense weak. Now including molecular viscosity and diffusivity, the equations of motion are

$$\frac{\partial \mathbf{v}}{\partial t} + (\mathbf{f} + 2\omega) \times \mathbf{v} \ = -\nabla B + b\mathbf{k} + \nu \nabla^2 \mathbf{v} \ , \tag{11a}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} = \frac{\partial b}{\partial t} + \div(b\mathbf{v} \ ) = \dot{Q} = J + \kappa \nabla^2 b, \tag{11b}$$

$$\div \mathbf{v} \ = 0, \tag{11c}$$

Multiply the momentum equation by $\mathbf{v}$ and integrate over a volume to give

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \frac{1}{2} \mathbf{v}^{\ 2} \right\rangle = \langle wb \rangle - \varepsilon, \tag{12}$$

where angle brackets denote a volume average and

$$\varepsilon = -\nu \left\langle \mathbf{v} \ \cdot \nabla^2 \mathbf{v} \ \right\rangle = \nu \left\langle \omega^2 \right\rangle , \tag{13}$$

after integrating by parts. The dissipation, ε, is a positive definite quantity.

Write the buoyancy equation as

$$\frac{\mathrm{D}bz}{\mathrm{D}t} = z\frac{\mathrm{D}b}{\mathrm{D}t} + b\frac{\mathrm{D}z}{\mathrm{D}t} = z\dot{Q} + bw, \tag{14}$$

whence

$$\frac{\mathrm{d}}{\mathrm{d}t} \langle bz \rangle = \left\langle z\dot{Q} \right\rangle + \langle bw \rangle . \tag{15}$$

Subtracting (15) from (12) gives the energy equation

$$\frac{\mathrm{d}}{\mathrm{d}t} \left\langle \frac{1}{2} \mathbf{v}^{\ 2} - bz \right\rangle = -\left\langle z\dot{Q} \right\rangle - \varepsilon. \tag{16}$$

In a steady state:

$$\left\langle z\dot{Q} \right\rangle = -\varepsilon < 0. \tag{17}$$

This is an analogue of our earlier results. It says that if we want to have a dissipative, statistically steady flow there has to be a negative correlation between heating and z. Put simply, the heating has to be below the cooling. But note that the heating and cooling include the diffusive terms.

With diffusion only

Take $\dot{Q} = \kappa \nabla^2 b$ whence (17) becomes

$$\frac{\kappa}{V} \int z \nabla^2 b \, \mathrm{d}V = -\varepsilon \tag{18}$$

The horizontal part of the integral vanishes so that

$$\frac{\kappa}{H} \int_{-H}^{0} z \frac{\partial^2 \overline{b}}{\partial z^2} \, \mathrm{d}z = -\varepsilon. \tag{19}$$

where an overbar is a horizontal average. Integrating the LHS by parts gives

$$\frac{\kappa}{H} \left[\overline{b}(0) - \overline{b}(-H) \right] = \varepsilon. \tag{20}$$

The LHS is bounded by the surface buoyancy gradient, so the KE dissipation goes to zero as $\kappa \to 0$.

The result is (at least from a physicist's point of view) quite rigorous. It can also be extended to a nonlinear equation of state (Nycander, 2010). It tells us that the dissipation of kinetic energy in a fluid diminishes with the diffusivity, and that if $\kappa = 0$ then dissipation vanishes. It doesn't say there is no flow at all, but it is hard to envision a flow in a finite domain that does not dissipate kinetic energy. The result is often characterized as saying that the flow is *non turbulent.*

3 Buoyancy and Mixing Driven Scaling Theories

Now we talk about scaling, becoming a bit less rigorous. Interestingly the scaling, dating from Rossby (1965), predates the rigorous theories, and it also provides much stronger bounds. However, it is a scaling and not a rigorous result and therefore open to dispute.

3.1 Equations of motion

A non-rotating Boussinesq fluid heated and cooled from above obeys the equations.

$$\frac{\mathrm{D}\mathbf{v}}{\mathrm{D}t} = -\nabla \phi + \nu \nabla^2 \mathbf{v} \; + b \, \mathbf{k}, \tag{21}$$

$$\frac{\mathrm{D}b}{\mathrm{D}t} = \kappa \nabla^2 b \tag{22}$$

$$\nabla \cdot \mathbf{v} \; = 0. \tag{23}$$

with boundary conditions

$$b(x, y, 0, t) = g(x, y), \tag{24}$$

For algebraic simplicity consider the two-dimensional version of these equations, in y and z . We can define a streamfunction

$$v = -\frac{\partial \psi}{\partial z}, \quad w = \frac{\partial \psi}{\partial y}, \qquad \zeta = \nabla_x^2 \psi = \left(\frac{\partial^2 \psi}{\partial y^2} + \frac{\partial^2 \psi}{\partial z^2} \right), \tag{25}$$

Taking the curl of the momentum equation gives

$$\frac{\partial \nabla^2 \psi}{\partial t} + J(\psi, \nabla^2 \psi) = \frac{\partial b}{\partial y} + \nu \nabla^4 \psi \tag{26a}$$

$$\frac{\partial b}{\partial t} + J(\psi, b) = \kappa \nabla^2 b \tag{26b}$$

where $J(a, b) \equiv (\partial_y a)(\partial_z b) - (\partial_z a)(\partial_y b)$.

Non-dimensionalization and scaling

We non-dimensionalize (26) by formally setting

$$b = \Delta b \, \widehat{b}, \qquad \psi = \Psi \widehat{\psi}, \qquad y = L\widehat{y}, \qquad z = H\widehat{z}, \qquad t = \frac{LH}{\Psi}\widehat{t}, \tag{27}$$

where the hatted variables are non-dimensional, Δb is the temperature difference across the surface, L is the horizontal size of the domain, and Ψ, and ultimately the vertical scale H, are to be determined. Substituting (27) into (26) gives

$$\frac{\partial \widehat{\nabla}^2 \widehat{\psi}}{\partial \widehat{t}} + \widehat{J}(\widehat{\psi}, \nabla^2 \widehat{\psi}) = \frac{H^3 \Delta b}{\Psi^2} \frac{\partial \widehat{b}}{\partial \widehat{y}} + \frac{\nu L}{\Psi H} \widehat{\nabla}^4 \widehat{\psi}, \tag{28a}$$

$$\frac{\partial \widehat{b}}{\partial \widehat{t}} + \widehat{J}(\widehat{\psi}, \widehat{b}) = \frac{\kappa L}{\Psi H} \widehat{\nabla}^2 \widehat{b}, \tag{28b}$$

where $\widehat{\nabla}^2 = (H/L)^2 \partial^2/\partial \widehat{y}^2 + \partial^2/\partial \widehat{z}^2$ and the Jacobian operator is similarly non-dimensional. If we now use (28b) to choose Ψ as

$$\Psi = \frac{\kappa L}{H}, \tag{29}$$

so that $t = H^2 \widehat{t}/\kappa$, then (28) becomes

$$\frac{\partial \widehat{\nabla}^2 \widehat{\psi}}{\partial \widehat{t}} + \widehat{J}(\widehat{\psi}, \widehat{\nabla}^2 \widehat{\psi}) = Ra \; \sigma \alpha^5 \frac{\partial \widehat{b}}{\partial \widehat{y}} + \sigma \widehat{\nabla}^4 \widehat{\psi}, \tag{30}$$

$$\frac{\partial \widehat{b}}{\partial \widehat{t}} + \widehat{J}(\widehat{\psi}, \widehat{b}) = \widehat{\nabla}^2 \widehat{b}, \tag{31}$$

and the non-dimensional parameters that govern the behaviour of the system are

$$Ra = \left(\frac{\Delta b L^3}{\nu \kappa} \right), \qquad \text{(the Rayleigh number)}, \tag{32a}$$

$$\sigma = \frac{\nu}{\kappa}, \qquad \text{(the Prandtl number)}, \tag{32b}$$

$$\alpha = \frac{H}{L}, \qquad \text{(the aspect ratio)}. \tag{32c}$$

3.2 Rossby's Scaling

For steady non-turbulent flows, and also perhaps for statistically steady flows, then we can demand that the buoyancy term in (30) is $\mathcal{O}(1)$. If it is smaller then the flow is not buoyancy driven, and if it is larger there is nothing to balance it. Our demand can be satisfied only if the vertical scale of the motion adjusts appropriately and, for $\sigma = \mathcal{O}(1)$, this suggests the scalings

$$H = L\sigma^{-1/5}Ra^{-1/5} = \left(\frac{\kappa^2 L^2}{\Delta b}\right)^{1/5}, \qquad \Psi = Ra^{1/5}\sigma^{-4/5}\nu = (\kappa^3 L^3 \Delta b)^{1/5}. \qquad (33a,b)$$

The vertical scale H arises as a consequence of the analysis, and the vertical size of the domain plays no direct role. [For $\sigma \gg 1$ we might expect the nonlinear terms to be small and if the buoyancy term balances the viscous term in (30) the right-hand sides of (33) are multiplied by $\sigma^{1/5}$ and $\sigma^{-1/5}$. For seawater, $\sigma \approx 7$ using the molecular values of κ and ν. If small scale turbulence exists, then the eddy viscosity will likely be similar to the eddy diffusivity and $\sigma \approx 1$.] Numerical experiments (Figs. 1 and 2, taken from Ilicak & Vallis 2012) do provide some support for this scaling, and a few simple and robust points that have relevance to the real ocean emerge, as follows.

- Most of the box fills up with the densest available fluid, with a boundary layer in temperature near the surface required in order to satisfy the top boundary condition. The boundary gets thinner with decreasing diffusivity, consistent with (33). This is a diffusive prototype of the oceanic thermocline.
- The horizontal scale of the overturning circulation is large, being at or near the scale of the box.
- The downwelling regions (the regions of convection) are of smaller horizontal scale than the upwelling regions, especially as the Rayleigh number increases.

3.3 The importance of mechanical forcing

The above results do not, strictly speaking, prohibit there from being a thermal circulation, with fluid sinking at high latitudes and rising at low, even for zero diffusivity. However, in the absence of any mechanical forcing, this circulation must be laminar, even at high Rayleigh number, meaning that flow is not allowed to break in such a way that energy can be dissipated — a very severe constraint that most flows cannot satisfy. The scalings (33) further suggest that the magnitude of the circulation in fact scales (albeit nonlinearly) with the size molecular diffusivity, and if these scalings are correct the circulation will in fact diminish as $\kappa \to 0$. For small diffusivity, the solution most likely to be adopted by the fluid is for the flow to become confined to a very thin layer at the surface, with no abyssal motion at all, which is completely unrealistic vis-à-vis the observed ocean. Thus, the deep circulation of the ocean cannot be considered to be wholly forced by buoyancy gradients at the surface.

Suppose we add a mechanical forcing, $\mathbf{F}$, to the right-hand side of the momentum equation (11a); this might represent wind forcing at the surface, or tides. The kinetic

energy budget becomes

$$\varepsilon = \langle wb \rangle + \langle \mathbf{F} \cdot \mathbf{v} \rangle = H^{-1}\kappa[\overline{b}(0) - \overline{b}(-H)] + \langle \mathbf{F} \cdot \mathbf{v} \rangle \,. \tag{34}$$

In this case, even for $\kappa = 0$, there is a source of energy and therefore turbulence (i.e., a dissipative circulation) can be maintained. The turbulent motion at small scales then provides a mechanism of mixing and so can effectively generate an 'eddy diffusivity' of buoyancy. *Given* such an eddy diffusivity, wind forcing is no longer necessary for there to be an overturning circulation. Therefore, it is useful to think of mechanical forcing as having two distinct effects.

1. The wind provides a stress on the surface that may directly drive the large-scale circulation, including the overturning circulation.

2. Both tides and the wind provide a mechanical source of energy to the system that allows the flow to become turbulent and so provides a source for an eddy diffusivity and eddy viscosity.

In either case, we may conclude that the presence of mechanical forcing is necessary for there to be an overturning circulation in the world's oceans of the kind observed.

4 The Relative Scale of Convective Plumes and Diffusive Upwelling

Why is the downwelling region narrower than the upwelling? The answer is that high Rayleigh number convection is much more efficient than diffusional upwelling, so that the the convective buoyancy flux can match the diffusive flux only if the convective plumes cover a much smaller area than diffusion. (Tom Haine explained this to me.) Suppose that the basin is initially filled with water of an intermediate temperature, and that surface boundary conditions of a temperature decreasing linearly from low latitudes to high latitudes are imposed. The deep water will be convectively unstable, and convection at high latitudes (where the surface is coldest) will occur, quickly filling the abyss with dense water. After this initial adjustment, the deep, dense water at lower latitudes will be slowly warmed by diffusion, but at the same time surface forcing will maintain a cold high latitude surface, thus leading to high latitude convection. A steady state or statistically steady state is eventually reached with the deep water having a slightly higher potential density than the surface water at the highest latitudes, and so with continual convection, but convection that takes place only at the highest latitudes.

To see this more quantitatively consider the respective efficiencies of the convective heat flux and the diffusive heat flux. Consider an idealized re-arrangement of two parcels, initially with the heavier one on top as illustrated in Fig. 3. The potential energy released by the re-arrangement, ΔP is given by

$$\Delta P = P_{\text{final}} - P_{\text{initial}} \tag{35}$$

$$= g\left[(\rho_1 z_2 + \rho_2 z_1) - (\rho_1 z_1 + \rho_2 z_2)\right] \tag{36}$$

$$= g(z_2 - z_1)(\rho_1 - \rho_2) = \rho_0 \Delta b \Delta z \tag{37}$$

where $\Delta z = z_2 - z_1$ and $\Delta b = g(\rho_1 - \rho_2)/\rho_0$.

The kinetic energy gained by this re-arrangement, ΔK is given by $\Delta K = \rho_0 w^2$ and equating this to (35) gives

$$w^2 = -\Delta b \Delta z. \tag{38}$$

Note that if the heavier fluid is initially on top then $\rho_2 > \rho_1$ and, as defined, $\Delta b < 0$. The vertical convective buoyancy flux per unit area, B_c, is given by $B_c = w\Delta b$ and using (38) we find

$$B_c = (-\Delta b)^{3/2}(\Delta z)^{1/2}. \tag{39}$$

The upwards diffusive flux, B_d, per unit area is given by

$$B_d = \kappa \frac{\Delta b}{H} \tag{40}$$

where H is the thickness of the layer over which the flux occurs. In a steady state the total diffusive flux must equal the convective flux so that, from (39) and (40),

$$(-\Delta b)^{3/2}(\Delta z)^{1/2}\delta_A = \kappa \frac{\Delta b}{H}, \tag{41}$$

where δ_A is the fractional area over which convection occurs. Thus If we set $\Delta z = H$, we get

$$\delta_A = \frac{\kappa}{(\Delta b)^{1/2} H^{3/2}} \tag{42}$$

This is a small number, although it is not quite right yet — we don't really know H. Let us use (33a), namely $H = (\kappa^2 L^2/\Delta b)^{1/5}$ then

$$\begin{aligned} \delta_A &= \frac{\kappa}{(\Delta b)^{1/2}(\kappa^2 L^2/\Delta b)^{3/10}} \\ &= \left(\frac{\kappa^2}{\Delta b L^3}\right)^{1/5} = (Ra\ \sigma)^{-1/5}. \end{aligned} \tag{43}$$

For geophysically relevant situations this is a very small number, usually smaller than 10^{-5}. Although the details of the above calculation may be questioned (for example, the use of the same buoyancy difference and vertical scale in the convection and the diffusion), the physical basis for the result is clear: for realistic choices of the diffusivity the convection is much more efficient than the diffusion and so will occur over a much smaller area. This result almost certainly transcends the limitations of its derivation.

References

Ilicak, M. & Vallis, G. K., 2012. Simulations and scaling of horizontal convection. *Tellus A*, **64**, 1–17.

Nycander, J., 2010. Horizontal convection with a non-linear equation of state: generalization of a theorem of Paparella and Young. *Tellus A*, **62**, 134–137.

Paparella, F. & Young, W. R., 2002. Horizontal convection is non-turbulent. *J. Fluid Mech.*, **466**, 205–214.

Rossby, H. T., 1965. On thermal convection driven by non-uniform heating from below: an experimental study. *Deep-Sea Res.*, **12**, 9–16.

Sandström, J. W., 1908. Dynamische Versuche mit Meerwasser (Dynamical experiments with seawater). *Annal. Hydrogr. Marit. Meteorol.*, **36**, 6–23.

Sandström, J. W., 1916. Meteologische Studien im Schwedischen Hochgebirge (Meteological studies in the Swedish high mountains). *Goteborgs Kungl. Vetenskaps-och Vitterhets-Samhalles, Handingar*, **27**, 1–48.

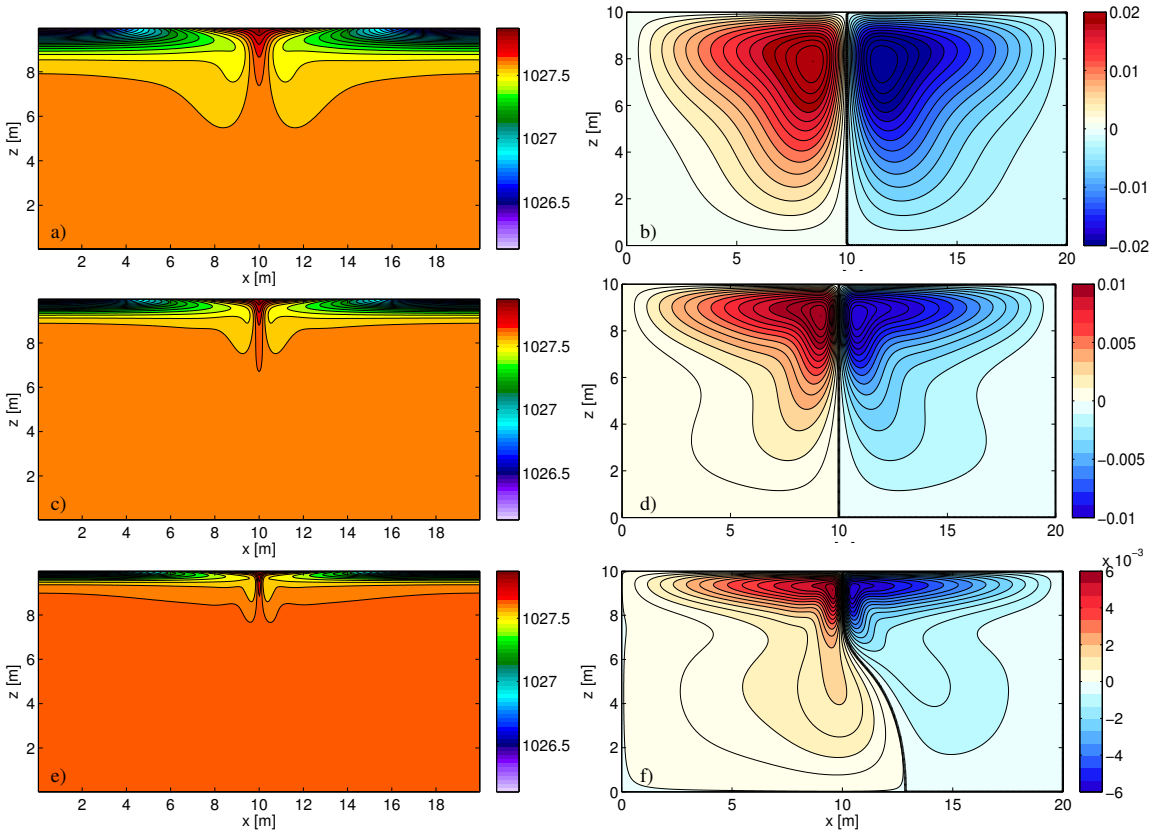


Figure 1: Temperature (left) and streamfunction (right) fields. From the top, the Rayleigh numbers are 10^6, 10^7, 10^8.

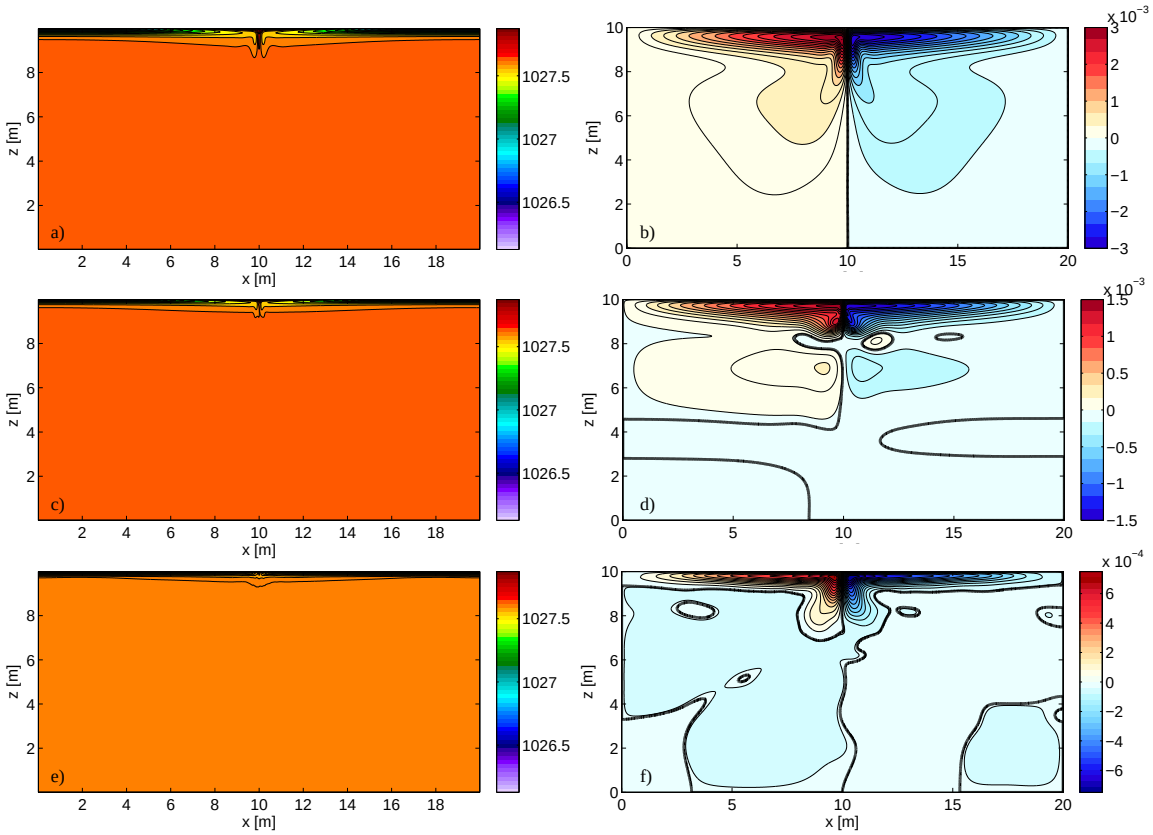


Figure 2: Temperature (left) and streamfunction (right) fields. From the top, the Rayleigh numbers are 10^9, 10^{10}, 10^{11}.

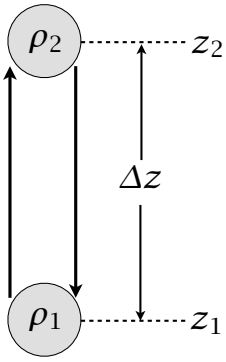


Figure 3: Two fluid parcels, of density ρ_1 and ρ_2 and initially at positions z_1 and z_2 respectively, are interchanged. If $\rho_2 > \rho_1$ then the final potential energy is lower than the initial potential energy, with the difference being converted into kinetic energy.

Lecture 9: The Deep Ocean Circulation

Geoff Vallis; notes by Ashley Payne and Erica Rosenblum

June 26

In this lecture we try to understand the processes that give rise to a deep meridional overturning circulation. We'll present a zonally-averaged model of the meridional overturning circulation of the ocean, following Nikurashin & Vallis (2011, 2012). It is a quantitative model, and might even be called a theory, depending on what one's definition of theory is.

1 A Model of the Wind-driven Overturning Circulation

The model is motivated by the plot of the stratification shown in Fig. 1, and the schematic of water mass properties of the Atlantic shown in Fig. 3. The following features are apparent.

1. Two main masses of water, known as North Atlantic Deep Water (NADW) and Antarctic Bottom Water (AABW). Both are interhemispheric. NADW appears to outcrop in high northern latitudes and high southern latitudes, and AABW just at high Southern.

2. Isopycnals are flat over most of the ocean, and slope with a fairly uniform slope in the Southern Ocean.

3. The circulation is along isopycnals in much of the interior. There is some water mass transformation between AABW and NADW, but most of it occurs near the surface.

2 A Theory for the MOC in a Single Hemisphere

Let us first imagine there is a wall at the equator, and make a model of the circulation in the Southern Hemisphere (Figs. 4, 5, 6); that is, essentially of AABW. The model will have the following features, or bugs if you are being critical.

1. Zonally averaged.

2. Simple geometry. A zonally re-entrant channel at high latitudes, with an enclosed basin between it and the equator.

3. We solve the equations of motion separately in the two regions and match the solutions at the boundary.

4. Mesoscale eddies are parameterized with a very simple down-gradient scheme.

5. There are no wind-driven gyres.

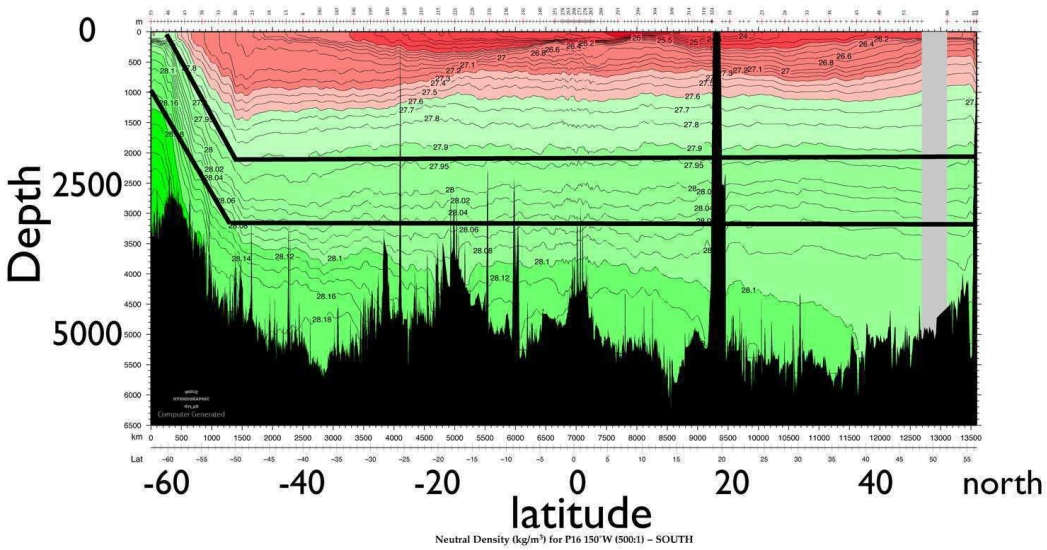


Figure 1: Stratification in the Pacific at 150° W

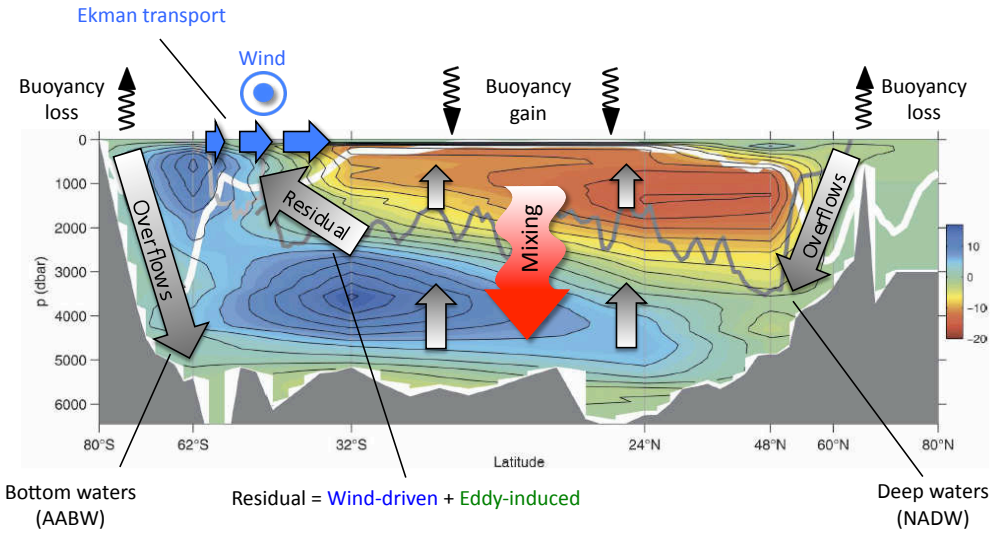


Figure 2: A schematic of deep ocean circulation.

2.1 Equations of motion

With quasi-geostrophic scaling the zonally-averaged zonal momentum and buoyancy equations are

$$\frac{\partial \overline{u}}{\partial t} - f_0\,\overline{v} = -\frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial \tau}{\partial z}, \tag{1}$$

$$\frac{\partial \overline{b}}{\partial t} + N^2\overline{w} = -\frac{\partial}{\partial y}\overline{v'b'} + \kappa_v \frac{\partial^2 \overline{b}}{\partial z^2}. \tag{2}$$

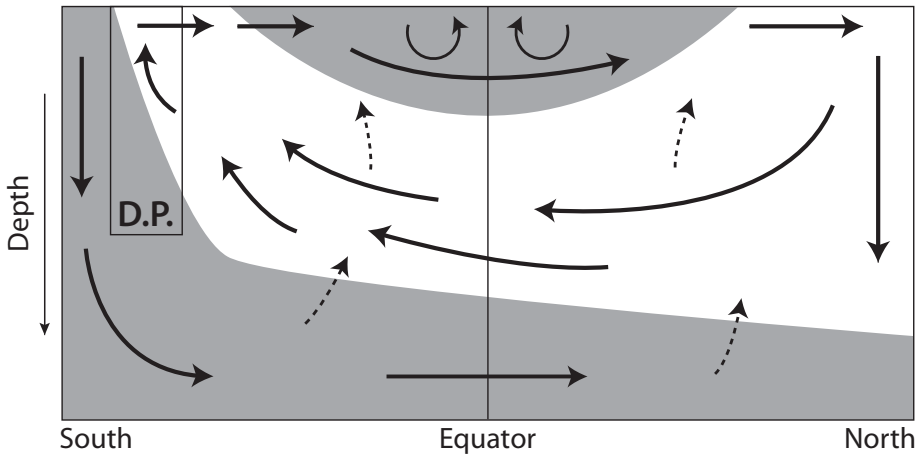


Figure 3: A simpler schematic of deep ocean circulation

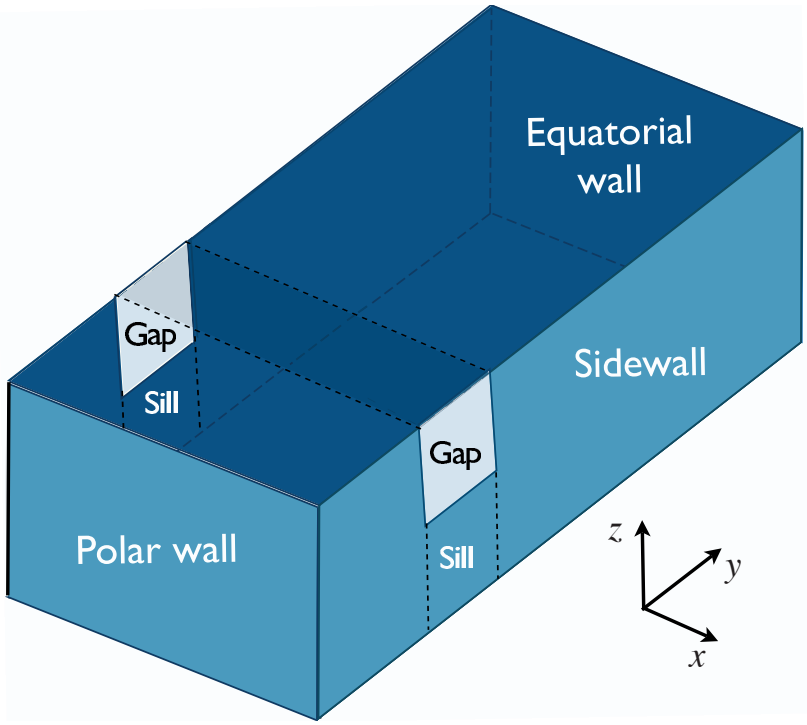


Figure 4: Idealized geometry of the Southern Ocean: a re-entrant channel, partially blocked by a sill, is embedded within a closed rectangular basin; thus, the channel has periodic boundary conditions. The channel is a crude model of the Antarctic Circumpolar Current, with the area over the sill analogous to the Drake Passage.

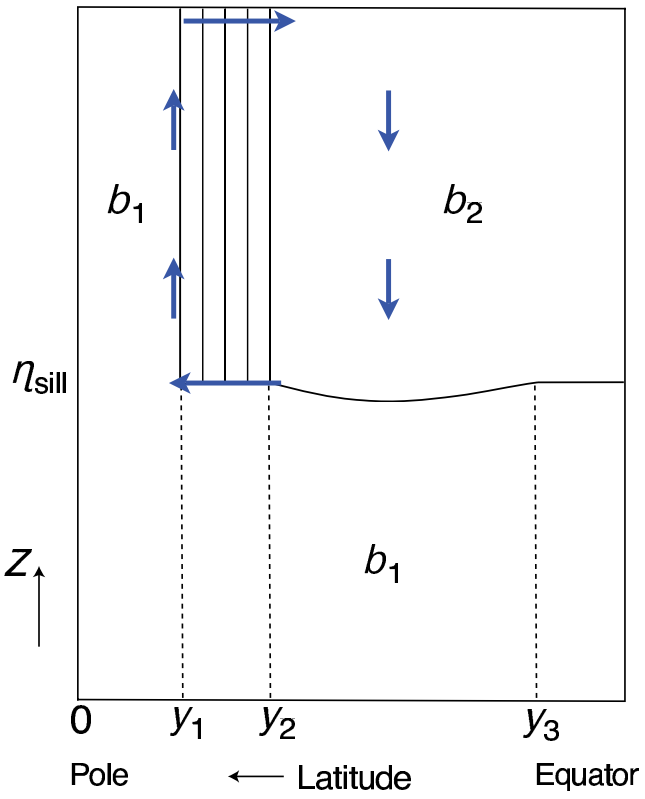


Figure 5: Cross-section of a structure of the single-hemisphere ocean model. There is a channel between y_1 and y_2. The arrows indicate the fluid flow driven by the equatorward Ekman transport in the channel, and the solid lines are isopycnals.

where b is buoyancy ('temperature') and $N^2 = \partial_z b_0$. To these we add the thermal wind relation and mass continuity:

$$f_0 \frac{\partial \overline{u}}{\partial z} = -\frac{\partial \overline{b}}{\partial y}, \qquad \frac{\partial \overline{v}}{\partial y} + \frac{\partial \overline{w}}{\partial z} = 0.$$

Define a *residual flow* such that

$$\overline{v}^* = \overline{v} - \frac{\partial}{\partial z}\left(\frac{1}{N^2}\overline{v'b'}\right), \qquad \overline{w}^* = \overline{w} + \frac{\partial}{\partial y}\left(\frac{1}{N^2}\overline{v'b'}\right).$$

whence

$$\frac{\partial \overline{u}}{\partial t} - f_0\,\overline{v}^* = \overline{v'q'} + \frac{\partial \tau}{\partial z} \tag{3a}$$

$$\frac{\partial \overline{b}}{\partial t} + N^2 \overline{w}^* = \kappa_v \frac{\partial^2 \overline{b}}{\partial z^2}. \tag{3b}$$

These are the so-called transformed Eulerian mean (TEM) equations. The theory of them is extensive and suffice it to say here that $\overline{v}^*$ and $\overline{w}^*$ more nearly represent the trajectories

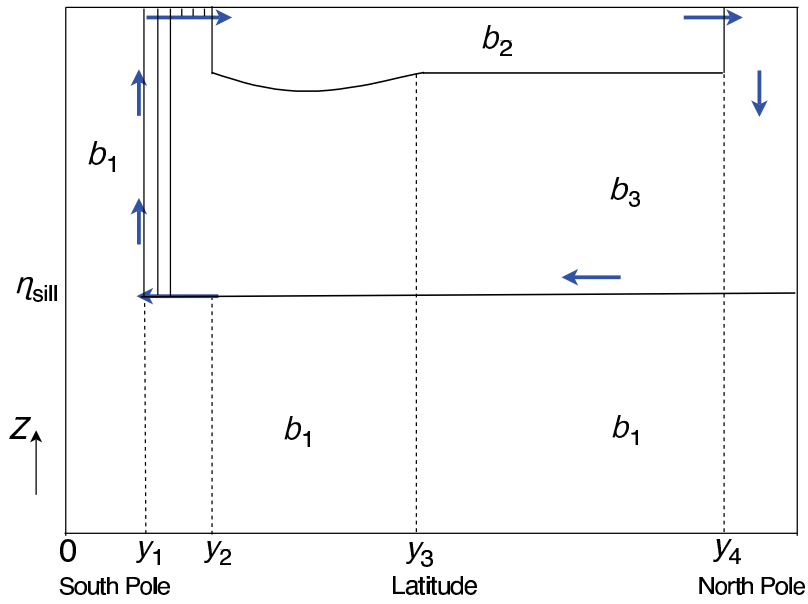


Figure 6: As for Fig. 5, but now for a two-hemisphere ocean with a source of dense water, b_3, at high northern latitudes.

of fluid parcels. Note that there are no fluxes in the buoyancy equation and that only the PV flux, $\overline{v'q'}$, need be parameterized. If we now put back some of the terms we omitted, our complete equations are

$$\frac{\partial \overline{u}}{\partial t} - f\overline{v}^* = \overline{v'q'} + \frac{\partial \tau}{\partial z} \tag{4a}$$

$$\frac{\partial \overline{b}}{\partial t} + \overline{v}^* \frac{\partial \overline{b}}{\partial y} + \overline{w}^* \frac{\partial \overline{b}}{\partial z} = \kappa_v \frac{\partial^2 \overline{b}}{\partial z^2}. \tag{4b}$$

where $(\overline{v}^* , \overline{w}^*) = (-\partial\psi/\partial z, \partial\psi/\partial y)$ and $f\partial\overline{u}/\partial z = -\partial\overline{b}/\partial y$. The stress τ in only non-zero near the top (wind-stress) and bottom (Ekman drag), and τ integrates to zero. We'll look for steady state solutions and drop the * notation so that all variables are residuals and zonal averages.

Equations in the channel

We parameterize

$$\overline{v'q'} = -K_e \frac{\partial \overline{q}}{\partial y}. \tag{5}$$

where, approximately, for the large-scale ocean

$$\overline{q} \approx f \frac{\partial}{\partial z}\left(\frac{\overline{b}}{\overline{b}_z}\right), \quad \text{so that} \quad \frac{\partial \overline{q}}{\partial y} \approx f \frac{\partial}{\partial z}\left(\frac{\overline{b}_y}{\overline{b}_z}\right) = -f\frac{\partial S}{\partial z} \tag{6}$$

where $S = -\overline{b}_y/\overline{b}_z$ is the slope of the isopycnals. Thus

$$\overline{v'q'} = fK_e \frac{\partial S}{\partial z}. \tag{7}$$

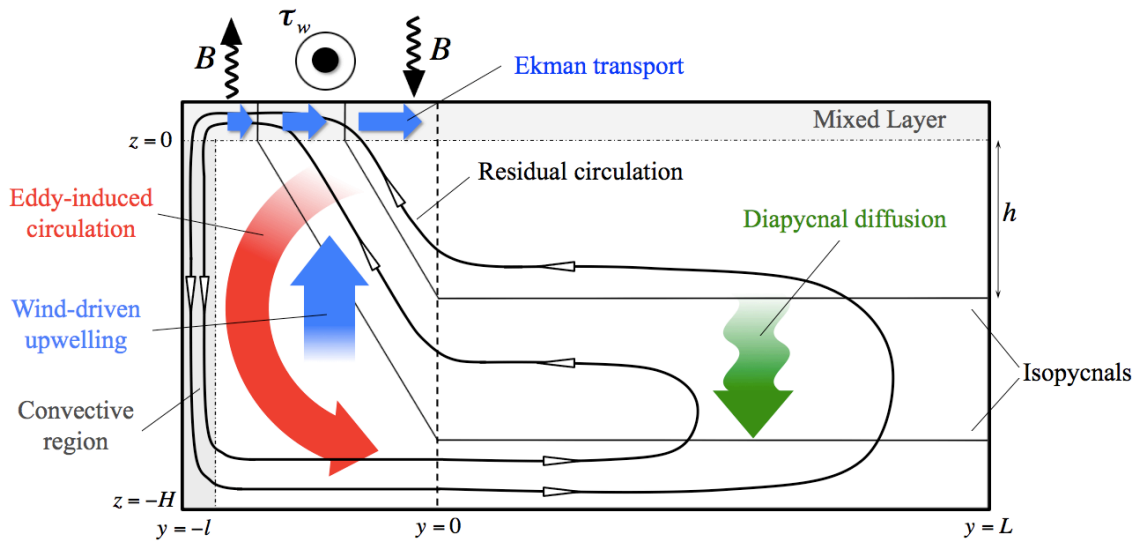


Figure 7: Schematic of the single hemisphere meridional overturning circulation crudely representing AABW. Thin black lines are the isopycnals, thick black line is a overturning streamfunction, dashed vertical line is the northern edge of the channel, shaded gray areas are the convective region and the surface mixed layer.

and the momentum equation becomes

$$-f\overline{v} = fK_e\frac{\partial S_b}{\partial z} + \frac{\partial \tau}{\partial z}. \tag{8}$$

Since $\overline{v}^* = -\partial\psi/\partial z$ we integrate this from the top to a level z and obtain

$$\psi = -\frac{\tau_w}{f} + K_e S. \tag{9}$$

We have assumed $\psi = 0$ and $S = 0$ at the top, and $\tau = \tau_w$ at the top (base of mixed layer) and $\tau = 0$ in the interior.

The buoyancy equation in terms of streamfunction is

$$\overline{\mathbf{v}}\cdot\nabla\overline{b} = \kappa_v\frac{\partial^2\overline{b}}{\partial z^2} \qquad \Longrightarrow \qquad \frac{\partial\psi}{\partial y}\frac{\partial\overline{b}}{\partial z} - \frac{\partial\psi}{\partial z}\frac{\partial\overline{b}}{\partial y} = \kappa_v\frac{\partial^2\overline{b}}{\partial z^2} \tag{10}$$

or

$$\frac{\partial\psi}{\partial y} + S\frac{\partial\psi}{\partial z} = \frac{\partial_z^2\overline{b}}{\partial_z\overline{b}}\,. \tag{11}$$

The boundary condition on ψ for this will be supplied by the basin! The other boundary condition we will need is the buoyancy distribution at the top, and so we specify

$$\overline{b}(y, z = 0) = b_0(y). \tag{12}$$

Equations in the basin

In the basin the slope of the isopycnals is assumed zero and (11) becomes the conventional upwelling diffusive balance,

$$w\frac{\partial \overline{b}}{\partial z} = \kappa \frac{\partial^2 \overline{b}}{\partial z^2} \qquad \text{or} \qquad \frac{\partial \psi}{\partial y}\frac{\partial \overline{b}}{\partial z} = \kappa_v \frac{\partial^2 \overline{b}}{\partial z^2}. \tag{13}$$

Integrate this from the edge of the channel, $y = 0$, to the northern edge, $y = L$, and obtain

$$\psi|_{y=0} = -\kappa_v L \frac{\overline{b}_{zz}}{\overline{b}_z} \ . \tag{14}$$

2.2 Scaling

Let hats denote non-dimensional values and let

$$z = h\widehat{z}, \qquad y = l\widehat{y}, \qquad \tau_w = \tau_0 \widehat{\tau}_w, \tag{15}$$

$$f = f_0 \widehat{f}, \qquad \psi = \frac{\tau_0}{f_0}\widehat{\psi}, \qquad S = \frac{h}{l}\widehat{S}, \tag{16}$$

where h is a characteristic vertical scale such that $S \sim h/l$. It will emerge as part of the solution. If we have scaled properly then variables with hats on are of order one. The nondimensional equations of motion are then

$$\text{Buoyancy evolution:} \qquad \partial_{\widehat{y}}\widehat{\psi} + \widehat{S}\partial_{\widehat{z}}\widehat{\psi} = \epsilon \left(\frac{l}{L}\right) \frac{\partial_{\widehat{z}\widehat{z}}\widehat{b}}{\partial_{\widehat{z}}\widehat{b}}, \tag{17a}$$

$$\text{Momentum balance:} \qquad \widehat{\psi} = -\frac{\widehat{\tau}}{\widehat{f}} + \Lambda \widehat{S}, \tag{17b}$$

$$\text{Boundary condition:} \qquad \widehat{\psi}|_{\widehat{y}=0} = -\epsilon \frac{\partial_{\widehat{z}\widehat{z}}\widehat{b}}{\partial_{\widehat{z}}\widehat{b}}, \tag{17c}$$

where

$$\Lambda = \frac{\text{Eddies}}{\text{Wind}} = \frac{K_e}{\tau_0/f_0}\frac{h}{l} \sim 1. \qquad \text{and} \qquad \epsilon = \frac{\text{Mixing}}{\text{Wind}} = \frac{\kappa_v}{\tau_0/f_0}\frac{L}{h} \sim 0.1 - 1. \tag{18a,b}$$

These are the two important nondimensional numbers in the problem and we can obtain estimates of their values by using some observed values for the other parameters. Thus, with $\kappa_v = 10^{-5}$ m^2 s^{-1}, $K_e = 10^3$ m^2 s^{-1}, $\tau_0 = 0.1$ N m^{-2}, $f_0 = 10^{-4}$ s^{-1}, $\rho_0 = 10^3$ kg m^{-3}, $L = 10,000$ km, $l = 1,000$ km, and $h = 1$ km we find

$$\Lambda = 1, \qquad \epsilon = 0.1, \qquad \text{and} \qquad \frac{l}{L} = 0.1. \tag{19}$$

Note that Λ and ϵ are not independent of each other for they both depend on the vertical scale of stratification h which is a part of the solution, and for that we must look at some limiting cases.

The weak diffusiveness limit

Suppose that mixing is small and that $\epsilon \ll 1$. We can then *require* that $\Lambda = 1$ in order that the eddy-induced circulation nearly balance the wind-driven circulation (because the diffusive term is small), whence the vertical scale h is given by

$$\frac{h}{l} = \frac{\tau_0/f_0}{K_e}. \tag{20}$$

If K_e does become small then h becomes large, meaning that the isopycnals are near vertical. Using the above value for h we find that

$$\epsilon = \frac{\kappa_v K_e}{(\tau_0/f_0)^2}\frac{L}{l} \tag{21}$$

This is an appropriate measure of the strength of the diapycnal diffusion in the ocean. Using (17c) we see that $\widehat{\psi} \sim \epsilon$ so that the dimensional strength of the circulation goes as

$$\Psi = \epsilon\frac{\tau_0}{f_0} = \kappa_v \frac{K_e}{\tau_0/f_0}\frac{L}{l}. \tag{22}$$

Another way to obtain this is to note that for weak diffusion the balance in the dimensional momentum equation is between wind forcing and eddy effects (because they must nearly cancel) so that

$$\frac{\tau_w}{f} \sim K_e S, \tag{23}$$

which may be written as

$$\frac{h}{l} \sim \frac{\tau_w}{K_e f}. \tag{24}$$

Advective-diffusive balance in the basin gives

$$\frac{\partial\psi}{\partial y}\frac{\partial\overline{b}}{\partial z} = \kappa_v \frac{\partial^2\overline{b}}{\partial z^2} \qquad \Longrightarrow \qquad \Psi = \frac{\kappa_v L}{h} \tag{25}$$

and (24) and (25) together give (22).

The strong diffusiveness limit

This limit may be appropriate for the abyssal ocean and in any case it is worth doing, so let us take $\epsilon \gg 1$ and the circulation in the basin will in some sense be strong. As before the nondimensional strength of the circulation is given by

$$\widehat{\psi} = \mathcal{O}(\epsilon) \gg 1. \tag{26}$$

The fact that $\widehat{\psi} \neq \mathcal{O}(1)$ means we haven't scaled things in an ideal fashion, but let's proceed anyway. Dimensionally

$$\Psi = \epsilon\frac{\tau_0}{f_{\ 0}} \qquad \text{or} \qquad \Psi = \frac{\kappa_v L}{h} \tag{27a,b}$$

but h and ϵ are both different than before.

Now, if $\widehat{\psi} \sim \epsilon \gg 1$ the diffusion driven circulation in the basin cannot be matched by a purely wind-driven circulation in the channel, since the latter is $\mathcal{O}(1)$. We can only match the circulation with an eddy-driven circulation and therefore we require

$$\Lambda = \mathcal{O}(\epsilon). \tag{28}$$

In particular, if we set $\Lambda = \epsilon$ then

$$\epsilon = \Lambda = \sqrt{\left(\frac{K_e \kappa_v L}{(\tau_0/f)^2 l}\right)}. \tag{29}$$

This is the square root of the expression for ϵ in the weak diffusiveness limit. Using (29) and (27a) we find

$$h = \sqrt{\frac{\kappa_v}{K_e} L l} \quad \text{and} \quad \Psi = \sqrt{\frac{K_e \kappa_v L}{l}}. \tag{30a,b}$$

Discussion of limits

If diffusion is weak the stratification itself is set by the eddies. Thus, upwelling-diffusion gives $\psi \sim \kappa_v L/h$, with h being set by the wind as in (24), thus giving a circulation strength that is linearly proportional to diffusivity, as in (22). In the strong diffusion case the diffusion itself affects the stratification, and so we get a weaker dependence of the circulation strength on κ_v. In this limit diapycnal mixing deepens the isopycnals in the basin away from the channel, so that the isopycnals are steeper in the channel. This steepening is balanced by the slumping effects of baroclinic instability, and wind only has a secondary effect. From an asymptotic perspective in the small ϵ limit the residual circulation is zero to lowest order. At next order it follows the isopycnals except in the mixed layer.

Instead of varying diffusivity we can think of the wind changing. In the weak wind limit the circulation is diffusively driven and independent of the wind strength. In the strong wind limit the circulation actually *decreases* as the wind increases. This is because the wind steepens the isopycnals so the diffusive term ($\sim \kappa_v \overline{b}_{zz}$) gets smaller and hence the circulation weaker.

3 An Interhemispheric Circulation

We now introduce another 'water mass' into the mix — *North Atlantic Deep Water*, or NADW. We will construct a model of similar type to what we did in the previous lecture, but now we will divide the ocean into three regions, namely

1. a southern channel, say from 50° S to 70° S
2. a basin region, say from from 50° S to 60° N
3. a northern convective region

(Fig. 8). The idea will be to write down the dynamics in these three regions and match them at the boundaries. The main difference, and it is an important one, between this

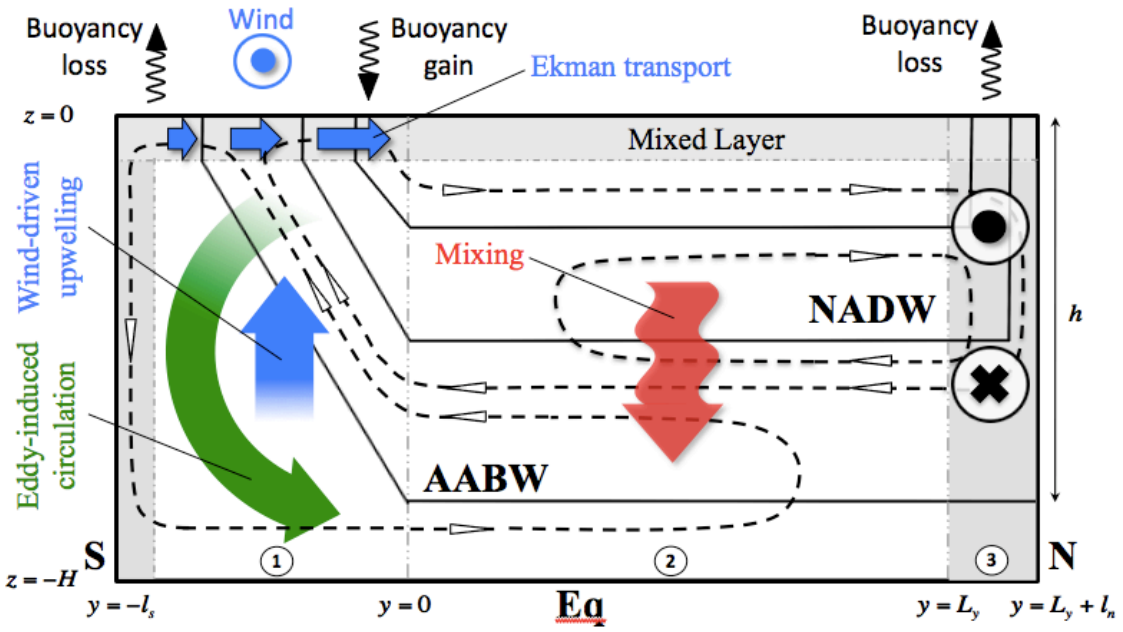


Figure 8: Schematic of the interhemispheric MOC. Thin solid black lines are the isopycnals, dashed lines with arrows are the streamlines, dashed vertical lines are the boundaries between adjacent regions, shaded gray areas are the convective regions at high latitudes and the surface mixed layer,and the red arrow represents downward diffusive heat flux. Labels 1, 2, and 3 (in circles) denote the circumpolar channel, ocean basin, and isopycnal outcrop regions.

model and the previous one is the presence of an interhemispheric cell that is primarily wind driven, and sits on top of the lower cell. It is convenient to write down the equations of motion for each region separately, with the first two regions being similar to those of the previous section.

3.1 Equations of motion

In the equations below use restoring conditions at the top, but a specified buoyancy would work too.

Region 1, the southern channel

In the channel the buoyancy equation takes its full advective-diffusive form (although we later find that in some circumstances diffusion is unimportant). The momentum equation has a wind-driven component and and eddy-driven component, as before. In dimensional form the equations are

$$\text{Buoyancy advection:} \qquad J(\psi_1, b_1) = \kappa_v \frac{\partial^2 b_1}{\partial z^2} \tag{31a}$$

$$\text{Momentum:} \qquad \psi_1 = -\frac{\tau(y)}{f} - K_e S \tag{31b}$$

$$\text{Surface boundary, at } z = 0: \qquad -\kappa \frac{\partial b_1}{\partial z} = \lambda(b^*(y) - b_1) \tag{31c}$$

$$\text{Buoyancy match to interior:} \qquad b_1(z)|_{y=0} = b_2(z) \tag{31d}$$

$$\text{Streamfunction match to interior} \qquad \psi|_{y=0} = -\kappa_v L \frac{\overline{b}_{zz}}{\overline{b}_z} \tag{31e}$$

These are more-or-less the same as those for the channel in the previous section, although we have explicitly added a buoyancy condition and we use a restoring condition on temperature.

Region 2, the basin

In the basin the isopycnals are flat and the buoyancy equation is an upwelling-diffusive balance. We won't need the momentum equation, and so we have

$$\text{Upwelling diffusive:} \qquad \frac{(\psi_3 - \psi_1)}{L} \frac{\partial b_2}{\partial z} = \kappa_v \frac{\partial^2 b_2}{\partial z^2} \tag{32a}$$

$$\text{Surface boundary:} \qquad -\kappa \frac{\partial b_2}{\partial z}\bigg|_{z=0} = \lambda(b^* - b_2) \tag{32b}$$

The upwelling diffusive balance is just $w\partial_z b = \partial_z^2 b$ with flat isopycnals, with ψ_1 and ψ_3 being the streamfunctions at the southern and northern ends of the basin, respectively. If $\psi_1 \neq \psi_2$ there is a net convergence and hence an upwelling. If $\kappa_v = 0$ then either there is no upwelling or no vertical buoyancy gradient. However, there can be an interhemispheric flow; the properties of the water mass do not change, and we expect the meridional flow to occur in a western boundary current.

Region 3, the northern convective region

In this region the values of buoyancy at the surface (i.e., $b_3(y, z = 0)$) are mapped on to the flat isopycnals in the interior (i.e, $b_2(z)$). We assume this matching occurs by convection. That is, the surface waters convect downward to the level of neutral buoyancy and then move meridionally. By thermal wind the outcropping isopycnals give rise to a zonal flow, with the total zonal transport being determined by the meridional temperature gradient and the depth to which flow convects, which is a function of such things as the winds, eddy strength and diapycnal mixing in the Southern channel. The zonal flow is thus

$$u_3(y, z) = -\frac{1}{f} \int_{-h}^{z} \frac{\partial b_3}{\partial y} \, \mathrm{d}z' + C \tag{33}$$

where C is determined by the requirement that $\int_{-h}^{0} u_3 \, \mathrm{d}z = 0$. When the relatively shallow eastward moving zonal flow collides with the eastern wall it subducts and returns. When the deep westward flow collides with the western wall it may move equatorward in a frictional deep western boundary current. It is the upper, northward moving branch of the deep

western boundary current that feeds the eastward moving flow. The total volume transport in these zonal flows thus translates to a meridional streamfunction that has the value

$$\psi_3(z) = \int_{-h}^{z} \mathrm{d}z' \int_{L_y}^{L_n} u_3 \,\mathrm{d}y \tag{34}$$

where L_y is the latitude of the southern edge of the convecting region and $L_n = L_y + l_n$ is the northern edge of the domain. Summing up, the equations in the convective region are

$$\text{Convective matching:} \qquad b_2(z) \Rightarrow b_3(y, z=0) \tag{35}$$

$$\text{Thermal wind:} \qquad f u_3(y,z) = -\int_{-h}^{z} \frac{\partial b_3}{\partial y}\,\mathrm{d}z' + C \tag{36}$$

$$\text{Mass Continuity:} \qquad \psi_3(z) = \iint u_3 \,\mathrm{d}y\,\mathrm{d}z \tag{37}$$

We now discuss how all this fits together.

3.2 Scaling and Dynamics

Our main focus is on the upper cell, since the lower cell has essentially the same dynamics as in the single hemisphere case. We proceed by writing down some parametric expressions for the streamfunctions in the three domains.

$$\Psi_1 = \left(\frac{\tau_0}{f_1} - K_e \frac{h}{l_s}\right) L_x, \tag{38a}$$

$$\Psi_2 = \Psi_3 - \Psi_1 = \frac{\kappa_v}{h} L_x L_y, \tag{38b}$$

$$\Psi_3 = \frac{\Delta b h^2}{f_3}. \tag{38c}$$

We don't like these equations because when doing scaling we don't like having additive expressions but for now we damn the torpedoes. The four unknowns are Ψ_1, Ψ_2, Ψ_3 and h and there are four equations (note that (38b) is two equations). If we combine them we obtain

$$\frac{\Delta b h^2}{f_3} - \left(\frac{\tau_0}{f_1} - K_e \frac{h}{l_s}\right) L_x = \frac{\kappa_v}{h} L_x L_y \tag{39}$$

This expression is very similar to one obtained by Gnanadesikan (1999).

With no northern source

Suppose that $\Delta b = 0$ and that there is no deep water formation in the North Atlantic. If also κ_v is small then we obtain $h/l_s = (\tau_0/f_1)/K_e$, which is essentially the same as (20), obtained previously. If κ_v is large then we find $h^2 = \kappa_v L l_s / K_e$; that is, we recover (30a). Pretty much everything is the same as it was section 2.

With no southern channel

If there is no southern channel then $\Psi_1 = 0$ and we have

$$\frac{\Delta b h^2}{f_3} = \frac{\kappa_v}{h} L_x L_y \tag{40}$$

and

$$h^3 = \kappa_v \left(\frac{f_3 L_x L_y}{\Delta b} \right) \quad \text{and} \quad \Psi_3 = \Psi_2 = (\kappa_v L_x L_y)^{2/3} \left(\frac{\Delta b}{f_3} \right)^{1/3}. \tag{41}$$

These are classical expressions for the thickness of a diffusive thermocline and the strength of a diffusively-driven overturning circulation, going back to Robinson and Stommel. This is also the same as the strong diffusivity limit.

With all three regions

This is the new bit. The weak diffusivity limit is the interesting case, as the strong diffusivity limit is really just the case with no southern channel.

In this case the upwelling is weak and $|\Psi_3| \approx \Psi_1|$ and

$$\frac{\Delta b h^2}{f_3} - \left(\frac{\tau_0}{f_1} - K_e \frac{h}{l_s} \right) L_x = 0. \tag{42}$$

In this case the basin is just a 'pass-through' region: water formed in the North Atlantic just passes through the basin without change, and upwells in the Southern Ocean. For the moment let us also assume that K_e is small and then

$$\frac{\Delta b h^2}{f_3} = \frac{\tau_0}{\rho_0 f_1} L_x, \tag{43}$$

which results in a depth scale h for the stratification,

$$h = \left(\frac{\tau_0 f_3 L_x}{f_1 \Delta b} \right)^{1/2} \tag{44}$$

Putting in the numbers, we find $h \approx 320\,\mathrm{m}$. Furthermore, the strength of the circulation is just determined by the wind stress,

$$\Psi_1 = \Psi_3 = \left(\frac{\tau_0 L_x}{f_1} \right) \tag{45}$$

which is about $10\,\mathrm{Sv}$.

In the more general case we solve (42) to give

$$h = \left(\frac{\tau_0 f_3 L_x}{f_1 \Delta b} \right)^{1/2} \left(-\alpha + \sqrt{1 + \alpha^2} \right) \tag{46}$$

where α is the nondimensional number given by the ratio of the wind to eddy effects

$$\alpha = \frac{1}{2} \frac{K_e}{l_s} \left(\frac{L_x f_1 f_3}{\tau_0 \Delta b} \right)^{1/2} = \frac{1}{2} \frac{\Psi^*}{\overline{\Psi}}. \tag{47}$$

where

$$\overline{\Psi} = \frac{\tau_0}{f_1} L_x, \qquad \Psi^* = -K_e h \frac{L_x}{l_s} = -k_e \frac{L_x}{l_s} \left(\frac{\tau_0 f_3 L_x}{f_1 \Delta b} \right)^{1/2} . \tag{48}$$

If we put in numbers then $\alpha \approx 0.08$, $\Psi^* \approx 1.6\,\mathrm{Sv}$ and $\overline{\Psi} \approx 10\,\mathrm{Sv}$. That is, the wind-induced circulation is the dominant factor in the meridional overturning circulation, and if we take α to small then we have

$$h \approx \left(\frac{\tau_0 f_3 L_x}{f_1 \Delta b} \right)^{1/2}, \qquad \Psi \approx \frac{\tau_0}{f_1} L_x \tag{49}$$

Discussion

Although there can be no certainties when eddy diffusivities are present, the use of representative parameters suggests that the eddy-induced circulation is indeed smaller than the wind-driven circulation in the Southern Ocean. That is, putting in numbers, we find $\alpha \approx 0.08$ with $\Psi^* \approx 1.6\,\mathrm{Sv}$ and $\overline{\Psi} \approx 10\,\mathrm{Sv}$. This suggests that, for typical oceanic parameters, the strength of the eddy-induced circulation on isopycnals corresponding to the middepth overturning cell is only about 10-20% of the wind-driven circulation. Thus, rather than the residual circulation vanishing as is sometimes assumed, the middepth residual circulation is comparable to the wind-driven circulation and acts to pull $O(10)$ Sv of deep water formed at high northern latitudes in the North Atlantic back up to the surface. As a result, the depth scale of stratification h is not linearly proportional to the wind stress τ, as one would obtain from the vanishing residual circulation argument with a simple eddy parameterization, but rather it scales with τ as $\tau^{1/2}$ and is dependent on Δb which is the buoyancy range for isopycnals which are shared between the circumpolar channel and the isopycnal outcrop region in the Northern Hemisphere.

In summary, in the limit of weak diapycnal mixing, relevant to the present middepth ocean, the strength of the middepth overturning circulation is primarily determined by the Ekman transport in the Southern Ocean. The rest of the ocean is essentially forced to adjust and produce the amount of deep water demanded by the Ekman transport and the associated wind-driven upwelling in the Southern Ocean. For instance, during the transient adjustment, the Ekman transport in the circumpolar channel, in conjunction with the surface buoyancy flux, pulls dense waters up from the deep ocean, converts them into light waters at the surface, and pumps these waters into, or just below, the main thermocline in the ocean basin. The rate at which these light waters are then imported into the deep water formation region in the North Atlantic, converted back into dense waters, and exported to the ocean basin at middepth, is controlled by the meridional pressure gradient set up by the outcropping isopycnals in the north. Hence, light waters pumped into the ocean basin by the Ekman transport in the south accumulate in, or just below, the main thermocline, therefore deepening the middepth isopycnals and increasing the transport of light water into the deep water formation region in the north until the transports in the north and south match. The established interhemispheric balance sets the depth of the isopycnals in the ocean basin and thus stratification throughout the entire ocean.

In the case when deep waters are not produced in the north, as observed in the Pacific Ocean, light waters pumped into the ocean basin by the Southern Ocean wind will deepen

the mid-depth isopycnal in the ocean basin, and thus steepen their slopes in the Southern Ocean, until the eddy-induced circulation in the Southern Ocean cancels the wind-driven circulation resulting in a zero residual circulation and water mass transformation.

4 Theory of the Main Thermocline (And why it is not really like the tropopause)

Our goal in this lecture is to give a sense of the structure of the thermocline, and to draw out similarities and differences (mainly differences) with the tropopause.

A simple kinematic model

The fact that cold water with polar origins upwells into a region of warmer water suggests that we consider the simple one-dimensional advective–diffusive balance,

$$w\frac{\partial T}{\partial z} = \kappa\frac{\partial^2 T}{\partial z^2}, \tag{50}$$

where w is the vertical velocity, κ is a diffusivity and T is temperature. In mid-latitudes, where this might hold, w is positive and the equation represents a balance between the upwelling of cold water and the downward diffusion of heat. If w and κ are given constants, and if T is specified at the top ($T = T_T$ at $z = 0$) and if $\partial T/\partial z = 0$ at great depth ($z = -\infty$) then the temperature falls exponentially away from the surface according to

$$T = (T_T - T_B)\mathrm{e}^{wz/\kappa} + T_B, \tag{51}$$

here T_B is a constant. This expression cannot be used to estimate how the thermocline depth scales with either w or κ, because the magnitude of the overturning circulation depends on κ. However, it is reasonable to see if the observed ocean is broadly consistent with this expression. The diffusivity κ can be measured; it is an eddy diffusivity, maintained by small-scale turbulence, and measurements produce values that range between $10^{-5}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ in the main thermocline and $10^{-4}\,\mathrm{m}^2\,\mathrm{s}^{-1}$ in abyssal regions over rough topography and in and near continental margins, with still higher values locally.

The vertical velocity is too small to be measured directly, but various estimates based on deep water production suggest a value of about $10^{-7}\,\mathrm{m\ s}^{-1}$. Using this and the smaller value of κ in (51) gives an e-folding vertical scale, κ/w, of just 100 m, beneath which the stratification is predicted to be very small (i.e., nearly uniform potential density). Using the larger value of κ increases the vertical scale to 1000 m, which is probably closer to the observed value for the total thickness of the thermocline (look at Fig. 9), but using such a large value of κ in the main thermocline is not supported by the observations. Similarly, the deep stratification of the ocean is rather larger than that given by (50), except with values of diffusivity on the large side of those observed.

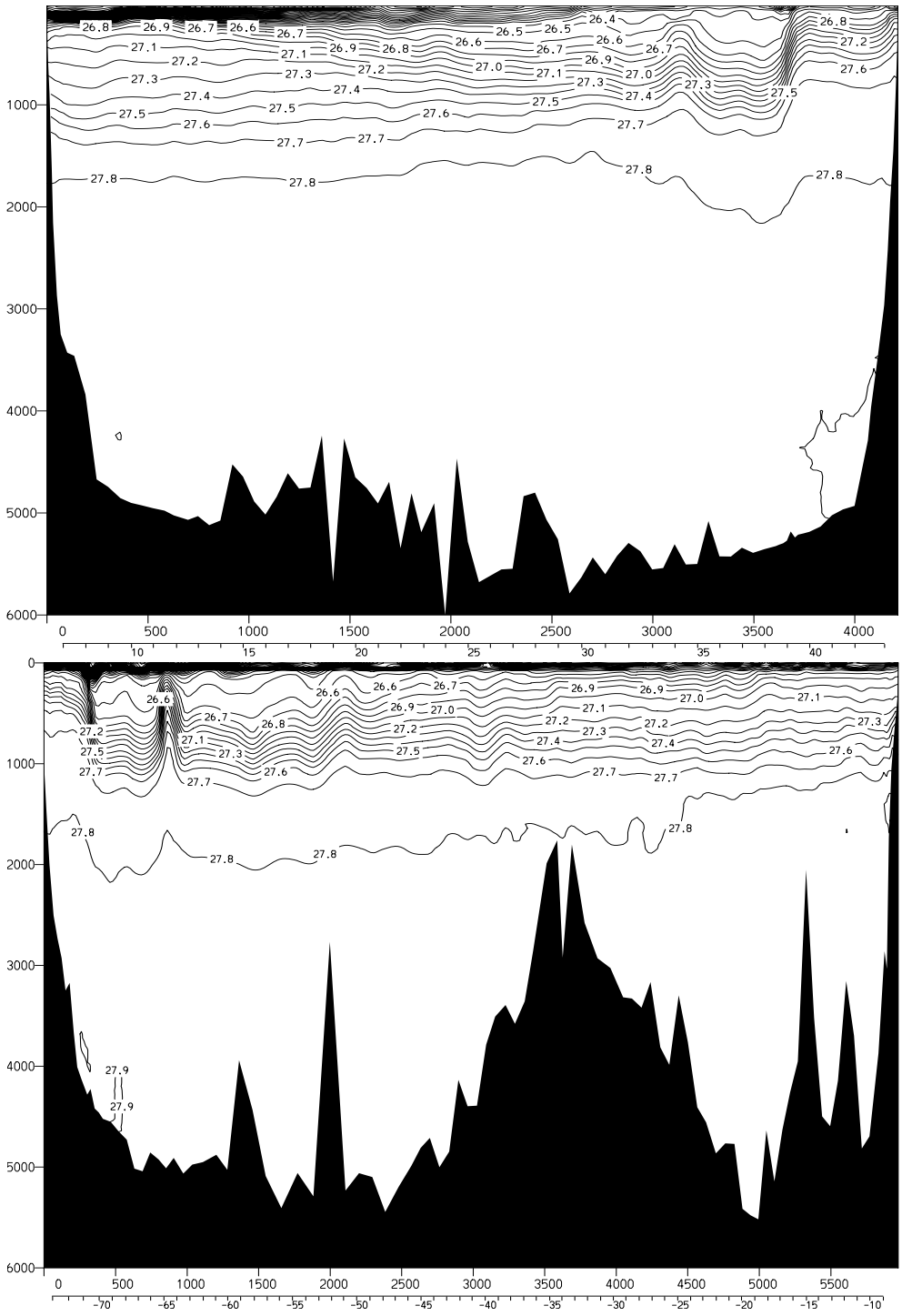


Figure 9: Sections of potential density (σ_θ) in the North Atlantic. Upper panel: meridional section at 53° W , from 5° N to 45° N , across the subtropical gyre. Lower panel: zonal section at 36° N , from about 75° W to 10° W .

5 Scaling and Simple Dynamics of the Main Thermocline

The Rossby number of the large-scale circulation is small and the scale of the motion large, and the flow obeys the planetary-geostrophic equations:

$$\mathbf{f}\times\mathbf{u} = -\nabla\phi, \qquad \frac{\partial\phi}{\partial z} = b, \tag{52a,b}$$

$$\div\mathbf{v} = 0, \qquad \frac{\mathrm{D}b}{\mathrm{D}t} = \kappa\frac{\partial^2 b}{\partial z^2}. \tag{53a,b}$$

5.1 An advective scale

If there is upwelling ($w > 0$) from the abyss, and Ekman downwelling ($w < 0$) at the surface, there is some depth D_a at which $w = 0$. By cross-differentiating (52a) we obtain $\beta v = -f \div [z]\mathbf{u}$, or

$$\beta v = f\frac{\partial w}{\partial z} \qquad \rightarrow \qquad \beta V = f\frac{W}{D_a} = f\frac{W_E}{D_a}. \tag{54}$$

Thermal wind

$$\mathbf{f}\times\frac{\partial\mathbf{u}}{\partial z} = -\nabla b \qquad \rightarrow \qquad \frac{U}{D_a} = \frac{1}{f}\frac{\Delta b}{L}, \tag{55}$$

where Δb is the scaling value of variations of buoyancy in the horizontal. Assuming the vertical scales are the same in (54) and (55) and that $V \sim U$ then

$$D_a = W_E^{1/2}\left(\frac{f^2 L}{\beta\,\Delta b}\right)^{1/2}. \tag{56}$$

5.2 A diffusive scale

The estimate (56) cares nothing about the thermodynamic equation, so let's now include some and construct a scaling from from advective–diffusive balance in the thermodynamic equation, the linear geostrophic vorticity equation, and thermal wind balance:

$$w\frac{\partial b}{\partial z} = \kappa\frac{\partial^2 b}{\partial z^2}, \qquad \beta v = f\frac{\partial w}{\partial z}, \qquad f\frac{\partial\mathbf{u}}{\partial z} = \mathbf{k}\times\nabla b, \tag{57a,b,c}$$

with corresponding scales

$$\frac{W}{\delta} = \frac{\kappa}{\delta^2}, \qquad \beta U = \frac{fW}{\delta}, \qquad \frac{U}{\delta} = \frac{\Delta b}{fL}, \tag{58a,b,c}$$

where δ is the vertical scale. Because there is now one more equation than in the advective scaling theory we cannot take the vertical velocity as a given, otherwise the equations would be overdetermined. We therefore take it to be the abyssal upwelling velocity, which then becomes part of the *solution,* rather than being imposed. From (58) we obtain the diffusive vertical scale,

$$\delta = \left(\frac{\kappa f^2 L}{\beta\,\Delta b}\right)^{1/3}. \tag{59}$$

With $\kappa = 10^{-5}\,\mathrm{m}^{\,2}\,\mathrm{s}^{\,-2}$ and with the other parameters taking the values given following (56), (59) gives $\delta \approx 150\,\mathrm{m}$ and, using (58a), $W \approx 10^{-7}\,\mathrm{m\ s^{-1}}$, which is an order of magnitude smaller than the Ekman pumping velocity W_E.

6 The Internal thermocline

6.1 The M equation

The planetary-geostrophic equations can be written as a single partial differential equation in a single variable, although the resulting equation is of quite high order and is nonlinear. We write the equations of motion as

$$-fv = -\frac{\partial \phi}{\partial x}, \qquad fu = -\frac{\partial \phi}{\partial y}, \qquad b = \frac{\partial \phi}{\partial z}, \tag{60a,b,c}$$

$$\div \mathbf{v} \ = 0, \qquad \frac{\partial b}{\partial t} + \mathbf{v} \ \cdot \nabla b = \kappa \nabla^2 b, \tag{61a,b}$$

where we take $f = \beta y$. Cross-differentiating the horizontal momentum equations and using (61a) gives the linear geostrophic vorticity relation $\beta v = f \partial w / \partial z$ which, using (60a) again, may be written as

$$\frac{\partial \phi}{\partial x} + \frac{\partial}{\partial z}\left(-\frac{f^2}{\beta} w \right) = 0. \tag{62}$$

This equation is the divergence in (x, z) of $(\phi, -f^2 w/\beta)$ and is automatically satisfied if

$$\phi = M_z \qquad \text{and} \qquad \frac{f^2 w}{\beta} = M_x. \tag{63a,b}$$

where the subscripts on M denote derivatives. Then straightforwardly

$$u = -\frac{\partial_y \phi}{f} = -\frac{M_{zy}}{f}, \qquad v = \frac{\partial_x \phi}{f} = \frac{M_{zx}}{f}, \qquad b = \partial_z \phi = M_{zz}. \tag{64a,b,c}$$

The thermodynamic equation, (61b) becomes

$$\frac{\partial M_{zz}}{\partial t} + \left(\frac{-M_{zy}}{f} M_{zzx} + \frac{M_{zx}}{f} M_{zzy} \right) + \frac{\beta}{f^2} M_x M_{zzz} = \kappa M_{zzzz} \tag{65}$$

or

$$\frac{\partial M_{zz}}{\partial t} + \frac{1}{f} J(M_z, M_{zz}) + \frac{\beta}{f^2} M_x M_{zzz} = \kappa M_{zzzz}. \tag{66}$$

where J is the usual horizontal Jacobian. This is the M *equation,*[1] somewhat analogous to the potential vorticity equation in quasi-geostrophic theory in that it expresses the entire dynamics of the system in a single, nonlinear, advective–diffusive partial differential equation, although note that M_{zz} is materially conserved (in the absence of diabatic effects) by the three-dimensional flow. Because of the high differential order and nonlinearity of the system analytic solutions of (66) are very hard to find, and from a numerical perspective it is easier to integrate the equations in the form (60) and (61) than in the form (66). Nevertheless, it is possible to move forward by approximating the equation to one or two dimensions, or by a priori assuming a boundary-layer structure.

[1] Welander (1971).

A one-dimensional model

Let us consider an illustrative one-dimensional model (in z) of the thermocline. Merely setting all horizontal derivatives in (66) to zero is not very useful, for then all the advective terms on the left-hand side vanish. Rather, we look for steady solutions of the form $M = M(x, z)$, and the M equation then becomes

$$\frac{\beta}{f^2} M_x M_{zzz} = \kappa M_{zzzz}, \tag{67}$$

which represents the advective–diffusive balance

$$w \frac{\partial b}{\partial z} = \kappa \frac{\partial^2 b}{\partial z^2}. \tag{68}$$

(We must also suppose that the value of κ varies meridionally in the same manner as does β/f^2; without this technicality M would be a function of y , violating our premise.) If the ocean surface is warm and the abyss is cold, then (67) represents a balance between the upward advection of cold water and the downward diffusion of warm water. The horizontal advection terms vanish because the zonal velocity, u, and the meridional buoyancy gradient, b_y, are each zero. Let us further consider the special case

$$M = (x - x_e) W(z), \tag{69}$$

where the domain extends from $0 \le x \le x_e$, so satisfying $M = 0$ on the eastern boundary. Equation (67) becomes the ordinary differential equation

$$\frac{\beta}{f^2} W W_{zzz} = \kappa W_{zzzz}, \tag{70}$$

where W has the dimensions of velocity squared. We non-dimensionalize this by setting

$$z = H\widehat{z}, \qquad \kappa = \widehat{\kappa}(H W_S), \qquad W = \left(\frac{f^2 W_S}{\beta} \right) \widehat{W}, \tag{71a,b,c}$$

where the hatted variables are non-dimensional and W_S is a scaling value of the dimensional vertical velocity, w (e.g., the magnitude of the Ekman pumping velocity W_E). Equation (70) becomes

$$\widehat{W} \widehat{W}_{\widehat{z}\widehat{z}\widehat{z}} = \widehat{\kappa} \widehat{W}_{\widehat{z}\widehat{z}\widehat{z}\widehat{z}}, \tag{72}$$

The parameter $\widehat{\kappa}$ is a non-dimensional measure of the strength of diffusion in the interior, and the interesting case occurs when $\widehat{\kappa} \ll 1$; in the ocean, typical values are $H = 1\,\mathrm{km}$, $\kappa = 10^{-5}\,\mathrm{m\ s^{-2}}$ and $W_S = W_E = 10^{-6}\,\mathrm{m\ s^{-1}}$ so that $\widehat{\kappa} \approx 10^{-2}$, which is indeed small. (It might appear that we could completely scale away the value of κ in (70) by scaling W appropriately, and if so there would be no meaningful way that one could say that κ was small. However, this is a chimera, because the value of κ would still appear in the boundary conditions.)

The time-dependent form of (72), namely

$$\widehat{W}_{\widehat{z}\widehat{z}t} + \widehat{W} \widehat{W}_{\widehat{z}\widehat{z}\widehat{z}} = \widehat{\kappa} \widehat{W}_{\widehat{z}\widehat{z}\widehat{z}\widehat{z}} \tag{73}$$

is similar to Burger's equation

$$V_t + VV_z = \nu V_{zz} \tag{74}$$

which is known to develop fronts. (In the inviscid Burger's equation, $\mathrm{D}V/\mathrm{D}t = 0$, where the advective derivative is one-dimensional, and therefore the velocity of a given fluid parcel is preserved on the line. Suppose that the velocity of the fluid is positive but diminishes in the positive z-direction, so that a fluid parcel will catch-up with the fluid parcel in front of it. But since the velocity of a fluid parcel is fixed, there are two values of velocity at the same point, so a singularity must form. In the presence of viscosity, the singularity is tamed to a front.) Thus, we might similarly expect (72) to produce a front, but because of the extra derivatives the argument is not as straightforward and it is simplest to obtain solutions numerically.

Equation (72) is fourth order, so four boundary conditions are needed, two at each boundary. Appropriate ones are a prescribed buoyancy and a prescribed vertical velocity at each boundary, for example

$$\begin{aligned} &\widehat{W} = \widehat{W}_E, \quad -\widehat{W}_{\widehat{z}\widehat{z}} = B_0, && \text{at top} \\ &\widehat{W} = 0, \qquad -\widehat{W}_{\widehat{z}\widehat{z}} = 0, && \text{at bottom,} \end{aligned} \tag{75}$$

where $\widehat{W}_E$ is the (non-dimensional) vertical velocity at the base of the top Ekman layer, which is negative for Ekman pumping in the subtropical gyre, and B_0 is a constant, proportional to the buoyancy difference across the domain. We obtain solutions numerically by Newton's method. The solutions here are obtained using about 1000 uniformly spaced grid points to span the domain, taking just a few seconds of computer time. Because of the boundary layer structure of the solutions employing a non-uniform grid would be even more efficient for this problem, but there is little point in designing a streamlined hat to reduce the effort of walking. These are shown in Figs. 10 and 11. The solutions do indeed display fronts, or boundary layers, for small diffusivity. If the wind forcing is zero (Fig. 11), the boundary layer is at the top of the fluid. If the wind forcing is non-zero, an internal boundary layer — a front — forms in the fluid interior with an adiabatic layer above and below. In the real ocean, where wind forcing is of course non-zero, the frontal region is known as the *internal thermocline.*

6.2 Boundary-layer analysis

The reasoning and the numerical solutions of the above sections suggest that the internal thermocline has a boundary-layer structure whose thickness decreases with κ. If the Ekman pumping at the top of the ocean is non-zero, the boundary layer is internal to the fluid. To learn more, let us perform a boundary layer analysis.

One-dimensional model

Let us now *assume* a steady two-layer structure of the form illustrated in Fig. 12, and that the dynamics are governed by (72) in a domain that extends from 0 to -1. The buoyancy thus varies rapidly only in an internal boundary layer of non-dimensional thickness $\widehat{\delta}$ located at $\widehat{z} = -h$; above and below this the buoyancy is assumed to be only very slowly

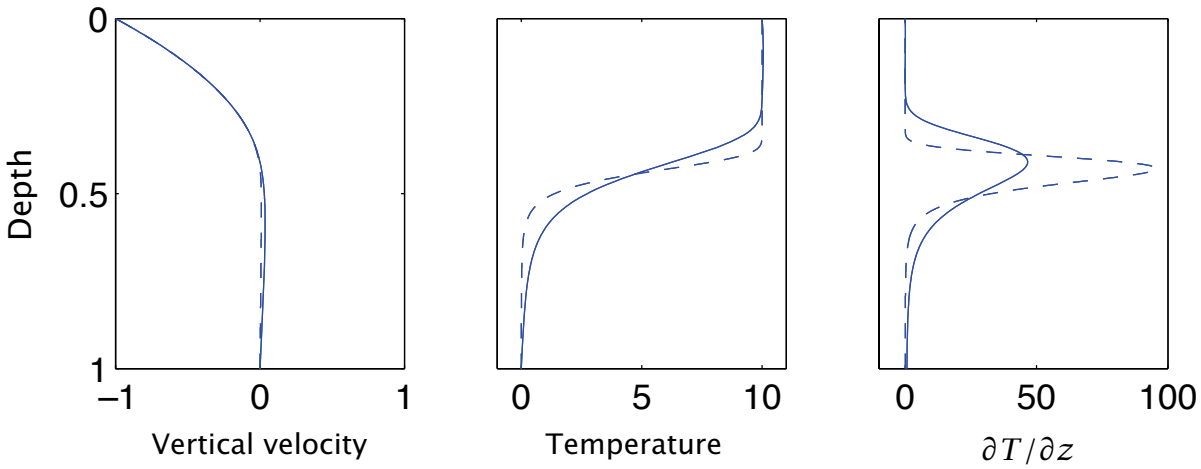


Figure 10: Solution of the one-dimensional thermocline equation, (72), with boundary conditions (75), for two different values of the diffusivity: $\widehat{\kappa} = 3.2 \times 10^{-3}$ (solid line) and $\widehat{\kappa} = 0.4 \times 10^{-3}$ (dashed line), in the domain $0 \le \widehat{z} \le -1$.

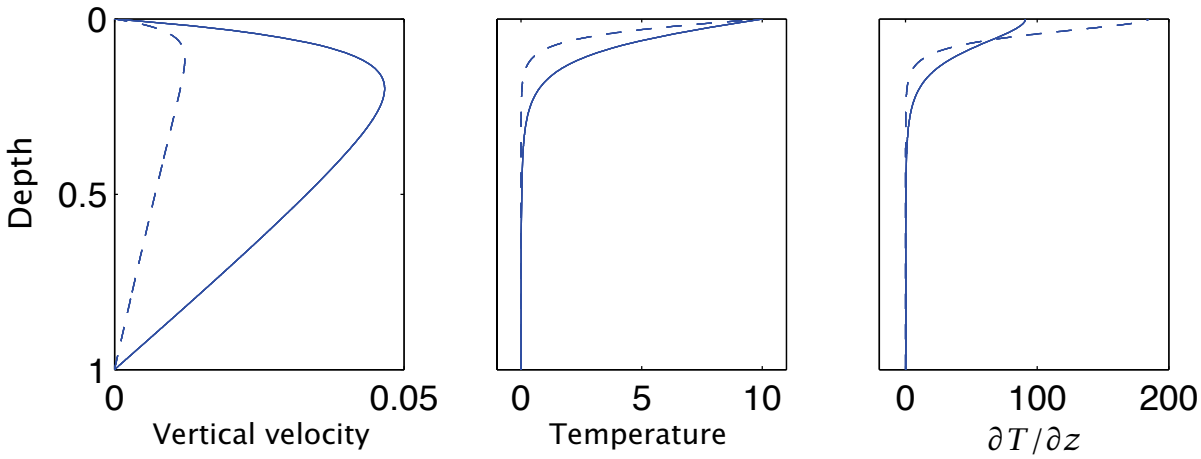


Figure 11: As for Fig. 10, but with no imposed Ekman pumping velocity at the upper boundary ($\widehat{W}_E = 0$), again for two different values of the diffusivity.

varying. Following standard boundary layer procedure we introduce a stretched boundary layer coordinate ζ where

$$\widehat{\delta}\zeta = \widehat{z} + h. \tag{76}$$

That is, ζ is the distance from $\widehat{z} = -h$, scaled by the boundary layer thickness $\widehat{\delta}$, and within the boundary layer ζ is an order-one quantity. We also let

$$\widehat{W}(\widehat{z}) = \widehat{W}_I(\widehat{z}) + \widetilde{W}(\zeta), \tag{77}$$

where $\widehat{W}_I$ is the solution away from the boundary layer and $\widetilde{W}$ is the boundary layer correction. Because the boundary layer is presumptively thin, $\widehat{W}_I$ is effectively constant through

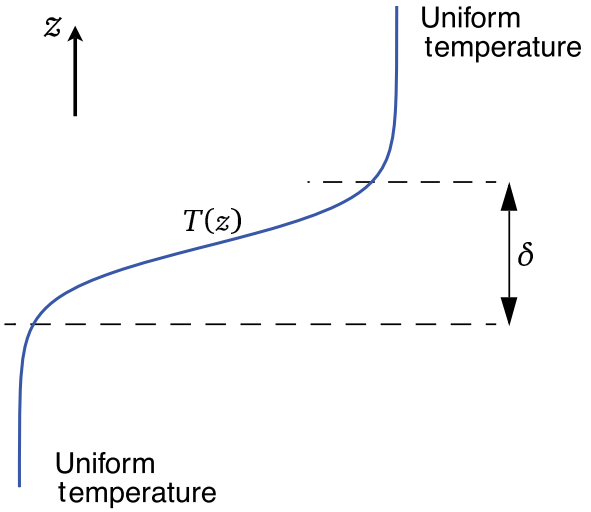


Figure 12: The simplified boundary-layer structure of the internal thermocline. In the limit of small diffusivity the internal thermocline forms a boundary layer, of thickness δ in the figure, in which the temperature and buoyancy change rapidly.

it and, furthermore, for $\widehat{z} < -h$, $\widehat{W}$ vanishes in the limit as $\kappa = 0$. We thus take $\widehat{W}_I = 0$ throughout the boundary layer. (The small diffusively-driven upwelling below the boundary layer is part of the boundary layer solution, not the interior solution.) Now, buoyancy varies rapidly in the boundary layer but it remains an order-one quantity throughout. To satisfy this we explicitly scale $\widetilde{W}$ in the boundary layer by writing

$$\widetilde{W}(\zeta) = \widehat{\delta}^2 B_0 A(\zeta), \tag{78}$$

where B_0 is defined by (75) and A is an order-one field. The derivatives of W are

$$\frac{\partial \widehat{W}}{\partial \widehat{z}} = \frac{1}{\widehat{\delta}} \frac{\partial \widetilde{W}}{\partial \zeta} = \widehat{\delta} B_0 \frac{\partial A}{\partial \zeta}, \qquad \frac{\partial^2 \widehat{W}}{\partial \widehat{z}^2} = B_0 \frac{\partial^2 A}{\partial \zeta^2}, \tag{79}$$

so that $\widehat{W}_{\widehat{z}\widehat{z}}$ is an order-one quantity. Far from the boundary layer the solution must be able to match the external conditions on temperature and velocity, (75); the buoyancy condition on $W_{\widehat{z}\widehat{z}}$ is satisfied if

$$A_{\zeta\zeta} \to \begin{cases} 1 & \text{as } \zeta \to +\infty \\ 0 & \text{as } \zeta \to -\infty. \end{cases} \tag{80}$$

On vertical velocity we require that $W \to (\widehat{z}/h + 1) W_E$ as $\zeta \to +\infty$, and $W \to$ constant as $\zeta \to -\infty$. The first matches the Ekman pumping velocity above the boundary layer, and the second condition produces the abyssal upwelling velocity, which as noted vanishes for $\kappa \to 0$.

Substituting (77) and (78) into (72) we obtain

$$B_0 A A_{\zeta\zeta\zeta} = \frac{\widehat{\kappa}}{\widehat{\delta}^3} A_{\zeta\zeta\zeta\zeta}. \tag{81}$$

Because all quantities are presumptively $\mathcal{O}(1)$, (81) implies that $\widehat{\delta} \sim (\widehat{\kappa}/B_0)^{1/3}$. We restore the dimensions of δ by using $\kappa = \widehat{\kappa}(HW_S)$ and $\Delta b = B_0 L f^2 W_S/(\beta H^2)$, where Δb is the dimensional buoyancy difference across the boundary layer [note that $b = M_{zz} = (x-1)W_{zz} \sim LW_{zz} \sim LB_0 f^2 W_S/(\beta H^2)$ using (71)]. The dimensional boundary layer thickness, δ, is then given by

$$\delta \sim \left(\frac{\kappa f^2 L}{\Delta b\, \beta}\right)^{1/3}, \tag{82}$$

which is the same as the heuristic estimate (59). The dimensional vertical velocity scales as

$$W \sim \frac{\kappa}{\delta} \sim \kappa^{2/3} \left(\frac{\Delta b\, \beta}{f^2 L}\right)^{1/3}, \tag{83}$$

this being an estimate of strength of the upwelling velocity at the base of the thermocline and, more generally, the strength of the diffusively-driven component of meridional overturning circulation of the ocean.

The qualitative features of these models transcend their detailed construction, and in particular:

- the thickness of the internal thermocline increases with increasing diffusivity, and decreases with increasing buoyancy difference across it, and as the diffusivity tends to zero the thickness of the internal thermocline tends to zero.
- the strength of the upwelling velocity, and hence the strength of the meridional overturning circulation, increases with increasing diffusivity and increasing buoyancy difference.

The three-dimensional equations

We now apply boundary layer techniques to the three-dimensional M equation.[2] The main difference is that the depth of the boundary layer is now a function of x and y , so that the stretched coordinate ζ is given by

$$\widehat{\delta}\zeta = z + h(x, y). \tag{84}$$

[The coordinates (x, y, z) in this subsection are non-dimensional, but we omit their hats to avoid too cluttered a notation.] Just as in the one-dimensional case we rescale M in the boundary layer and write

$$M = B_0 \widehat{\delta}^2 \widehat{A}(x, y, \zeta), \tag{85}$$

where the scaling factor $\widehat{\delta}^2$ again ensures that the temperature remains an order-one quantity. In the boundary layer the derivatives of M become

$$\frac{\partial M}{\partial z} = \frac{1}{\widehat{\delta}} \frac{\partial A}{\partial \zeta}, \tag{86}$$

and

$$\frac{\partial M}{\partial x} = \widehat{\delta}^2 B_0 \left(\frac{\partial A}{\partial \zeta}\frac{\partial \zeta}{\partial x} + \frac{\partial A}{\partial x}\right) = \widehat{\delta}^2 B_0 \left(\frac{\partial A}{\partial \zeta}\frac{1}{\widehat{\delta}}\frac{\partial h}{\partial x} + \frac{\partial A}{\partial x}\right). \tag{87}$$

[2] Following Samelson (1999).

Substituting these into (65) we obtain, omitting the time-derivative,

$$\begin{aligned}\widehat{\delta}\left[\frac{1}{f}\left(A_{\zeta x}A_{\zeta\zeta y}-A_{\zeta y}A_{\zeta\zeta x}\right)+\frac{\beta}{f^2}A_xA_{\zeta\zeta\zeta}\right]+\frac{\beta}{f^2}h_xA_\zeta A_{\zeta\zeta\zeta}\\ +\frac{1}{f}\left[h_x\left(A_{\zeta\zeta}A_{\zeta\zeta y}-A_{\zeta y}A_{\zeta\zeta\zeta}\right)+h_y\left(A_{\zeta x}A_{\zeta\zeta\zeta}-A_{\zeta\zeta}A_{\zeta\zeta x}\right)\right] \\ =\frac{\kappa}{B_0\widehat{\delta}^2}A_{\zeta\zeta\zeta\zeta},\end{aligned} \tag{88}$$

where the subscripts on A and h denote derivatives. If $h_x = h_y = 0$, that is if the base of the thermocline is flat, then (88) becomes

$$\frac{1}{f}\left[A_{\zeta x}A_{\zeta\zeta y}-A_{\zeta y}A_{\zeta\zeta x}\right]+\frac{\beta}{f^2}A_xA_{\zeta\zeta\zeta}=\frac{\kappa}{B_0\widehat{\delta}^3}A_{\zeta\zeta\zeta\zeta}. \tag{89}$$

Since all the terms in this equation are, by construction, order one, we immediately see that the non-dimensional boundary layer thickness $\widehat{\delta}$ scales as

$$\widehat{\delta}\sim\left(\frac{\kappa}{B_0}\right)^{1/3}, \tag{90}$$

just as in the one-dimensional model. On the other hand, if h_x and h_y are order-one quantities then the dominant balance in (88) is

$$\frac{1}{f}\left[h_x(A_{\zeta\zeta}A_{\zeta\zeta y}-A_{\zeta y}A_{\zeta\zeta\zeta})+h_y(A_{\zeta x}A_{\zeta\zeta\zeta}-A_{\zeta\zeta}A_{\zeta\zeta x})\right]=\frac{\kappa}{B_0\widehat{\delta}^2}A_{\zeta\zeta\zeta\zeta} \tag{91}$$

and

$$\widehat{\delta}\sim\left(\frac{\kappa}{B_0}\right)^{1/2}, \tag{92}$$

confirming the heuristic scaling arguments. Thus, if the isotherm slopes are fixed independently of κ (for example, by the wind stress), then as $\kappa\to 0$ an internal boundary layer will form whose thickness is proportional to $\kappa^{1/2}$. We expect this to occur at the base of the main thermocline, with purely advective dynamics being dominant in the upper part of the thermocline, and determining the slope of the isotherms (i.e., the form of h_x and h_y). Interestingly, the balance in the three-dimensional boundary layer equation does not in general correspond locally to $wT_z \approx \kappa T_{zz}$. Both at $\mathcal{O}(1)$ and $\mathcal{O}(\delta)$ the horizontal advective terms in (88) are of the same asymptotic size as the vertical advection terms. In the boundary layer the thermodynamic balance is thus $\mathbf{u}\cdot\nabla_z T + wT_z \approx \kappa T_{zz}$, whether the isotherms are sloping or flat. We might have anticipated this, because the vertical velocity passes through zero within the boundary layer.

References

Gnanadesikan, A., 1999. A simple predictive model for the structure of the oceanic pycnocline. *Science*, **283**, 2077–2079.

Nikurashin, M. & Vallis, G. K., 2011. A theory of deep stratification and overturning circulation in the ocean. *J. Phys. Oceanogr.*, **41**, 485–502.

Nikurashin, M. & Vallis, G. K., 2012. A theory of the interhemispheric meridional overturning circulation and associated stratification. *J. Phys. Oceanogr.*, **42**, 1652–1667.

Samelson, R. M., 1999. Geostrophic circulation in a rectangular basin with a circumpolar connection. *J. Phys. Oceanogr*, **29**, 3175-3184

Welander, P., 1971. The Thermocline Problem. *Phil. Trans. Royal Soc. London, A*, **270**, 415–421.

Lecture 10a: The Hadley Cell

Geoff Vallis; notes by Jim Thomas and Geoff J. Stanley

June 27

In this short lecture we take a look at the general circulation of the atmosphere, and in particular the Hadley cell. Look again at the zonally averaged circulation in the top panel of (lec.2:fig.2). The centre two circulations are the Hadley cells. Deep tropical convection lifts air near at the Intertropical Convergence Zone (ITCZ) near the equator. At the tropopause its vertical motion is inhibited by strong static stability, so it begins a poleward migration which extends as far as some critical latitude ϑ_H. In this lecture we will attempt to explain *why* the Hadley cell terminates at ϑ_H, and not some other latitude. We will outline three possibilities.

1. The Hadley cell is terminated in order to satisfy certain thermodynamic constraints, described in section 1.
2. The Hadley cell is terminated by the onset of baroclinic instability, described in section 2.
3. The Hadley cell is terminated by the effects of the breaking of Rossby waves, described in section 3.

Almost certainly none of these models describes the real Hadley cell in anything other than an approximate way, but this does not mean they are not useful.

1 A Zonally Symmetric Steady Model of the Hadley cell

We begin with a a model of the zonally symmetric circulation – that is, the circulation has no eddies, in fact no variation at all in the zonal direction. A parcel of air moving polewards away from the boundary layer will then conserve its axial angular momentum, as shown in Figure 1. To construct a mathematical model, following Schneider & Lindzen (1977) and Held & Hou (1980), we suppose the following.

1. The circulation is steady.
2. The polewards moving air conserves its axial angular momentum, whereas the zonal flow associated with the near-surface, equatorwards moving flow is frictionally retarded and weak.
3. The circulation is in thermal wind balance.
4. The flow is symmetric about the equator. Seasons can in fact be added to such a model.

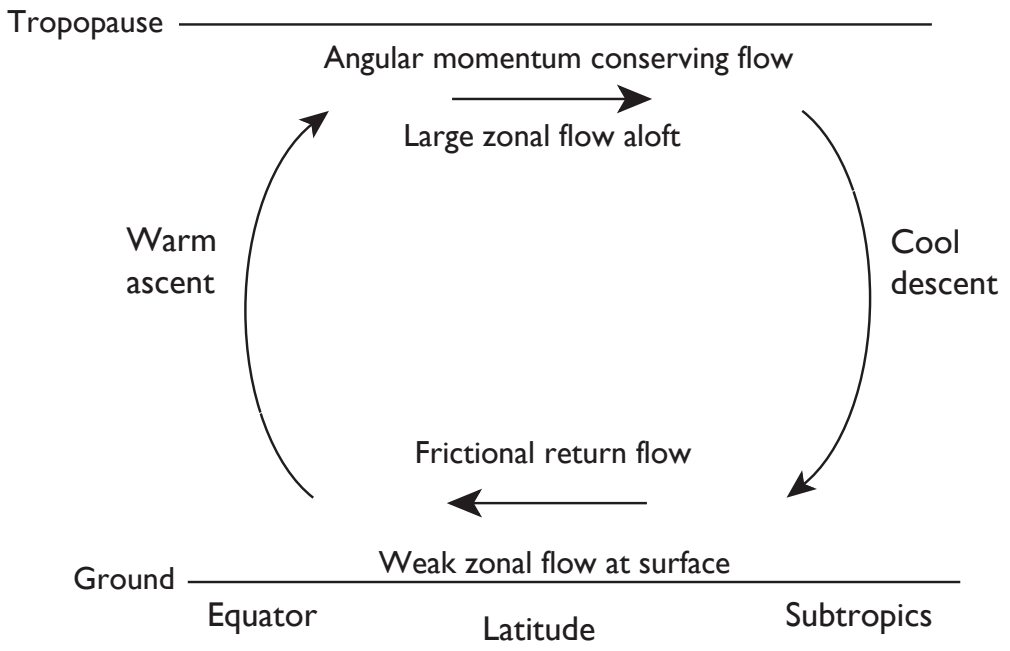


Figure 1: A simple model of the Hadley cell. Rising air near the equator moves polewards near the tropopause, descending in the subtropics and returning near the surface. The polewards moving air conserves its axial angular momentum, leading to a zonal flow that increases away from the equator. By the thermal wind relation the temperature of the air falls as it moves polewards, and to satisfy the thermodynamic budget it sinks in the subtropics. The return flow at the surface is frictionally retarded and small.

1.1 Angular momentum conservation

Momentum equation:

$$\frac{\partial \overline{u}}{\partial t} - (f + \overline{\zeta})\overline{v} + \overline{w}\frac{\partial \overline{u}}{\partial z} = -\frac{1}{a\cos^2\vartheta}\frac{\partial}{\partial \vartheta}(\cos^2\vartheta \overline{u'v'}) - \frac{\partial \overline{u'w'}}{\partial z}, \tag{1}$$

where $\overline{\zeta} = -(a\cos\vartheta)^{-1}\partial_\vartheta(\overline{u}\cos\vartheta)$ and the overbars represent zonal averages. We simplify this to

$$(f + \overline{\zeta})\overline{v} = 0. \tag{2}$$

It is easy to show that this is equivalent to

$$2\Omega\sin\vartheta = \frac{1}{a}\frac{\partial \overline{u}}{\partial \vartheta} - \frac{\overline{u}\tan\vartheta}{a}. \tag{3}$$

and the solution is (c.f. Fig. 2)

$$\overline{u} = \Omega a\frac{\sin^2\vartheta}{\cos\vartheta} \equiv U_M. \tag{4}$$

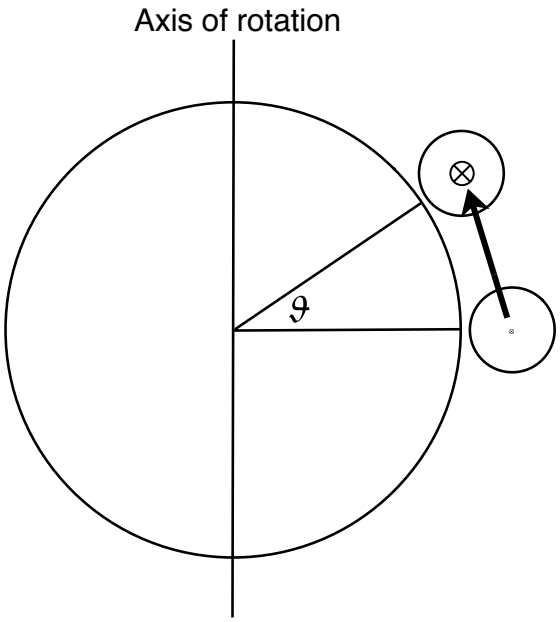


Figure 2: If a ring of air at the equator moves polewards it moves closer to the axis of rotation. If the parcels in the ring conserve their angular momentum their zonal velocity must increase; thus, if $m = (\overline{u} + \Omega a \cos\vartheta) a \cos\vartheta$ is preserved and $\overline{u} = 0$ at $\vartheta = 0$ we recover (4).

Temperature field

Thermal wind balance:

$$2\Omega \sin\vartheta \frac{\partial u}{\partial z} = -\frac{1}{a}\frac{\partial b}{\partial \vartheta}, \tag{5}$$

where $b = g\,\delta\theta/\theta_0$ is the buoyancy and $\delta\theta$ is the deviation of potential temperature from a constant reference value θ_0. (Be reminded that θ is potential temperature, whereas ϑ is latitude.) Vertically integrating from the ground to the height H where the outflow occurs and substituting (4) for u yields

$$\frac{1}{a\theta_0}\frac{\partial \theta}{\partial \vartheta} = -\frac{2\Omega^2 a}{gH}\frac{\sin^3\vartheta}{\cos\vartheta}, \tag{6}$$

where $\theta = H^{-1}\int_0^H \delta\theta\,\mathrm{d}z$ is the vertically averaged potential temperature. If the latitudinal extent of the Hadley cell is not too great we can make the small-angle approximation, and replace $\sin\vartheta$ by ϑ and $\cos\vartheta$ by one, then integrating (6) gives

$$\theta = \theta(0) - \frac{\theta_0 \Omega^2 y^4}{2gHa^2}, \tag{7}$$

where $y = a\vartheta$ and $\theta(0)$ is the potential temperature at the equator, as yet unknown. Away from the equator, the zonal velocity given by (4) increases rapidly polewards and the temperature correspondingly drops. How far polewards is this solution valid? And what determines the value of the integration constant $\theta(0)$? To answer these questions we turn to thermodynamics.

1.2 Thermodynamics

In the above discussion, the temperature field is slaved to the momentum field in that it seems to follow passively from the dynamics of the momentum equation. Nevertheless, the thermodynamic equation must still be satisfied. Let us assume that the thermodynamic forcing can be represented by a Newtonian cooling to some specified radiative equilibrium temperature, θ_E; this is a severe simplification, especially in equatorial regions where the release of heat by condensation is important. The thermodynamic equation is then

$$\frac{\mathrm{D}\theta}{\mathrm{D}t} = \frac{\theta_E - \theta}{\tau}, \tag{8}$$

where τ is a relaxation time scale, perhaps a few weeks. Let us suppose that θ_E falls monotonically from the equator to the pole, and that it increases linearly with height, and a simple representation of this is

$$\frac{\theta_E(\vartheta, z)}{\theta_0} = 1 - \frac{2}{3}\Delta_H P_2(\sin\vartheta) + \Delta_V \left(\frac{z}{H} - \frac{1}{2}\right), \tag{9}$$

where Δ_H and Δ_V are non-dimensional constants that determine the fractional temperature difference between the equator and the pole, and the ground and the top of the fluid, respectively. P_2 is the second Legendre polynomial. At $z = H/2$, or for the vertically averaged field, this approximates to

$$\theta_E = \theta_{E0} - \Delta\theta \left(\frac{y}{a}\right)^2, \tag{10}$$

where θ_{E0} is the equilibrium temperature at the equator, $\Delta\theta$ determines the equator–pole radiative-equilibrium temperature difference, and

$$\theta_{E0} = \theta_0(1 + \Delta_H/3), \qquad \Delta\theta = \theta_0\Delta_H. \tag{11}$$

Now, let us suppose that the solution (7) is valid between the equator and a latitude ϑ_H where $v = 0$, so that within this region the system is essentially closed. Conservation of potential temperature then requires that the solution (7) must satisfy

$$\int_0^{Y_H} \theta \,\mathrm{d}y = \int_0^{Y_H} \theta_E \,\mathrm{d}y, \tag{12}$$

where $Y_H = a\vartheta_H$ is as yet undetermined. Polewards of this, the solution is just $\theta = \theta_E$. Now, we may demand that the solution be continuous at $y = Y_H$ (without temperature continuity the thermal wind would be infinite) and so

$$\theta(Y_H) = \theta_E(Y_H). \tag{13}$$

The constraints (12) and (13) determine the values of the unknowns $\theta(0)$ and Y_H. A little algebra gives

$$Y_H = \left(\frac{5\Delta\theta g H}{3\Omega^2\theta_0}\right)^{1/2}, \tag{14}$$

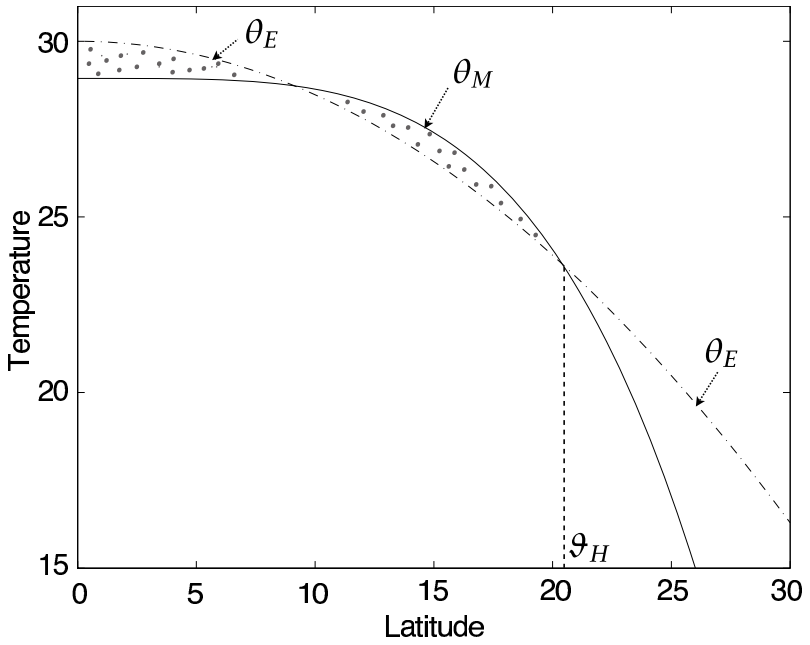


Figure 3: The radiative equilibrium temperature (θ_E, dashed line) and the angular-momentum-conserving solution (θ_M, solid line) as a function of latitude. The two dotted regions have equal areas. The parameters are: $\theta_{EO} = 303\,\mathrm{K}$, $\Delta\theta = 50\,\mathrm{K}$, $\theta_0 = 300\,\mathrm{K}$, $\Omega = 7.272 \times 10^{-5}\,\mathrm{s}^{-1}$, $g = 9.81\,\mathrm{m\ s}^{-2}$, $H = 10\,\mathrm{km}$. These give $R = 0.076$ and $Y_H/a = 0.356$, corresponding to $\vartheta_H = 20.4^\circ$.

and

$$\theta(0) = \theta_{E0} - \left(\frac{5\Delta\theta^2 gH}{18a^2\Omega^2\theta_0}\right). \tag{15}$$

A useful non-dimensional number that parameterizes these solutions is

$$R \equiv \frac{gH\Delta\theta}{\theta_0\Omega^2a^2} = \frac{gH\Delta_H}{\Omega^2a^2}, \tag{16}$$

which is the square of the ratio of the speed of shallow water waves to the rotational velocity of the Earth, multiplied by the fractional temperature difference from equator to pole. Typical values for the Earth's atmosphere are a little less than 0.1. In terms of R we have

$$Y_H = a\left(\frac{5}{3}R\right)^{1/2}, \tag{17}$$

and

$$\theta(0) = \theta_{E0} - \left(\frac{5}{18}R\right)\Delta\theta\,. \tag{18}$$

The solution, (7) with $\theta(0)$ given by (18) is plotted in Fig. 3. Perhaps the single most important aspect of the model is that it predicts that the Hadley cell has a *finite* meridional extent, *even for an atmosphere that is completely zonally symmetric.* The baroclinic instability that does occur in mid-latitudes is not necessary for the Hadley cell to terminate

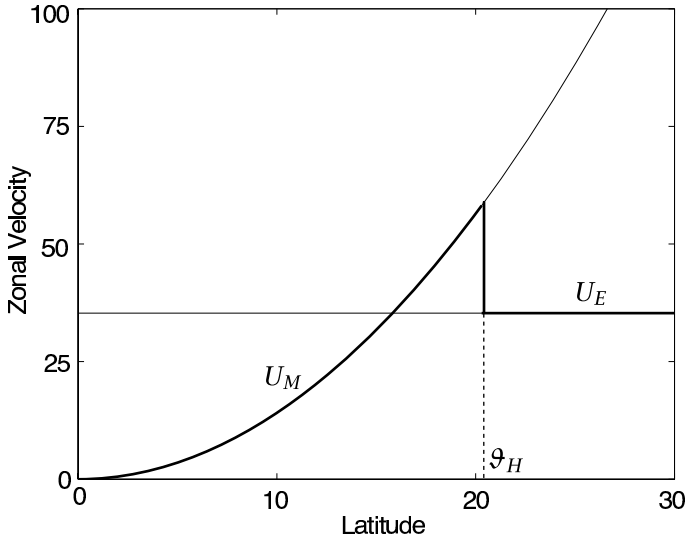


Figure 4: The zonal wind corresponding to the radiative equilibrium temperature (U_E,) and the angular-momentum-conserving solution (U_M) as a function of latitude, given (19) and (20) respectively. The zonal wind (in the model) follows the thick solid line: $u = U_m$ for $\vartheta < \vartheta_H$ ($y < Y_H$), and $u = U_E$ for $\vartheta > \vartheta_H$ ($y > Y_H$), and so has a discontinuity at ϑ_H.

in the subtropics, although it may be an important factor, or even the determining factor, in the real world.

1.3 Zonal wind

The angular-momentum-conserving zonal wind is given by (4), which in the small-angle approximation becomes

$$U_M = \Omega \frac{y^2}{a}. \tag{19}$$

This relation holds for $y < Y_H$. The zonal wind corresponding to the radiative-equilibrium solution is given using thermal wind balance and (10), which leads to

$$U_E = \Omega a R, \tag{20}$$

and this holds polewards of Y_H, or ϑ_H, as sketched in Fig. 4.

2 Baroclinic Instability and Termination of the Hadley Cell

One mechanism that could halt the Hadley cell is baroclinic instability. Having assumed that the surface winds are weak, and knowing the upper level zonal velocity from (4), the shear $\partial U_M/\partial z$ is determined by the height of the tropopause H, which we suppose to be a constant. At some latitude ϑ_C the shear will become baroclinically unstable at which point any assumption of zonal symmetry will break down and the Hadley cell will terminate. What model of baroclinic instability should we use to calculate this? The Eady model has

no critical shear — all shears are unstable — but it has no beta-effect and beta is almost certainly important. The Charney model has beta, but it too has no critical shear. However, small shears give rise to shallow, weak instabilities that may not be important. Thus, we are led to the two-level Phillips model of baroclinic instability, because it accounts for the β effect.

In the Phillips model, a flow becomes unstable when it reaches a critical velocity difference between upper and lower levels given by

$$U = U_1 - U_2 = \frac{1}{4}\beta L_d^2 \tag{21}$$

where $L_d = NH/f$ is the baroclinic deformation radius, and on the sphere $\beta = 2\Omega\cos\phi/a$. Both β and the f hiding in L_d make this U grow towards the equator and decay towards the pole.

Now, from the Hadley cell solution

$$U = \Omega a \frac{\sin^2\phi}{\cos\phi} \tag{22}$$

so that, modulo constant factors, the Hadley cell terminates when

$$\frac{\sin^4\phi_c}{\cos^2\phi_c} = \frac{N^2H^2}{\Omega^2a^2}, \tag{23}$$

or, with a small angle approximation,

$$\vartheta_H \approx \left(\frac{N^2H^2}{\Omega^2a^2}\right)^{1/4} \sim (NH)^{1/2}. \tag{24}$$

As we discussed previously, both theory and modelling suggest that the tropopause will move higher as Global Warming progresses. This model shows that such an increase in H should be accompanied by a poleward expansion of the Hadley cell, perhaps by 1° – 2° over the 21^{st} century. But perhaps even more significant will be the changes in N^2, which is essentially set by the moist adiabatic lapse rate. A warmer atmosphere will hold more moisture by Clausius-Clapeyron (assuming no major changes in the relative humidity) which reduces the moist adiabatic lapse rate and reduces N^2. Thus the Hadley cell might in fact shrink equatorward based on this reasoning. However, the value of N in (24) should be evaluated at ϑ_H where the baroclinic instability occurs. This is not determined by the moist adiabatic lapse rate, and indeed model results suggest that subtropic static stability may increase with global warming, which would lead to an expansion of the Hadley cell.

3 Effect of Rossby-wave breaking

We will conclude this lecture with an outline of a third model for the extent of the Hadley cell. Recall we had reduced the zonal momentum equation (1) to a balance of two terms; let us now include a third for the momentum balance within the Hadley cell:

$$(f + \overline{\zeta})\overline{v} = -\frac{\partial}{\partial y}\left(\overline{u'v'}\right). \tag{25}$$

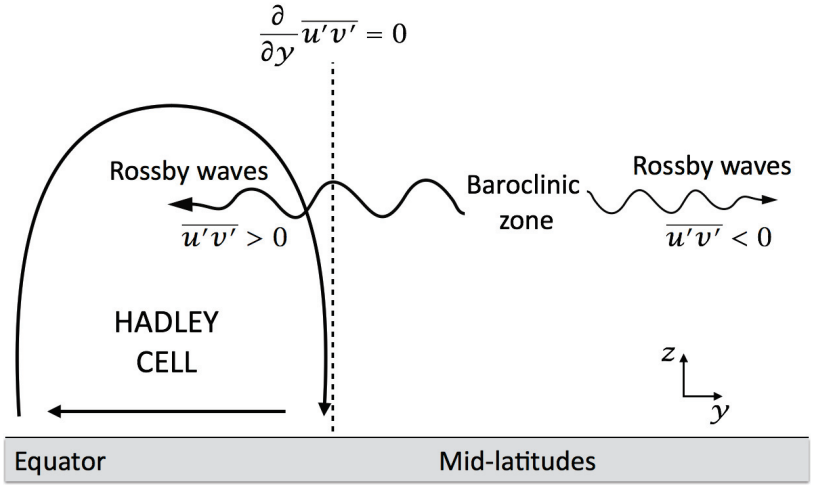


Figure 5: A schematic for the mechanism described in section 3. Rossby waves are generated through baroclinic instability at mid-latitudes, accelerating the flow eastwards: $\partial(\overline{u'v'})/\partial y > 0$. Some propagate equatorwards, and deposit westward momentum, $\partial(\overline{u'v'})/\partial y < 0$, near the critical latitude inside the Hadley cell. At some latitude the Rossby wave momentum flux is neither convergent nor divergent, $\partial(\overline{u'v'})/\partial y = 0$, corresponding to the edge of the Hadley cell.

However, at the edge of the Hadley cell we have $\overline{v} = 0$, and thus $\partial_y\left(\overline{u'v'}\right) = 0$. This is not necessarily the latitude where the flow is baroclinically unstable (Section 2). Rather, baroclinic instability, occurring at some latitude possibly poleward of here, generates Rossby waves; some of these propagate equatorwards and attenuate as they approach a critical latitude where the mean zonal wind matches the Rossby wave's phase speed (see the discussion of lecture 10b). Recalling our previous discussion, angular momentum conservation initiates a situation with weak winds in low-latitudes and strongly eastward winds in mid-latitudes. Thus a Rossby wave generated at mid-latitude has a phase speed somewhat less than the peak eastward wind speed, but certainly still positive for realistic parameters. This Rossby wave, then, will encounter a critical latitude equatorward of which it cannot flow. The wave breaks near this critical latitude and accelerates the zonal wind westward. This acceleration means that the next Rossby wave will encounter its critical latitude slightly more polewards. We thus have a situation in which the Rossby wave momentum flux convergence $\partial(\overline{u'v'})/\partial y$ is positive in the mid-latitudes and negative in the low-latitudes, requiring a zero crossing $\partial(\overline{u'v'})/\partial y = 0$ at some latitude in between, shown schematically in Figure 5. This, as was argued through (25), is the edge of the Hadley cell. Note that this edge is equatorward of where the baroclinic instability occurs (which was taken to be the edge in Section 2). The precise latitude will be established through a feedback between the eastward acceleration by angular momentum conservation and westward acceleration by Rossby wave breaking.

References

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

Schneider, E. K. & Lindzen, R. S., 1977. Axially symmetric steady-state models of the basic state for instability and climate studies. Part I: linearized calculations. *J. Atmos. Sci.*, **34**, 263–279.

Lecture 10b: Rossby Waves and Surface Winds

Geoff Vallis; notes by Jim Thomas and Geoff Stanley

June 27

In this our third lecture we stay with the atmosphere and introduce some dynamics. Our first goal is to understand why there are surface winds, and in particular why there are surface westerlies (Fig. 1). A full explanation of this would require a discussion of baroclinic instability and take up a couple of lectures in itself. We'll skip all that and carry out explicit derivations only for the barotropic vorticity equation, with the reader filling in the gaps phenomenologically. We do note that there are westerly winds aloft in the atmosphere because of the thermal wind relation, $f\partial u/\partial z = \partial b/\partial z$, where b is buoyancy which is like temperature. Thus, a temperature gradient between the equator and the pole implies that the zonal wind increases with height. But this doesn't of itself mean that the surface winds are non-zero – we will need momentum fluxes for that. By the same token, momentum fluxes are not needed to have westerly winds aloft.

We begin with a few basic equations.

1 Momentum Equation

The zonally-averaged momentum, in Cartesian geometry has the form

$$\frac{\partial \overline{u}}{\partial t} - (f + \overline{\zeta})\overline{v} = \frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial \tau}{\partial z} \tag{1}$$

where $f = f_0 + \beta y$ In mid-latitudes we usually neglect the mean advection terms ($\overline{\zeta}\,\overline{v}$ here) which in midlatitudes are small. If we multiply by density and integrate vertically then, in a steady state the terms on the left-hand side both vanish, whence

$$\tau_s = \int_z \rho\overline{u'v'}\,\mathrm{d}z \tag{2}$$

where τ_s is the surface stress, which is roughly proportional to the surface wind: $\tau_s \approx r\overline{u}_s$ where r is a constant. Thus

$$\overline{u}_s \approx \frac{1}{r}\int_z \rho\overline{u'v'}\,\mathrm{d}z. \tag{3}$$

In other words, the surface winds arise because of the eddy convergence of momentum in the atmosphere. Where does this come from? It turns out that it arises from the sphericity of the Earth which gives rise to differential rotation and Rossby waves, as we shall see.

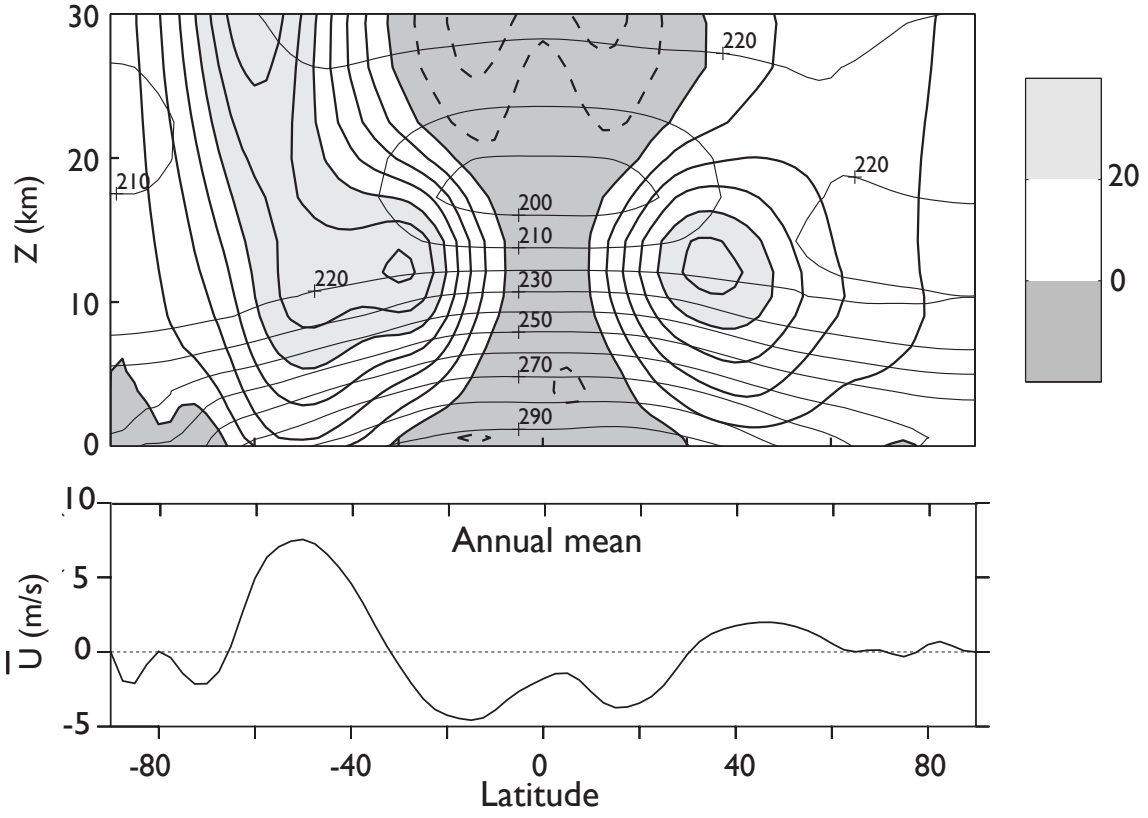


Figure 1: (a) Annual mean, zonally averaged zonal wind (heavy contours and shading) and the zonally averaged temperature (lighter contours). (b) Annual mean, zonally averaged zonal winds at the surface. The wind contours are at intervals of $5\,\mathrm{m\ s^{-1}}$ with shading for eastward winds above $20\,\mathrm{m\ s^{-1}}$ and for all westward winds, and the temperature contours are labelled. The ordinate of (a) and (c) is $Z = -H\log(p/p_R)$, where p_R is a constant, with scale height $H = 7.5\,\mathrm{km}$.

2 Rossby Waves: A Brief Tutorial

The inviscid, adiabatic potential vorticity equation is

$$\frac{\partial q}{\partial t} + \mathbf{u}\ \cdot \nabla q = 0, \tag{4}$$

where $q(x, y, z, t)$ is the potential vorticity and $\mathbf{u}\ (x, y, z, t)$ is the horizontal velocity. The velocity is related to a streamfunction by $u = -\partial\psi/\partial y,\ v = \partial\psi/\partial x$ and the potential vorticity is some function of the streamfunction, which might differ from system to system. Two examples, one applying to a continuously stratified system and the second to a single layer system, are

$$q = f + \zeta + \frac{\partial}{\partial z}\left(S(z)\frac{\partial\psi}{\partial z}\right), \qquad q = \zeta + f - {k_d}^2\psi. \tag{5a,b}$$

We deal mainly with the second. If the basic state is a zonal flow and purely a function of y then

$$q = \overline{q}(y,z) + q'(x,y,t), \qquad \psi = \overline{\psi}(y,z) + \psi'(x,y,z,t) \tag{6}$$

whence

$$\frac{\partial q'}{\partial t} + \overline{\mathbf{u}} \cdot \nabla \overline{q} + \overline{\mathbf{u}} \cdot \nabla q' + \mathbf{u}\,' \cdot \nabla \overline{q} + \mathbf{u}\,' \cdot \nabla q' = 0. \tag{7}$$

Linearizing gives

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0. \tag{8}$$

2.1 Rossby waves in a single layer

In the single-layer case we have $q = \beta y + \nabla^2\psi - {k_d}^2\psi$. If we linearize this around a zonal flow then $\psi = \overline{\psi} + \psi'$ and

$$\overline{\psi} = -\overline{u}y \qquad \overline{q} = \beta y + \overline{u}k_d^2 y \tag{9}$$

and

$$q' = \nabla^2\psi' - {k_d}^2\psi' \tag{10}$$

and (8) becomes

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)(\nabla^2\psi' - \psi' {k_d}^2) + \frac{\partial \psi'}{\partial x}(\beta + Uk_d^2) = 0 \tag{11}$$

Substituting $\psi' = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}$ we obtain the dispersion relation,

$$\omega = \frac{k(UK^2 - \beta)}{K^2 + {k_d}^2} = Uk - k\frac{\beta + U{k_d}^2}{K^2 + {k_d}^2}. \tag{12}$$

We will simplify by taking $U = 0$ whence

$$\omega = -\frac{\beta}{K^2 + {k_d}^2}. \tag{13}$$

The corresponding components of phase speed and group velocity are

$$c_p^x \equiv \frac{\omega}{k} = -\frac{\beta}{K^2 + {k_d}^2}, \qquad c_p^y \equiv \frac{\omega}{l} = \frac{k}{l}\left(\frac{\beta}{K^2 + {k_d}^2}\right) \tag{14a,b}$$

and

$$c_g^x \equiv \frac{\partial \omega}{\partial k} = \frac{\beta(k^2 - l^2 - {k_d}^2)}{\left(K^2 + {k_d}^2\right)^2}, \qquad c_g^y \equiv \frac{\partial \omega}{\partial l} = \frac{2\beta kl}{\left(K^2 + {k_d}^2\right)^2}, \tag{15a,b}$$

which $K^2 = k^2 + l^2$.

3 Momentum Transport in Rossby Waves

It turns out that Rossby waves will transport momentum from place to place, and this is why we have surface winds! (Well, at least it is an explication of why we have surface winds. Other explications that don't involve Rossby waves can be given (Vallis, 2006), but they are all really the same explanation.)

Let us suppose that some mechanism is present that excites Rossby waves in mid-latitudes. This mechanism is in fact baroclinic instability, but we don't really need to know that. We expect that Rossby waves will be generated there, propagate away and break and dissipate. To the extent that the waves are quasi-linear and do not interact, then just away from the source region each wave has the form

$$\psi = \mathrm{Re}\, C \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)} = \mathrm{Re}\, C \mathrm{e}^{\mathrm{i}(kx+ly-kct)}, \tag{16}$$

where C is a constant, with dispersion relation

$$\omega = ck = \overline{u}k - \frac{\beta k}{k^2+l^2} \equiv \omega_R, \tag{17}$$

taking $k_d = 0$ and provided that there is no meridional shear in the zonal flow. The meridional component of the group velocity is given by

$$c_g^y = \frac{\partial \omega}{\partial l} = \frac{2\beta kl}{(k^2+l^2)^2}. \tag{18}$$

Now, the direction of the group velocity must be *away* from the source region; this is a radiation condition, demanded by the requirement that Rossby waves transport energy *away* from the disturbance. Thus, northwards of the source kl is positive and southwards of the source kl is negative. That the product kl can be positive or negative arises because for each k there are two possible values of l that satisfy the dispersion relation (17), namely

$$l = \pm \left(\frac{\beta}{\overline{u}-c} - k^2 \right)^{1/2}, \tag{19}$$

assuming that the quantity in parentheses is positive.

The velocity variations associated with the Rossby waves are

$$u' = -\mathrm{Re}\, C\, \mathrm{i} l \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \qquad v' = \mathrm{Re}\, C\, \mathrm{i} k \mathrm{e}^{\mathrm{i}(kx+ly-\omega t)}, \tag{20a,b}$$

and the associated momentum flux is

$$\overline{u'v'} = -\frac{1}{2} C^2 kl. \tag{21}$$

Thus, given that the sign of kl is determined by the group velocity, northwards of the source the momentum flux associated with the Rossby waves is southward (i.e., $\overline{u'v'}$ is negative), and southwards of the source the momentum flux is northward (i.e., $\overline{u'v'}$ is positive). That is, the momentum flux associated with the Rossby waves is *toward* the source region. Momentum converges in the region of the stirring, producing net eastward flow there and westward flow to either side (Fig. 2).

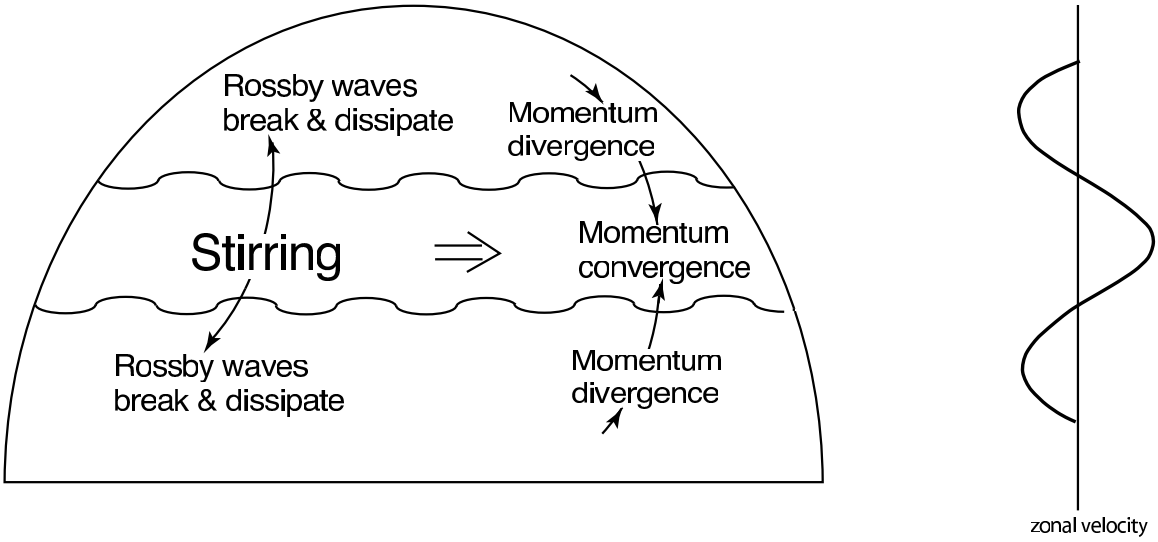


Figure 2: Generation of zonal flow on a β-plane or on a rotating sphere. Stirring in mid-latitudes (by baroclinic eddies) generates Rossby waves that propagate away from the disturbance. Momentum converges in the region of stirring, producing eastward flow there and weaker westward flow on its flanks.

Another way of describing the same effect is to note that if kl is positive then lines of constant phase ($kx + ly =$ constant) are tilted north-west/south-east, as in Fig. 3 and the momentum flux associated with such a disturbance is negative ($\overline{u'v'} < 0$). Similarly, if kl is negative then the constant-phase lines are tilted north-east/south-west and the associated momentum flux is positive ($\overline{u'v'} > 0$). The net result is a convergence of momentum flux into the source region. In physical space this is reflected by having eddies that are shaped like a boomerang, as in Fig. 3.

Pseudomomentum and wave–mean-flow interaction

The kinematic relation between vorticity flux and momentum flux for non-divergent two-dimensional flow is

$$v\zeta = \frac{1}{2}\frac{\partial}{\partial x}\left(v^2 - u^2\right) - \frac{\partial}{\partial y}(uv). \tag{22}$$

After zonal averaging this gives

$$\overline{v'\zeta'} = -\frac{\partial \overline{u'v'}}{\partial y}, \tag{23}$$

noting that $\overline{v} = 0$ for two-dimensional incompressible (or geostrophic) flow.

Now, the barotropic zonal momentum equation is (for horizontally non-divergent flow)

$$\frac{\partial u}{\partial t} + \frac{\partial u^2}{\partial x} + \frac{\partial uv}{\partial y} - fv = -\frac{\partial \phi}{\partial x} + F_u - D_u, \tag{24}$$

where F_u and D_u represent the effects of any forcing and dissipation. Zonal averaging, with

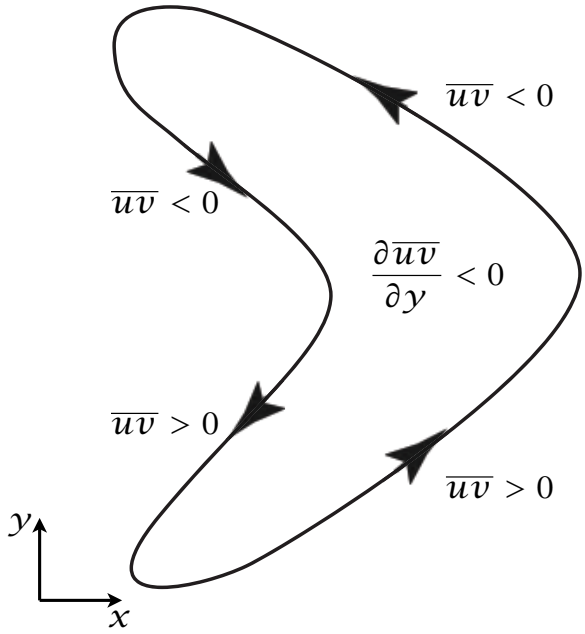


Figure 3: The momentum transport in physical space, caused by the propagation of Rossby waves away from a source in mid-latitudes. The ensuing boomerang-shaped eddies are responsible for a convergence of momentum, as indicated in the idealization pictured.

$\overline{v} = 0$, gives

$$\frac{\partial \overline{u}}{\partial t} = -\frac{\partial \overline{u'v'}}{\partial y} + \overline{F}_u - \overline{D}_u, \tag{25}$$

or, using (23),

$$\frac{\partial \overline{u}}{\partial t} = \overline{v'\zeta'} + \overline{F}_u - \overline{D}_u. \tag{26}$$

Thus, the zonally averaged wind is maintained by the zonally averaged vorticity flux. On average there is little if any direct forcing of horizontal momentum and we may set $\overline{F}_u = 0$, and if the dissipation is parameterized by a linear drag (26) becomes

$$\frac{\partial \overline{u}}{\partial t} = \overline{v'\zeta'} - r\overline{u}, \tag{27}$$

where the constant r is an inverse frictional time scale.

Now consider the maintenance of this vorticity flux. The barotropic vorticity equation is

$$\frac{\partial \zeta}{\partial t} + \mathbf{u} \cdot \nabla\zeta + v\beta = F_\zeta - D_\zeta, \tag{28}$$

where F_ζ and D_ζ are forcing and dissipation of vorticity. Linearize about a mean zonal flow to give

$$\frac{\partial \zeta'}{\partial t} + \overline{u}\frac{\partial \zeta'}{\partial x} + \gamma v' = F'_\zeta - D'_\zeta, \tag{29}$$

where

$$\gamma = \beta - \frac{\partial^2 \overline{u}}{\partial y^2} \tag{30}$$

is the meridional gradient of absolute vorticity. Multiply (29) by ζ'/γ and zonally average, assuming that $\overline{u}_{yy}$ is small compared to β or varies only slowly, to form the pseudomomentum equation,

$$\frac{\partial \mathcal{A}}{\partial t} + \overline{v'\zeta'} = \frac{1}{\gamma}(\overline{\zeta' F'_\zeta} - \overline{\zeta' D'_\zeta}), \tag{31a}$$

$$\mathcal{A} = \frac{1}{2\gamma}\overline{\zeta'^2} \tag{31b}$$

is a wave activity density, equal to the (negative of) the pseudomomentum for this problem. The parameter γ is positive if the average absolute vorticity increases monotonically northwards, and this is usually the case in both Northern and Southern Hemispheres.

3.1 An aside on wave activity and stability

Suppose the flow is unforced and inviscid (common conditions that we impose in stability problems). Then the wave activity equation above becomes

$$\frac{\partial \mathcal{A}}{\partial t} + \overline{v'\zeta'} = 0. \tag{32}$$

This condition holds even in the presence of shear. Integrating between quiescent latitudes gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\int \mathcal{A}\,\mathrm{d}y = 0. \tag{33}$$

The quantity $\widehat{A} \equiv \int \mathcal{A}\,\mathrm{d}y$ is wave activity, something that is quadratic in wave amplitude and is conserved. $\mathcal{A}$ itself is a wave activity density. Energy is not normally a wave activity, because it grows if the flow is unstable, whereas a wave activity does not.

Now suppose that γ is positive everywhere. In this case the conservation of $\widehat{A}$ prevents $\overline{\zeta'^2}$ from growing! Thus, for a wave to grow, $\beta - \overline{u}_{yy}$ must change sign somewhere in the domain. We have derived the *Rayleigh-Kuo* criterion for barotropic instability. Note that there is no mention of normal modes, although we have still (in this derivation) assumed linearity.

4 Wave–mean-flow interaction, acceleration and non-acceleration

In the absence of forcing and dissipation, (27) and (31a) imply an important relationship between the change of the mean flow and the pseudomomentum, namely

$$\frac{\partial \overline{u}}{\partial t} + \frac{\partial \mathcal{A}}{\partial t} = 0. \tag{34}$$

We have now essentially derived a special case of the *non-acceleration* result. If the waves are steady and inviscid, then from (31a) $\overline{v'\zeta'} = 0$. Then from (34) the mean flow does not

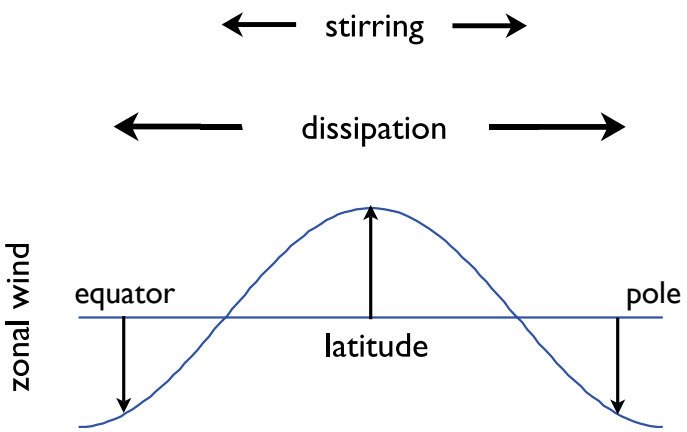


Figure 4: Mean flow generation by a meridionally confined stirring. Because of Rossby wave propagation away from the source region, the distribution of pseudomomentum dissipation is broader than that of pseudomomentum forcing, and the sum of the two leads to the zonal wind distribution shown, with positive (eastward) values in the region of the stirring. See also Fig. 6.

accelerate. We need to do a bit more work in the stratified case, but the essence of the result is the same.

Now if for some reason $\mathcal{A}$ increases, perhaps because a wave enters an initially quiescent region because of stirring elsewhere, then mean flow must decrease. However, because the vorticity flux integrates to zero, the zonal flow cannot decrease everywhere. Thus, if the zonal flow decreases in regions away from the stirring, it must *increase* in the region of the stirring. In the presence of forcing and dissipation this mechanism can lead to the production of a statistically steady jet in the region of the forcing, since (27) and (31a) combine to give

$$\frac{\partial \overline{u}}{\partial t} + \frac{\partial \mathcal{A}}{\partial t} = -r\overline{u} + \frac{1}{\gamma}(\overline{\zeta' F'_\zeta} - \overline{\zeta' D'_\zeta}), \tag{35}$$

and in a statistically steady state

$$r\overline{u} = \frac{1}{\gamma}(\overline{\zeta' F'_\zeta} - \overline{\zeta' D'_\zeta}). \tag{36}$$

The terms on the right-hand side represent the stirring and dissipation of vorticity, and integrated over latitude their sum will vanish, or otherwise the pseudomomentum budget cannot be in a steady state. However, let us suppose that forcing is confined to mid-latitudes. In the forcing region, the first term on the right-hand side of (36) will be larger than the second, and an eastward mean flow will be generated. Away from the direct influence of the forcing, the dissipation term will dominate and westward mean flows will be generated, as sketched in Fig. 4. Thus, *on a β -plane or on the surface of a rotating sphere an eastward mean zonal flow can be maintained by a vorticity stirring that imparts no net momentum to the fluid.* In general, stirring in the presence of a vorticity gradient will give rise to a mean flow, and on a spherical planet the vorticity gradient is provided by differential rotation.

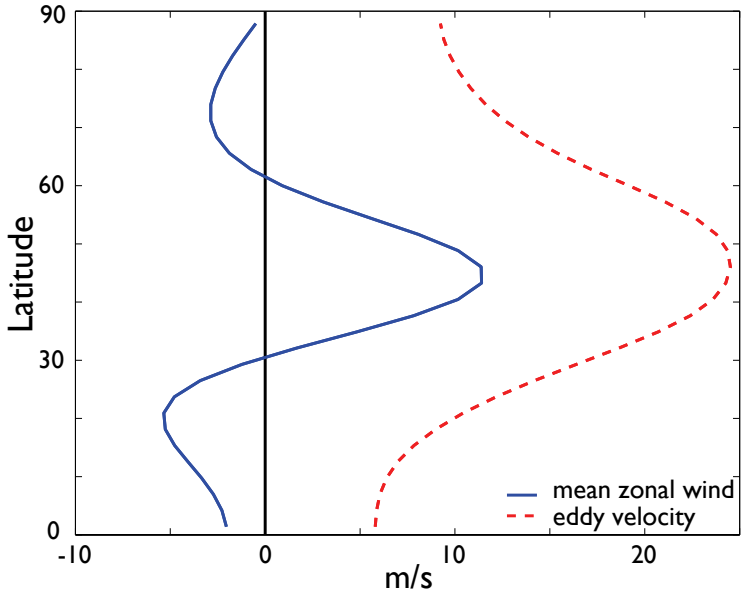


Figure 5: The time and zonally averaged wind (solid line) obtained by an integration of the barotropic vorticity equation on the sphere. The fluid is stirred in mid-latitudes by a random wavemaker that is statistically zonally uniform, acting around zonal wavenumber 8, and that supplies no net momentum. Momentum converges in the stirring region leading to an eastward jet with a westward flow to either side, and zero area-weighted spatially integrated velocity. The dashed line shows the r.m.s. (eddy) velocity created by the stirring.

It is crucial to the generation of a mean flow that the dissipation has a broader latitudinal distribution than the forcing: if all the dissipation occurred in the region of the forcing then from (36) no mean flow would be generated. However, Rossby waves are generated in the forcing region, and these propagate meridionally before dissipating thus broadening the dissipation distribution and allowing the generation of a mean flow.

5 Rossby Waves in an Inhomogeneous Medium

Consider the horizontal problem with infinite deformation radius and linearized equation of motion

$$\left(\frac{\partial}{\partial t} + \overline{u}(y)\frac{\partial}{\partial x}\right) q' + v'\frac{\partial \overline{q}}{\partial y} = 0, \tag{37}$$

where $q' = \nabla^2\psi'$, $v' = \partial\psi'/\partial x$ and $\partial\overline{q}/\partial y = \beta - \overline{u}_{yy}$. If $\overline{u}$ and $\partial\overline{q}/\partial y$ do not vary in space then we may seek wavelike solutions in the usual way and obtain the dispersion relation

$$\omega \equiv ck = \overline{u}k - \frac{\partial\overline{q}/\partial y}{k}k^2 + l^2 \tag{38}$$

where k and l are the x - and y -wavenumbers.

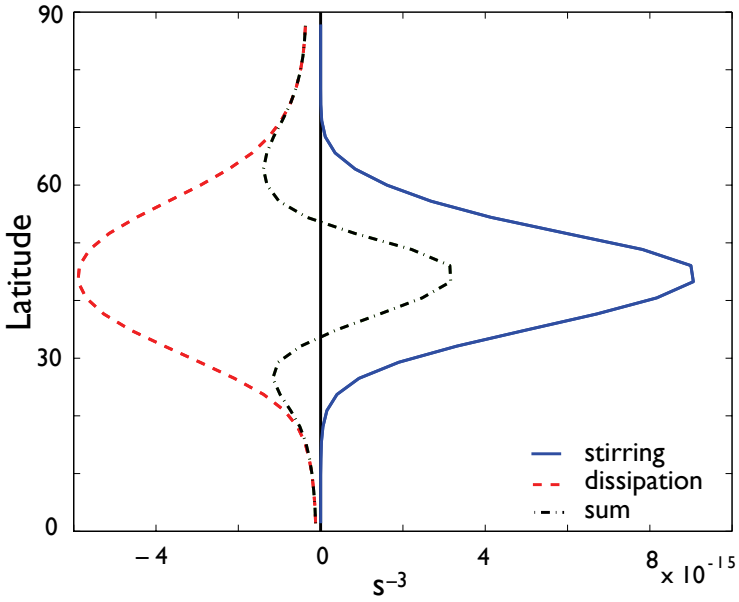


Figure 6: The pseudomomentum stirring (solid line, $\overline{F'_\zeta \zeta'}$), dissipation (dashed line, $\overline{D'_\zeta \zeta'}$) and their sum (dot–dashed), for the same integration as Fig. 5. Because Rossby waves propagate away from the stirred region before breaking, the distribution of dissipation is broader than the forcing, resulting in an eastward jet where the stirring is centred, with westward flow on either side.

If the parameters do vary in the y-direction then we seek a solution of the form $\psi' = \widetilde{\psi}(y)\exp[\mathrm{i}k(x-ct)]$ and obtain

$$\frac{\partial^2 \widetilde{\psi}}{\partial y^2} + l^2(y)\widetilde{\psi} = 0, \quad \text{where} \quad l^2(y) = \frac{\partial \overline{q}/\partial y}{\overline{u}-c} - k^2 \tag{39a,b}$$

If the parameter variation is sufficiently small, occurring on a spatial scale longer than the wavelength of the waves, then we may expect that the disturbance will propagate locally as a plane wave. The solution is then of WKB form namely

$$\widetilde{\psi}(y) = A_0 l^{-1/2} \exp\left(\mathrm{i}\int l \,\mathrm{d}y\right). \tag{40}$$

where A_0 is a constant. The phase of the wave in the y-direction, θ, is evidently given by $\theta = \int l \,\mathrm{d}y$, so that the local wavenumber is given by $\mathrm{d}\theta/\mathrm{d}y = l$. The group velocity is, as before,

$$c_g^x = \overline{u} + \frac{(k^2-l^2)\partial \overline{q}/\partial y}{(k^2+l^2)^2}, \qquad c_g^y = \frac{2kl\,\partial \overline{q}/\partial y}{(k^2+l^2)^2}. \tag{41a,b}$$

The group velocity can now vary spatially, although it is only allowed to vary slowly.

5.1 Wave amplitude

As a Rossby wave propagates its amplitude is not necessarily constant because, in the presence of a shear, the wave may exchange energy with the background state. It goes like

$l^{-1/2}(y)$. This variation can be understood from somewhat more general considerations. As we saw earlier in the simple one-layer case (and discussed more in the appendix) an inviscid, adiabatic wave will conserve its wave activity meaning that

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot \mathcal{F} = 0, \tag{42}$$

where $\mathcal{A}$ is the wave amplitude and $\mathcal{F}$ is the flux, and $\mathcal{F} = \mathbf{c}_g \mathcal{A}$. In the stratified case we have

$$\mathcal{A} = \frac{\overline{q'^2}}{2\partial \overline{q}/\partial y}, \qquad \mathcal{F} = -\overline{u'v'}\,\mathbf{j} + \frac{f_0}{N^2}\overline{v'b'}\,\mathbf{k}, \tag{43}$$

with $\mathcal{F}$ is the Eliassen–Palm (EP) flux, and in the 2D case there is no buoyancy and the $\mathbf{k}$ component is zero. If the waves are steady then $\nabla \cdot \mathcal{F} = 0$, and in the two-dimensional case under consideration this means that $\partial \overline{u'v'}/\partial y = 0$.

Thus, $u'v' = kl|\widetilde{\psi}|^2 =$ constant, and since k is constant the amplitude of a wave varies like

$$|\widetilde{\psi}| = \frac{A_0}{\sqrt{l(y)}} \tag{44}$$

as in the WKB solution. The energy of the wave then varies like

$$\text{Energy} = (k^2 + l^2)\frac{A_0^2}{l}. \tag{45}$$

6 Rossby Wave Propagation in a Slowly Varying Medium

The linear equation of motion is, in terms of streamfunction,

$$\left(\frac{\partial}{\partial t} + \overline{u}(y,z)\frac{\partial}{\partial x}\right)\left[\nabla^2\psi' + \frac{{f_0}^2}{\rho_R}\frac{\partial}{\partial z}\left(\frac{\rho_R}{N^2}\frac{\partial \psi'}{\partial z}\right)\right] + \frac{\partial \psi'}{\partial x}\frac{\partial \overline{q}}{\partial y} = 0. \tag{46}$$

We suppose that the parameters of the problem vary slowly in y and/or z but are uniform in x and t. The frequency and zonal wavenumber are therefore constant. We seek solutions of the form $\psi' = \widetilde{\psi}(y,z)\mathrm{e}^{\mathrm{i}k(x-ct)}$ and find (if, for simplicity, N^2 and ρ_R are constant)

$$\frac{\partial^2 \widetilde{\psi}}{\partial y^2} + \frac{{f_0}^2}{N^2}\frac{\partial^2 \widetilde{\psi}}{\partial z^2} + n^2(y,z)\widetilde{\psi} = 0 \tag{47a}$$

where

$$n^2(y,z) = \frac{\partial \overline{q}/\partial y}{\overline{u} - c} - k^2. \tag{47b}$$

The value of n^2 must be positive in order that waves can propagate, and so waves cease to propagate when they encounter either

1. *A turning line,* where $n^2 = 0$, or
2. *A critical line,* where $\overline{u} = c$ and n^2 becomes infinite.

The bounds may usefully be expressed as a condition on the zonal flow:

$$0 < \overline{u} - c < \frac{\partial \overline{q}/\partial y}{k^2}. \tag{48}$$

If the length scale over which the parameters of the problem vary is much longer than the wavelengths themselves we can expect the solution to look locally like a plane wave and a WKB analysis can be employed. In the purely horizontal problem we assume a solution of the form $\psi' = \widetilde{\psi}(y)\mathrm{e}^{\mathrm{i}k(x-ct)}$ and find

$$\frac{\partial^2 \widetilde{\psi}}{\partial y^2} + l^2(y)\widetilde{\psi} = 0, \qquad l^2(y) = \frac{\partial \overline{q}/\partial y}{\overline{u} - c} - k^2. \tag{49}$$

The solution is of the form

$$\widetilde{\psi}(y) = Al^{-1/2} \exp\Big(\pm \mathrm{i} \int l \, \mathrm{d}y \Big). \tag{50}$$

Thus, $l(y)$ is the local y -wavenumber, and the amplitude of the solution varies like $l^{-1/2}$. At a critical line the amplitude of the wave will go to zero although the energy may become very large, and since the wavelength is small the waves may break. At a turning line the amplitude and energy will both be large, but since the wavelength is long the waves will not necessarily break. A similar analysis may be employed for vertically propagating Rossby waves.

6.1 Two examples

(i) Waves with a turning latitude

A turning line arises where $l = 0$. The line arises if the potential vorticity gradient diminishes to such an extent that $l^2 < 0$ and the waves then cease to propagate in the y -direction. This may happen even in unsheared flow as a wave propagates polewards and the magnitude of beta diminishes.

As a wave packet approaches a turning latitude then l goes to zero so the amplitude, and the energy, of the wave approach infinity. This may happen as a wave propagates polewards and β diminishes. However, the wave will never reach the turning latitude because the meridional component of the group velocity is zero, as can be seen from the expressions for the group velocity, (41). As a wave approaches the turning latitude $c_g^x \to (\beta - \overline{u}_{yy})/k^2$ and $c_g^y \to 0$, so the group velocity is purely zonal and indeed as $l \to 0$

$$\frac{c_g^x - \overline{u}}{c_g^y} = \frac{k}{2l} \to \infty. \tag{51}$$

Because the meridional wavenumber is small the wavelength is large, so we do not expect the waves to break. Rather, we intuitively expect that a wave packet will turn — hence the eponym 'turning latitude' — and be reflected.

To illustrate this, consider waves propagating in a background state that has a beta effect that diminishes polewards but no horizontal shear. To be concrete suppose that $\beta = 5$ at $y = 0$, diminishing linearly to $\beta = 0$ at $y = 0$, and that $\overline{u} - c = 1$ everywhere.

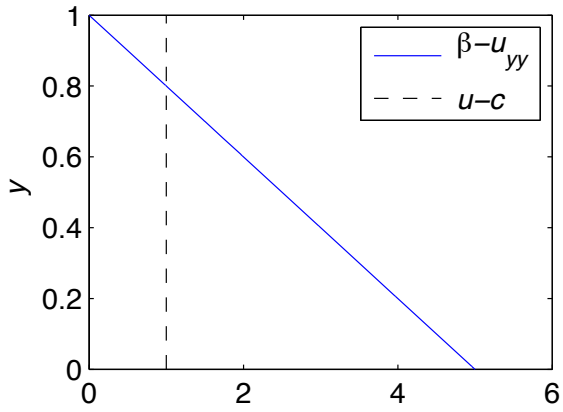


Figure 7: Parameters for the first example considered in section 6.1, with all variables nondimensional. The zonal flow is uniform with $u = 1$ and $c = 0$ (so that $\overline{u}_{yy} = 0$) and β diminishes linearly as y increases polewards as shown. With zonal wavenumber $k = 1$ there is a turning latitude at $y = 0.8$, and the wave properties are illustrated in Fig. 8.

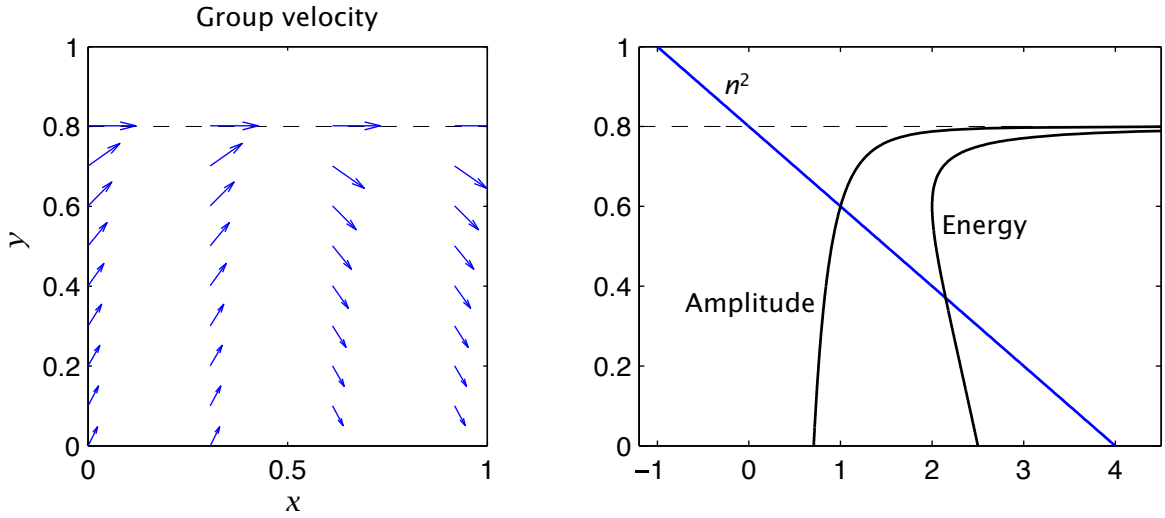


Figure 8: Left: The group velocity evaluated using (41) for the parameters illustrated in Fig. 7, which give a turning latitude at $y = 0.8$. For $x < 0.5$ we choose positive values of n, and a northward group velocity, whereas for $x > 0.5$ we choose negative values of n. Right panel: Values of refractive index squared (n^2), the energy and the amplitude of a wave. n^2 is negative for $y > 0.8$. See text for more description.

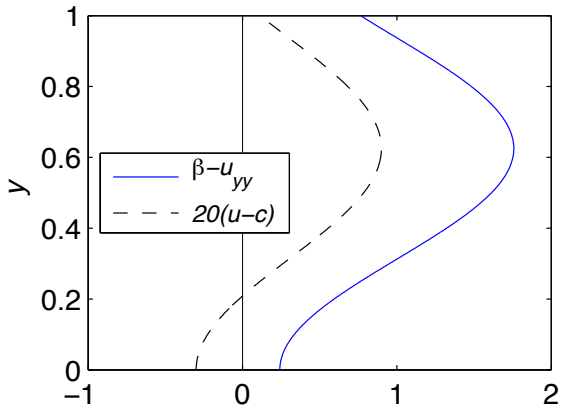


Figure 9: Parameters for the second example considered in section 6.1, with all variables nondimensional. The zonal flow has a broad eastward jet and β is constant. There is a critical line at $y = 0.2$, and with zonal wavenumber $k = 5$ the wave properties are illustrated in Fig. 10.

There is no critical line but depending on the x -wavenumber there may be a turning line, and if we choose $k = 1$ then the turning line occurs when $\beta = 1$ and so at $y = 0.8$. Note that the turning latitude depends on the value of the x -wavenumber — if the zonal wavenumber is larger then waves will turn further south. The parameters are illustrated in Fig. 7.

For a given zonal wavenumber ($k = 1$ in this example) the value of l^2 is computed using (39b), and the components of the group velocity using (41), and these are illustrated in Fig. 8. Note that we may choose either a positive or a negative value of l, corresponding to northward or southward oriented waves, and we illustrate both in the figure. The value of l^2 becomes zero at $y = 0.8$, and this corresponds to a turning latitude. The values of the wave amplitude and energy are computed using (44) and (45) (with an arbitrary amplitude at $y = 0$) and these both become infinite at the turning latitude.

(ii) Waves with a critical latitude

A critical line occurs when $\overline{u} = c$, corresponding to the upper bound of c, and from (39) we see that at a critical line the meridional wavenumber approaches infinity. From (41) we see that both the x - and y -components of the group velocity are zero — a wave packet approaching a critical line just stops. Specifically, as l becomes large

$$c_g^x - \overline{u} \to 0, \qquad c_g^y \to 0, \qquad \frac{c_g^x - \overline{u}}{c_g^y} \to -\frac{l}{k} \to -\infty. \tag{52}$$

From (44) the amplitude of the wave packet also approaches zero, but its energy approaches infinity. Since the wavelength is very small we expect the waves to *break* and deposit their momentum, and this situation commonly arises when Rossby waves excited in midlatitudes propagate equatorward and encounter a critical latitude in the subtropics.

To illustrate this let us construct background state that has an eastward jet in midlatitudes becoming westward at low latitudes, with β constant chosen to be large enough so that

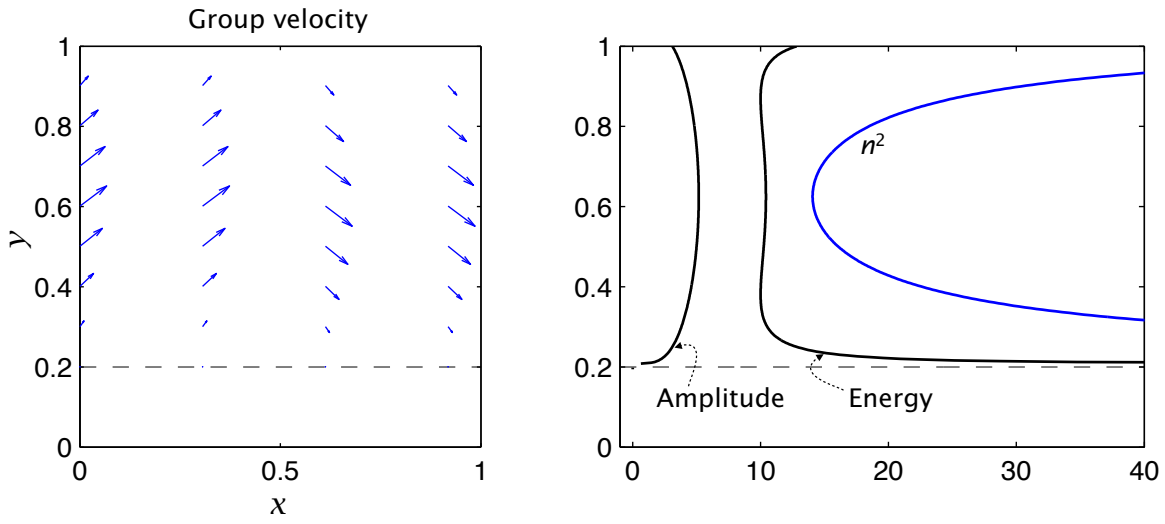


Figure 10: Left: The group velocity evaluated using (41) for the parameters illustrated in Fig. 7, which give a critical line at $y = 0.2$. For $x < 0.5$ we choose positive values of n, and a northward group velocity, whereas for $x > 0.5$ we choose negative values of n. Right panel: Values of refractive index squared, the energy and the amplitude of a wave. The value of n^2 becomes infinite at the critical line. See text for more description.

$\beta - \overline{u}_{yy}$ is positive everywhere. (Specifically, we choose $\beta = 1$ and $\overline{u} = -0.03\sin(8\pi y/5 + \pi/2) - 0.5$), but the precise form is not important.) If $c = 0$ then there is a critical line when $\overline{u}$ passes through zero, which in this example occurs at $x = 0.2$. (The value of $\overline{u} - c$ is small at $y = 1$, but no critical line is actually reached.) These parameters are illustrated in Fig. 9. We also choose $k = 5$, which results in a positive value for l^2 everywhere.

As in the previous example, we compute the value of l^2 using (39b) and the components of the group velocity using (41), and these are illustrated in Fig. 10, with northward propagating waves shown for $x < 0.5$ and southward propagating waves for $x > 0.5$. The value of l^2 increases considerably at the northern and southern edges of the domain, and is actually infinite at the critical line at $y = 0.2$. Using (44) the amplitude of the wave diminishes as the critical line approaches, but the energy increases rapidly, suggesting that the linear approximation will break down. The waves will actually stall before reaching the critical layer, because both the x and the y components of the group velocity become very small. Also, because the wavelength is so small we may expect the waves to break and deposit their momentum, but a full treatment of waves in the vicinity of a critical layer requires a nonlinear analysis.

The situation illustrated in this example is of particular relevance to the maintenance of the zonal wind structure in the troposphere. Waves are generated in midlatitude and propagate equatorward and on encountering a critical layer in the subtropics they break, deposit westward momentum and retard the flow, as the reader who braves the next section will discover explicitly.

7 Rossby Wave Absorption near a Critical Layer

We noted in the last section that as a wave approaches a critical latitude the meridional wavenumber l becomes very large, but the group velocity itself becomes small. These observations suggest that the effects of friction might become very large and that the wave would deposit its momentum, thereby accelerating or decelerating the mean flow, and if we are willing to make one or two approximations we can construct an explicit analytic model of this phenomena. Specifically, we will need to choose a simple form for the friction and assume that the background properties vary slowly, so that we can use a WKB approximation. Note that we have to include some form of dissipation, otherwise the Eliassen–Palm flux divergence is zero and there is no momentum deposition by the waves.

7.1 A model problem

Consider horizontally propagating Rossby waves obeying the linear barotropic vorticity equation on the beta-plane (vertically propagating waves may be considered using similar techniques). The equation of motion is

$$\left(\frac{\partial}{\partial t}+\overline{u}\frac{\partial}{\partial x}\right)\nabla^2\psi+\beta^*\frac{\partial\psi}{\partial x}=-r\nabla^2\psi, \tag{53}$$

where $\beta^*=\beta-\overline{u}_{yy}$. The parameter r is a drag coefficient that acts directly on the relative vorticity. It is not a particularly realistic form of dissipation but its simplicity will serve our purpose well. We shall assume that r is small compared to the Doppler-shifted frequency of the waves and seek solutions of the form

$$\psi'(x,y,t)=\widetilde{\psi}(y)\mathrm{e}^{\mathrm{i}(k(x-ct))}. \tag{54}$$

Substituting into (53) we find, after a couple of lines of algebra, that $\widetilde{\psi}$ satisfies, analogously to (39),

$$\frac{\partial^2\widetilde{\psi}}{\partial y^2}+l^2(y)\widetilde{\psi}=0,\quad\text{where}\quad l^2(y)=\frac{\beta^*}{\overline{u}-c-\mathrm{i}r/k}-k^2. \tag{55a,b}$$

Evidently, as with the inviscid case, if the zonal wind has a lateral shear then l is a function of y. However, l now has an imaginary component so that the wave decays away from its source region. We can already see that if $\overline{u}=c$ the decay will be particularly strong.

7.2 WKB solution

Let us suppose that the zonal wavenumber is small compared to the meridional wavenumber l, which will certainly be the case approaching a critical layer. If $r\ll k(\overline{u}-c)$ then the meridional wavenumber is given by

$$l^2(y)\approx\left[\frac{\beta^*(\overline{u}-c+\mathrm{i}r/k)}{(\overline{u}-c)^2+r^2/k^2}\right]\approx\frac{\beta^*}{\overline{u}-c}\left[1+\frac{\mathrm{i}r}{k(\overline{u}-c)}\right] \tag{56}$$

whence

$$l(y)\approx\left(\frac{\beta^*}{\overline{u}-c}\right)^{1/2}\left[1+\frac{\mathrm{i}r}{2k(\overline{u}-c)}\right]. \tag{57}$$

The streamfunction itself is then given by, in the WKB approximation,

$$\widetilde{\psi} = Al^{-1/2} \exp\left(\pm \mathrm{i} \int^y l \, \mathrm{d}y'\right). \tag{58}$$

But now the wave will decay as it moves away from its source and deposit momentum into the mean flow, as we now calculate.

The momentum flux, F_k, associated with the wave with x -wavenumber of k is given by

$$F_k(y) = \overline{u'v'} = -\mathrm{i}k\left(\psi \frac{\partial \psi^*}{\partial y} - \psi^* \frac{\partial \psi}{\partial y}\right), \tag{59}$$

and using (57) and (58) in (59) we obtain

$$F_k(y) = F_0 \exp\left(\pm \mathrm{i} \int_0^y (l - l^*) \, \mathrm{d}y'\right) = F_0 \exp\left(\int_0^y \frac{\pm r \beta^{*1/2}}{k(\overline{u} - c)^{3/2}} \, \mathrm{d}y'\right). \tag{60}$$

In deriving this expression we use that fact that the amplitude of $\widetilde{\psi}$ (i.e., $l^{-1/2}$) varies only slowly with y so that when calculating $\partial \widetilde{\psi}/\partial y$ its derivative may be ignored. In (60) F_0 is the value of the flux at $y = 0$ and the sign of the exponent must be chosen so that the group velocity is directed away from the wave source region. Clearly, if $r = 0$ then the momentum flux is constant.

The integrand in (60) is the attenuation rate of the wave and it has a straightforward physical interpretation. Using the real part of (57) in (41b), and assuming $|l| \gg |k|$, the meridional component of the group velocity is given by

$$c_g^y = \frac{2kl \, \beta^*}{(k^2 + l^2)^2} \approx \frac{2k \, \beta^*}{l^3} = \frac{2k(\overline{u} - c)^{3/2}}{\beta^{*1/2}}. \tag{61a,b}$$

Thus, we have

$$\text{Wave attenuation rate} = \frac{r\beta^{*1/2}}{k(\overline{u} - c)^{3/2}} = \frac{2 \times \text{Dissipation rate} = 2r}{\text{Meridional group velocity}, c_g^y}. \tag{62}$$

As the group velocity diminishes the dissipation has more time to act and so the wave is preferentially attenuated, a result that we discuss more in the next subsection.

How does this attenuation affect the mean flow? The mean flow is subject to many waves and so obeys the equation

$$\frac{\partial \overline{u}}{\partial t} = -\sum_k \frac{\partial F_k}{\partial y} + \text{viscous terms}. \tag{63}$$

Because the amplitude varies only slowly compared to the phase, the amplitude of $\partial F_k/\partial y$ varies mainly with the attenuation rate (62) and is largest near a critical layer. Consider a Rossby wave propagating away from some source region with a given frequency and x -wavenumber. Because k is negative a Rossby wave always carries westward (or negative) momentum with it. That is, F_k is always negative and increases (becomes more positive) as the wave is attenuated; that is to say, if $r \neq 0$ then $\partial F_k/\partial y$ is positive and from (63)

the mean flow is accelerated *westward* as the wave dissipates. This acceleration will be particularly strong if the wave approaches a critical layer where $\overline{u} = c$. Indeed, such a situation arises when Rossby waves, generated in mid-latitudes, propagate equatorward. As the waves enter the subtropics $\overline{u} - c$ becomes smaller and the waves dissipate, producing a westward force on the mean flow, even though a true critical layer may never be reached. Globally, momentum is conserved because there is an equal and opposite (and therefore eastward) wave force at the wave source producing an eddy-driven jet, as discussed in the previous lecture.

7.3 Interpretation using wave activity

We can derive and interpret the above results by thinking about the propagation of wave activity. For barotropic Rossby waves, multiply (53) by ζ/β^* and zonally average to obtain the wave activity equation,

$$\frac{\partial \mathcal{A}}{\partial t} + \frac{\partial \mathcal{F}}{\partial y} = -\alpha \mathcal{A}, \tag{64}$$

where $\mathcal{A} = \overline{\zeta'^2}/2\beta^*$ is the wave activity density, $\partial \mathcal{F}/\partial y = \overline{v'\zeta'}$ is its flux divergence, and $\alpha = 2r$. Referring as needed to the discussion in sections A.2 and A.3, the flux obeys the group velocity property so that

$$\frac{\partial \mathcal{A}}{\partial t} + \frac{\partial}{\partial y}(\mathbf{c}_g \mathcal{A}) = -\alpha \mathcal{A}. \tag{65}$$

Let us suppose that the wave is in a statistical steady state and that the spatial variation of the group velocity occurs on a longer spatial scale than the variations in wave activity density, consistent with the WKB approximation. We then have

$$c_g^y \frac{\partial \mathcal{A}}{\partial y} = -\alpha \mathcal{A}. \tag{66}$$

which integrates to give

$$\mathcal{A}(y) = \mathcal{A}_0 \exp\left(- \int^y \frac{\alpha}{c_g^y} \, \mathrm{d}y' \right). \tag{67}$$

That is, the attenuation rate of the wave activity is the dissipation rate of wave activity divided by the group velocity, as in (60) and (62) (note that $\alpha = 2r$).

Appendix A: Various properties of Rossby Waves

In this appendix we derive various properties of Rossby waves useful in wave–mean-flow interaction theory, assuming a good knowledge of stratified quasi-geostrophic theory. We use the Boussinesq approximation throughout. This material was not presented in the lectures at Walsh.

A.1 The Eliassen–Palm Flux

The eddy flux of potential vorticity may be expressed in terms of vorticity and buoyancy fluxes as

$$v'q' = v'\zeta' + f_0\, v' \frac{\partial}{\partial z}\left(\frac{b'}{N^2}\right). \tag{68}$$

The second term on the right-hand side can be written as

$$\begin{aligned} f_0\, v' \frac{\partial}{\partial z}\left(\frac{b'}{N^2}\right) &= f_0\, \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - f_0\, \frac{\partial v'}{\partial z}\frac{b'}{N^2} \\ &= f_0\, \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - f_0\, \frac{\partial}{\partial x}\left(\frac{\partial \psi'}{\partial z}\right)\frac{b'}{N^2} \\ &= f_0\, \frac{\partial}{\partial z}\left(\frac{v'b'}{N^2}\right) - \frac{{f_0}^2}{2N^2}\frac{\partial}{\partial x}\left(\frac{\partial \psi'}{\partial z}\right)^2, \end{aligned} \tag{69}$$

using $b' = f_0\, \partial\psi'/\partial z$.

Similarly, the flux of relative vorticity can be written

$$v'\zeta' = -\frac{\partial}{\partial y}(u'v') + \frac{1}{2}\frac{\partial}{\partial x}(v'^2 - u'^2) \tag{70}$$

Using (69) and (70), (68) becomes

$$v'q' = -\frac{\partial}{\partial y}(u'v') + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}v'b'\right) + \frac{1}{2}\frac{\partial}{\partial x}\left((v'^2 - u'^2) - \frac{b'^2}{N^2}\cdot\right) \tag{71}$$

Thus the meridional potential vorticity flux, in the quasi-geostrophic approximation, can be written as the divergence of a vector: $v'q' = \nabla\cdot\mathcal{E}$ where

$$\mathcal{E} \equiv \frac{1}{2}\left((v'^2 - u'^2) - \frac{b'^2}{N^2}\right)\mathbf{i} - (u'v')\,\mathbf{j} + \left(\frac{f_0}{N^2}v'b'\right)\mathbf{k}. \tag{72}$$

A particularly useful form of this arises after zonally averaging, for then (71) becomes

$$\overline{v'q'} = -\frac{\partial}{\partial y}\overline{u'v'} + \frac{\partial}{\partial z}\left(\frac{f_0}{N^2}\overline{v'b'}\right). \tag{73}$$

The vector defined by

$$\mathcal{F} \equiv -\overline{u'v'}\,\mathbf{j} + \frac{f_0}{N^2}\overline{v'b'}\,\mathbf{k} \tag{74}$$

is called the (quasi-geostrophic) *Eliassen–Palm (EP) flux* (Eliassen & Palm (1961)), and its divergence, given by (73), gives the poleward flux of potential vorticity:

$$\overline{v'q'} = \nabla_x\mathcal{F}, \tag{75}$$

where $\nabla_x \equiv (\partial/\partial y, \partial/\partial z)\cdot$ is the divergence in the meridional plane. Unless the meaning is unclear, the subscript x on the meridional divergence will be dropped.

A.2 The Eliassen–Palm relation

On dividing by $\partial\overline{q}/\partial y$ and using (75), the enstrophy equation becomes

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot \mathcal{F} = \mathcal{D}, \tag{76a}$$

where

$$\mathcal{A} = \frac{\overline{q'^2}}{2\partial\overline{q}/\partial y}, \qquad \mathcal{D} = \frac{\overline{D'q'}}{\partial\overline{q}/\partial y}. \tag{76b}$$

Equation (76a) is known as the *Eliassen–Palm relation,* and it is a conservation law for the *wave activity density* $\mathcal{A}$. The conservation law is exact (in the linear approximation) if the mean flow is constant in time. It will be a good approximation if $\partial\overline{q}/\partial y$ varies slowly compared to the variation of $\overline{q'^2}$.

If we integrate (76b) over a meridional area A bounded by walls where the eddy activity vanishes, and if $\mathcal{D} = 0$, we obtain

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_A \mathcal{A}\,\mathrm{d}A = 0. \tag{77}$$

The integral is a wave activity — a quantity that is quadratic in the amplitude of the perturbation and that is conserved in the absence of forcing and dissipation. In this case $\mathcal{A}$ is the negative of the *pseudomomentum,* for reasons we will encounter later. ('Wave action' is a particular form of wave activity; it is the energy divided by the frequency and it is a conserved property in many wave problems.) Note that neither the perturbation energy nor the perturbation enstrophy are wave activities of the linearized equations, because there can be an exchange of energy or enstrophy between mean and perturbation — indeed, this is how a perturbation grows in baroclinic or barotropic instability! This is already evident from an enstrophy equation. Or, in general, take the linearized PV equation with $D' = 0$ and multiply by q' to give the enstrophy equation

$$\frac{1}{2}\frac{\partial q'^2}{\partial t} + \frac{1}{2}\overline{\mathbf{u}} \cdot \nabla q'^2 + \mathbf{u}\,' q' \cdot \nabla\overline{q} = 0, \tag{78}$$

where here the overbar is an average (although it need not be a zonal average). Integrating this over a volume V gives

$$\frac{\mathrm{d}\hat{Z}'}{\mathrm{d}t} \equiv \frac{\mathrm{d}}{\mathrm{d}t}\int_V \frac{1}{2}q'^2\,\mathrm{d}V = -\int_V \mathbf{u}\,' q' \cdot \nabla\overline{q}\,\mathrm{d}V. \tag{79}$$

The right-hand side does not, in general, vanish and so $\hat{Z}'$ is not in general conserved.

A.3 The group velocity property for Rossby waves

The vector $\mathcal{F}$ describes how the wave activity propagates. In the case in which the disturbance is composed of plane or almost plane waves that satisfy a dispersion relation, then $\mathcal{F} = \mathbf{c}_g\mathcal{A}$, where $\mathbf{c}_g$ is the group velocity and (76a) becomes

$$\frac{\partial \mathcal{A}}{\partial t} + \nabla \cdot (\mathcal{A}\mathbf{c}_g) = 0. \tag{80}$$

This is a useful property, because if we can diagnose $\mathbf{c}_g$ from observations we can use (76a) to determine how wave activity density propagates. Let us demonstrate this explicitly for the pseudomomentum in Rossby waves, that is for (76a).

The Boussinesq quasi-geostrophic equation on the β -plane, linearized around a uniform zonal flow and with constant static stability, is

$$\frac{\partial q'}{\partial t} + \overline{u}\frac{\partial q'}{\partial x} + v'\frac{\partial \overline{q}}{\partial y} = 0, \tag{81}$$

where $q' = [\nabla^2 + ({f_0}^2/N^2)\partial^2/\partial z^2]\psi'$ and, if $\overline{u}$ is constant, $\partial\overline{q}/\partial y = \beta$. Thus, we have

$$\left(\frac{\partial}{\partial t} + \overline{u}\frac{\partial}{\partial x}\right)\left[\nabla^2\psi' + \frac{\partial}{\partial z}\left(\frac{{f_0}^2}{N^2}\frac{\partial \psi'}{\partial z}\right)\right] + \beta\frac{\partial \psi'}{\partial x} = 0. \tag{82}$$

Seeking solutions of the form

$$\psi' = \mathrm{Re}\,\widetilde{\psi}\mathrm{e}^{\mathrm{i}(kx+ly+mz-\omega t)}, \tag{83}$$

we find the dispersion relation,

$$\omega = \overline{u}k - \frac{\beta k}{\kappa^2}. \tag{84}$$

where $\kappa^2 = (k^2 + l^2 + m^2{f_0}^2/N^2)$, and the group velocity components:

$$c_g^y = \frac{2\beta k l}{\kappa^4}, \qquad c_g^z = \frac{2\beta k m {f_0}^2/N^2}{\kappa^4}. \tag{85}$$

Also, if $u' = \mathrm{Re}\,\widetilde{u}\exp[\mathrm{i}(kx + ly + mz - \omega t)]$, and similarly for the other fields, then

$$\begin{aligned} \widetilde{u} &= -\mathrm{Re}\ \mathrm{i}l\widetilde{\psi}, & \widetilde{v} &= \mathrm{Re}\ \mathrm{i}k\widetilde{\psi}, \\ \widetilde{b} &= \mathrm{Re}\,\mathrm{i}mf_0\,\widetilde{\psi}, & \widetilde{q} &= -\mathrm{Re}\,\kappa^2\widetilde{\psi}, \end{aligned} \tag{86}$$

The wave activity density is then

$$\mathcal{A} = \frac{1}{2}\frac{\overline{q'^2}}{\beta} = \frac{\kappa^4}{4\beta}|\widetilde{\psi}^2|, \tag{87}$$

where the additional factor of 2 in the denominator arises from the averaging. Using (86) the EP flux, (74), is

$$\mathcal{F}^y = -\overline{u'v'} = \frac{1}{2}kl|\widetilde{\psi}^2|, \qquad \mathcal{F}^z = \frac{f_0}{N^2}\overline{v'b'} = \frac{{f_0}^2}{2N^2}km|\widetilde{\psi}^2|. \tag{88}$$

Using (85), (87) and (88) we obtain

$$\mathcal{F} = (\mathcal{F}^y, \mathcal{F}^z) = \mathbf{c}_g\mathcal{A}. \tag{89}$$

If the properties of the medium are slowly varying, so that a (spatially varying) group velocity can still be defined, then this is a useful expression to estimate how the wave activity propagates in the atmosphere and in numerical simulations.

A.4 Energy flux in Rossby waves

Start with

$$\frac{\partial}{\partial t}\left(\nabla^2 - k_d^2\right)\psi + \beta\frac{\partial \psi}{\partial x} = 0. \tag{90}$$

To obtain an energy equation multiply (90) by $-\psi$ and obtain

$$\frac{1}{2}\frac{\partial}{\partial t}\left((\nabla\psi)^2 + k_d^2\psi^2\right) - \nabla\cdot\left(\psi\nabla\frac{\partial\psi}{\partial t} + \mathbf{i}\frac{\beta}{2}\psi^2\right) = 0, \tag{91}$$

where $\mathbf{i}$ is the unit vector in the x direction. The first group of terms are the energy itself, or more strictly the energy density. (An energy density is an energy per unit mass or per unit volume, depending on the context.) The term $(\nabla\psi)^2/2 = (u^2+v^2)/2$ is the kinetic energy and $k_d^2\psi^2/2$ is the potential energy, proportional to the displacement of the free surface, squared. The second term is the energy flux, so that we may write

$$\frac{\partial E}{\partial t} + \nabla\cdot\mathbf{F} = 0. \tag{92}$$

where $E = (\nabla\psi)^2/2 + k_d\ ^2\psi^2$ and $\mathbf{F} = -\left(\psi\nabla\partial\psi/\partial t + \mathbf{i}\beta\psi^2\right)$. We haven't yet used the fact that the disturbance has a dispersion relation, and if we do so we may expect that the energy moves at the group velocity. Let us now demonstrate this explicitly.

We assume a solution of the form

$$\psi = A(x)\cos(\mathbf{k}\cdot\mathbf{x} - \omega t) = A(x)\cos\left(kx + ly - \omega t\right) \tag{93}$$

where $A(x)$ is assumed to vary slowly compared to the nearly plane wave. (Note that $\mathbf{k}$ is the wave vector, to be distinguished from $\mathbf{k}$, the unit vector in the z -direction.) The kinetic energy in a wave is given by

$$KE = \frac{A^2}{2}\left(\psi_x^2 + \psi_y^2\right) \tag{94}$$

so that, averaged over a wave period,

$$\overline{KE} = \frac{A^2}{2}(k^2+l^2)\frac{\omega}{2\pi}\int_0^{2\pi/\omega}\sin^2(\mathbf{k}\cdot\mathbf{x} - \omega t)\,\mathrm{d}t. \tag{95}$$

The time-averaging produces a factor of one half, and applying a similar procedure to the potential energy we obtain

$$\overline{KE} = \frac{A^2}{4}(k^2+l^2), \qquad \overline{PE} = \frac{A^2}{4}k_d^2, \tag{96}$$

so that the average total energy is

$$\overline{E} = \frac{A^2}{4}(K^2 + k_d\ ^2), \tag{97}$$

where $K^2 = k^2 + l^2$.

The flux, $\mathbf{F}$, is given by

$$\mathbf{F} = -\left(\psi\nabla\frac{\partial\psi}{\partial t} + \mathbf{i}\frac{\beta}{2}\psi^2\right) = -A^2\cos^2(\mathbf{k}\cdot\mathbf{x} - \omega t)\left(\mathbf{k}\omega - \mathbf{i}\frac{\beta}{2}\right), \tag{98}$$

so that evidently the energy flux has a component in the direction of the wavevector, $\mathbf{k}$, and a component in the x -direction. Averaging over a wave period straightforwardly gives us additional factors of one half:

$$\overline{\mathbf{F}} = -\frac{A^2}{2}\left(\mathbf{k}\omega + \mathbf{i}\frac{\beta}{2}\right). \tag{99}$$

We now use the dispersion relation $\omega = -\beta k/(K^2 + {k_d}^2)$ to eliminate the frequency, giving

$$\overline{\mathbf{F}} = \frac{A^2\beta}{2}\left(\mathbf{k}\frac{k}{K^2 + {k_d}^2} - \mathbf{i}\frac{1}{2}\right), \tag{100}$$

and writing this in component form we obtain

$$\overline{\mathbf{F}} = \frac{A^2\beta}{4}\left[\mathbf{i}\left(\frac{k^2 - l^2 - {k_d}^2}{K^2 + {k_d}^2}\right) + \mathbf{j}\left(\frac{2kl}{K^2 + k_d^2}\right)\right] \tag{101}$$

Comparison of (101) with (15) and (97) reveals that

$$\overline{\mathbf{F}} = \mathbf{c}_g\overline{E} \tag{102}$$

so that the energy propagation equation, (92), when averaged over a wave, becomes

$$\frac{\partial\overline{E}}{\partial t} + \nabla\cdot\mathbf{c}_g\overline{E} = 0. \tag{103}$$

This is an important result, and more general than our derivation implies. One immediate implication is that if there is a disturbance that generates waves, *the group velocity is directed away from the disturbance.*

Most of the time in waves, energy is not conserved because it can be extracted from the flow.

Appendix B: The WKB Approximation for Linear Waves

We are concerned with finding solutions to an equation of the form

$$\frac{\mathrm{d}^2\xi}{\mathrm{d}z^2} + m^2(z)\xi = 0, \tag{104}$$

where $m^2(z)$ is positive for wavelike solutions. If m is constant the solution has the harmonic form

$$\xi = \mathrm{Re}\, A_0\mathrm{e}^{\mathrm{i}mz} \tag{105}$$

where A_0 is a complex constant. If m varies only 'slowly' with z — meaning that the variations occur on a scale much longer than $1/m$ — one might reasonably expect that the

harmonic solution above would provide a reasonable first approximation; that is, we expect the solution to locally look like a plane wave with local wavenumber $m(z)$. However, we might also expect that the solution would not be *exactly* of the form $\exp(\mathrm{i}m(z)z)$, because the phase of ξ is $\theta(z) = mz$, so that $\mathrm{d}\theta/\mathrm{d}z = m + z\mathrm{d}m/\mathrm{d}z \neq m$. Thus, in (105) m is not the wavenumber unless m is constant. Nevertheless, this argument suggests that we seek solutions of a similar form to (105), and we find such solutions by way of a perturbation expansion below. We note that the condition that variations in m, or in the wavelength m^{-1}, occur only slowly may be expressed as

$$\frac{m}{|\partial m/\partial z|} \gg m^{-1} \quad \text{or} \qquad \left|\frac{\partial m}{\partial z}\right| \ll m^2. \tag{106}$$

This condition will generally be satisfied if variations in the background state, or in the medium, occur on a scale much longer than the wavelength.

B.1 Solution by perturbation expansion

To explicitly recognize the rapid variation of m we rescale the coordinate z with a small parameter ϵ; that is, we let $\widehat{z} = \epsilon z$ where $\widehat{z}$ varies by $\mathcal{O}(1)$ over the scale on which m varies. Eq. (104) becomes

$$\epsilon^2 \frac{\mathrm{d}^2\xi}{\mathrm{d}\widehat{z}^2} + m^2(\widehat{z})\xi = 0, \tag{107}$$

and we may now suppose that all variables are $\mathcal{O}(1)$. If m were constant the solution would be of the form $\xi = A\exp(m\widehat{z}/\epsilon)$ and this suggests that we look for a solution to (107) of the form

$$\xi(z) = \mathrm{e}^{g(\widehat{z})/\epsilon}, \tag{108}$$

where $g(\widehat{z})$ is some as yet unknown function. We then have, with primes denoting derivatives,

$$\xi' = \frac{1}{\epsilon} g' \mathrm{e}^{g/\epsilon}, \qquad \xi'' = \left(\frac{1}{\epsilon^2} g'^2 + \frac{1}{\epsilon} g''\right) \mathrm{e}^{g/\epsilon}. \tag{109a,b}$$

Using these expressions in (107) yields

$$\epsilon g'' + g'^2 + m^2 = 0, \tag{110}$$

and if we let $g = \int h \, \mathrm{d}\widehat{z}$ we obtain

$$\epsilon \frac{\mathrm{d}h}{\mathrm{d}\widehat{z}} + h^2 + m^2 = 0. \tag{111}$$

To obtain a solution of this equation we expand h in powers of the small parameter ϵ,

$$h(\widehat{z};\epsilon) = h_0(\widehat{z}) + \epsilon h_1(\widehat{z}) + \epsilon^2 h_2(\widehat{z}) + \cdots . \tag{112}$$

Substituting this in (111) and setting successive powers of ϵ to zero gives, at first and second order,

$$h_0^2 + m^2 = 0, \qquad 2h_0 h_1 + \frac{\mathrm{d}h_0}{\mathrm{d}\widehat{z}} = 0. \tag{113a,b}$$

The solutions of these equations are

$$h_0 = \pm \mathrm{i} m, \qquad h_1(\widehat{z}) = -\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}\widehat{z}} \ln \frac{m(\widehat{z})}{m_0}. \tag{114a,b}$$

where m_0 is a constant. Now, ignoring higher-order terms, (108) may be written in terms of h_0 and h_1 as

$$\xi(\widehat{z}) = \exp\left(\int h_0 \, \mathrm{d}\widehat{z}/\epsilon\right) \exp\left(\int h_1 \, \mathrm{d}\widehat{z}\right), \tag{115}$$

and, using (114) and with z in place of $\widehat{z}$, we obtain

$$\xi(z) = A_0 m^{-1/2} \exp\left(\pm \mathrm{i} \int m \, \mathrm{d}z\right). \tag{116}$$

where A_0 is a constant, and this is the WKB solution to (104). In general

$$\xi(z) = B_0 m^{-1/2} \exp\left(\mathrm{i} \int m \, \mathrm{d}z\right) + C_0 m^{-1/2} \exp\left(-\mathrm{i} \int m \, \mathrm{d}z\right). \tag{117}$$

or

$$\xi(z) = D_0 m^{-1/2} \cos\left(\int m \, \mathrm{d}z\right) + E_0 m^{-1/2} \sin\left(\int m \, \mathrm{d}z\right). \tag{118}$$

A property of (116) is that the derivative of the phase is just m; that is, m is indeed the local wavenumber. Note that a crucial aspect of the derivation is that m varies slowly, so that there is a small parameter, ϵ, in the problem. Having said this, it is often the case that WKB theory can provide qualitative guidance even when there is little scale separation between the variation of the background state and the wavelength. Asymptotics often works when it seemingly shouldn't.

References

Eliassen, A. & E. Palm, 1961. On the transfer of energy in stationary mountain waves *Geofysiske Publikasjoner*, **22(3)**, 1–23.

Vallis, G. K., 2006. *Atmospheric and Oceanic Fluid Dynamics*. Cambridge University Press, 745 pp.

The Role of Mixed Layer Instabilities in Submesoscale Turbulence

Jörn Callies

1 Introduction

The upper ocean can support energetic flows at scales smaller than the order 100 km mesoscale eddies. Sharp surface fronts associated with strong along-front currents emerge in high-resolution numerical simulations [8, 16] and are observed in the wintertime mid-latitude ocean [7]. These submesoscale flows are associated with large vertical fluxes of both physical and biogeochemical tracers and may thereby regulate the uptake of heat and carbon from the atmosphere [8, 17, 11]. What drives these submesoscale flows?

Two mechanisms have been proposed: surface frontogenesis [19, 28] and mixed layer instabilities [3]. The process underlying surface frontogenesis can be understood with quasi-geostrophic (QG) dynamics [32]. In the interior of the ocean, when a strain field increases a horizontal buoyancy gradient, an ageostrophic circulation develops according to the omega equation [13] and acts to oppose the increase of the buoyancy gradient. Light water downwells on the dense side and dense water upwells on the light side of the buoyancy gradient. At the surface, however, the vertical velocity must vanish and the ageostrophic circulation cannot act to oppose the increase of the buoyancy gradient in the same way—the mesoscale strain field is left to create strong surface fronts.

The simplest model of these dynamics is the surface QG model [10, 32, 1, 12]. It assumes an infinitely deep ocean with constant interior potential vorticity (PV),

$$\nabla^2\psi + \frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial\psi}{\partial z}\right) = 0, \tag{1}$$

where ψ is the geostrophic streamfunction, f is the Coriolis frequency, and N the buoyancy frequency, such that the dynamics are completely determined by the advection of buoyancy at the surface,

$$\frac{\partial b}{\partial t} + \mathrm{J}(\psi, b) = 0, \tag{2}$$

where $b = f\partial\psi/\partial z$ is buoyancy and

$$\mathrm{J}(a,b) = \frac{\partial a}{\partial x}\frac{\partial b}{\partial y} - \frac{\partial a}{\partial y}\frac{\partial b}{\partial x} \tag{3}$$

is the Jacobian operator. This supplies the boundary condition for the elliptic problem (1). Straining by mesoscale eddies creates sharp buoyancy gradients and strong associated flows at the surface. Secondary instabilities lead to eventually fully turbulent dynamics [12],

for which Kolmogorov-like dimensional arguments predict the energy spectrum in the submesoscale range to scale like $E_{k_\mathrm{h}} \sim {k_\mathrm{h}}^{-5/3}$ in an inertial range in which surface buoyancy variance is cascaded to small scales [1]. If non-QG dynamics are taken into account, ageostrophic advection of buoyancy further accelerates frontogenesis and leads to frontal collapse, the formation of true discontinuities [14]. In this case, the submesoscale energy spectrum is modified to $E_{k_\mathrm{h}} \sim {k_\mathrm{h}}^{-2}$ [4].

As opposed to what is assumed in surface QG dynamics, however, the upper ocean does not have a nearly constant PV. Instead, there is typically a weakly stratified mixed layer with low PV overlying a strongly stratified thermocline with high PV. There is a sharp step-like increase in PV at the base of the mixed layer. This PV step is dynamically important, because it supports edge waves that have the potential to interact with surface edge waves and thus produce a baroclinic instability in the mixed layer [3]. In winter, when mixed layers are deep, this mixed layer instability occurs at scales of 1–10 km, so it has the potential to energize submesoscale turbulence.

The importance of mixed layer instabilities is hinted at by the observation that submesoscale turbulence undergoes a seasonal cycle. Both modeling [22] and observations [7] suggest that submesoscale turbulence is energized in winter and suppressed in summer. Mixed layer instabilities are expected to undergo a strong seasonal cycle, following the seasonal cycle of mixed layer depth itself. In the frontogenetic picture, on the other hand, submesoscale turbulence is driven by mesoscale eddies that do not exhibit a strong seasonal cycle.

The goal of this report is to understand the dynamics of mixed layer instabilities and of the turbulent dynamics that emerge when these instabilities grow to finite amplitude. We formulate a simple QG model consisting of two constant-PV layers representing the mixed layer and the thermocline, coupled at a deformable interface. Despite its simplicity, this model captures both mesoscale and mixed layer instabilities and thereby a number of fundamental aspects of submesoscale dynamics. The model also allows straightforward comparison to surface QG dynamics, in which no mixed layer is present.

We use QG scaling to formulate the dynamics of our model, which requires small Rossby and Froude numbers [24]. Typical mesoscale Rossby and Froude numbers are on the order 0.1 and increase slowly with wavenumber if the submesoscale is energetic [7]. To leading order, submesoscale flows are thus expected to follow QG dynamics. Higher-order effects can become important at submesoscales, however, and we take up the discussion of non-QG effects in the conclusions.

2 Model formulation

Consider two layers of constant PV on an f-plane, with constant stratification and constant mean shear (Fig. 1). The upper layer represents the mixed layer, which has a mean depth h, stratification N_m, and mean shear Λ_m. The lower layer represents the thermocline and has stratification N_t and mean shear Λ_t. The total depth is H. The layers are coupled by a deformable interface; a rigid lid condition is applied at the surface; a flat bottom condition is applied at the bottom. The presence of a bottom at the base of the thermocline is not realistic, but we will see that the bottom layer will still represent important properties of the thermocline. The approximation that the stratification is discontinuous at the base of

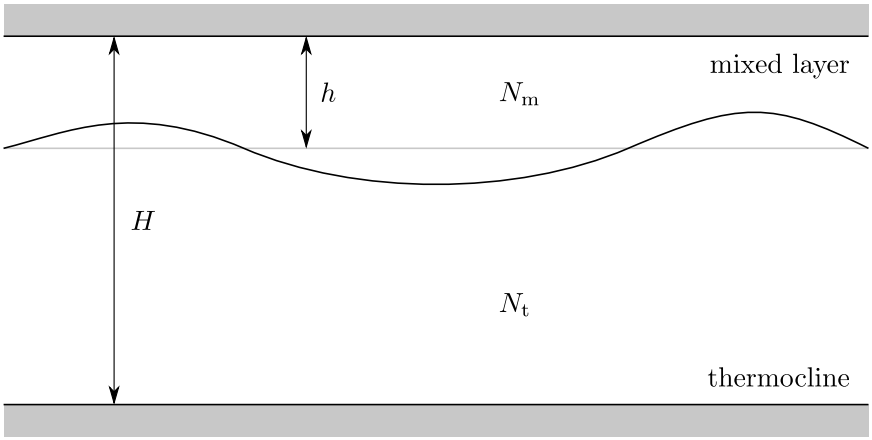


Figure 1: Schematic of the model setup in a vertical–horizontal plane. There are rigid surfaces at $z = 0$ and $z = -H$ and a deformable interface at the mean depth $z = -h$, separating layers of constant stratifications, N_m in the mixed layer and N_t in the thermocline.

the mixed layer is appropriate at horizontal scales larger than the deformation radius Nd/f associated with the transition depth d [30]. The transition at the base of the mixed layer is typically quite sharp, so this deformation radius is much smaller than the submesoscales we are interested in here.

The uniform PV within the two layers simplifies the dynamics dramatically. PV conservation within the layers is trivial, like in the classic Eady problem [10]. The flow in the interior of the layers is obtained by solving (1), with the boundary conditions supplied by the distribution of buoyancy at the surface and bottom and by matching conditions at the interface between the mixed layer and the thermocline.

At the rigid boundaries at the surface and bottom, where the vertical velocity w vanishes, the advection of buoyancy anomalies b is given by

$$\frac{\partial b}{\partial t} + \mathrm{J}(\psi, b) = 0. \tag{4}$$

To ensure that pressure is continuous at the interface, we require that the streamfunction ψ is continuous. For mass conservation, we require that the vertical velocity w is also continuous. These conditions are applied at $z = -h$, consistent with QG scaling. The conservation equations for buoyancy just above and below the interface at $z = -h$,

$$\frac{\partial b^+}{\partial t} + \mathrm{J}(\psi, b^+) + wN_\mathrm{m}^2 = 0, \qquad \frac{\partial b^-}{\partial t} + \mathrm{J}(\psi, b^-) + wN_\mathrm{t}^2 = 0, \tag{5}$$

can then be combined to eliminate w. This gives

$$\frac{\partial \theta_1}{\partial t} + \mathrm{J}(\psi_1, \theta_1) = 0, \tag{6}$$

where ψ_1 denotes the streamfunction at $z = -h$. This is a conservation equation for

$$\theta_1 = f\left(\frac{b^+}{N_\mathrm{m}^2} - \frac{b^-}{N_\mathrm{t}^2}\right). \tag{7}$$

The quantity θ_1 is nothing but the PV

$$q = \nabla^2\psi + f\frac{\partial}{\partial z}\left(\frac{b}{N^2}\right) \tag{8}$$

integrated across the interface. While there are no PV anomalies within the two layers, the displacement of the interface between the two layers of constant PV induces a PV anomaly that, according to (6) and consistent with QG dynamics, is advected by the geostrophic flow at $z = -h$. The conservation equation (6) has been used to study the dynamics of the tropopause, which is similarly an interface between weakly stratified fluid in the troposphere and strongly stratified fluid in the stratosphere [10, 27, 15, 12].

The model can be considered as consisting of three PV sheets:

$$q = \theta_0\delta(z) + \theta_1\delta(z+h) + \theta_2\delta(z+H), \tag{9}$$

where δ is Dirac's delta function and $\theta_0 = -fb/N_{\mathrm{m}}^2$ at $z = 0$ and $\theta_2 = fb/N_{\mathrm{t}}^2$ at $z = -H$. PV is advected by the geostrophic flow, so

$$\frac{\partial\theta_i}{\partial t} + \mathrm{J}(\psi_i, \theta_i) = 0, \tag{10}$$

where $i = 0, 1, 2$ and ψ_i is the streamfunction at the level corresponding to θ_i. This formulation is simply an extension of Bretherton's representation of boundary conditions [5] to include an interior PV sheet due to deflection of an interface between layers of different stratification.

Note that the statement that θ_1 is only advected by the geostrophic flow does not imply that the vertical velocity vanishes at the interface, just like the fact that PV anomalies in the QG system are advected only by the geostrophic flow does not imply that the vertical velocity vanishes. The vertical velocity is implicit in the dynamics and can be solved for using the omega equation.

To complete the dynamics, we require an inversion relation that allows us to obtain the streamfunctions ψ_i from the conserved quantities θ_i. This relation can be written as a linear equation for the Fourier coefficients, denoted by subscripts k, l:

$$\boldsymbol{\theta}_{k,l} = L\boldsymbol{\psi}_{k,l}, \qquad \boldsymbol{\theta} = (\theta_0, \theta_1, \theta_2)^{\mathsf{T}}, \qquad \boldsymbol{\psi} = (\psi_0, \psi_1, \psi_2)^{\mathsf{T}}. \tag{11}$$

The matrix L is determined by solving

$$-k_{\mathrm{h}}^2\psi_{k,l} + \frac{\partial}{\partial z}\left(\frac{f^2}{N^2}\frac{\partial\psi_{k,l}}{\partial z}\right) = 0 \tag{12}$$

in each layer. For example, the first column of L is determined by setting $\boldsymbol{\psi} = (1, 0, 0)^{\mathsf{T}}$, solving (12) for $\psi(z)$, and subsequently calculating

$$\theta_0 = -\frac{f^2}{N_{\mathrm{m}}^2}\frac{\partial\psi}{\partial z}(0), \qquad \theta_1 = \frac{f^2}{N_{\mathrm{m}}^2}\frac{\partial\psi}{\partial z}(-h^+) - \frac{f^2}{N_{\mathrm{t}}^2}\frac{\partial\psi}{\partial z}(-h^-), \qquad \theta_2 = \frac{f^2}{N_{\mathrm{t}}^2}\frac{\partial\psi}{\partial z}(-H). \tag{13}$$

This procedure gives

$$L = fk_{\mathrm{h}}\begin{pmatrix} -\frac{\coth\mu_{\mathrm{m}}}{N_{\mathrm{m}}} & \frac{\operatorname{csch}\mu_{\mathrm{m}}}{N_{\mathrm{m}}} & 0 \\ \frac{\operatorname{csch}\mu_{\mathrm{m}}}{N_{\mathrm{m}}} & -\frac{\coth\mu_{\mathrm{m}}}{N_{\mathrm{m}}} - \frac{\coth\mu_{\mathrm{t}}}{N_{\mathrm{t}}} & \frac{\operatorname{csch}\mu_{\mathrm{t}}}{N_{\mathrm{t}}} \\ 0 & \frac{\operatorname{csch}\mu_{\mathrm{t}}}{N_{\mathrm{t}}} & -\frac{\coth\mu_{\mathrm{t}}}{N_{\mathrm{t}}} \end{pmatrix}, \tag{14}$$

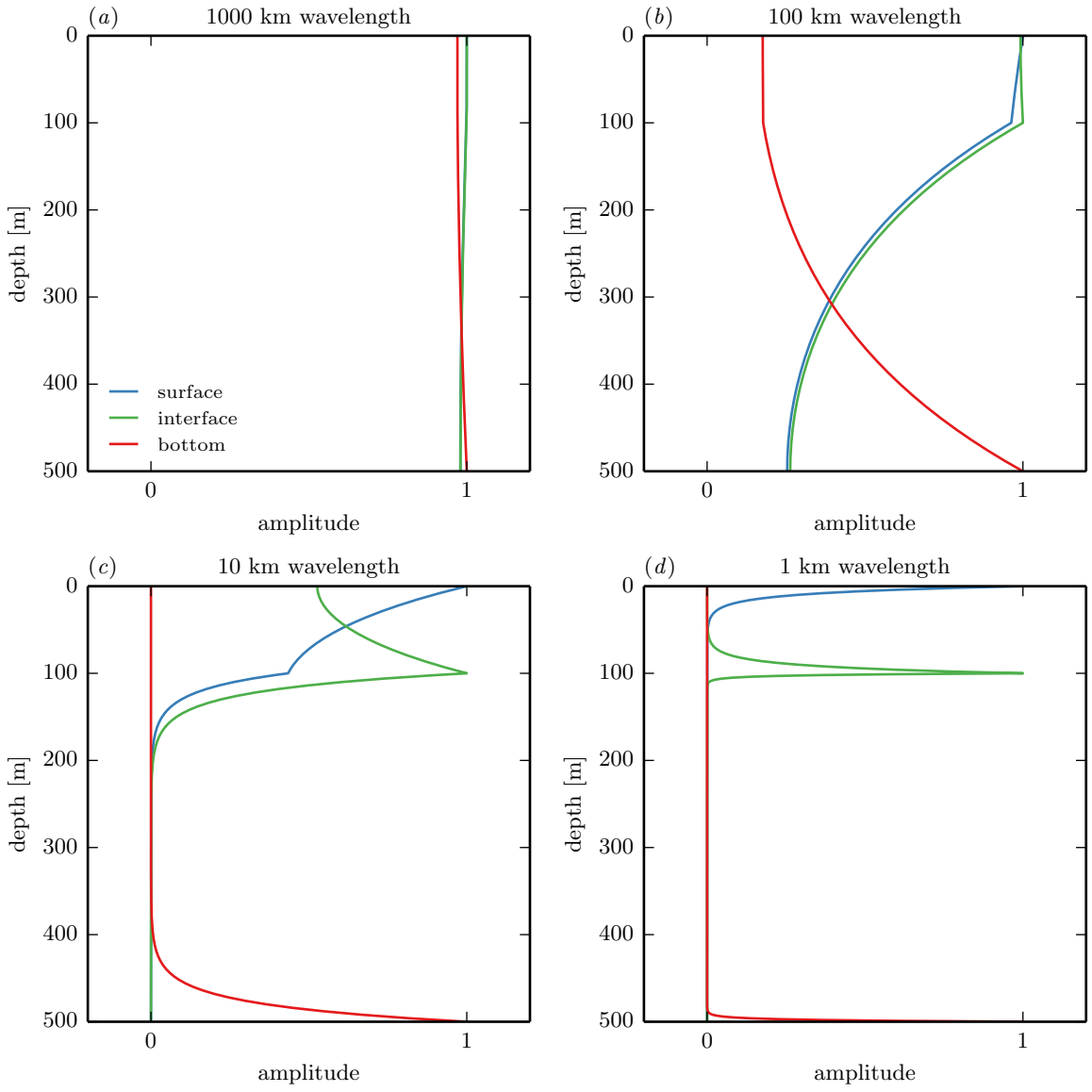


Figure 2: Vertical structure of streamfunction amplitude associated with anomalies of θ_0 (surface), θ_1 (interface), and θ_2 (bottom). Shown are the vertical profiles for θ_i anomalies with different horizontal wavenumbers $k_\mathrm{h} = 2\pi/\lambda$. The wavelength λ is given in the panel titles.

where $\mu_\mathrm{m} = N_\mathrm{m} k_\mathrm{h} h/f$ and $\mu_\mathrm{t} = N_\mathrm{t} k_\mathrm{h}(H-h)/f$ are nondimensional wavenumbers. This 3×3 matrix can easily be inverted.

This model can be generalized to an arbitrary number of layers of constant stratification and shear, which may be a useful way to approximate more realistic stratification and shear profiles. This is discussed in Appendix A. The model can also be extended to include a density jump at the interface, as is sometimes present at the base of the mixed layer. The formulation is given in Appendix B, but we here restrict ourselves to the case of a continuous density profile.

To build intuition for the dynamics of the model, we illustrate the vertical structure of flow associated with anomalies of the conserved quantities at the surface, the interface, and the bottom (Fig. 2). We apply parameters that will be used throughout the report, typical of the wintertime midlatitude ocean (Tab. 1). At the largest scales, for anomalies with wavelength $\lambda = 1000$ km or $k_\mathrm{h} \ll f/N_\mathrm{t} H$, the flow is nearly barotropic, irrespective of which level the anomaly is at. At smaller scales, $\lambda = 100$ km or $k_\mathrm{h} \sim f/N_\mathrm{t} H \ll f/N_\mathrm{m} h$, there is significant decay in the thermocline while the flow is nearly barotropic in the mixed layer. Anomalies at the surface and interface still induce significant flow at the bottom and vice versa. At $\lambda = 10$ km or $k_\mathrm{h} \sim f/N_\mathrm{m} h$, on the other hand, anomalies at the surface or interface induce very little flow at the bottom and vice versa. The flow is also not barotropic in the mixed layer anymore, but surface anomalies still induce significant flow at the interface and vice versa. At $\lambda = 1$ km or $k_\mathrm{h} \gg f/N_\mathrm{m} h$, all levels are decoupled: anomalies on any of the levels induce very little flow at the other levels.

The dependence of the vertical flow structure on the horizontal scale of the anomalies illuminates the qualitative dynamics of the model. At the smallest scales, all three levels are independent and follow surface QG dynamics. At scales $k_\mathrm{h} \sim f/N_\mathrm{m} h$, around the mixed layer deformation radius, surface and interface anomalies can interact, allowing for phase locking and instability in the mixed layer. Bottom anomalies, on the other hand are independent, so there is no deep instability at these scales. At scales $k_\mathrm{h} \sim f/N_\mathrm{t} H$, around the deep deformation radius, surface or interface anomalies can interact with bottom anomalies, so there is potential for a deep instability at these scales. At the largest scales, the flow is essentially barotropic and follows two-dimensional dynamics.

3 Linear stability analysis

We now analyse the linear stability of the model formulated above. While this linear analysis is not directly applicable to the strongly nonlinear turbulent regime, it reveals some key characteristics of the dynamics that will help us understand the nonlinear regime.

Blumen analyzed short-wave instabilities in the atmosphere also using a model consisting of two coupled constant-PV layers [2]. He performed a linear stability analysis equivalent to what will be presented in the following. For completeness, we review the linear stability in the context of the upper ocean dynamics, at the cost of being somewhat redundant with Blumen's study.

We consider the linear stability of normal-mode perturbations to a zonal flow with constant vertical shear Λ_m in the mixed layer and Λ_t in the thermocline (Fig. 3). The system is Galilean invariant, so we can arbitrarily set the mean zonal flow to zero at the surface. The linearized conservation equations for the perturbations from this mean state

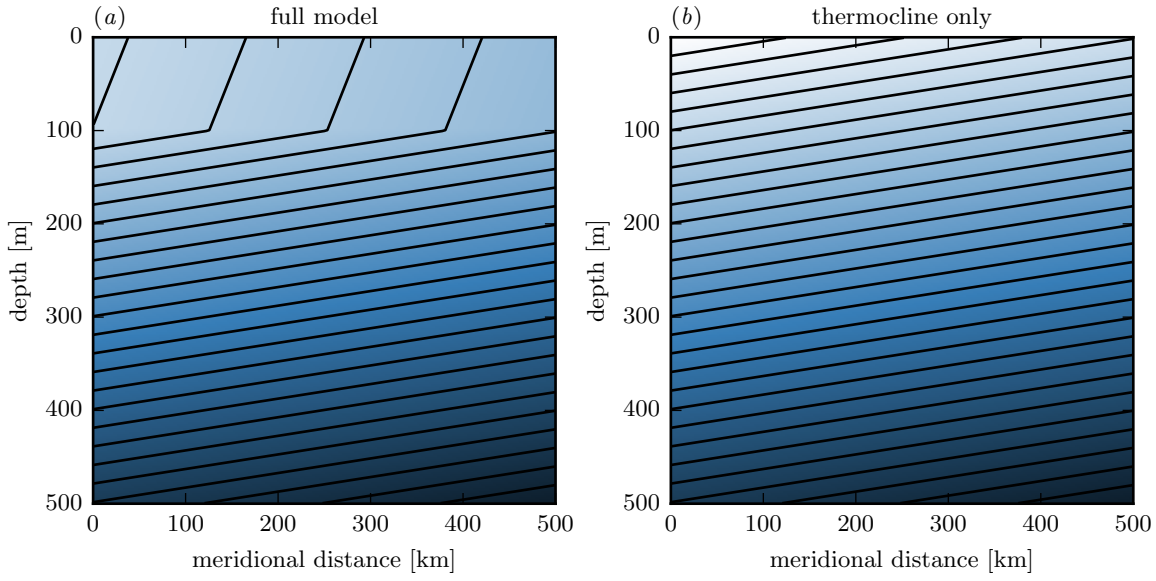


Figure 3: Mean buoyancy structure in the meridional–vertical plane for (a) the full model and (b) the thermocline only case. The contours show isopycnals; light shading indicates more buoyant fluid.

are

$$\frac{\partial \boldsymbol{\theta}_{k,l}}{\partial t} + ikU\boldsymbol{\theta}_{k,l} + ik\Gamma\boldsymbol{\psi}_{k,l} = 0, \tag{15}$$

where the mean zonal flows and mean meridional PV gradients at the respective levels are denoted by the diagonal elements of the matrices U and Γ:

$$U = \operatorname{diag}\big(0, -\Lambda_\mathrm{m} h, -\Lambda_\mathrm{m} h - \Lambda_\mathrm{t}(H-h)\big), \tag{16}$$

$$\Gamma = \operatorname{diag}\big(f^2\Lambda_\mathrm{m}/N_\mathrm{m}^2, -f^2\Lambda_\mathrm{m}/N_\mathrm{m}^2 + f^2\Lambda_\mathrm{t}/N_\mathrm{t}^2, -f^2\Lambda_\mathrm{t}/N_\mathrm{t}^2\big). \tag{17}$$

Using the inversion relation (11), we can replace the $\boldsymbol{\theta}_{k,l}$, such that

$$L\frac{\partial \boldsymbol{\psi}_{k,l}}{\partial t} + ikUL\boldsymbol{\psi}_{k,l} + ik\Gamma\boldsymbol{\psi}_{k,l} = 0. \tag{18}$$

Assuming that modes vary harmonically in time with (complex) frequency ω turns this into the generalized eigenvalue problem

$$(UL + \Gamma)\boldsymbol{\psi}_{k,l} = cL\boldsymbol{\psi}_{k,l}, \tag{19}$$

where the eigenvalue is $c = \omega/k$. The real part of c is the zonal phase speed; the imaginary part gives the growth rate $\sigma = k\,\mathrm{Im}\,c$.

Being a third-order system, (19) can be solved analytically, but the solutions are rather complicated and give little useful insight. We instead explore the characteristics of the solutions numerically for the set of parameters given above. We then explain the stability properties and parameter dependencies by considering simplifications of the model.

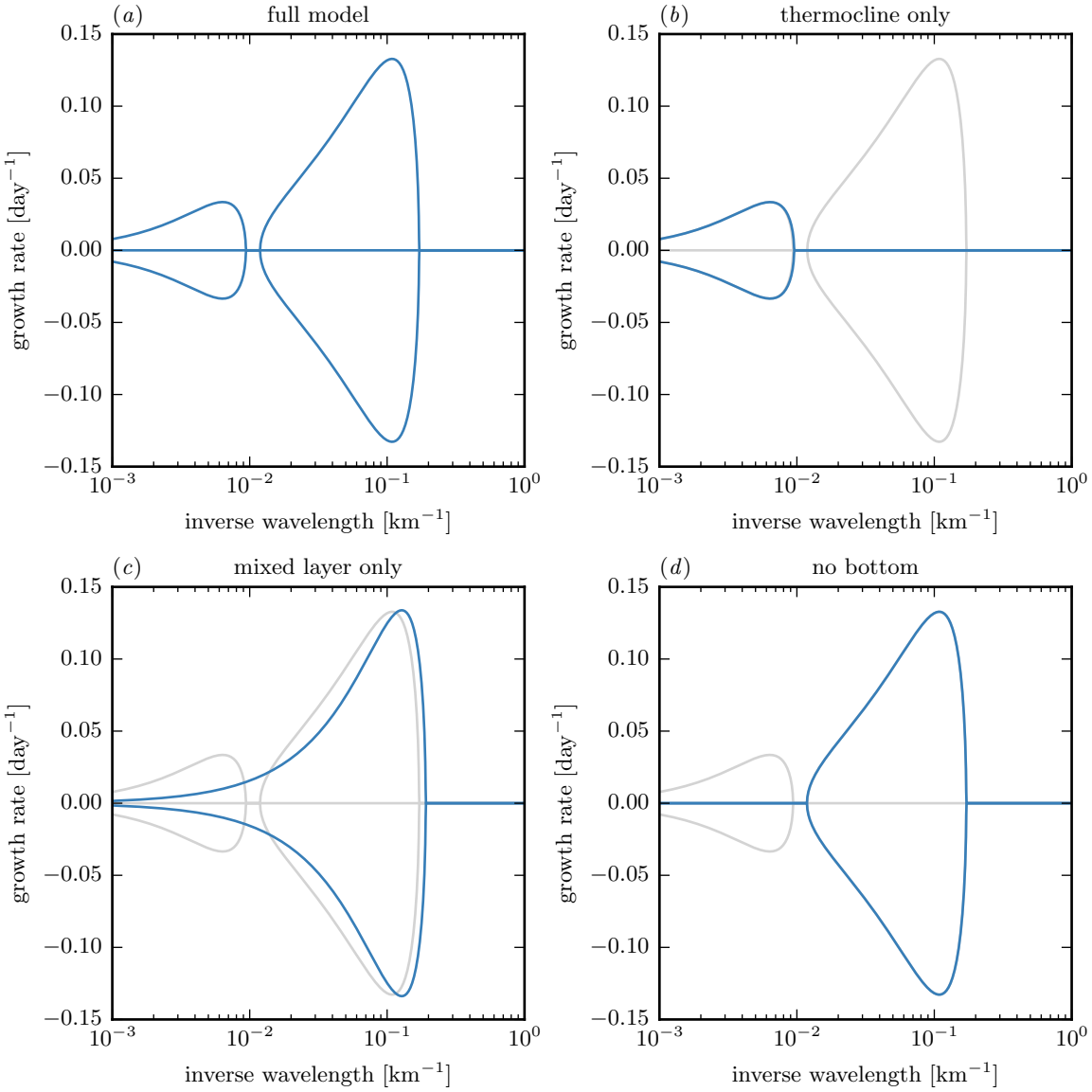


Figure 4: Growth rates from the linear stability analysis for (*a*) the full model incorporating a mixed layer, (*b*) an Eady model representing the thermocline only, (*c*) an Eady model representing the mixed layer only, and (*d*) a model like the full model but without a bottom. Growth rates are shown in blue, the growth rates of the full model are overlaid for reference in gray.

Parameter	Symbol	Value
Mixed layer depth	h	100 m
Total depth	H	500 m
Mixed layer stratification	N_m	2×10^{-3} s^{-1}
Thermocline stratification	N_t	8×10^{-3} s^{-1}
Mixed layer shear	Λ_m	10^{-4} s^{-1}
Thermocline shear	Λ_t	10^{-4} s^{-1}
Coriolis frequency	f	10^{-4} s^{-1}
Domain size	a	500 km

Table 1: Parameters used throughout this report unless otherwise noted. These are typical of the wintertime midlatitude ocean.

3.1 Full model

The linear stability analysis reveals that there are two lobes of instability: one at the mesoscale and one at the submesoscale (Fig. 4*a*). The maximum growth rates occur at $l = 0$ and zonal wavelengths of about 160 km and 10 km. The submesoscale instability has a peak growth rate much larger than the mesoscale instability in this case with equal shear in the two layers. The growing modes are conjugate to decaying modes. The growth rates are similar to what Boccaletti et al. found in a linear QG stability analysis of a realistic mean state of the wintertime eastern subtropical North Pacific [3]. The magnitudes are slightly smaller here, because the shear is slightly weaker.

The phase speeds of the linear modes give clues to the dynamics in different ranges of scales (Fig. 5*a*). For each wavenumber, there are three modes. Growing and decaying modes, being conjugate to each other, have the same phase speeds. This is the familiar phase locking of counter-propagating waves in baroclinic instability (branches 'b' and 'e'). Where these growing and decaying modes exist, there is an additional neutral mode (branches 'a' and 'd'). Where there are no growing modes, all three neutral modes have distinct phase speeds—no phase locking occurs. We will discuss the dynamics of the various branches by considering approximations to the full model.

But first consider the spatial structure of the modes corresponding to the peak growth rates. The perturbation streamfunctions show that the mesoscale mode is deep and spans the entire water column (Fig. 6*a*), whereas the submesoscale mode is almost completely confined to the mixed layer, with only weak leaking into the thermocline below 100 m depth (Fig. 6*b*). Both modes show the familiar pattern of unstable modes tilted into the shear, which is necessary to extract potential energy from the mean.

3.2 Thermocline only

We start explaining the instability properties of the full model by comparing it to the classic Eady model representing the thermocline only (Fig. 3*b*). This amounts to setting $h = 0$ or $N_\mathrm{m} = N_\mathrm{t}$ in the full model. In this case, the system reduces to two variables, the inversion

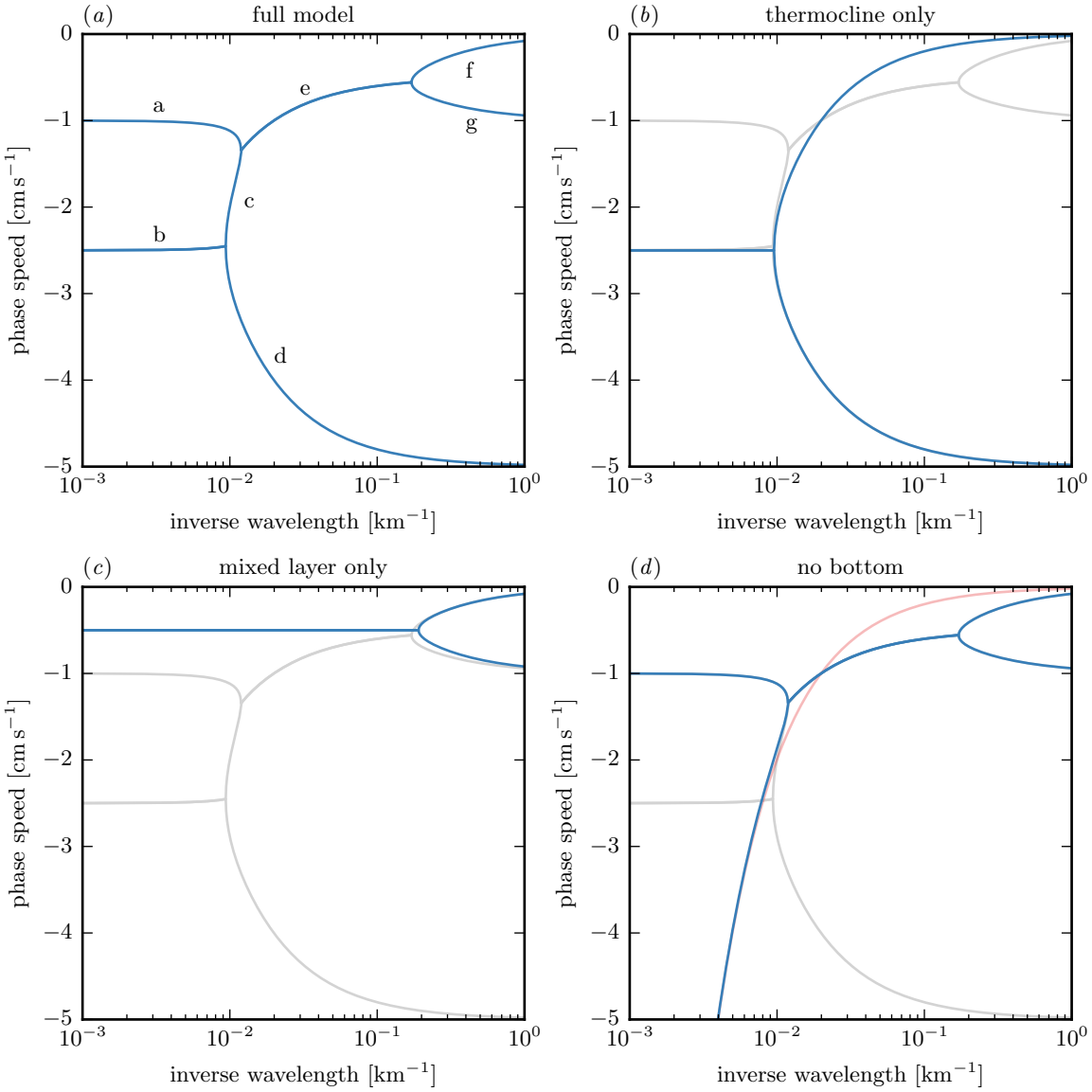


Figure 5: Phase speeds from the linear stability analysis for (*a*) the full model incorporating a mixed layer, (*b*) an Eady model representing the thermocline only, (*c*) an Eady model representing the mixed layer only, and (*d*) a model like the full model but without a bottom. Phase speeds are shown in blue, the growth rates of the full model are overlaid for reference in gray. The faint red line shows the phase speed of a surface edge wave.

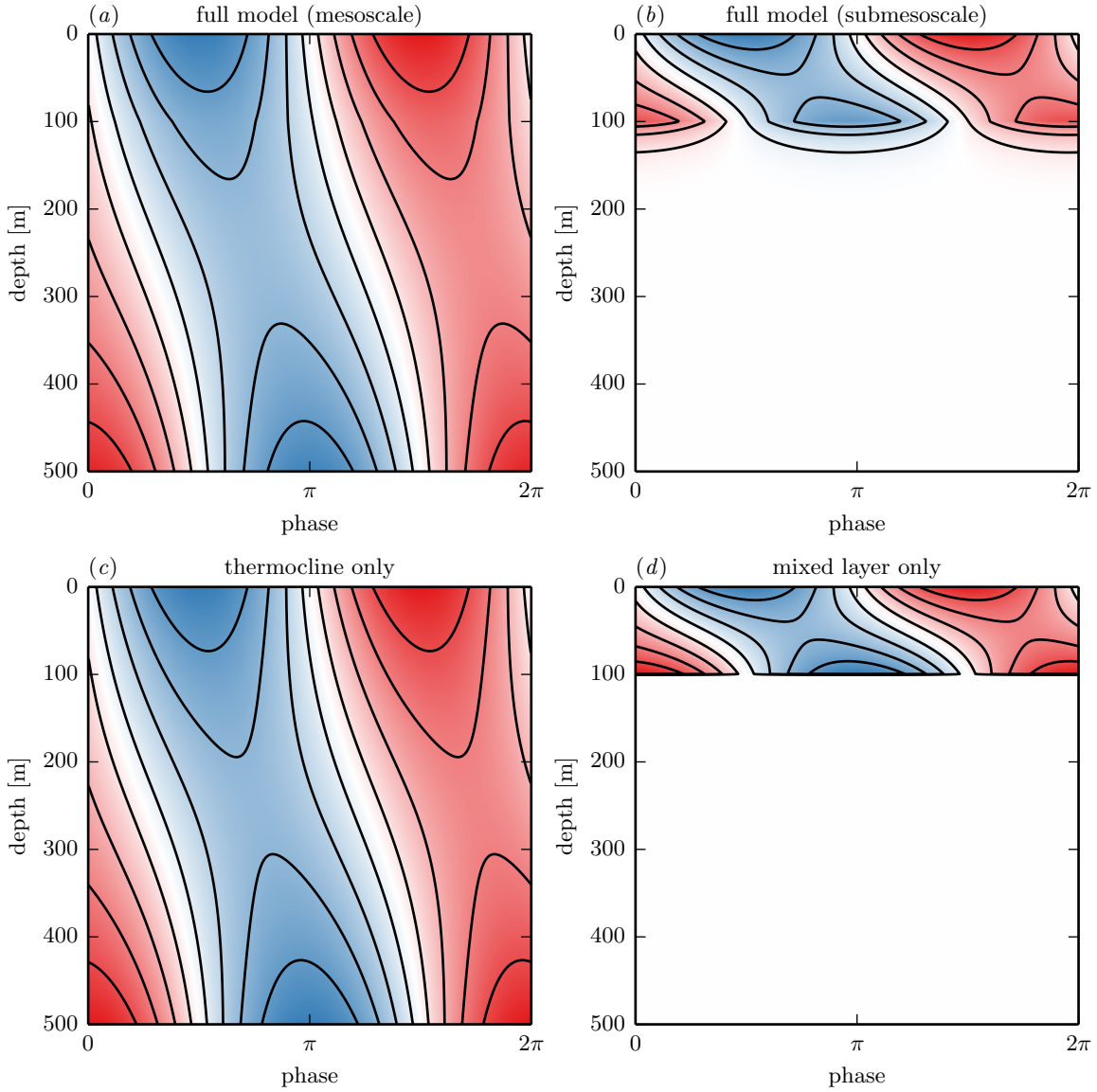


Figure 6: Perturbation streamfunction as obtained from the linear stability analysis of (a) the most unstable mesoscale mode of the full model, (b) the most unstable submesoscale mode of the full model, (c) the most unstable mode of an Eady model representing the thermocline only, and (d) the most unstable mode of an Eady model representing the mixed layer only. Red and blue shading represents positive and negative values.

matrix is

$$L = fk_{\mathrm{h}} \begin{pmatrix} -\frac{\coth \mu_{\mathrm{t}}}{N_{\mathrm{t}}} & \frac{\operatorname{csch} \mu_{\mathrm{t}}}{N_{\mathrm{t}}} \\ \frac{\operatorname{csch} \mu_{\mathrm{t}}}{N_{\mathrm{t}}} & -\frac{\coth \mu_{\mathrm{t}}}{N_{\mathrm{t}}} \end{pmatrix}, \tag{20}$$

and

$$U = \operatorname{diag}\left(0, -\Lambda_{\mathrm{t}} h\right), \tag{21}$$

$$\Gamma = \operatorname{diag}\left(f^2 \Lambda_{\mathrm{t}} / N_{\mathrm{t}}^2, -f^2 \Lambda_{\mathrm{t}} / N_{\mathrm{t}}^2\right). \tag{22}$$

The solution is [10, 33]

$$c = -\frac{\Lambda_{\mathrm{t}} h}{2} \pm \frac{i \Lambda_{\mathrm{t}} h}{\mu_{\mathrm{m}}} \left(\mu_{\mathrm{t}} \coth \mu_{\mathrm{t}} - 1 - \frac{\mu_{\mathrm{t}}^2}{4}\right)^{\frac{1}{2}}, \tag{23}$$

where $\mu_{\mathrm{t}} = N_{\mathrm{t}} k_{\mathrm{h}} H / f$ is the nondimensional wavenumber.

The Eady model has a baroclinic instability near the deformation radius $N_{\mathrm{t}} H / f$. The maximum growth rate $\sigma = 0.31 f \Lambda_{\mathrm{t}} / N_{\mathrm{t}}$ occurs at $l = 0$ and $\mu_{\mathrm{t}} = 1.6$, which corresponds to a zonal wavelength $\lambda = 3.9 N_{\mathrm{t}} H / f$. The growth curve for this thermocline only model traces out the mesoscale lobe of the full model almost perfectly (Fig. 4*b*). The short-wave cutoff in the Eady model at $\lambda = 2.6 N_{\mathrm{t}} H / f$ nearly coincides with the short-wave cutoff of the mesoscale instability in the full model. The phase speed of the phase-locked waves $-\Lambda_{\mathrm{t}} H / 2$ very nearly matches the phase speed of the unstable mesoscale mode of the full model (Fig. 5*b*). The critical level, where the phase speed matches the mean flow, is at mid-depth. The split at the short-wave cutoff into surface and bottom modes also features in the full model. In the thermocline only model, the surface and bottom modes are very nearly Eady edge waves that do not sense the other boundary. The bottom mode of the thermocline only model almost perfectly matches that of the full model (branch 'd'). The surface mode of the thermocline only model traces out branch 'c' of the full model, but then the full model transitions to dynamics associated with the mixed layer that are not present in the thermocline only model. The mesoscale instability of the full model therefore follows Eady dynamics. The presence of the mixed layer only modifies the characteristics of the instability slightly. The spatial structure of the most unstable mesoscale mode is also well captured by the thermocline only model (Fig. 6*c*).

3.3 Mixed layer only

Shifting our attention to the submesoscale instability, we now consider an Eady model representing the mixed layer with a rigid bottom at its base. A priori, we can see from (7) that the full model converges to these dynamics as the thermocline stratification goes to infinity. The thermocline then acts like a rigid bottom at the base of the mixed layer and the conservation of θ_1 turns into a conservation of the buoyancy at the base of the mixed layer $b(-h^+)$.

The Eady model for the mixed layer captures some key properties of the submesoscale instability, but misses others. The solution is the same as (23), with Λ_{t} replaced by Λ_{m} and μ_{t} replaced by $\mu_{\mathrm{m}} = N_{\mathrm{m}} k_{\mathrm{h}} h / f$. The peak growth rate $\sigma = 0.31 f \Lambda_{\mathrm{m}} / N_{\mathrm{m}}$ occurs at $\mu_{\mathrm{m}} = 1.6$, which corresponds to a zonal wavelength $\lambda = 3.9 N_{\mathrm{m}} h / f$. This is a reasonable

approximation of the maximum submesoscale growth rate of the full model (Fig. 4*c*). The mixed layer only model captures the bulk shape of the submesoscale lobe of the full model; its short-wave cutoff at $\lambda = 2.6 N_\mathrm{m} H/f$ is close to the short-wave cutoff of the full model. But the mixed layer only model does not have a long-wave cutoff, unlike the submesoscale instability of the full model. The phase speed structure of the mixed layer only model, with its two phase-locked modes in the unstable range and the split into a surface mode and one propagating on the base of the mixed layer, corresponds to a similar mode structure of the full model (Fig. 5*c*). The spatial structure of the most unstable Eady mode resembles that of the most unstable submesoscale mode of the full model (Fig. 6*d*). It tilts into the shear in the mixed layer, but does not leak into the thermocline.

3.4 No bottom

The submesoscale instability is better approximated if modes are allowed to penetrate into the thermocline. To isolate the submesoscale instability, we consider again the layered model but let the thermocline be infinitely deep. That eliminates bottom edge waves, so no mesoscale instability occurs.

Eady considered the atmospheric analogue to this system to relax the assumption that a rigid lid is placed at the tropopause [10]. The system with no bottom again reduces to two variables; the inversion matrix is

$$L = f k_\mathrm{h} \begin{pmatrix} -\frac{\coth \mu_\mathrm{m}}{N_\mathrm{m}} & \frac{\operatorname{csch} \mu_\mathrm{m}}{N_\mathrm{m}} \\ \frac{\operatorname{csch} \mu_\mathrm{m}}{N_\mathrm{m}} & -\frac{\coth \mu_\mathrm{m}}{N_\mathrm{m}} - \frac{1}{N_\mathrm{t}} \end{pmatrix} \tag{24}$$

and

$$U = \operatorname{diag}\big(0, -\Lambda_\mathrm{m} h\big), \tag{25}$$

$$\Gamma = \operatorname{diag}\big(f^2 \Lambda_\mathrm{m}/N_\mathrm{m}^2, -f^2 \Lambda_\mathrm{m}/N_\mathrm{m}^2 + f^2 \Lambda_\mathrm{t}/N_\mathrm{t}^2\big). \tag{26}$$

In the case $\Lambda = \Lambda_\mathrm{m} = \Lambda_\mathrm{t}$, the solution is [10, 2]

$$c = -\frac{\Lambda h}{2}\left(1 + \frac{\alpha}{\mu_\mathrm{m}}\right) \pm \frac{i \Lambda h}{\mu_\mathrm{m}} \left[\frac{(1-\alpha^2)(\mu_\mathrm{m} - \tanh \mu_\mathrm{m})}{\tanh \mu_\mathrm{m} + \alpha} - \frac{1}{4}(\mu_\mathrm{m} - \alpha)^2\right]^{\frac{1}{2}}, \tag{27}$$

where $\mu_\mathrm{m} = N_\mathrm{m} k_\mathrm{h} h/f$ and $\alpha = N_\mathrm{m}/N_\mathrm{t}$. This solution converges to the mixed layer only solution if $\alpha \ll 1$ and $\alpha \ll \mu_\mathrm{m}$, which is equivalent to $N_\mathrm{t} \gg N_\mathrm{m}$ and $k_\mathrm{h} \gg f/N_\mathrm{t} h$. This shows that large thermocline stratification can act like a rigid bottom, but only for scales that are not too large. Modes of large horizontal scale still manage to penetrate into the thermocline, by which the large-scale dynamics are altered.

The growth rates and phase speeds of this reduced model almost perfectly match the growth rates and phase speeds of the full model at scales smaller than about 100 km (Fig. 4*d* and 5*d*). This model now captures the long-wave cutoff of the submesoscale instability. At large scales, where $\mu_\mathrm{m} \ll \alpha$ and $\mu_\mathrm{m} \ll 1$ or equivalently $k_\mathrm{h} \ll f/N_\mathrm{t} h$ and $k_\mathrm{h} \ll f/N_\mathrm{m} h$, the dynamics split into modes that are barotropic and baroclinic in the mixed layer. The barotropic mode behaves like a surface edge wave, with phase speed $-f\Lambda/N_\mathrm{t} k_\mathrm{h}$, which does not sense the mixed layer (Fig. 5*d*). The baroclinic mode is baroclinic in the mixed layer

and remains shallow for large scales—its critical level is the base of the mixed layer, its phase speed is $-\Lambda h$. The vastly different phase speeds of these two modes prevent phase-locking, so no instability occurs at large scales. This stabilization is analogous to that by the β-effect [25, 21, 33]. For the unstable modes, this reduced model with no bottom also captures the deepening of the critical level as the scale gets larger, $-\Lambda h(1 + f/N_\mathrm{t} k_\mathrm{h} h)/2$, which is due to the increasing penetration of the unstable mode into the thermocline.

The longwave cutoff in this constant-shear case depends on the ratio $N_\mathrm{m}/N_\mathrm{t}$. In the more general case $\Lambda_\mathrm{m} \neq \Lambda_\mathrm{t}$, it also depends on the ratio $\Lambda_\mathrm{m}/\Lambda_\mathrm{t}$. No longwave cutoff occurs if $\Lambda_\mathrm{t} = 0$, as found by Rivest et al., who considered the atmospheric case with no shear in the stratosphere [27]. There is also no longwave cutoff as $N_\mathrm{m}/N_\mathrm{t} \to 0$, which is the Eady limit. The instability itself requires a reversal of the PV gradient, so the condition for instability is $\Lambda_\mathrm{m}/N_\mathrm{m}^2 > \Lambda_\mathrm{t}/N_\mathrm{t}^2$. This condition is typically satisfied in the ocean.

3.5 Summary

We are now in a position to understand all branches in the phase speed diagram of the full model. Branch 'a' is a mode that is baroclinic in the mixed layer and does not penetrate much into the thermocline. It does not sense the bottom. Branch 'b' is the unstable branch corresponding to the deep, Eady-like instability. Branch 'c' is a mode that is nearly barotropic in the mixed layer and behaves like a surface edge wave in the thermocline. It does not interact much with the bottom. Branch 'd' is a bottom edge wave that is independent of the surface and interface. Branch 'e' is the unstable branch corresponding to the mixed layer instability. The instability is significantly modified by the modes' penetration into the thermocline, but the scale and growth rate of the most unstable mode still scale with the mixed layer deformation radius and the Eady growth rate. Branches 'f' and 'g' are edge waves propagating on the surface and the interface that do not interact with any of the other edge waves.

4 Nonlinear dynamics

We now turn to the nonlinear dynamics that arise when the perturbations grow to finite amplitude. Before considering the combined effect of the deep mesoscale and mixed layer instabilities, we first consider them separately. We start with the thermocline only case, in which only the deep instability is present. We subsequently contrast that case with the case with no bottom, in which only the submesoscale instability is present. We finally consider the full model, in which both instabilities occur.

We solve the full nonlinear perturbation equations

$$\frac{\partial \boldsymbol{\theta}'}{\partial t} + U\frac{\partial \boldsymbol{\theta}'}{\partial x} + \Gamma\frac{\partial \boldsymbol{\psi}'}{\partial x} + \mathrm{J}(\boldsymbol{\psi}', \boldsymbol{\theta}') = -r\nabla^{-2}\boldsymbol{\theta}' - \nu\nabla^{2n}\boldsymbol{\theta}', \tag{28}$$

where the Jacobian operator is understood to act element-wise. We introduce hypoviscosity with coefficient r, which provides a large-scale drag, and hyperviscosity with coefficient ν and order n, which helps ensure numerical stability and absorbs enstrophy that is cascaded to small scales.

We solve these equations in a 500 km × 500 km domain that is doubly periodic in the perturbation quantities. A fully dealiased pseudo-spectral code with a resolution 512×512 is used. The time derivatives are discretized using a forth-order Runge-Kutta scheme. The hypoviscosity coefficient is $r = 10^{-16}\ \mathrm{m^{-2}\,s^{-1}}$; the hyperviscosity is of order $n = 10$ and the coefficient is $\nu = 2.5 \times 10^{46}\ \mathrm{m^{20}\,s^{-1}}$. All calculations are started with a white noise of small amplitude in $\boldsymbol{\theta}'$.

4.1 Thermocline only

We start with the familiar Eady model representing the thermocline. Since the dissipative terms are weak in the linear equations, the instability grows until it reaches finite amplitude, when the nonlinear terms become important. Secondary instabilities set in and the flow quickly evolves into a fully turbulent regime. The perturbations grow in scale until they reach a scale where hypoviscosity is strong enough to damp the flow significantly. Thereby, the flow comes into statistical equilibrium, which is the time period considered in what follows.

A snapshot from the equilibrated state exhibits a patchy surface buoyancy field (Fig. 7*a*). There are strong buoyancy gradients. The strongest vortices visible are those at a scale of about 200 km. Smaller-scale vortices are present, but successively weaker. They result from a roll-up instability that features prominently in the evolution of the flow [12]. The submesoscale dynamics of this setup are decoupled surface QG dynamics at the surface and bottom that are stirred by the meosscale thermocline instability [28].

As typical for turbulent flows, a continuum of scales is energized. This is quantified by the kinetic and potential energy spectra,

$$K_{k,l} = \frac{1}{2}\left(|u_{k,l}|^2 + |v_{k,l}|^2\right), \qquad P_{k,l} = \frac{1}{2}\frac{|b_{k,l}|^2}{N^2}, \tag{29}$$

which we average azimuthally in wavenumber space, because the statistics are very nearly isotropic. We also average in time to characterize the statistics of the equilibrated state. The surface and bottom spectra of both kinetic and potential energy peak at a wavelength of about 200 km and fall off roughly like $k_\mathrm{h}^{-5/3}$ (Fig. 8), as predicted by surface QG turbulence theory [1]. Since smaller-scale modes decay more rapidly in the vertical than larger-scale modes, the spectra are steeper in the interior. At 100 m depth, the mesoscale energy levels are similar to those at the surface, but submesoscales energy levels are much lower.

A useful diagnostic of turbulent dynamics is the spectral energy budget. While the dynamics are completely determined by the advection of conserved quantities at the surface and bottom, we first consider the energy budget over the entire depth range. We will present a lower-order energy diagnostic for the no bottom case below.

The equations for the spectral perturbation potential and kinetic energies are

$$\frac{\partial P_{k,l}}{\partial t} = \mathrm{Re}\left[\frac{f\Lambda}{N^2} v^*_{k,l} b_{k,l} - w^*_{k,l} b_{k,l} - \frac{1}{N^2} b^*_{k,l}\, \mathrm{J}_{k,l}(\psi', b')\right] - \left(r k_\mathrm{h}^{-2} + \nu k_\mathrm{h}^{2n}\right) P_{k,l} \tag{30}$$

$$\frac{\partial K_{k,l}}{\partial t} = \mathrm{Re}\left[-f\frac{\partial}{\partial z}\left(w^*_{k,l}\psi_{k,l}\right) + w^*_{k,l} b_{k,l} + \psi^*_{k,l}\, \mathrm{J}_{k,l}(\psi', \nabla^2\psi')\right] - \left(r k_\mathrm{h}^{-2} + \nu k_\mathrm{h}^{2n}\right) K_{k,l} \tag{31}$$

where P is potential energy, K is kinetic energy, the asterisks denote complex conjugates, and Re denotes the real part. The first term on the right-hand side of the potential energy

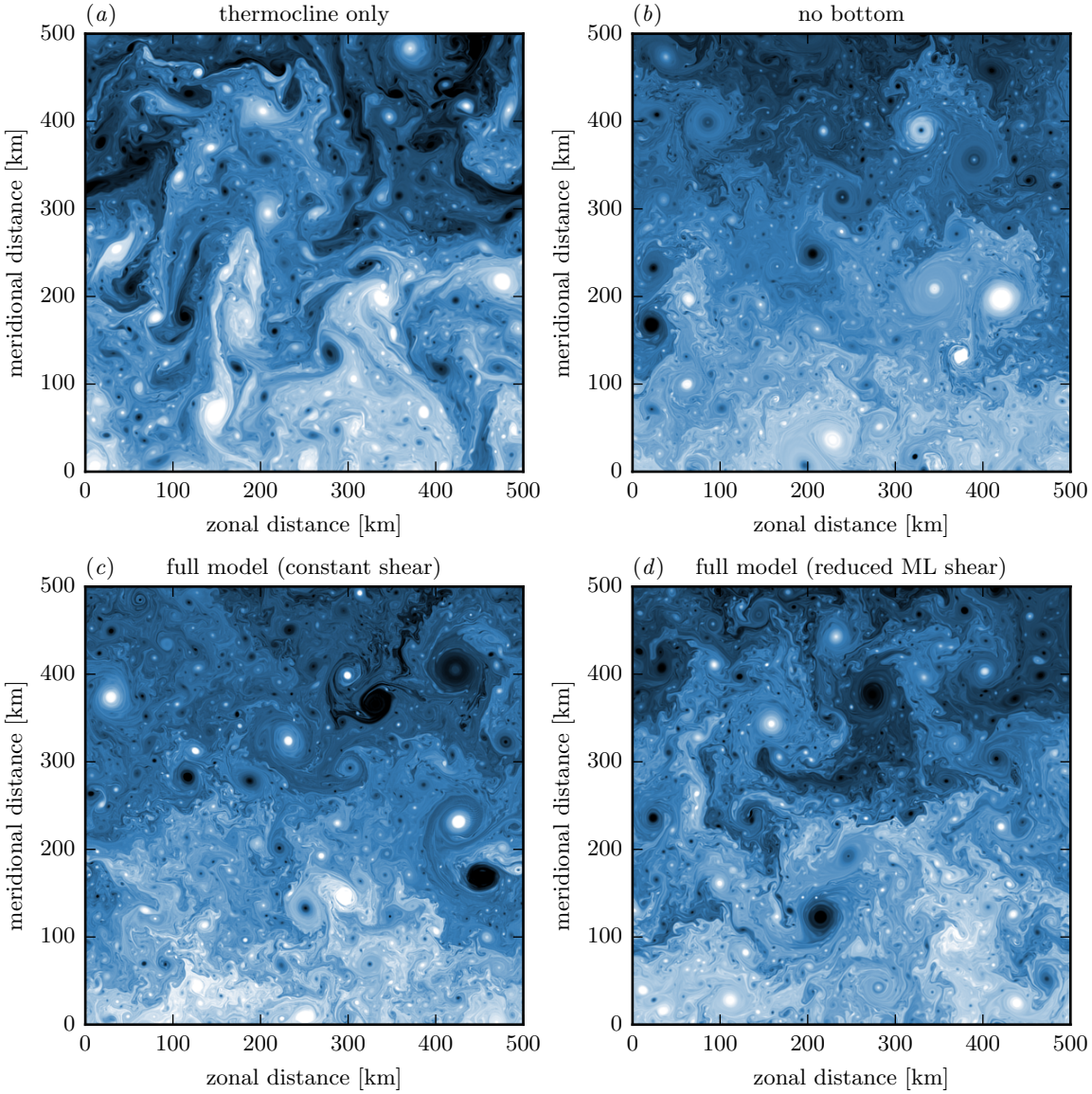


Figure 7: Snapshots of surface buoyancy (mean plus anomalies) from the equilibrated states of the (a) thermocline only simulations, (b) the no bottom simulation, (c) the full model simulation with equal mixed layer and thermocline shears, and (d) the full model with reduced mixed layer shear. The color scale extends from white (more buoyant) through blue to black (less buoyant) and extends between $\pm f \Lambda_{\mathrm{m}} L$.

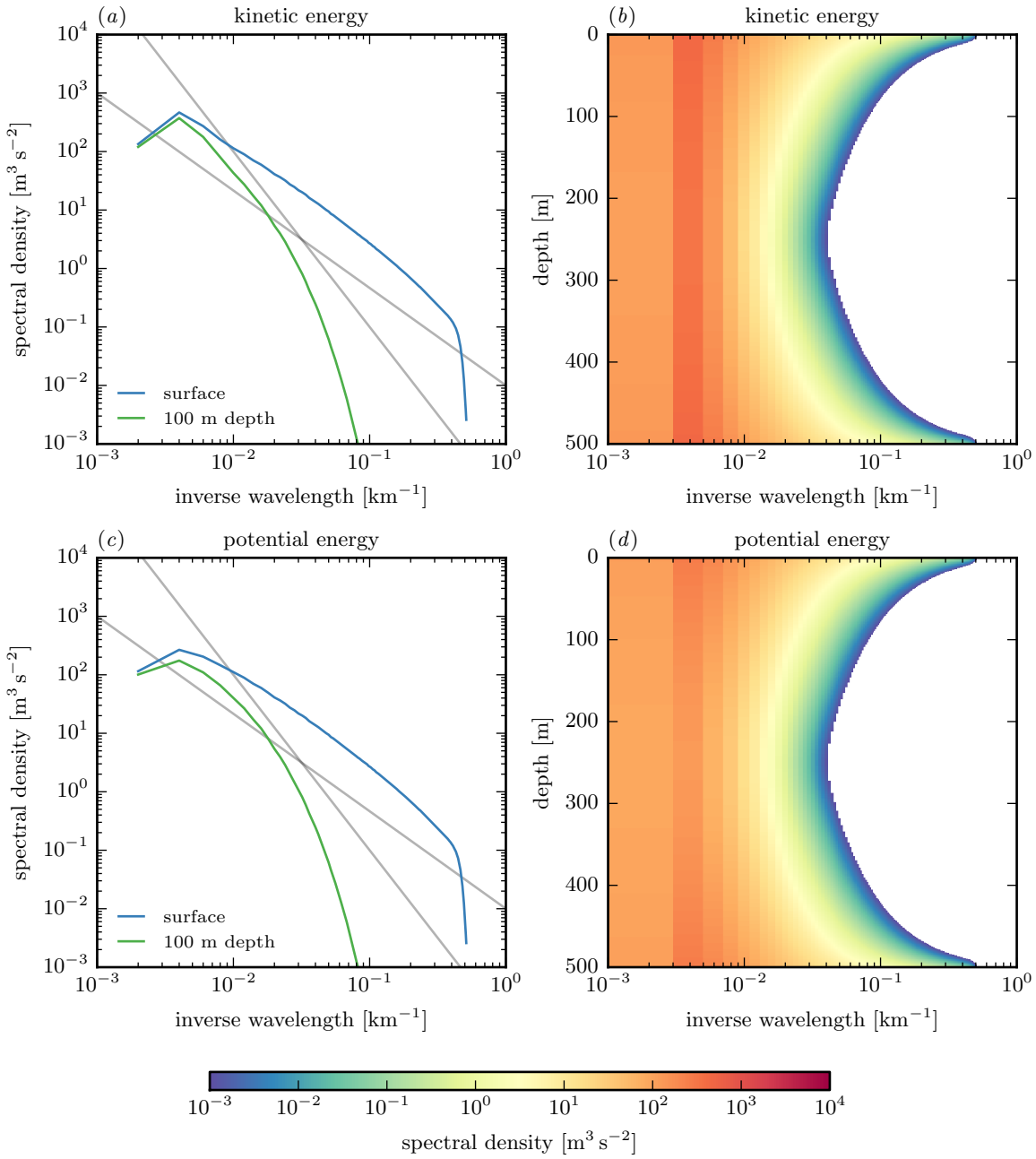


Figure 8: Wavenumber spectra of kinetic and potential energy from the thermocline only simulation. (*a*) Kinetic energy spectra at the surface and 100 m depth, (*b*) spectral density of kinetic energy in the wavenumber–depth plane, (*c*) potential energy spectra at the surface and 100 m depth, (*d*) spectral density of potential energy in the wavenumber–depth plane. In panels (*b*) and (*d*), no values below 10^{-3} m^3 s^{-2} are shown.

equation represents the extraction of potential energy from the mean flow. The second term represents the conversion from potential to kinetic energy. This term appears as a source term in the kinetic energy budget. The third term in the potential energy budget represents spectral transfer by triadic interactions. The sum of this term over all wavenumbers vanishes. An equivalent spectral transfer term appears in the kinetic energy budget (third term). Kinetic energy can also be distributed vertically by pressure fluxes, represented by the first term in the kinetic energy budget. The vertical integral of this term vanishes. The viscosity terms act as sinks for both potential and kinetic energy—hypoviscosity acting at large scales, hyperviscosity at small scales. We present these budgets again averaged azimuthally in wavenumber space and over time.

The extraction of potential energy from the mean is dominated by the largest, most energetic eddies (Fig. 9*a*). The extraction is independent of depth, because $q' = 0$ and therefore

$$0 = \mathrm{Re}\, v^*_{k,l} q_{k,l} = \mathrm{Re}\, \frac{\partial}{\partial z} \left(\frac{f}{N^2} v^*_{k,l} b_{k,l} \right). \tag{32}$$

Potential energy is transferred downscale by triadic interactions and deposited near the deformation radius and in wedges near the surface and the bottom that reach to much smaller scales (Fig. 9*b*). Where potential energy is deposited by scale interactions, it is converted into kinetic energy (Fig. 9*c*). Near the deformation radius, this conversion is due to the mesoscale instability that produces vertical buoyancy fluxes. In the wedges near the surface and bottom, the conversion is due to secondary instabilities present in the surface QG cascades, which occur independently at the surface and the bottom [28]. The kinetic energy thus created is transferred back to large scales (Fig. 9*d*). The bulk of the energy is dissipated through hypoviscosity at the scales of the largest, most energetic eddies (Fig. 9*e*). The energy dissipation through hyperviscosity is small, which reflects the fundamental property of geostrophic turbulence that energy is trapped at large scales and viscous energy dissipation vanishes as the viscosity goes to zero [18, 9].

4.2 No bottom

Now consider the case with no bottom, which has a mixed layer instability only. This instability, too, grows to finite amplitude and the flow becomes turbulent. There is a turbulent spin-up phase, in which the eddies that are initially of the size of the instability grow larger until they reach a statistical equilibrium with hypoviscosity.

A snapshot of surface buoyancy reveals that the flow's structure is quite different from the thermocline only case (Fig. 7*b*). There are many more coherent vortices. They are prominent at a scale of about 100 km, but many smaller-scale coherent vortices exist. These do not appear to be present in observations of sea surface temperature or realistic regional ocean models, a point we will come back to in the discussion. If we focus on the filamentary sea, however, there are again strong fronts, superimposed by submesoscale structure, which is realistic.

The energy spectra reflect this nearly frontal structure at the surface (Fig. 10). The kinetic energy spectra fall off slightly more steeply than $k^{-5/3}$ at the scales of the linear instability and like $k^{-5/3}$ at scales smaller than the linear short-wave cutoff, both at the surface and at the base of the mixed layer at 100 m depth (Fig. 10*a*). The mixed layer

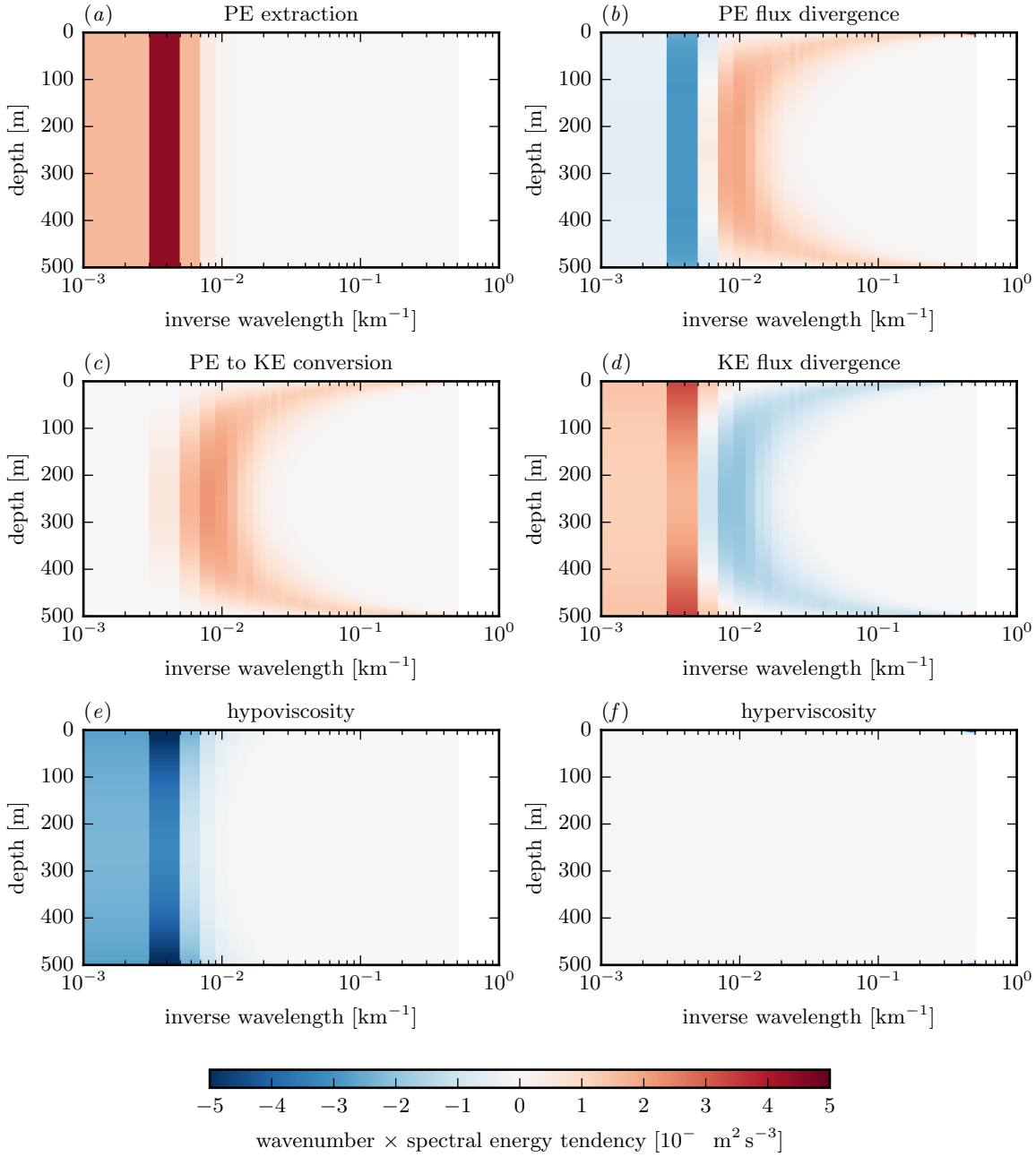


Figure 9: Spectral energy budget for the thermocline only simulation. The terms are (*a*) potential energy extraction from the mean, (*b*) spectral potential energy flux divergence, (*c*) potential to kinetic energy conversion, (*d*) kinetic energy flux divergence, including spectral flux and pressure flux, (*e*) hypoviscosity on both kinetic and potential energy, and (*f*) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

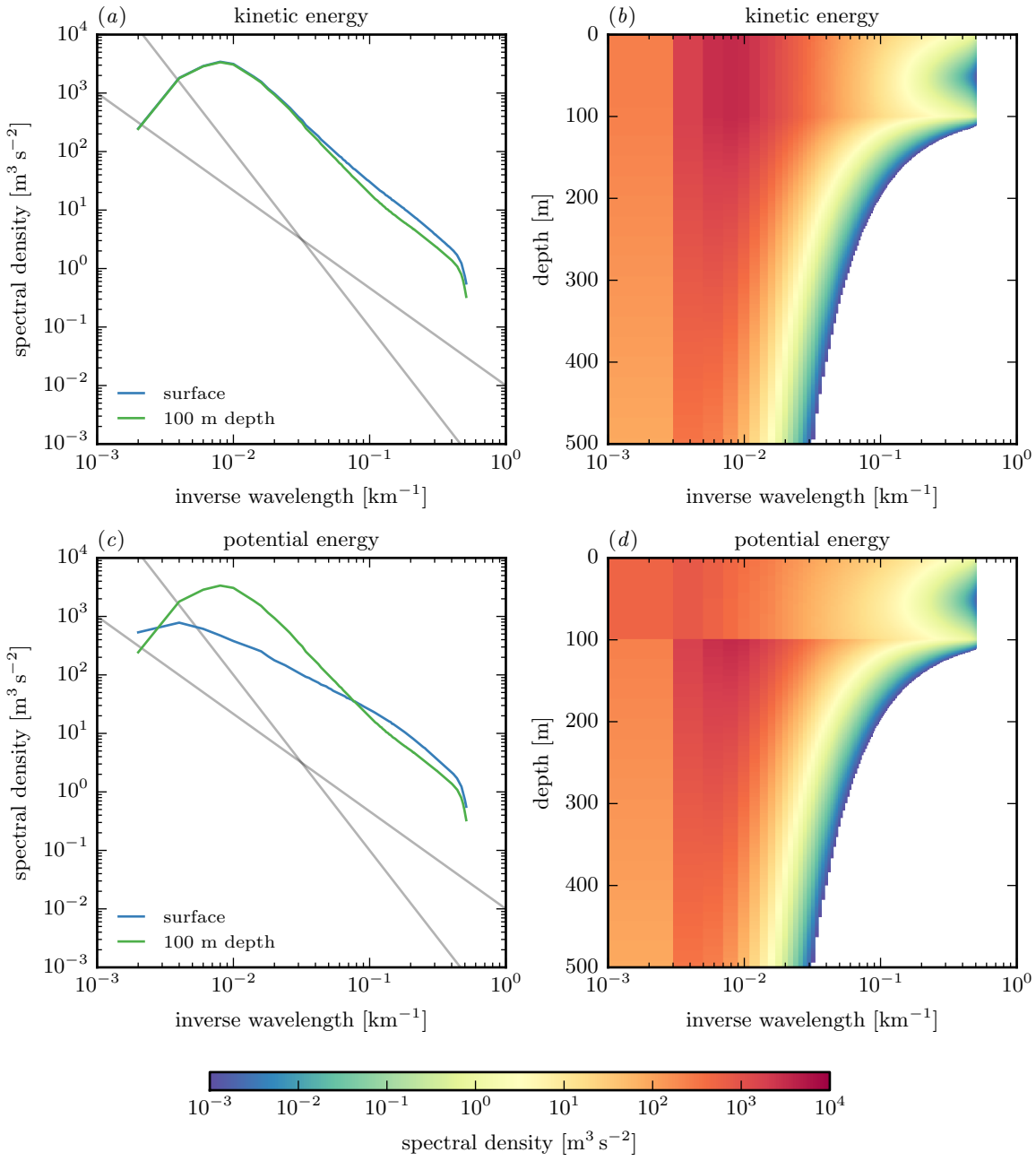


Figure 10: Wavenumber spectra of kinetic and potential energy from the no bottom simulation. (*a*) Kinetic energy spectra at the surface and 100 m depth, (*b*) spectral density of kinetic energy in the wavenumber–depth plane, (*c*) potential energy spectra at the surface and 100 m depth, (*d*) spectral density of potential energy in the wavenumber–depth plane. In panels (*b*) and (*d*), no values below 10^{-3} $\mathrm{m^3\,s^{-2}}$ are shown.

instabilities energize the entire depth of the mixed layer. This is in sharp contrast to the thermocline only simulation, in which the surface QG dynamics energize a thin wedge close to the surface only.

It should be noted that the equilibrated flow in the no bottom case is much more energetic than in the thermocline only case and more energetic than is realistic. While the equilibration by hypoviscosity is unrealistic, we will see that the enhanced energy levels are due to more efficient extraction of mean potential energy in the weakly stratified mixed layer, which is a dynamical property of the system that does not depend on how the flow is equilibrated. We will discuss reasons for this excess in energy in the conclusions.

Below the base of the mixed layer, The potential energy spectra are the same as the kinetic energy spectra (Fig. 10c). In the mixed layer, the potential energy spectra are significantly flatter than the kinetic energy spectra. This is in contrast to observations that show rough equipartition between kinetic and potential energy [6, 7].

The vertical structure of the energy shows that the mixed layer instabilities also energize the thermocline below (Fig. 10b,d). At the instability scale, the flow does not reach much into the thermocline. But as the horizontal scale of the flow increases, so does the vertical scale. The flow exhibits the familiar property of geostrophic turbulence that it barotropizes as it increases its horizontal scale [31].

The energy transfer into the thermocline can further be examined in the spectral energy budget (Fig. 11). Potential energy is again extracted at the scale of the largest, most energetic eddies, but the extraction is now confined to the mixed layer (Fig. 11a). Potential energy is transferred from the extraction scale to the scale of the mixed layer instability (Fig. 11b). The mixed layer instability converts potential energy into kinetic energy in the mixed layer, at the instability scale (Fig. 11c). The kinetic energy created by the instability again enters an inverse cascade, but now it is not only transferred to large horizontal scales, but also vertically into the thermocline (Fig. 11d). The deposition of kinetic energy at the scale of the largest eddies is well distributed across the mixed layer and upper thermocline. The vertical distribution of damping by hypoviscsosity also reflects the fact that the flow is not confined to the mixed layer at the scale of the largest eddies, where hypoviscosity acts (Fig. 11e). Hyperviscosity acts only at the smallest resolved scales (Fig. 11f). While small, it does affect the other terms in the budget. We neglect this effect, because it would tend to zero if the resolution was increased and the hyperviscosity coefficient decreased.

These energy pathways are reminiscent of the phenomenology of two-layer baroclinic turbulence. The turbulent dynamics of a two-layer system can be understood in terms of a dual cascade [26, 29]. Baroclinic energy is extracted from the mean at the scale of the largest, most energetic eddies. The barotropic flow dominates at these scales and transfers the baroclinic energy downscale. The baroclinic mode behaves like a passive tracer at these scales. Around the deformation radius, the instability converts baroclinic energy into barotropic energy. The barotropic energy then enters an inverse cascade, which gets halted at some scale by drag.

Can the turbulent dynamics induced by mixed layer instabilities be understood in similar terms? To phrase the analysis in these terms, we need a concept of modes, which the energy can be partitioned into. In our system, the vertically integrated energy can be written

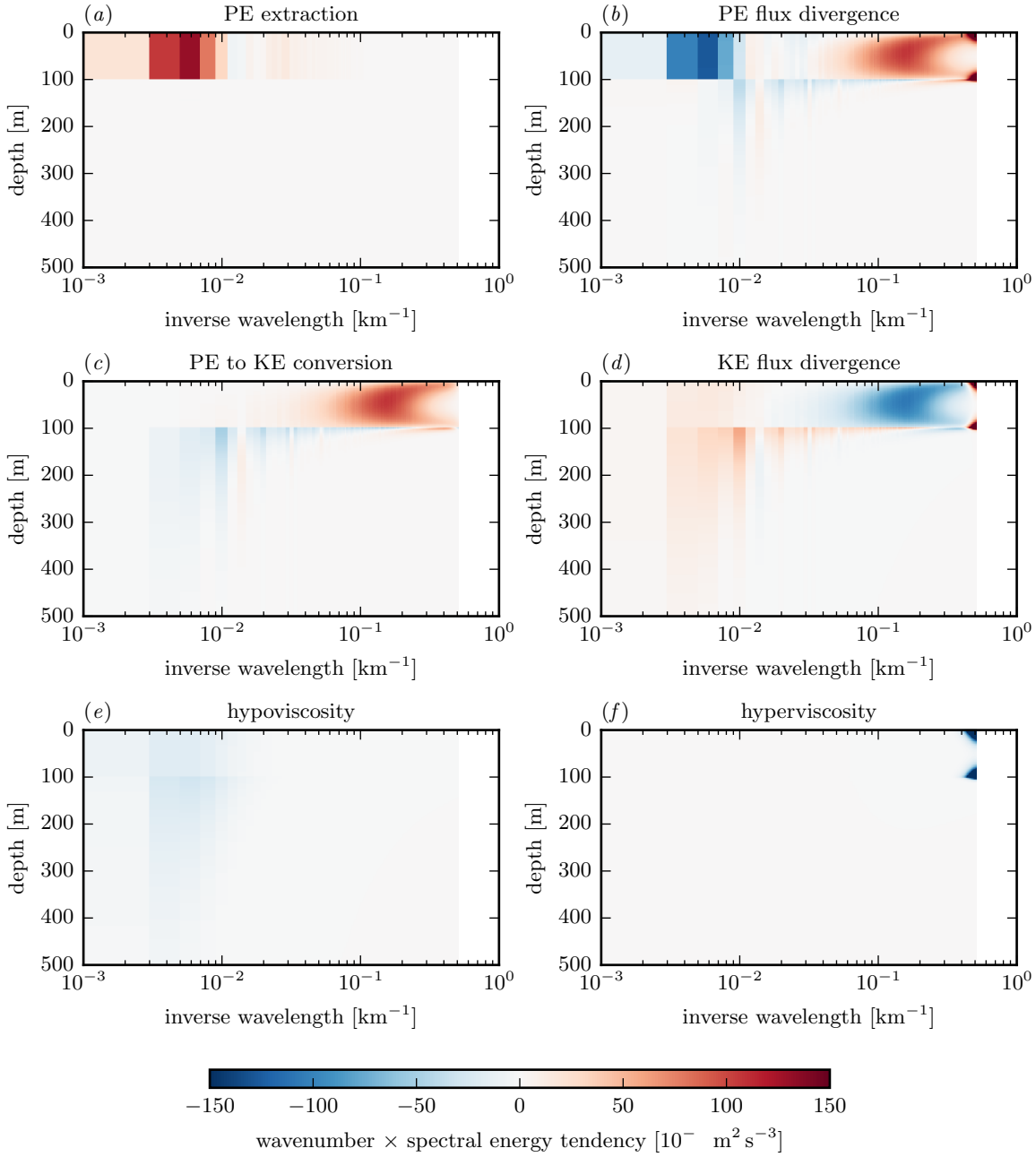


Figure 11: Spectral energy budget for the no bottom simulation. The terms are (a) potential energy extraction from the mean, (b) spectral potential energy flux divergence, (c) potential to kinetic energy conversion, (d) kinetic energy flux divergence, including spectral flux and pressure flux, (e) hypoviscosity on both kinetic and potential energy, and (f) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

entirely in terms of the quantities at the interfaces:

$$E = -\frac{1}{2}\boldsymbol{\psi}^\dagger \boldsymbol{\theta} = -\frac{1}{2}\boldsymbol{\psi}^\dagger L \boldsymbol{\psi}, \tag{33}$$

where we dropped the k, l subscripts to denote Fourier transforms. Since L is real and symmetric, it is diagonalizable by a unitary matrix,

$$L = S^\dagger D S, \tag{34}$$

where D is diagonal and consists of the real eigenvalues of L, $D_{ii} = \lambda_i$. The energy can now be written as

$$E = -\frac{1}{2}(S\boldsymbol{\psi})^\dagger D (S\boldsymbol{\psi}) = -\frac{1}{2}\sum_i \lambda_i \left|(S\boldsymbol{\psi})_i\right|^2. \tag{35}$$

This defines the modes $(S\boldsymbol{\psi})_i$ that are orthogonal with respect to the energy norm, i.e. the energy can be partitioned into contributions by these modes. The structure of the modes depends on wavenumber, because L and therefore S does.

For the no bottom case, with L given by (24), the eigenvalues of L are

$$\lambda_{0,1} = f k_\mathrm{h} \left(\frac{\coth \mu_\mathrm{m}}{N_\mathrm{m}} + \frac{1}{2N_\mathrm{t}} \pm \sqrt{\frac{\operatorname{csch}^2 \mu_\mathrm{m}}{N_\mathrm{m}^2} + \frac{1}{4N_\mathrm{t}^2}} \right) \tag{36}$$

and S, of which the columns constitute the eigenvectors, is

$$S = \begin{pmatrix} \frac{1}{\sqrt{1+\left(\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_0}{f k_\mathrm{h}} \sinh\mu_\mathrm{m}\right)^2}} & \frac{1}{\sqrt{1+\left(\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_1}{f k_\mathrm{h}} \sinh\mu_\mathrm{m}\right)^2}} \\ \frac{\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_0}{f k_\mathrm{h}}}{\sqrt{1+\left(\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_0}{f k_\mathrm{h}} \sinh\mu_\mathrm{m}\right)^2}} & \frac{\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_1}{f k_\mathrm{h}}}{\sqrt{1+\left(\cosh\mu_\mathrm{m} + \frac{N_\mathrm{m}\lambda_1}{f k_\mathrm{h}} \sinh\mu_\mathrm{m}\right)^2}} \end{pmatrix}. \tag{37}$$

For large scales, $k_\mathrm{h} \ll f/N_\mathrm{t} h$ or $\mu_\mathrm{m} \ll N_\mathrm{m}/N_\mathrm{t}$, this reduces to

$$\lambda_0 = -\frac{f k_\mathrm{h}}{2N_\mathrm{t}}, \qquad \lambda_1 = -\frac{2f^2}{N_\mathrm{m}^2 h}. \tag{38}$$

and simply

$$S = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 \\ 1 & -1 \end{pmatrix}. \tag{39}$$

This indicates that at large scales the first mode is barotropic in the mixed layer. It behaves like a surface QG mode penetrating into the thermocline. The streamfunction is proportional to k_h times the conserved quantity [12],

$$(S\boldsymbol{\psi})_0 = -\frac{f k_\mathrm{h}}{2N_\mathrm{t}} (S\boldsymbol{\theta})_0. \tag{40}$$

The second mode at large scales is baroclinic in the mixed layer. The relation between the streamfunction and the conserved quantity is

$$(S\boldsymbol{\psi})_1 = -\frac{2f^2}{N_\mathrm{m}^2 h} (S\boldsymbol{\theta})_1, \tag{41}$$

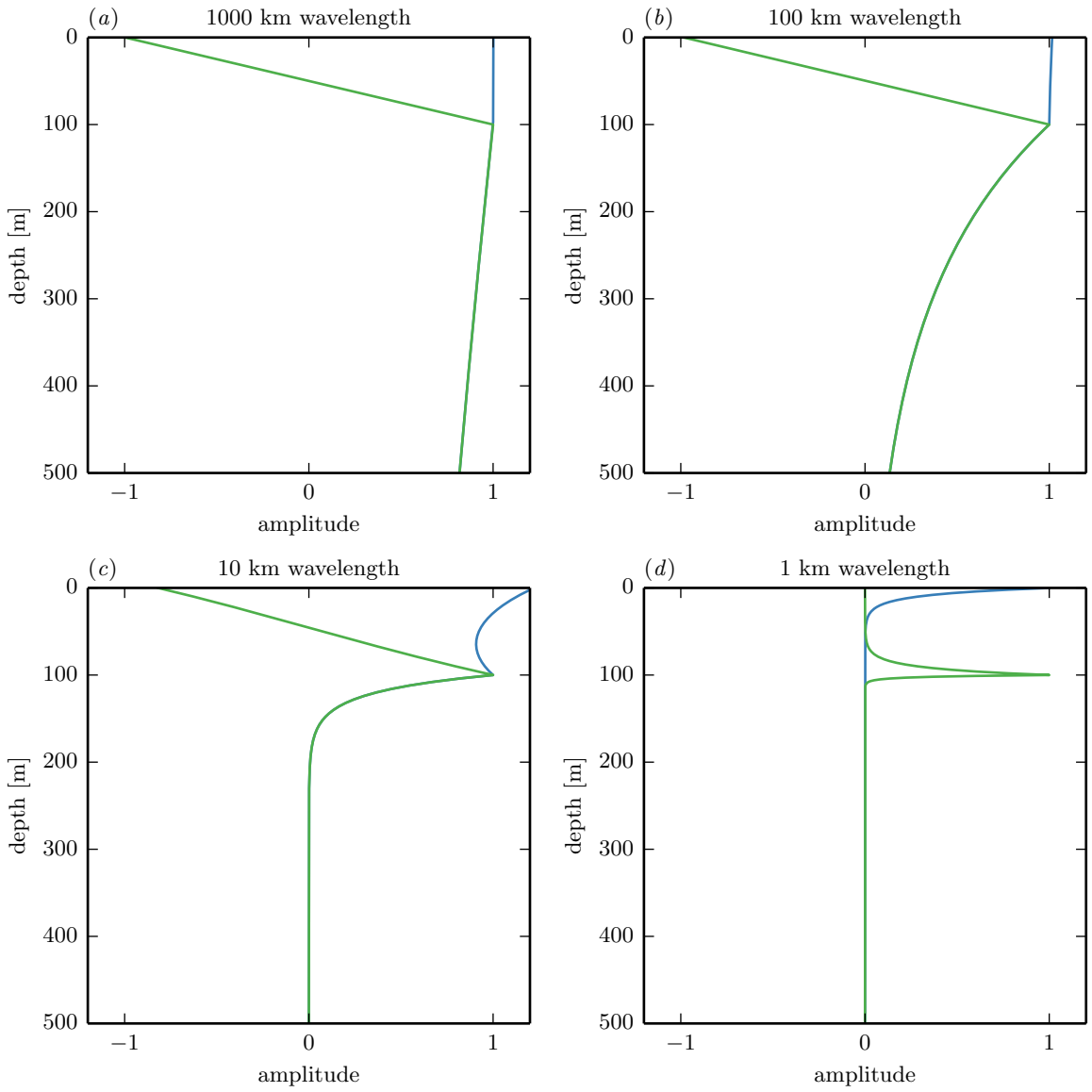


Figure 12: Vertical structure of the streamfunction corresponding to orthogonal modes in the no bottom case for different wavenumbers $k_\mathrm{h} = 2\pi/\lambda$, with the wavelength λ given in the panel titles. For panels (a), (b), and (c), the modes are normalized to unity at the interface at 100 m depth; for panel (d), the modes are normalized to have a maximum value of unity. Mode 0 is shown in blue, mode 1 in green. In panels (a), (b), and (c), the two modes coincide below the interface.

which is independent of k_h, as expected for a baroclinic mode.

This description of the modes as barotropic and baroclinic mixed layer modes only applies to large scales. At smaller scales, the modes have more complicated vertical structure (Fig. 12). At scales smaller than the mixed layer deformation radius, the modes morph into decoupled modes that are localized in the vertical at the surface and the interface. But for the cascade dynamics to be discussed, the mode structure at large scales is what is most important.

We can now consider the energy budget of these modes. We start from the vertically integrated spectral energy budget, written in terms of the conserved quantities and corresponding streamfunctions:

$$\frac{\partial E}{\partial t} = -\boldsymbol{\psi}^\dagger \frac{\partial \boldsymbol{\theta}}{\partial t}. \tag{42}$$

Again, the subscripts k, l are dropped. Using the diagonalization, this can be written as

$$\frac{\partial E}{\partial t} = -\boldsymbol{\psi}^\dagger L \frac{\partial \boldsymbol{\psi}}{\partial t} = -(S\boldsymbol{\psi})^\dagger D \frac{\partial}{\partial t}(S\boldsymbol{\psi}) = -\sum_i \lambda_i (S\boldsymbol{\psi})_i^* \frac{\partial}{\partial t}(S\boldsymbol{\psi})_i. \tag{43}$$

The budget therefore splits into

$$\frac{\partial E_i}{\partial t} = -\lambda_i (S\boldsymbol{\psi})_i^* \frac{\partial}{\partial t}(S\boldsymbol{\psi})_i. \tag{44}$$

The terms on the right-hand side of this budget can be obtained from

$$-\boldsymbol{\psi}^\dagger \frac{\partial \boldsymbol{\theta}}{\partial t} = -(S\boldsymbol{\psi})^\dagger S \frac{\partial \boldsymbol{\theta}}{\partial t}, \tag{45}$$

into which the spectral form of the evolution equation (28) is substituted. We further expand the nonlinear terms in (28) into

$$\begin{aligned} \mathrm{J}(\boldsymbol{\psi}, \boldsymbol{\theta}) &= \mathrm{J}(S^\dagger(S\boldsymbol{\psi}), S^\dagger(S\boldsymbol{\theta})) && (46) \\ &= \mathrm{J}(S^\dagger P_0(S\boldsymbol{\psi}), S^\dagger P_0(S\boldsymbol{\theta})) + \mathrm{J}(S^\dagger P_0(S\boldsymbol{\psi}), S^\dagger P_1(S\boldsymbol{\theta})) \\ &\quad + \mathrm{J}(S^\dagger P_1(S\boldsymbol{\psi}), S^\dagger P_0(S\boldsymbol{\theta})) + \mathrm{J}(S^\dagger P_1(S\boldsymbol{\psi}), S^\dagger P_1(S\boldsymbol{\theta})), && (47) \end{aligned}$$

where P_0 and P_1 are the projections onto the respective modes,

$$P_0 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \qquad P_1 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \tag{48}$$

This allows us to separate out the nonlinear interactions of the modes with themselves and with each other. The first term in (47), for example, represents the advection of the barotropic mode by the barotropic mode, to use the naming convention introduced above. The second term represents the advection of the baroclinic mode by the barotropic mode, and so on.

In terms of the orthogonal modes, the energy budget is very similar to that of a baroclinic two-layer system [20]. The extraction of potential energy from the mean flow is concentrated at the scale of the largest, most energetic eddies and creates mostly baroclinic energy (Fig. 13*b*). That input of baroclinic energy is compensated by a spectral transfer of

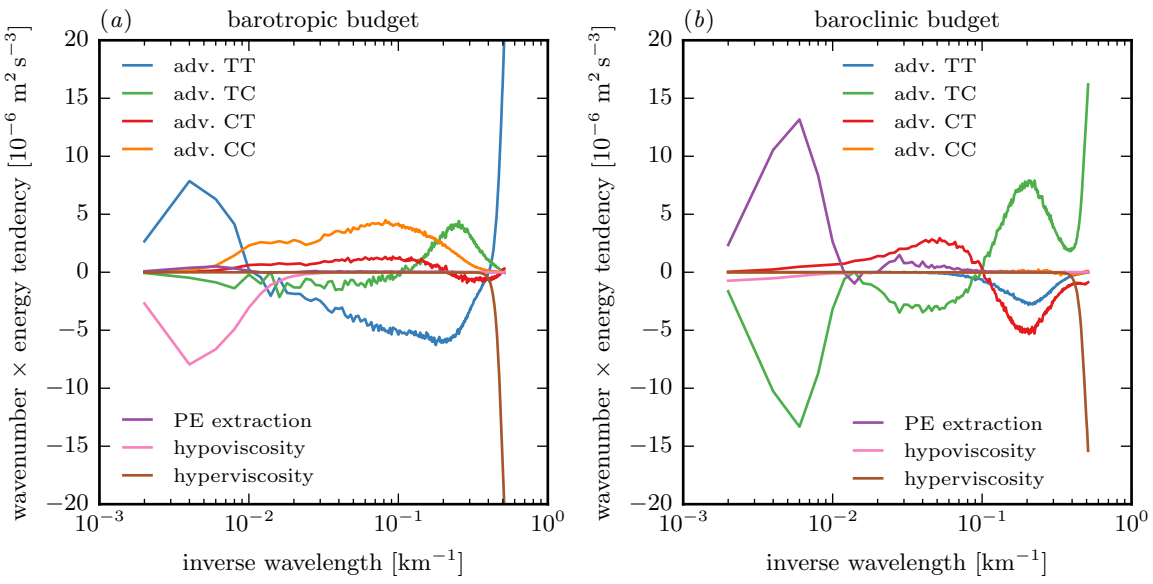


Figure 13: Modal energy budget for the no bottom case. The advective terms correspond to the contributions from the four terms in (47). The energy tendencies are multiplied by wavenumber to compensate for logarithmic shrinking.

baroclinic energy down to the instability scale, achieved by the advection of the baroclinic mode by the barotropic mode (Fig. 13*b*). The energy deposited around the instability scale is transferred to the barotropic mode by interactions between the two modes that amount to baroclinic instability (Fig. 13*b*). This energy enters the barotropic budget rather less localized in wavenumber space (Fig. 13*a*). This forcing of the barotropic mode is compensated by the upscale spectral transfer of the barotropic mode (Fig. 13*a*). The deposition of energy by the inverse cascade at the scale of the largest, most energetic eddies is compensated by hypodiffusion (Fig. 13*a*). Energy loss by hyperdiffusion again enters the budget, but is neglected because it is an artifact of finite resolution.

This model thus exhibits a dual cascade just like the classic two-layer system. Baroclinic energy is transferred downscale through advection by the barotropic mode, baroclinic instability converts baroclinic into barotropic energy, and barotropic energy is transferred back upscale in an inverse cascade. The difference is that the barotropic mode at large scales here behaves like a surface QG mode, instead of a truly barotropic or two-dimensional mode. The inverse cascade is therefore expected to yield a k_h^{-1} surface energy spectrum, which we find to emerge if the inertial range is wide enough (not shown). More importantly, the surface-QG-like behavior implies that in the inverse cascade, energy is transferred to successively larger vertical scales. This provides a pathway for mixed layer instabilities to energize the thermocline below.

4.3 Full model

We now consider the full model, which supports both deep and mixed layer instabilities. We analyze two cases, one with the same shear in mixed layer and thermocline—the setup

discussed in the linear stability analysis section—and one with reduced shear in the mixed layer.

In the constant shear case, the growth rate of the mixed layer instability is much larger than that of the deep instability. The mixed layer instability therefore grows to finite amplitude first and the evolution in the mixed layer is very similar to that of the no bottom case. Again, the eddies grow in size until they come into statistical equilibrium with hypoviscosity.

The equilibrated state of this case is very similar to that of the no bottom case, except near the bottom (Fig. 14). The energy levels and spectra at the surface and the base of the mixed layer are very similar. Near the bottom, a wedge in wavenumber–depth space is energized, just like in the thermocline only case. This is due to the surface QG dynamics at the bottom.

The energy budget is also similar to the case with no bottom (Fig. 15). The main energy pathway is again extraction of potential energy in the mixed layer, transfer to the mixed layer instability scale, conversion to kinetic energy, transfer back to large scales and into the thermocline, and dissipation by hypoviscosity. There is additional extraction in the thermocline, but that is weak compared to the extraction in the mixed layer. The dominant dynamics are therefore those described for the no bottom case. Interaction with the bottom is possible, but of secondary importance in this parameter regime.

A different picture emerges when the mixed layer shear is reduced. We choose the mixed layer shear such that the growth rates of the two instabilities are comparable, which from Eady scaling is expected to occur if $\Lambda_\mathrm{m}/N_\mathrm{m} = \Lambda_\mathrm{t}/N_\mathrm{t}$ is satisfied, so at $\Lambda_\mathrm{m} = 2.5 \times 10^{-5}$ s^{-1}. The results of the linear stability analysis for this mixed layer shear show that indeed the growth rates are comparable (Fig. 16). The horizontal scales of the instabilities and the overall structure of the dispersion curves have not changed. Reducing the mean shear by this amount means that the mean available potential energy is vertically constant.

This system with reduced mixed layer shear equilibrates to realistic and much lower energy levels than the constant shear case (Fig. 17). The energy levels are comparable to the thermocline only case. The vertical structure of energy in this case, however, is still different from the thermocline only case. The mixed layer instability, while not significantly increasing the mesoscale energy levels, does energize the mixed layer at submesoscales.

This can be explained by again considering the energy budget (Fig. 18). The potential energy extraction from the mean is now roughly constant vertically. The energy pathway induced by the mixed layer instability, however, is still active. Potential energy is transferred to the mixed layer deformation radius, where it is converted into kinetic energy and enters an inverse cascade. This energy cycle is stronger than that of surface QG turbulence in the thermocline only case (Fig. 9).

This difference between mixed layer dynamics and surface QG dynamics is also reflected in vertical velocities that are produced by the instabilities (Fig. 19). While the available potential energies are the same and the resulting surface energy levels comparable between this reduced mixed layer shear case and the thermocline only case, there are dramatically larger vertical velocities in the presence of a mixed layer. These enhanced vertical velocities extend significantly below the base of the mixed layer. The largest vertical velocities are located in the vicinity of fronts in the filamentary sea (Fig. 20). Coherent vortices, while associated with the large buoyancy gradients, induce relatively weak vertical motion. The large vertical velocities appear to be induced by the dynamic filamentary structure that

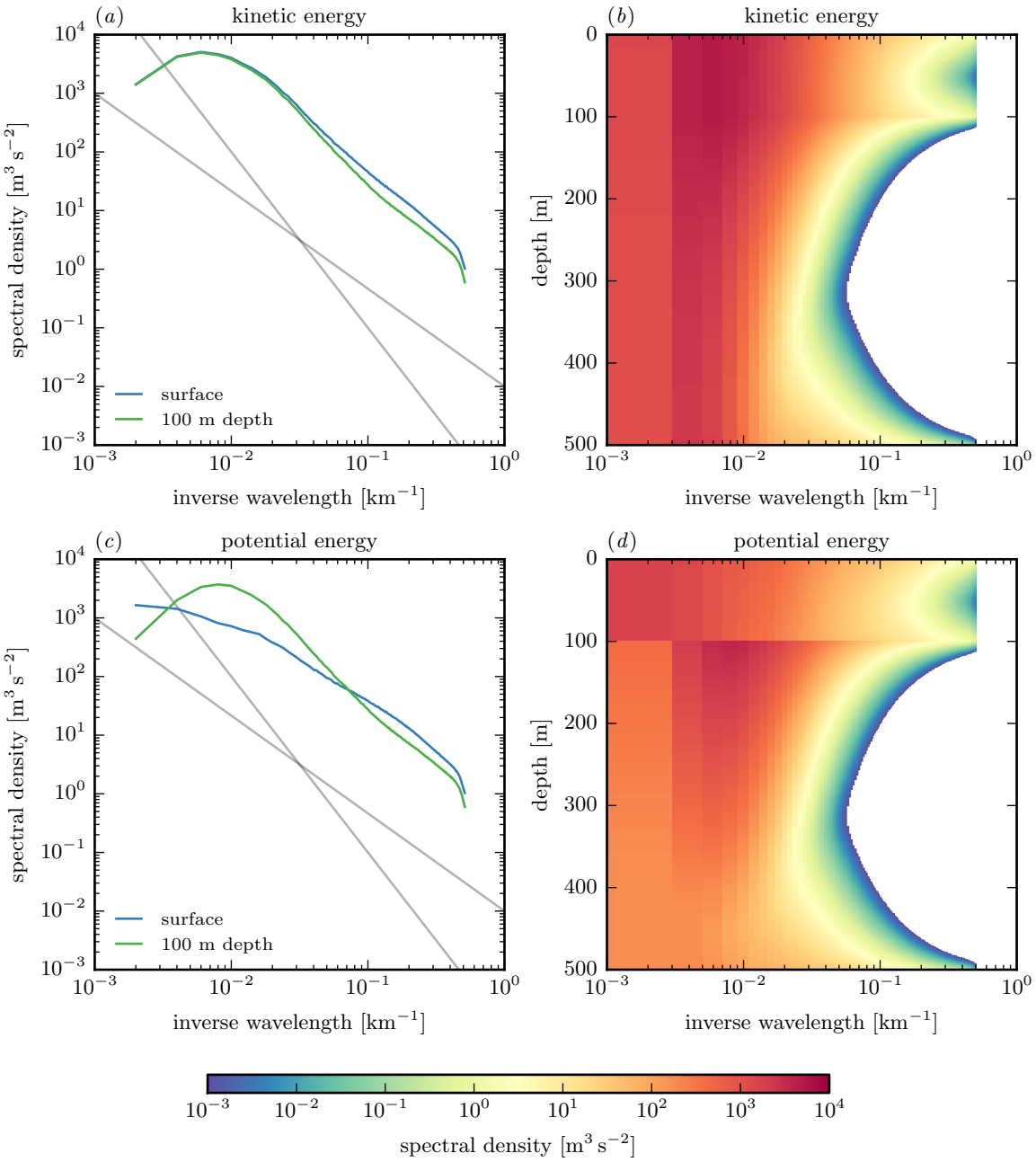


Figure 14: Wavenumber spectra of kinetic and potential energy from the full model simulation. (*a*) Kinetic energy spectra at the surface and 100 m depth, (*b*) spectral density of kinetic energy in the wavenumber–depth plane, (*c*) potential energy spectra at the surface and 100 m depth, (*d*) spectral density of potential energy in the wavenumber–depth plane. In panels (*b*) and (*d*), no values below 10^{-3} m^3 s^{-2} are shown.

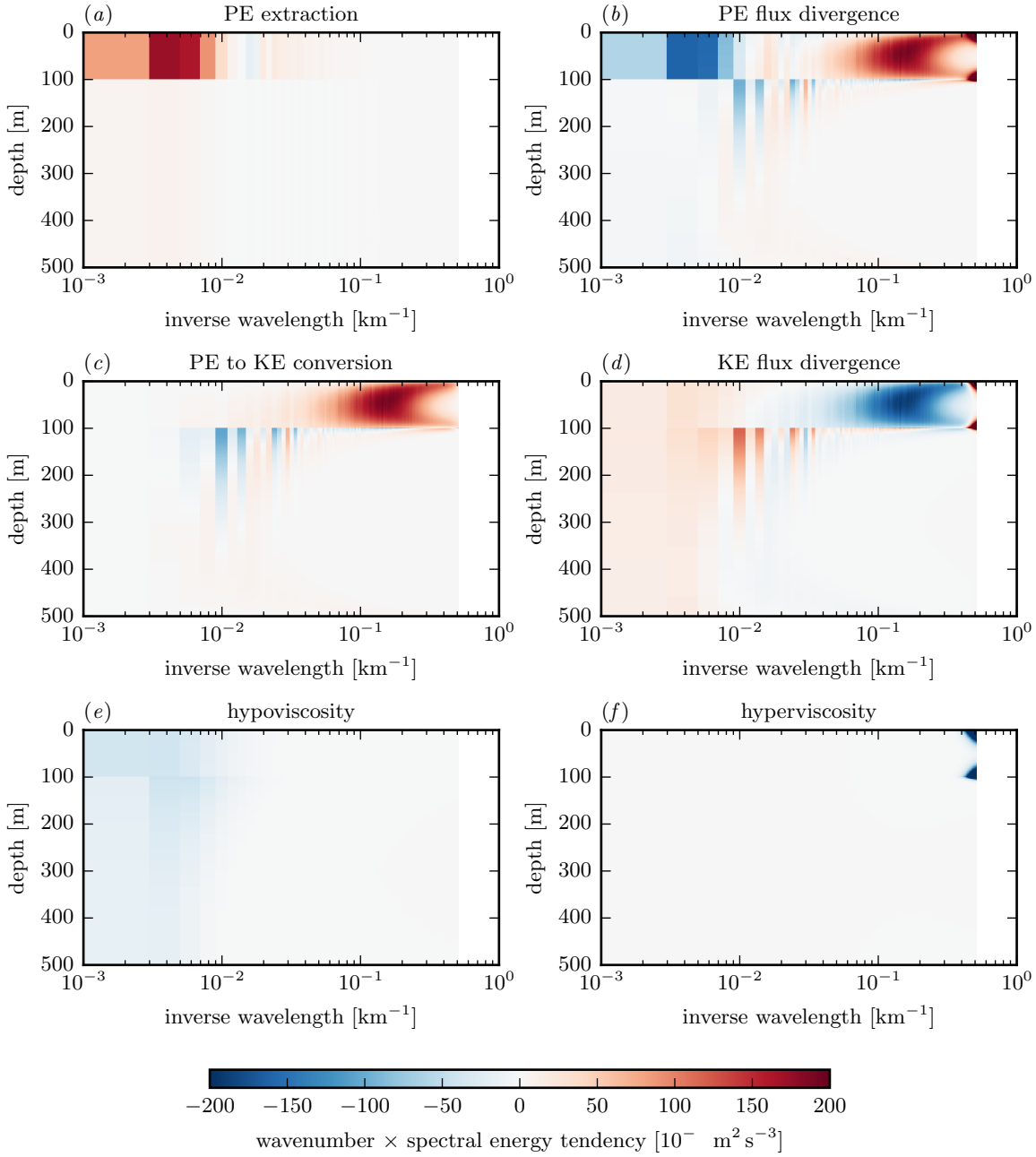


Figure 15: Spectral energy budget for the full model simulation. The terms are (a) potential energy extraction from the mean, (b) spectral potential energy flux divergence, (c) potential to kinetic energy conversion, (d) kinetic energy flux divergence, including spectral flux and pressure flux, (e) hypoviscosity on both kinetic and potential energy, and (f) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

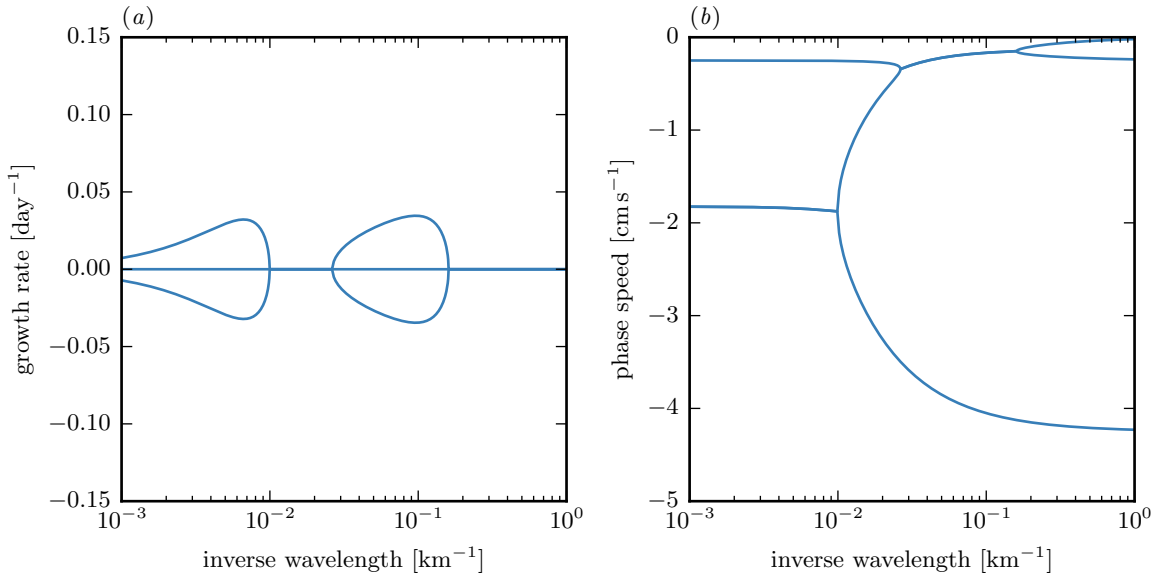


Figure 16: growth rates and phase speeds

undergoes mixed layer instabilities.

The root mean square vertical velocities are similar in structure to those found in primitive equation models [8]. A careful comparison is necessary to establish whether the QG dynamics described here reproduce the magnitude of the vertical velocities or whether non-QG effects significantly enhance vertical velocities. Such a comparison, however, is beyond the scope of the report.

5 Conclusions

Our analysis suggests that the presence of a mixed layer has a profound effect on submesoscale turbulence. The low stratification in the mixed layer, combined with geostrophic shear, provides a large amount of available potential energy that can be extracted through baroclinic instabilities in the mixed layer. The extraction is dominated by mesoscale eddies, but potential energy is subsequently cascaded down to the deformation radius of the mixed layer, where baroclinic instability converts it into kinetic energy. In the QG dynamics considered here, no energy is lost to small scales. The entire energy extracted from the mean in the mixed layer is converted to kinetic energy around the deformation radius of the mixed layer and subsequently transferred back to larger scales in an inverse cascade that also energizes the thermocline below. Through this process, mixed layer instabilities can energize the mesoscale eddy field.

These dynamics are significantly different from surface QG dynamics, which are often invoked to explain energetic submesoscales. Surface QG dynamics can only energize a thin surface layer, whereas mixed layer instabilities energize the entire mixed layer. The vertical and spectral structure found in the presence of a mixed layer is consistent with that found in wintertime observations in the Gulf Stream region, except for the potential energy spectra,

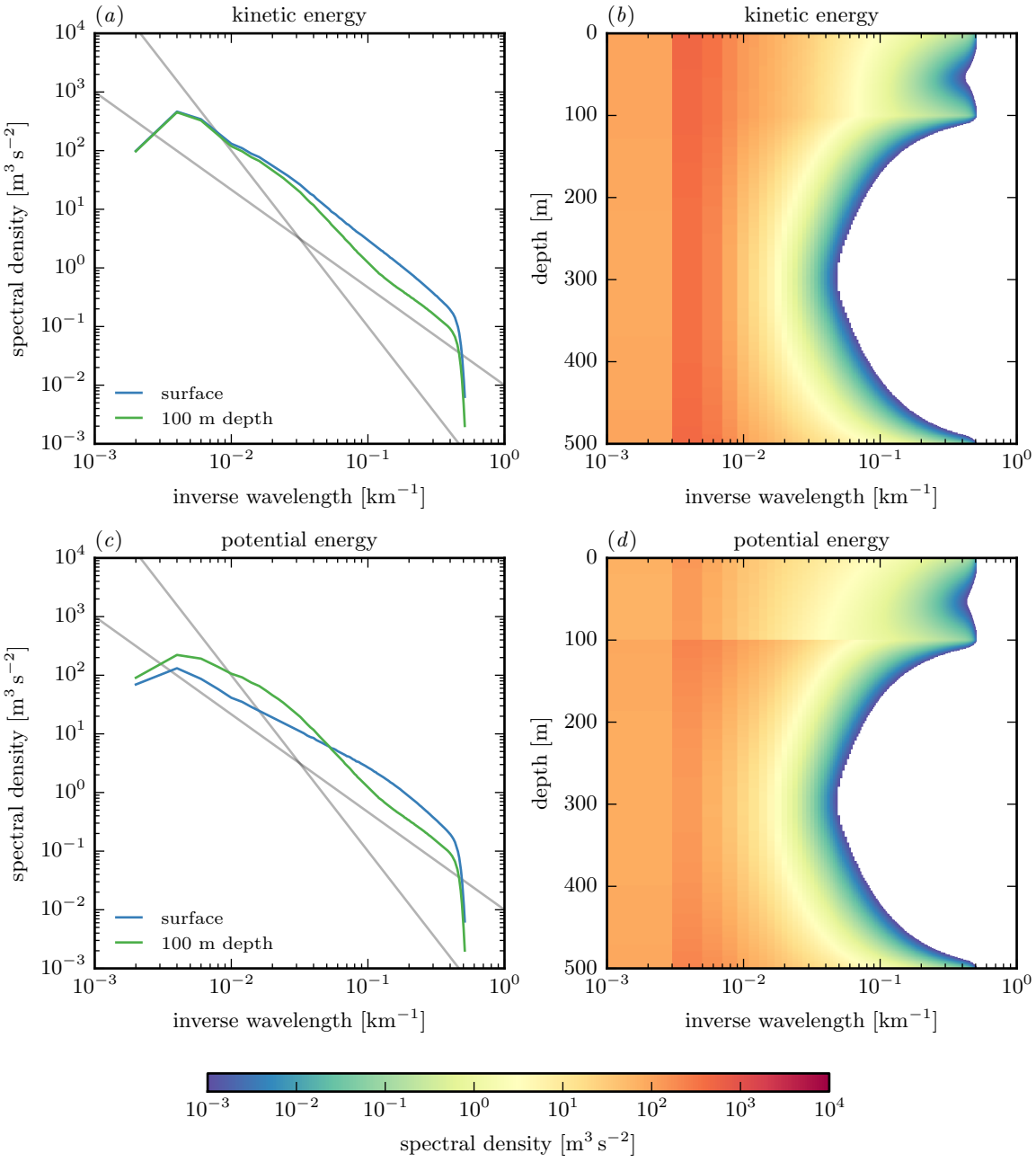


Figure 17: Wavenumber spectra of kinetic and potential energy from the full model simulation with reduced mixed layer shear. (*a*) Kinetic energy spectra at the surface and 100 m depth, (*b*) spectral density of kinetic energy in the wavenumber–depth plane, (*c*) potential energy spectra at the surface and 100 m depth, (*d*) spectral density of potential energy in the wavenumber–depth plane. In panels (*b*) and (*d*), no values below 10^{-3} m^3 s^{-2} are shown.

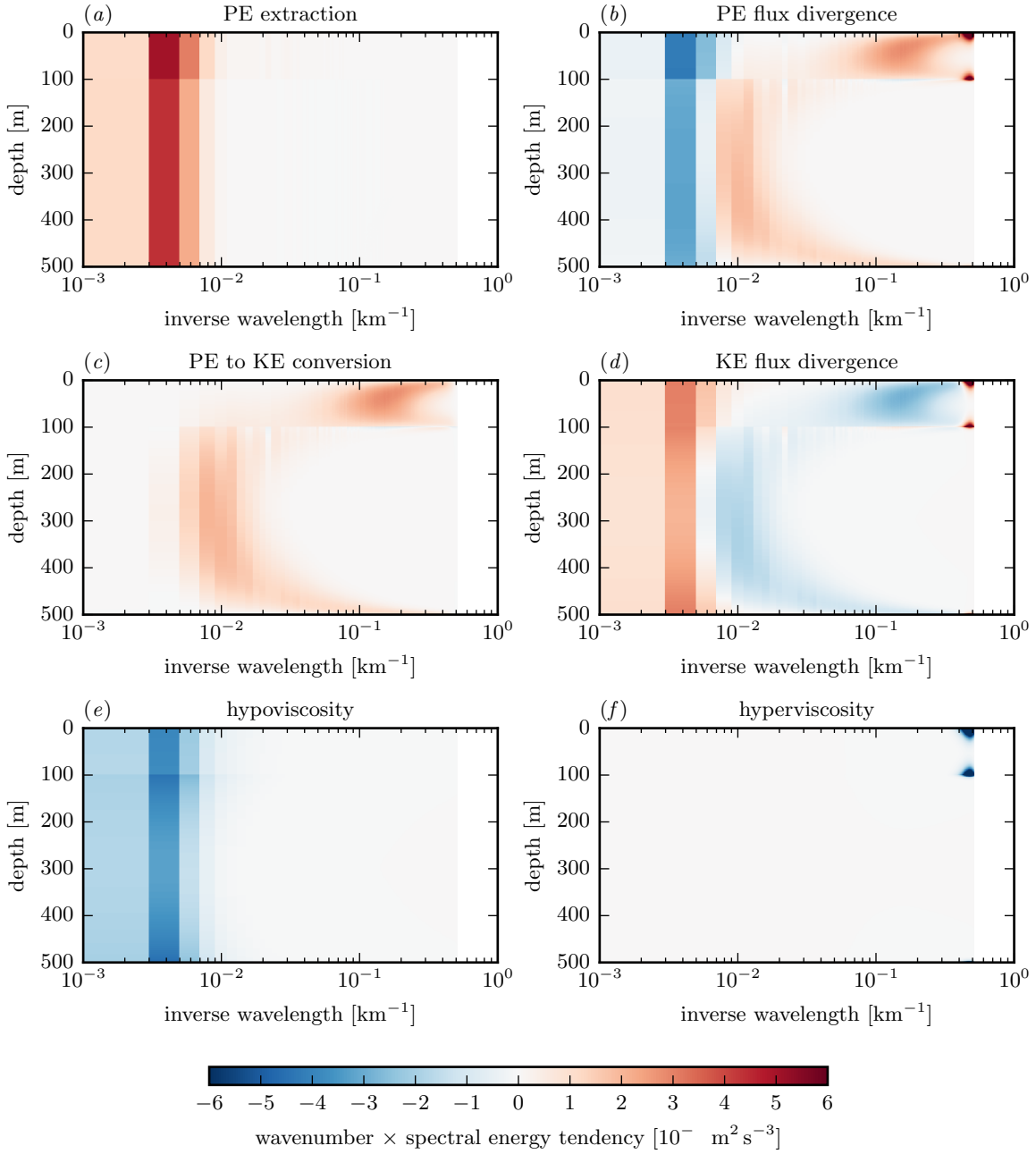


Figure 18: Spectral energy budget for the full model simulation with reduced mixed layer shear. The terms are (*a*) potential energy extraction from the mean, (*b*) spectral potential energy flux divergence, (*c*) potential to kinetic energy conversion, (*d*) kinetic energy flux divergence, including spectral flux and pressure flux, (*e*) hypoviscosity on both kinetic and potential energy, and (*f*) hyperviscosity on both kinetic and potential energy. All terms are multiplied by the wavenumber to compensate for logarithmic shrinking.

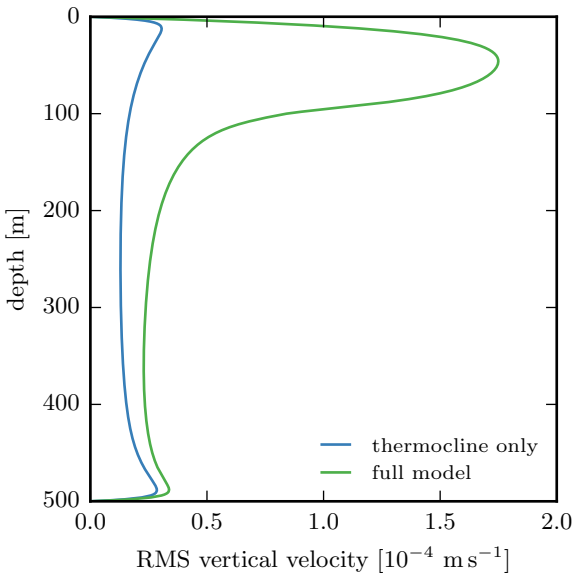


Figure 19: Profiles of root mean square vertical velocity for the thermocline only simulation and the full model simulation with reduced mixed layer shear.

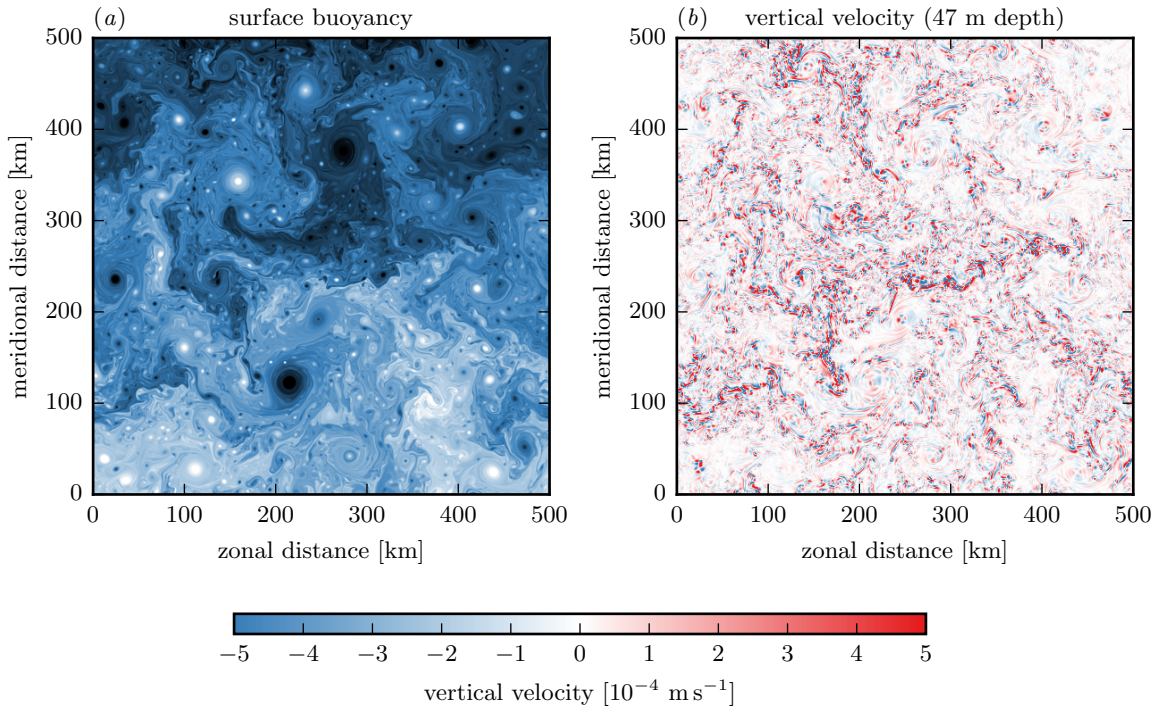


Figure 20: Concurrent surface buoyancy and vertical velocity snapshots from the reduced mixed layer shear simulation. (*a*) The surface buoyancy is the same as in Fig. 7*d*. (*b*) The vertical velocity snapshot is taken at 47 m, the depth of the maximum root mean square vertical velocity.

which are underestimated by our model.

The energization by mixed layer instabilities implies a seasonal cycle in submesoscale turbulence, which has been observed in both models and observations. Surface QG dynamics do not predict a seasonal cycle in submesoscale turbulence, because the dynamics are driven by mesoscale processes. Observational evidence for surface QG dynamics has yet to be found, but the dynamics may be important when mixed layers are shallow and surface buoyancy gradients strong.

An aspect of our model that is inconsistent with observations is that only the submesoscale appears to undergo a strong seasonal cycle, not the mesoscale, as implied by the dynamics of our simple QG model. The energization of the mesoscale by mixed layer instabilities appears to be much less effective in the real ocean. Our model is more consistent with the real ocean if the mixed layer shear is reduced. Then, the mesoscale surface energy levels do not depend on the presence of the mixed layer and the model equilibrates to realistic submesoscale energy levels.

This inconsistency with observations may point to shortcomings of the QG system in predicting the equilibrated energy levels. The lack of small-scale energy dissipation makes the mixed layer instabilities very effective in energizing the entire turbulent flow. If non-QG effects were allowed, a fraction of the energy extracted from the mean in the mixed layer would be dissipated at small scales. This energy leak to small scales is likely as the Rossby and Froude numbers can be large enough at submesoscales to only marginally allow QG scaling. Especially the host of coherent vortices forming in QG dynamics (Fig. 7) may not form in the real ocean, because they would be unstable to non-QG instabilities. The possibility of an energy leak to small scales has been demonstrated by Molemaker et al., who studied an Eady instability with $\mathrm{Ro} = \mathrm{Fr} = 0.5$ using the full Boussinesq equations [23]. While much of the energy extracted from the mean is still trapped at large scales, as predicted by QG dynamics, some is lost to dissipation at small scales. A small leak of energy in the instability may make a big difference in the cascade dynamics, because that energy is not transferred back to mesoscales, where it would further enhance the extraction of potential energy from the mean. We are currently investigating whether primitive equation models indeed equilibrate to lower energy levels than QG models at moderate Rossby and Froude numbers and will report the results elsewhere.

Another possible explanation for the lack of a strong seasonal cycle in the real ocean is that our setup overestimates the amount of mean potential energy available for extraction. QG dynamics do not allow for restratification, which would quickly increase the mixed layer stratification and thereby reduce the available potential energy. We also hold the mean shear fixed, which amounts to providing an infinite reservoir of available potential energy. It does appear, however, that strong wintertime atmospheric forcing, with heat fluxes on the order of several hundred watts per square meter, can effectively maintain both a weak mixed layer stratification and horizontal buoyancy gradients.

Acknowledgments

I would like to thank Glenn Flierl, Raf Ferrari, and Baylor Fox-Kemper, who made every imaginable effort to be available for stimulating discussions, both locally and remotely. I would also like to thank Malte Jansen, Bill Young, Ryan Abernathey, and Paola Cessi

for interesting discussions throughout the summer. I owe thanks to Antonello Provenzale, Glenn Flierl, and Raf Ferrari, our directors, for putting together such an interesting and stimulating program. Special thanks are due to Kerry Emanuel and Geoff Vallis for their engaging lectures and lunchtime discussions with us fellows. And, of course, I am grateful for the companionship of my fellow fellows who made this summer a great and memorable experience.

A Multi-layer model

A system of n layers of constant PV, of thickness h_i and stratification N_i, consists of $n+1$ conserved quantities that are advected by the geostrophic flow at their respective levels. Compared to the two-layer model considered in the main text, additional interface quantities analogous to θ_1 are present. The linear operator in the inversion relation (11) has tridiagonal structure:

$$L = fk_\mathrm{h} \begin{pmatrix} -\frac{\coth\mu_0}{N_0} & \frac{\operatorname{csch}\mu_0}{N_0} & & & & \\ \frac{\operatorname{csch}\mu_0}{N_0} & -\frac{\coth\mu_0}{N_0} - \frac{\coth\mu_1}{N_1} & \frac{\operatorname{csch}\mu_1}{N_1} & & & \\ & \ddots & \ddots & \ddots & & \\ & & \frac{\operatorname{csch}\mu_{n-1}}{N_{n-1}} & -\frac{\coth\mu_{n-1}}{N_{n-1}} - \frac{\coth\mu_n}{N_n} & \frac{\operatorname{csch}\mu_n}{N_n} & \\ & & & \frac{\operatorname{csch}\mu_n}{N_n} & -\frac{\coth\mu_n}{N_n} \end{pmatrix}, \quad (49)$$

where $\mu_i = N_i k_\mathrm{h} h_i / f$. It may be more efficient to solve the inversion relation numerically instead of calculating the inverse of this matrix, which will in general be full.

One can also include a PV gradient due to differential rotation. This can be done using a trick described by Lindzen [21]: instead of using linear shear and constant stratification in the layers, one can use parabolic shear or a modified stratification profile, which allows cancellation of the contribution from the β-effect and retaining constant PV within the layers. The PV gradient due to β is then included in the PV sheets at the interfaces.

B Density jump at layer interface

If there is a buoyancy jump g' at the interface, the matching conditions must be modified. To ensure a continuous pressure at the interface at $z = -h + \eta$, we require

$$\psi(-h^+) - \psi(-h^-) = -\frac{g'}{f}\eta. \quad (50)$$

Here, η is the perturbation of the interface between the constant-PV layers. The buoyancy equations (5) can be combined with the kinematic condition

$$w = \frac{\partial \eta}{\partial t} + \mathrm{J}(\psi, \eta), \quad (51)$$

applied at $z = -h^+$ and $z = -h^-$, to give

$$\frac{\partial \theta_1}{\partial t} + \mathrm{J}(\psi_1, \theta_1) = 0, \qquad \frac{\partial \theta_2}{\partial t} + \mathrm{J}(\psi_2, \theta_2) = 0, \quad (52)$$

where

$$\theta_1 = f\frac{b(-h^+)}{N_{\mathrm{m}}^2} + f\eta, \qquad \theta_2 = f\frac{b(-h^-)}{N_{\mathrm{t}}^2} + f\eta, \qquad \psi_1 = \psi(-h^+), \qquad \psi_2 = \psi(-h^-). \quad (53)$$

Together with the conservation of surface and bottom buoyancy,

$$\theta_0 = -f\frac{b(0)}{N_{\mathrm{m}}^2}, \qquad \theta_3 = f\frac{b(-H)}{N_{\mathrm{t}}^2}, \quad (54)$$

and the inversion relation obtained by solving (12) with the matching conditions above, the model is complete. It now consists of four conserved quantities.

References

[1] W. Blumen, *Uniform potential vorticity flow: Part I. Theory of wave interactions and two-dimensional turbulence*, J. Atmos. Sci., 35 (1978), pp. 774–783.

[2] ——, *On short-wave baroclinic instability*, J. Atmos. Sci., 36 (1979), pp. 1925–1933.

[3] G. Boccaletti, R. Ferrari, and B. Fox-Kemper, *Mixed layer instabilities and restratification*, J. Phys. Oceanogr., 37 (2007), pp. 2228–2250.

[4] J. P. Boyd, *The energy spectrum of fronts: Time evolution of shocks in Burgers' equation*, J. Atmos. Sci., 49 (1992), pp. 128–139.

[5] F. P. Bretherton, *Critical layer instability in baroclinic flows*, Q. J. R. Meteorol. Soc., 92 (1966), pp. 325–334.

[6] J. Callies and R. Ferrari, *Interpreting Energy and Tracer Spectra of Upper-Ocean Turbulence in the Submesoscale Range (1–200 km)*, J. Phys. Oceanogr., 43 (2013), pp. 2456–2474.

[7] J. Callies, R. Ferrari, J. M. Klymak, and J. Gula, *Seasonality in submesoscale turbulence*, Nat. Commun., 6 (2015).

[8] X. Capet, J. C. McWilliams, M. J. Molemaker, and A. F. Shchepetkin, *Mesoscale to submesoscale transition in the California Current System. Part I: flow structure, eddy flux, and observational tests*, J. Phys. Oceanogr., 38 (2008), pp. 29–43.

[9] J. G. Charney, *Geostrophic turbulence*, J. Atmos. Sci., 28 (1971), pp. 1087–1095.

[10] E. T. Eady, *Long waves and cyclone waves*, Tellus, 1 (1949), pp. 33–52.

[11] R. Ferrari, *A frontal challenge for climate models*, Science, 332 (2011), pp. 316–317.

[12] I. M. Held, R. T. Pierrehumbert, S. T. Garner, and K. L. Swanson, *Surface quasi-geostrophic dynamics*, J. Fluid Mech., 282 (1995), pp. 1–20.

[13] J. R. Holton, *An introduction to dynamic meteorology*, Elsevier Academic Press, Burlington, MA, 4 ed., 2004.

[14] B. J. Hoskins and F. P. Bretherton, *Atmospheric frontogenesis models: mathematical formulation and solution*, J. Atmos. Sci., 29 (1972), pp. 11–37.

[15] M. Juckes, *Quasigeostrophic dynamics of the tropopause*, J. Atmos. Sci., 51 (1994), pp. 2756–2768.

[16] P. Klein, B. L. Hua, G. Lapeyre, X. Capet, S. Le Gentil, and H. Sasaki, *Upper ocean turbulence from high-resolution 3D simulations*, J. Phys. Oceanogr., 38 (2008), pp. 1748–1763.

[17] P. Klein and G. Lapeyre, *The oceanic vertical pump induced by mesoscale and submesoscale turbulence*, Ann. Rev. Mar. Sci., 1 (2009), pp. 351–375.

[18] R. H. Kraichnan, *Inertial ranges in two-dimensional turbulence*, Phys. Fluids, 10 (1967), p. 1417.

[19] G. Lapeyre and P. Klein, *Dynamics of the upper oceanic layers in terms of surface quasigeostrophy theory*, J. Phys. Oceanogr., 36 (2006), pp. 165–176.

[20] V. D. Larichev and I. M. Held, *Eddy amplitudes and fluxes in a homogeneous model of fully developed baroclinic instability*, J. Phys. Oceanogr., 25 (1995), pp. 2285–2297.

[21] R. S. Lindzen, *The Eady problem for a basic state with zero pv gradient but $\beta \neq 0$*, J. Atmos. Sci., 51 (1994), pp. 3221–3226.

[22] J. A. Mensa, Z. Garraffo, A. Griffa, T. M. Özgökmen, A. Haza, and M. Veneziani, *Seasonality of the submesoscale dynamics in the Gulf Stream region*, Ocean Dyn., 63 (2013), pp. 923–941.

[23] M. J. Molemaker, J. C. McWilliams, and X. Capet, *Balanced and unbalanced routes to dissipation in an equilibrated Eady flow*, J. Fluid Mech., 654 (2010), pp. 35–63.

[24] J. Pedlosky, *Geophysical Fluid Dynamics*, Springer, New York, 2 ed., 1987.

[25] N. A. Phillips, *Energy transformations and meridional circulations associated with simple baroclinic waves in a two-level, quasi-geostrophic model*, Tellus, 6 (1954), pp. 273–286.

[26] P. B. Rhines, *The dynamics of unsteady currents*, in Sea, vol. VI, E. Goldberg, ed., Wiley, 1977, pp. 189–318.

[27] C. Rivest, C. A. Davis, and B. F. Farrell, *Upper-Tropospheric Synoptic-Scale Waves. Part I: Maintenance as Eady Normal Modes*, J. Atmos. Sci., 49 (1992), pp. 2108–2119.

[28] G. Roullet, J. C. McWilliams, X. Capet, and M. J. Molemaker, *Properties of steady geostrophic turbulence with isopycnal outcropping*, J. Phys. Oceanogr., 42 (2012), pp. 18–38.

[29] R. Salmon, *Two-layer quasi-geostrophic turbulence in a simple special case*, Geophys. Astrophys. Fluid Dyn., 10 (1978), pp. 25–52.

[30] K. S. Smith and E. Bernard, *Geostrophic turbulence near rapid changes in stratification*, Phys. Fluids, 25 (2013), p. 046601.

[31] K. S. Smith and G. K. Vallis, *The scales and equilibration of midocean eddies: freely evolving flow*, J. Phys. Oceanogr., 31 (2001), pp. 554–571.

[32] P. H. Stone, *Frontogenesis by horizontal wind deformation fields*, J. Atmos. Sci., 23 (1966), pp. 455–465.

[33] G. K. Vallis, *Atmospheric and Oceanic Fluid Dynamics*, Cambridge University Press, New York, 2006.

Understanding Eddy Saturation in the Southern Ocean using Mean Field Theory

Joe Fitzgerald

August 16, 2015

Abstract

The winds over the Southern Ocean are increasing with time, and the impact that this change in forcing will have on the structure of the Antarctic Circumpolar Current (ACC) is not well understood. The statistical mean state of the ACC is set by a competition of several physical processes including the direct effects of winds, as well as the effects of eddies produced by baroclinic instability. We formulate a highly idealized two-layer quasi-geostrophic model of the ACC using a zonally-reentrant channel geometry. We determine the dependence of the equilibrium shear on the imposed wind stress in this model and find that some eddy saturation occurs: the domain mean eddy diffusivity is found to increase with increasing wind stress, mitigating the wind-driven tendency to steepen the interfacial slope. We then formulate the mean-field dynamics of our idealized model in which eddy-eddy interactions are discarded but eddy-mean flow interactions are retained. We find that the mean-field model robustly reproduces the qualitative dependence of the shear on wind stress seen in the full model. The mean-field model also captures the sense of the dependence of the full dynamics on other model control parameters such as the bottom friction. Our results suggest that mean-field dynamics constitutes a simplified and useful theoretical framework within which progress toward a physical understanding of eddy diffusion in the ACC might be made.

1 Introduction

The Antarctic Circumpolar Current (ACC) is a strong eastward current circulating around the Antarctic continent with surface speeds on the order of 30 cm/s and strong eastward flow extending to depth. The ACC is driven by persistent westerly winds blowing over the surface of the Southern Ocean (SO) with characteristic velocities on the order of 10 m/s, corresponding to a surface wind stress of approximately 0.2 N/m^2 [10]. The ACC is characterized by a strong vertical shear of zonal velocity, with the large surface currents diminishing to a weak flow near the ocean bottom. This vertical shear is associated with the sloping of isopycnal surfaces across the ACC: isopycnals deepen moving away from the Antarctic continent, with the rate of deepening being proportional to the strength of the zonal mean shear. The SO is zonally reentrant in the Drake Passage latitude band near

60S, but connects with the zonally-bounded ocean basins equatorward of the ACC. Isopycnals in the ocean basins are relatively flat and are continuous with those in the ACC. An emerging theoretical paradigm [14, 15] relates the depth of each isopycnal in the basin interior to the slope of that isopycnal across the ACC. In this picture, isopycnal slopes in the ACC play an important role in setting the deep stratification of the ocean basins. The deep stratification in turn influences the large-scale meridional overturning circulation in the Atlantic basin, which largely flows along isopycnal surfaces. The overturning circulation redistributes carbon and heat on the planetary scale and thus plays a critical role in global climate [10]. These considerations indicate that understanding the physical processes controlling the strength and vertical structure of the ACC constitutes an important problem in climate dynamics.

Observational records indicate that the strength and structure of the winds over the SO are changing in time. These changes have been linked to Antarctic ozone depletion, and have led to stronger eastward surface winds at ACC latitudes [19, 20]. How the ACC will respond to this change in forcing is not well understood. Observational results have suggested that the SO carbon sink is weakening as a result of the wind increases due to an associated enhancement of the upwelling rate in the SO [8]. Some theoretical arguments [15] connect such increases in upwelling with steepening of ACC isopycnal slopes. However, observations have also suggested that isopycnal slopes in the ACC have remained essentially constant [3], and high-resolution numerical modeling results have recently demonstrated that SO overturning rate and isopycnal slopes may not covary as suggested by theoretical models [12].

The spatial structure of the zonal mean isopycnals in the depth-latitude ACC cross-section are determined by the interactions of several physical processes. These processes include the surface buoyancy fluxes, which set the buoyancy structure in the outcrop region; the surface Ekman transport, which tends to steepen isopycnals in the ACC region; and heat fluxes due to baroclinic eddies, which act to flatten isopycnals. Surface buoyancy fluxes are commonly taken to be constant in theoretical models of the ACC, although this assumption has been shown to have important impacts on the response of the overturning as winds are increased [2]. The Ekman response is well-understood and increases linearly with the imposed wind stress. The eddy response is less straightforward, and understanding how eddy effects vary with forcing constitutes a crucial step toward understanding the large scale stratification and overturning of the world oceans.

Recent work has shown that numerical model predictions of the response of the ACC mean state to changes in wind forcing is strongly dependent on how eddies are treated in those models. Coarse resolution simulations typically parameterize eddies using the Gent-McWilliams scheme and predict that the overturning and isopycnal slopes will respond strongly to increases in the wind stress [14, 15]. In contrast, high resolution simulations that permit or resolve eddies in the SO have predicted a dramatically reduced sensitivity of isopycnal slopes to increasing winds [13]. This reduction in sensitivity associated with the accurate simulation of SO eddies is referred to as eddy-saturation, and can produce isopycnal slopes that are entirely independent of wind stress in some cases.

The effects of eddies on the SO mean state are frequently studied using the theoretical framework of the transformed Eulerian mean (TEM) [9, 7]. The TEM description of the relationship between eddies, winds, and isopycnal slopes is typically phrased as follows

(following [2]). The steady-state Eulerian zonal mean budget buoyancy budget is

$$\overline{v}\partial_y\overline{b} + \overline{w}\partial_z\overline{b} + \partial_y\overline{v'b'} + \partial_z\overline{w'b'} = \mathcal{D} \tag{1}$$

in which (y, z) are the meridional and vertical Cartesian coordinates, (v, w) are the velocity components in the y and z directions, b is the buoyancy, $\mathcal{D}$ represents any diabatic forcing terms, and overbars and primes are used to indicate the zonal mean and deviations from the zonal mean. Defining the TEM eddy streamfunction by $\psi^\star = -\overline{w'b'}/\partial_y\overline{b}$, (1) can be rewritten as

$$J(\psi_{\text{res}}, \overline{b}) = \mathcal{D} - \partial_y\left[(1-\mu)\overline{v'b'}\right]. \tag{2}$$

Here we use the Jacobian notation $J(f, g) = \partial_y f\partial_z g - \partial_z f\partial_y g$ and have defined the quantities $\psi_{\text{res}} = \overline{\psi} + \psi^\star$ and $\mu = -(\partial_z b)\overline{w'b'}/(\partial_y\overline{b})\overline{v'b'}$, where $\overline{\psi}$ is the streamfunction corresponding to the Eulerian mean velocities $\overline{v}$ and $\overline{w}$. If the eddies act adiabatically so that their heat fluxes are aligned with isopycnals then $\mu = 1$ and in the absence of explicit diabatic terms we obtain

$$J(\psi_{\text{res}}, \overline{b}) = 0. \tag{3}$$

This result shows that the residual circulation, which is the circulation relevant to heat transport, is along isopycnals in an adiabatic ocean.

Many idealized models of the ACC are formulated in the geometry of a zonally-reentrant channel bounded by zonal walls. In such a model isopycnals will intersect the wall on the equatorward side, which constrains the residual circulation to vanish. This is because ψ_{res} vanishes on the boundary by the condition of no normal flow at the wall, and ψ_{res} is constant along isopycnals by (3). If $\mu = 0$ we can rewrite the eddy streamfunction as $\psi^\star = \overline{v'b'}/\partial_z\overline{b}$ and the condition of zero residual circulation gives

$$0 = \overline{\psi} + \psi^\star = \overline{\psi} + \frac{\overline{v'b'}}{\partial_z\overline{b}}. \tag{4}$$

The Eulerian mean streamfunction is given by the well-known Ekman overturning $\overline{\psi} = -\tau/(\rho_0 f)$ where τ is the surface wind stress, ρ_0 is a reference density and f is the Coriolis parameter. Assuming a Gent-McWilliams-like flux-gradient form for the eddy heat flux $\overline{v'b'} = -K\partial_y\overline{b}$ in (4) immediately yields the formula

$$s = -\frac{\partial_y\overline{b}}{\partial_z\overline{b}} = \frac{\tau}{\rho_0 f K} \tag{5}$$

which predicts the isopycnal slope s in terms of the external parameters τ, f, and ρ_0 as well as the eddy diffusion coefficient K. Although the real ACC is not bounded by zonal walls, it remains plausible that (5) holds at leading order, as observations of the SO indicate that the residual circulation is much weaker than either the eddy or Eulerian mean streamfunctions, implying that the two contributions to ψ_{res} cancel one another heavily [10].

Equation (5) summarizes the competing effects of winds and eddies in determining the isopycnal slopes. Increasing the wind stress τ steepens isopycnals according to the Ekman response, while strengthening the eddy fluxes by increasing K flattens isopycnals. If K is taken to be a constant, increasing τ produces a linear increase in the slope. On the other hand, if K is taken to be proportional to the wind stress, then s will be invariant

to changes in τ, i.e., the ACC will be completely eddy saturated. No accepted theoretical argument currently exists to predict K, and many different assumptions have been used to parameterize K and obtain a scaling for $s(\tau)$. Marshall and Radko [9] took K to be proportional to the isopycnal slope $K = \alpha s$ to obtain $s \sim \sqrt{\tau}$. Based on mixing-length arguments, Abernathey and Marshall [2] related K to the eddy kinetic energy (EKE) as $K \sim \sqrt{\text{EKE}}$. In their simulations EKE was found to depend linearly on τ, predicting $K \sim \sqrt{\tau}$. Meredith et al. [11] also suggested that K scales with EKE, but according to a more complex relationship having two limiting cases. In the limit of small EKE, they obtain $K \sim \text{EKE}^{3/2}$, while in the limit of large EKE they obtain $K \sim \sqrt{\text{EKE}}$. Assuming $\text{EKE} \sim \tau$, these results predict a family of scaling laws $s \sim \tau^\gamma$ with $-1/2 < \gamma < 1/2$. This range of scalings includes cases in which isopycnal slopes steepen with τ, are constant with τ, or decrease with τ.

Based on the above discussion it is clear that an improved understanding of eddy physics in the ACC is required to make progress on the problem of the SO response to changes in forcing. In this work, we use the two-layer quasigeostrophic (QG) equations to formulate a highly-idealized model of the ACC. Within the context of this model, we identify a reduced set of dynamics, the mean-field dynamics, which appears to reproduce the dependence of the mean state on the winds that we observe in the full dynamics. The mean-field dynamics is obtained by discarding eddy-eddy interactions from the full nonlinear equations of motion while retaining the interactions between the eddies and the mean flow. This approach has previously been successful in a variety of applications. O'Gorman and Schneider [16] applied mean-field theory to an atmospheric GCM and obtained a mean state and storm track statistics similar to those of the unapproximated GCM. Mean-field theory has also been successful in predicting the formation and structure of zonal jets emerging on the stochastically-driven barotropic β-plane [4, 17]. It is hoped that our application of the highly-simplified mean-field dynamics to the ACC will provide a more tractable theoretical framework within which progress toward understanding the eddy response to wind stress variations may be made in the future.

2 Model Formulation

2.1 Model Geometry and Equations of Motion

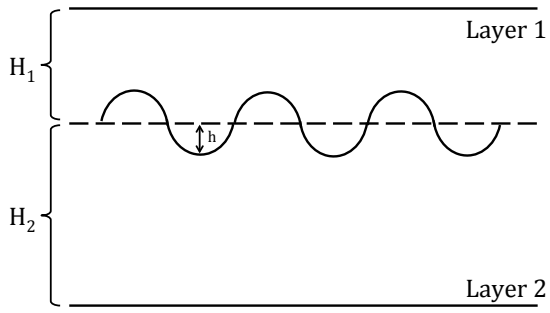


Figure 1: Schematic (x, z) cross-section of the two-layer model geometry.

We consider the motion of two fluid layers with thicknesses

$$h_1 = H_1 + h \tag{6}$$
$$h_2 = H_2 - h \tag{7}$$

where H_1 and H_2 are the (constant) mean thicknesses of layers 1 and 2 and $h(x, y, t)$ is the excess thickness of layer 1. We take x (increasing eastward) and y (increasing equatorward) to be our zonal and meridional coordinates on a zonally-reentrant β-plane channel of meridional width W and zonal extent L. A cross section of the model geometry is illustrated in Fig. 1. The local slope of the interface between the two layers is given by the gradient of h and is the analog of the isopycnal slope in the two-layer case. In the QG approximation, the dynamics of the two layers are formulated in terms of the upper and lower layer QG potential vorticities (PVs) q_1 and q_2, which are defined as

$$q_1 = \boldsymbol{\nabla}^2\psi_1 + \beta y - F_1(\psi_1 - \psi_2) \tag{8}$$
$$q_2 = \boldsymbol{\nabla}^2\psi_2 + \beta y + F_2(\psi_1 - \psi_2). \tag{9}$$

Here ψ_1 and ψ_2 denote the upper and lower layer geostrophic streamfunctions, which are related to the geostrophic velocity components via $u_{1,2} = -\partial\psi_{1,2}/\partial y$ and $v_{1,2} = \partial\psi_{1,2}/\partial x$. As we work on the β-plane, $\beta = \mathrm{d}f/\mathrm{d}y$ is a constant parameter. The parameters F_1 and F_2 are defined in terms of the mean layer depths as

$$F_1 = \frac{f_0^2}{g'H_1} \qquad\qquad F_2 = \frac{f_0^2}{g'H_2} \tag{10}$$

in which f_0 is the constant background Coriolis parameter and g' is the reduced gravity $g' = g(\rho_2 - \rho_1)/\rho_2$ where ρ_i is the constant density of the fluid in layer i. We take $\rho_2 > \rho_1$ so that the fluid is stably stratified. The difference $\psi_1 - \psi_2$ appearing in (8,9) is related to the interface deflection h through the relation

$$f_0(\psi_1 - \psi_2) = g'h \tag{11}$$

so that the zonal mean interfacial slope is given by

$$s \equiv -\frac{\partial\overline{h}}{\partial y} = \frac{f_0}{g'}(U_1 - U_2) \tag{12}$$

in which the overbar indicates a zonal average and we denote the zonal mean flow by $U_{1,2} = \overline{u}_{1,2}$. Note the negative sign in the definition of s: as h is the excess thickness in layer 1, $s < 0$ corresponds to h increasing with y. Equation (12) relates the baroclinic component of the mean flow to the interfacial slope, and indicates that for positive shear $U_1 > U_2$ the interface deepens as we move equatorward ($f_0 < 0$).

The PV fields evolve according to the dynamical equations

$$\partial_t q_1 + J(\psi_1, q_1) = -r(q_1 - \beta y) - \frac{1}{\rho_1 H_1}\partial_y\tau + \kappa\boldsymbol{\nabla}^2 q_1 \tag{13}$$
$$\partial_t q_2 + J(\psi_2, q_2) = -r(q_2 - \beta y) - dF_2\boldsymbol{\nabla}^2\psi_2 + \kappa\boldsymbol{\nabla}^2 q_2 \tag{14}$$

in which $J(f,g) = (\partial_x f)(\partial_y g) - (\partial_y f)(\partial_x g)$. Conservation of PV in (13) and (14) is modified by forcing, in the form of wind stress acting on layer 1, and by dissipation, in the form of Rayleigh drag, diffusion, and bottom friction. We take the wind stress to be of the idealized form

$$\frac{1}{\rho_1 H_1}\tau(y) = A\sin(l_g y) \tag{15}$$

where $l_g = \pi/W$ and A controls the strength of the forcing. The stress is thus zonally-symmetric with a half-sinusoidal profile across the channel, and acts to drive eastward flow in the channel. Rayleigh drag damps the dynamical part of the PV $q_i - \beta y$ in both layers with coefficient r. This Rayleigh drag on PV derives from Rayleigh drag on the upper and lower layer velocities as well as on the interface deflection h, all with equal coefficients. Bottom friction acts only on the relative vorticity in layer 2. We take the drag coefficient to be dF_2 so that d controls the strength of the bottom drag when F_2 is held constant. Diffusion of PV with coefficient κ is included for numerical stability of the mean-field dynamics, which we introduce in Section 2.5. We set $\kappa = 0$ when working with the full nonlinear equations of motion, and so we ignore diffusion when discussing the full dynamics.

The model boundary conditions are periodic in x and appropriate boundary conditions are applied in the meridional direction to ensure mass conservation. Mass conservation requires that the meridional velocity vanishes at the boundaries $y = 0, W$, including both the geostrophic (v) and ageostrophic (v_a) parts of the meridional velocity. Since $\overline{v}_i = \overline{\partial\psi_i/\partial x} = 0$, the $v_i = 0$ geostrophic condition implies that $v_i' = 0$, where the prime denotes the deviation from the zonal mean. The ageostrophic condition $v_{ai} = 0$ implies the boundary condition $U_1 = U_2 = 0$ on the zonal mean zonal flow, which we discuss in Section 2.2. Free-slip boundary conditions are applied to u_i'.

2.2 Zonal Mean Dynamics and Residual Circulation

Although the PV-evolution equations (13,14) are sufficient to determine the dynamics, it is illuminating to examine the zonal mean dynamics directly in terms of the velocities and interface deflection. The zonal mean equations of motion for the interface and zonal flow are

$$\partial_t U_1 = \overline{v_1' q_1'} + f_0 \overline{v}_1^\dagger - rU_1 + \frac{1}{\rho_1 H_1}\tau \tag{16}$$

$$\partial_t U_2 = \overline{v_2' q_2'} + f_0 \overline{v}_2^\dagger - rU_2 - dF_2 U_2 \tag{17}$$

$$\partial_t \overline{h} = -\overline{w}^\dagger - r\overline{h} \tag{18}$$

in which we have defined the residual meridional velocities

$$\overline{v}_1^\dagger = \overline{v}_{1a} + \frac{1}{H_1}\overline{v_1' h'} \tag{19}$$

$$\overline{v}_2^\dagger = \overline{v}_{2a} - \frac{1}{H_2}\overline{v_2' h'} \tag{20}$$

$$\overline{w}^\dagger = \overline{w}_1 + \partial_y \overline{v_1' h'} = \overline{w}_2 + \partial_y \overline{v_2' h'}. \tag{21}$$

The second equality in (21) follows from continuity at the interface of the component of velocity normal to the interface, which ensures that h is a material surface viewed from

either fluid layer. Note that (11) implies $\overline{v_1' h'} = \overline{v_2' h'}$ and that mass continuity implies $\overline{v}_1^\dagger = -(H_2/H_1)\overline{v}_2^\dagger$. The vertical and meridional residual velocities are related through $H_1 \partial_y \overline{v}_1^\dagger = -H_2 \partial_y \overline{v}_2^\dagger = \overline{w}^\dagger$. Note that (18) implies that in steady state the residual circulation $\overline{w}^\dagger$ balances the diabatic effect of Rayleigh drag on the interface. The residual circulation in our simplified model is thus strongly constrained by the imposed drag parameters, and vanishes by necessity in the limit $r \to 0$.

The boundary condition of no normal flow at $y = 0, W$ places a constraint on the zonal mean flow at the walls that can be understood from (16). The PV flux $\overline{v_1' q_1'} = 0$ and thickness flux $\overline{v_1' h'}$ vanish at the walls because $v' = 0$ there, and $\tau = 0$ at the walls by our choice of forcing structure. The budget then becomes

$$\partial_t U_1 = f_0 \overline{v}_{1a} - r U_1. \tag{22}$$

We require that the ageostrophic meridional flow $\overline{v}_{1a} = 0$ at the walls to prevent a net flux of mass into or out of the layer. This implies that U_1 decays exponentially at the walls, and that if $U_1 = 0$ at $t = 0$, then $U_1 = 0$ for all time. Similar considerations apply to U_2. As described in Section 2.1, we take $U_1 = U_2 = 0$ at the channel walls as our boundary conditions on the mean flow.

The upper layer mean zonal momentum equation (16) shows that in the absence of eddies, Rayleigh drag, or time-dependence, the wind stress drives an equatorward residual flow $\overline{v}_1^\dagger > 0$ whose Coriolis force balances the stress. This is the representation of the Ekman response in our simplified two-layer model. The residual return flow $\overline{v}_2^\dagger$ in the lower layer produces an eastward Coriolis force that must be balanced by bottom friction on the zonal flow U_2, thereby requiring eastward flow $U_2 > 0$ in the lower layer. In the presence of eddies, the eddy PV fluxes $\overline{v_1' q_1'}$ and $\overline{v_2' q_2'}$ act to transfer momentum from the upper layer to the lower layer. It is straightforward to show that $H_1 \langle v_1' q_1' \rangle = -H_2 \langle v_2' q_2' \rangle$, where angle brackets denote a mean over the entire domain. This shows that the rate at which eddies remove momentum from the upper layer is equal to the rate at which they deposit momentum into the lower layer. Eddies produced by a baroclinically unstable shear then act to barotropize the flow, reducing the shear and flattening the interface.

We now use the zonal mean equations (16,17,18) to cast the eddy saturation argument discussed in the introduction into the language of our model. We make the assumption that diabatic effects are weak and that the residual circulation is unimportant. The steady state zonal mean momentum budgets then become

$$0 = \overline{v_1' q_1'} + \frac{1}{\rho_1 H_1} \tau \tag{23}$$

$$0 = \overline{v_2' q_2'} - d F_2 U_2. \tag{24}$$

Further assuming that the interface term in the potential vorticity is dominant over the planetary and relative vorticity terms, we write the PV fluxes in terms of the thickness flux as

$$\overline{v_1' q_1'} \approx -\frac{f_0}{H_1} \overline{v_1' h'} \tag{25}$$

$$\overline{v_2' q_2'} \approx +\frac{f_0}{H_2} \overline{v_2' h'} = \frac{f_0}{H_2} \overline{v_1' h'}. \tag{26}$$

Finally, we assume a flux-gradient relationship for the thickness flux

$$\overline{v_1' h'} = -K \frac{\partial \overline{h}}{\partial y}. \tag{27}$$

Combining expressions (23), (25), and (27) immediately yields an expression for the zonal mean interfacial slope s in terms of the model parameters and the eddy diffusivity K:

$$s = -\frac{\partial \overline{h}}{\partial y} = \frac{\tau}{f_0 \rho_1 K}. \tag{28}$$

Equation (28) implies that if the eddy diffusivity K is constant, the slope of the interface scales linearly with the imposed wind stress. On the other hand, if K is linear in the wind stress $K = \alpha\tau$, then the slope of the interface does not depend at all on the winds. This highlights the crucial role of eddies in determining the response of the ocean to changing forcing conditions.

The interfacial slope s is related to the zonal mean shear $U_1 - U_2$ by (12). By applying the flux-gradient relation to the zonal mean momentum budget of the lower layer, we can also obtain explicit formulas for U_1 and U_2 individually. Combining equations (26), (27), and (28) yields

$$U_2 = \frac{\delta}{dF_2} \frac{\tau}{\rho_1 H_1} \tag{29}$$

where $\delta = H_1/H_2$ is the ratio of layer depths. Writing $U_1 = U_2 + (U_1 - U_2)$ and applying (28) then yields

$$U_1 = \left(\frac{\delta}{dF_2} + \frac{1}{F_1 K} \right) \frac{\tau}{\rho_1 H_1}. \tag{30}$$

Equation (29) shows that, under these assumptions, the lower layer flow is independent of the eddy diffusivity K and is determined solely by the winds and bottom friction. The upper layer flow U_1 weakens with increasing K as larger eddy diffusivities correspond to a shallower interfacial slope and reduced vertical shear.

2.3 Model Fixed Point

As we are interested in the behavior of the model equations (13) and (14) as the wind stress forcing is varied, it is useful to determine whether stable fixed point solutions exist over some range of forcing strengths. The model indeed has a single, zonally-symmetric fixed point that is stable for sufficiently weak wind stress. The fixed point becomes baroclinically unstable as the wind stress is increased beyond a threshold value. For sufficiently strong winds, the flow in the channel is turbulent.

To obtain the fixed point in terms of the model parameters, we first write

$$\overline{q}_1 - \beta y = \hat{Q}_1 \cos(l_g y) \qquad \overline{q}_2 - \beta y = \hat{Q}_2 \cos(l_g y) \tag{31}$$

$$U_1 = \hat{U}_1 \sin(l_g y) \qquad U_2 = \hat{U}_2 \sin(l_g y) \tag{32}$$

so that the steady state, zonally-symmetric PV budgets become (from (13) and (14))

$$0 = -r\hat{Q}_1 \cos(l_g y) - \partial_y \left(A \sin(l_g y) \right) \tag{33}$$

$$0 = -r\hat{Q}_2 \cos(l_g y) - dF_2 \partial_y \left(-\hat{U}_2 \sin(l_g y) \right). \tag{34}$$

Equations (33) and (34) can be solved for the steady state values of U_1 and U_2 by writing the PV in terms of the zonal mean flow. We obtain

$$U_1\Big|_{FP} = A\frac{F_2 + l_g^2\left(1+\frac{F_2 d}{r}\right)}{F_2 d(l_g^2+F_1)+r(l_g^2+F_1+F_2)}\sin(l_g y) \tag{35}$$

$$U_2\Big|_{FP} = A\frac{F_2}{F_2 d(l_g^2+F_1)+r(l_g^2+F_1+F_2)}\sin(l_g y). \tag{36}$$

The corresponding upper and lower layer PV gradients are

$$\frac{\partial \overline{q}_1}{\partial y}\Big|_{FP} = \beta + \frac{l_g^2 A}{r}\sin(l_g y) \tag{37}$$

$$\frac{\partial \overline{q}_2}{\partial y}\Big|_{FP} = \beta - \frac{l_g^2 A}{r}\frac{dF_2^2}{dF_2(l_g^2+F_1)+r(l_g^2+F_1+F_2)}\sin(l_g y). \tag{38}$$

Several important theoretical properties of our model are revealed by expressions (35)-(38). From (35) and (36) it is clear that the fixed point values of U_1, U_2, and $U_1 - U_2$ all increase linearly with the wind stress A. Equations (35) and (36) also show that the fixed point structure depends crucially on the Rayleigh drag r. As $r \to 0$, the fixed point value of U_1 tends to infinity while U_2 asymptotes to a constant finite value. This can be understood from the zonal mean momentum and thickness budgets (16) and (18). In the absence of Rayleigh drag, (18) constrains the residual circulation to vanish. In the absence of a residual circulation, eddies, or Rayleigh drag in the momentum equation, (16) shows that the momentum injection by wind stress must be balanced by acceleration of the upper level flow. This argument illustrates that no fixed point exists for $r = 0$: the solution is always time-dependent, with the upper layer flow steadily increasing until the shear becomes baroclinically unstable, even for very weak wind forcing. The fixed point PV gradients (37) and (38) reveal the important role played by bottom drag in our model. The upper layer PV gradient is always positive, while the PV gradient in layer 2 is positive for small A and negative for sufficiently large A. However, if the coefficient of bottom friction $d \to 0$, the PV gradient in layer 2 can never become negative. As a change in sign of the PV gradient must occur somewhere in the domain as a necessary condition for baroclinic instability, this demonstrates that bottom drag is required for baroclinic instability to occur in this model. We discuss the stability of the fixed point in more detail in the next section.

2.4 Model Parameters and Instability Properties

The model dynamics depends on a number of parameters, whose values we choose to be approximately representative of the SO. As our primary interest is in the response of the model dynamics to changing wind stress, we will normally vary the strength of the wind stress A while keeping other parameters fixed. We also conduct a number of sensitivity tests by varying the damping parameters r and d. We hold the following parameters constant in

all model integrations:

$$W = 1000 \text{ km} \quad (39)$$
$$L = 2000 \text{ km} \quad (40)$$
$$F_1 = 5.5 \times 10^{-4} \text{ km}^{-2} \quad (41)$$
$$F_2 = 1.1 \times 10^{-4} \text{ km}^{-2} \quad (42)$$
$$\beta = 1.2 \times 10^{-3} \text{ (km day)}^{-1}. \quad (43)$$

The zonal extent of the channel is thus twice its meridional width. The ratio $F_1/F_2 = H_2/H_1 = 5$, so that the lower layer has a mean thickness of five times that of the upper layer. The model dynamics does not depend on the values of H_1 and H_2 outside their appearance in the F_1 and F_2 parameters. Similarly, the dynamics does not depend explicitly on g' or f_0. The Rossby deformation radius R_d is given by $R_d = (F_1 + F_2)^{-1/2} = 38.8$ km, so that our channel extends over 50 Rossby radii in the zonal direction and approximately 25 in the meridional direction. The value of β was chosen to represent a channel centered at 60°S.

It remains to specify the values of the drag parameters r, d, and κ, as well as the strength of the wind forcing A. We take a constant value of $\kappa = 20$ km^2/day in the mean-field dynamics and $\kappa = 0$ in the fully-nonlinear dynamics. The value of κ was chosen to be as small as possible while still allowing the numerical model to remain stable. The Rayleigh drag r and bottom drag d are varied in sensitivity tests, but our reference values are taken to be

$$r = 1 \times 10^{-4} \text{ day}^{-1} \quad (44)$$
$$d = 500 \text{ km}^2/\text{day}. \quad (45)$$

For the reference bottom friction value, the drag on the lower layer velocity has coefficient $dF_2 = 0.05$ day^{-1}, so that the bottom drag acts approximately 500 times more rapidly than the bulk Rayleigh drag. In the mean-field dynamics, the PV diffusion acts on the largest scales with coefficient $\kappa l_g^2 \sim 2 \times 10^{-4}$ day^{-1}. Diffusion in the mean-field model is thus twice as strong as the Rayleigh drag at the largest scales, and even stronger at smaller scales. Diffusion plays an important role in the mean field equations (due to numerical constraints) while playing no role in the nonlinear equations. This difference in dissipation constitutes a limitation of our results, and resolving this issue is an important future direction of this work. However, preliminary integrations of the nonlinear model in which comparable diffusion was added to the dynamics produced results similar to those without diffusion for wind stress values well beyond the stability boundary of the fixed point. This suggests that comparisons between the mean-field and nonlinear dynamics remain informative in spite of the differences in the dissipation properties of the two models.

We vary the wind stress over the range $A = 0.1 - 7.0$ km/day^2. We take as our reference value $A = 1$ km/day^2. In our notation the wind stress A appears directly as an acceleration of the zonal wind, and so has units of acceleration. To compare with realistic values of the wind stress over the SO it is conventional to write the acceleration as $\tau_0/(\rho_1 H_1)$ where τ_0 has units of stress (N/m^2) and ρ_1 is the density of the upper layer. Choosing as a typical density $\rho_1 \sim 1035$ kg/m^3 and an upper layer depth $H_1 \sim 1000$ m, our reference acceleration corresponds to a surface stress $\tau_0 \sim 0.14$ N/m^2. This is near the conventional

value of $\tau_0 = 0.2$ N/m^2, so that $A = 1$ km/day^2 is reasonably representative of the SO wind stress. By varying A over a wide range in the vicinity of this reference value, we assess how the state of the two-layer model ocean responds to changes in wind stress.

The growth rate of baroclinic instability is shown in Fig. 2 as a function of zonal wavenumber for $A = 1.0$ km/day^2 and reference values of the drag parameters. Panel (a) shows the growth rate in the nonlinear model ($\kappa = 0$) and panel (b) shows the growth rate in the mean-field dynamics ($\kappa = 20$ km^2/day). For this forcing value the fixed point is unstable with or without diffusion, and the e-folding time of the fastest growing wave is on the order of a few days. The growth rate peaks near zonal wavenumbers 4 and 5. The Rossby radius $Rd = 38.8$ km corresponds to a wavenumber $k_d \sim 1/R_d = .025$ km^{-1} which is near zonal wavenumber 8. Baroclinic instability thus sets in at scales somewhat larger than the Rossby radius. In the absence of diffusion the stable high wavenumbers decay on the timescale of the Rayleigh drag. With diffusion, all growth rates are reduced, and higher wavenumbers decay more rapidly. For $\kappa = 0$, instability sets in at a minimum wind strength of $A \sim .015$ km/day^2 for the reference drag values. For $\kappa = 20$ km^2/day, baroclinic instability does not set in until $A = 0.25$ km/day^2. As we vary A over the range $A = 0.1 - 7.0$ km/day^2, all model integrations are in the unstable regime for the nonlinear model diffusion. The mean-field model is stable for the first few wind stress values and unstable thereafter.

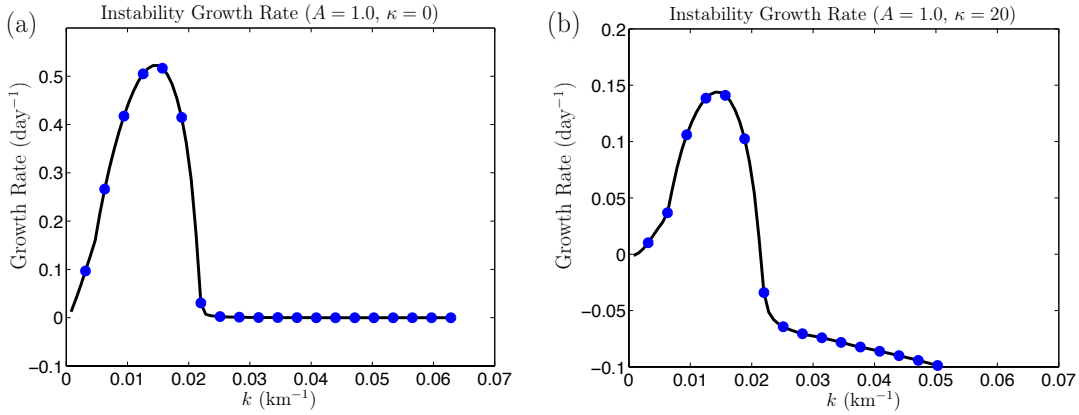


Figure 2: Growth rate of baroclinic instability as a function of wavenumber (a) without diffusion and (b) with diffusivity $\kappa = 20$ km^2/day. Blue circles indicate the quantized wavenumbers permitted by the periodic geometry. Reference values are used for all parameters.

2.5 Mean-Field Dynamics

Mean-field dynamics is a simplified theoretical model of the system (13,14) in which the nonlinearity associated with the advection of eddy quantities by the eddy velocity field is discarded but the nonlinearity associated with wave-mean flow interaction is retained. To derive the mean-field equations, we first rewrite the nonlinear equations of motion in the

Reynolds decomposition:

$$q_1 = \overline{q}_1 + q_1' = Q_1(y,t) + q_1'(x,y,t) \tag{46}$$
$$q_2 = \overline{q}_2 + q_2' = Q_2(y,t) + q_2'(x,y,t). \tag{47}$$

The mean PV fields Q_1 and Q_2 evolve according to the dynamical equations

$$\partial_t Q_1 = -\partial_y \overline{v_1' q_1'} - r\,(Q_1 - \beta y) - \frac{1}{\rho_1 H_1}\partial_y \tau + \kappa \partial_{yy} Q_1 \tag{48}$$
$$\partial_t Q_2 = -\partial_y \overline{v_2' q_2'} - r\,(Q_2 - \beta y) - dF_2 \partial_{yy}\overline{\psi}_2 + \kappa \partial_{yy} Q_2. \tag{49}$$

These equations are equivalent to the set of equations (16,17,18) except that we include the diffusion term for numerical stability of the approximated equations derived in what follows. The eddy PV fields q_1' and q_2' evolve according to

$$\partial_t q_1' = -U_1 \partial_x q_1' - v_1' \partial_y Q_1 - r q_1' + \kappa \boldsymbol{\nabla}^2 q_1' + \mathrm{EENL}_1 \tag{50}$$
$$\partial_t q_2' = -U_2 \partial_x q_2' - v_2' \partial_y Q_2 - r q_2' + \kappa \boldsymbol{\nabla}^2 q_2' - dF_2 \boldsymbol{\nabla}^2 \psi_2' + \mathrm{EENL}_2. \tag{51}$$

Here we use the notation EENL to denote the eddy-eddy nonlinearities, given by

$$\mathrm{EENL}_1 = -\partial_y \left(v_1' q_1' - \overline{v_1' q_1'} \right) - \partial_x \left(u_1' q_1' - \overline{u_1' q_1'} \right) \tag{52}$$
$$\mathrm{EENL}_2 = -\partial_y \left(v_2' q_2' - \overline{v_2' q_2'} \right) - \partial_x \left(u_2' q_2' - \overline{u_2' q_2'} \right). \tag{53}$$

Equations (48)-(51) are exact. To form the mean field equations, we simply drop the EENL terms, thereby discarding the effects of eddies advecting eddies. The discarded EENL terms are those responsible for the scattering of energy between different zonal wavenumber components of the flow. In the mean-field approximation, the dynamics is that of wave-mean flow interaction: the structure of the eddy field is shaped by its interaction with the mean flow, and the eddy fluxes in turn modify the mean flow to complete the two-way wave-mean coupling.

The mean-field dynamics has considerable conceptual and practical advantages over the full nonlinear equations of motion. As result of our approximation, the eddy equations of motion (50,51) become linear in eddy quantities. We write the eddy PV fields as Fourier series in zonal wavenumber components $k_n = 2\pi n/L$

$$q_1(x,y,t) = \mathrm{Re}\left[\sum_{n=1}^{\infty} \tilde{q}_{1,n}(y,t) e^{ik_n x} \right] \qquad q_2(x,y,t) = \mathrm{Re}\left[\sum_{n=1}^{\infty} \tilde{q}_{2,n}(y,t) e^{ik_n x} \right]. \tag{54}$$

Equations (50,51) then imply that each wavenumber component n evolves independently of the others, interacting only through their mutual interaction with the mean flow. This is in contrast to the full nonlinear equations in which Fourier components interact directly. Since wavenumbers do not interact directly in the mean-field dynamics, it is mathematically consistent to retain a single zonal wavenumber k in the expansions (54) and write

$$q_1(x,y,t) = \mathrm{Re}\left[\tilde{q}_1(y,t) e^{ikx} \right] \qquad q_2(x,y,t) = \mathrm{Re}\left[\tilde{q}_2(y,t) e^{ikx} \right]. \tag{55}$$

The resulting equations of motion for the single wave and the mean flow are

$$\partial_t \tilde{q}_1 = \left[-ikU_1 - r + \kappa(\partial_{yy} - k^2)\right]\tilde{q}_1 + \left[-ik\partial_y Q_1\right]\tilde{\psi}_1 \quad (56)$$

$$\partial_t \tilde{q}_2 = \left[-ikU_2 - r + \kappa(\partial_{yy} - k^2)\right]\tilde{q}_2 + \left[-ik\partial_y Q_2 - dF_2(\partial_{yy} - k^2)\right]\tilde{\psi}_2 \quad (57)$$

$$\partial_t Q_1 = (k/2)\partial_y \mathrm{Im}\left(\tilde{\psi}_1 \tilde{q}_1^\star\right) - rQ_1 - \frac{1}{\rho_1 H_1}\partial_y \tau + \kappa \partial_{yy} Q_1 \quad (58)$$

$$\partial_t Q_2 = (k/2)\partial_y \mathrm{Im}\left(\tilde{\psi}_2 \tilde{q}_2^\star\right) - rQ_2 + dF_2 \partial_y U_2 + \kappa \partial_{yy} Q_2 \quad (59)$$

in which we have used the relation $\overline{v_i' q_i'} = -(k/2)\mathrm{Im}(\tilde{\psi}_i \tilde{q}_i^\star)$ to write the eddy PV fluxes explicitly in terms of the Fourier coefficient. Stars indicate complex conjugation. In equations (56)-(59), the retained wavenumber k is a free parameter. We choose k based on the dominant wavenumber observed in nonlinear model integrations. The mean-field dynamics is also often referred to as the quasilinear approximation. Based on this nomenclature, we will refer to the closed dynamical system (56)-(59) as the quasilinear model, abbreviated "QL". To mirror our naming convention we refer to the fully nonlinear dynamics (13)-(14) by the abbreviation "NL". We emphasize that the difference between NL and QL is in the EENL terms only. As these terms play no role in the zonally-symmetric fixed point solution, the QL dynamics has the same fixed point and stability properties as NL, with the exception of the influence of PV diffusion which is included in QL but not in NL.

In the full nonlinear dynamics (13,14), a flow initialized with energy concentrated in single zonal wavenumber component would nonlinearly produce a full spectrum of waves through the EENL terms. In the mean-field dynamics this scattering of energy to other waves does not occur and the interaction of a single wave with the mean flow can be consistently investigated. The choice to retain a single wavenumber constitutes an extreme simplification of the dynamics, reducing the dynamical variables of the model from two-dimensional fields fluctuating in time $\psi'(x, y, t)$ to one-dimensional Fourier structures varying in time $\tilde{\psi}_k(y, t)$. In addition to practical simplification, the mean field equations also provide conceptual clarity by isolating the effects of wave-mean flow sector of the model physics from the EENL effects. This separation allows us to evaluate the importance of EENL processes in determining the model's response to forcing. In the remainder of this paper we will compare solutions of the nonlinear and mean-field equations to evaluate the extent to which the mean-field theory can reproduce the response of the nonlinear model to changes in wind forcing. The success or failure of the mean-field dynamics in mirroring the behavior of the full model has important implications for our understanding of how eddies influence the mean climate state.

2.6 Numerical Implementation

The NL equations (13)-(14) are integrated numerically on a staggered finite-difference grid with $nx = 128$ points in the zonal direction and $ny = 64$ points in the meridional direction. This corresponds to a model grid spacing of $dx = dy \approx 15$ km, so that we have several grid points within each Rossby radius $R_d \approx 40$ km. PV inversion is performed spectrally. Advection terms are calculated in flux form and a van Leer flux limiter is used for numerical stability. We use a second-order Runge-Kutta method for the marching scheme with a time step of $dt = 1/64$ days. The model is initialized near the fixed point solution for low values of

the wind stress. For large values of wind stress $U_{1,FP}$ is so large that initialization near the fixed point constrains the timestep due to the CFL condition, so we instead initialize with zonal mean fields equal to those of the fixed point but rescaled by a coefficient with value less than one. Small scale noise is added to break the symmetry and allow the instability to develop. Model integrations are typically carried out for 5000 days, which was found to be sufficient to obtain reasonable flow statistics. For some parameter values very long runs (50000 days) were carried out, and the statistics were not found to change very much compared to the shorter integrations. The sensitivity of the model results to the spatial resolution was also tested. Model integrations at doubled resolution (Fig. 12, Appendix) showed minor quantitative differences when compared with the default resolution, but the similarity of the two solutions demonstrates that the model dynamics are sufficiently well-resolved at 64×128 resolution.

The QL equations (56)-(59) are solved numerically using finite-differences on an $ny = 64$ point meridional grid. PV inversions were performed using finite differences. The same marching scheme, time step, and initialization procedure was used as in NL. As the QL model is inexpensive to numerically integrate (even compared with our highly simplified two-layer QG model) we integrate the QL model for 100000 days to obtain reliable statistics. In many cases the QL model exhibits long transient oscillations after the initial instability, but these were eliminated from the statistics by averaging only over the second half of the model integrations.

3 Results

3.1 Reference Case

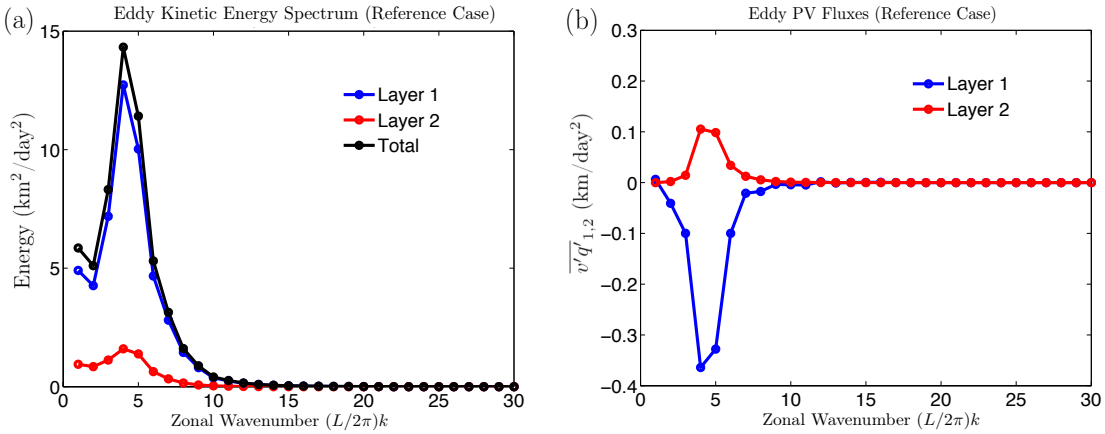


Figure 3: (a) EKE as a function of zonal wavenumber in the NL model (b) Eddy PV flux as a function of wavenumber in the NL model. Both EKE and PV flux show large peaks at zonal wavenumbers 4 and 5, indicating that these waves are dominant in the statistical mean state

We now compare the results of the NL and QL model integrations for the reference case $A = 1$ km/day^2, $r = 1 \times 10^{-4}$ day^{-1}, $d = 500$ km^2/day. Figure 3 (a) shows the time and

domain average EKE as a function of zonal wavenumber in the equilibrated NL state, where EKE is calculated as

$$EKE = \frac{H_1\langle u_1'^2 + v_1'^2\rangle + H_2\langle u_2'^2 + v_2'^2\rangle}{H_1 + H_2} = \frac{F_2\langle u_1'^2 + v_1'^2\rangle + F_1\langle u_2'^2 + v_2'^2\rangle}{F_1 + F_2}. \tag{60}$$

Angle brackets indicate a time and domain average. The contribution from each wavenumber was determined by expanding the velocity field in a Fourier series as in (54) and calculating the energy of each wavenumber component separately. The energy spectrum shows a clear peak at zonal wavenumber 4, with some energy at the neighboring wavenumbers and very little energy at small scales. Figure 3 (b) shows the time and domain mean eddy PV flux $\langle v'q'\rangle$ for the upper and lower layers as a function of wavenumber. In the upper layer the eddies flux PV poleward, decelerating the upper layer flow by (16). The eddy flux in the lower layer is in the opposite direction, with the eddies acting to drive the lower layer flow against bottom friction. The eddy flux spectrum indicates that zonal wavenumber 4 dominates the flux of PV, in addition to containing the most kinetic energy. Based on these observations, we take $k = 4$ as our QL wavenumber.

Figure 4 (a) and (c) show snapshots of the upper layer PV fields in QL and NL in statistical equilibrium. The zonal wavenumber 4 and 5 structures appearing in the energy spectrum are also visibly evident in the NL PV field. The NL and QL PV fields are qualitatively similar, although EENL effects in the NL model produce PV filaments that are not present in the QL model. The QL PV field is by design characterized by an exact wave-4 structure, since only $k = 4$ is included in our calculation.

Panels (b) and (d) in Fig. 4 show the zonal mean flow in the upper and lower layers for the QL and NL models. It is informative to compare these equilibrated turbulent structures to the fixed point solutions (35,36). For the reference parameter values, the upper and lower layer unstable fixed point jets have strengths $U_{1,FP} \sim 180$ km/day and $U_{2,FP} \sim 3$ km/day so that the fixed point baroclinic shear is $\mathcal{O}(175$ km/day). The shear maintained by the equilibrated turbulence is $\mathcal{O}(15$ km/day) at the channel center, indicating that the action of baroclinic eddies produces a dramatic barotropization of the flow. The magnitude of the upper layer flow is approximately 20 km/day ≈ 23 cm/s, which is realistic for the ACC.

The strengths of the zonal jets U_1 and U_2 are similar in the QL and NL models. Inspection of the upper layer flow in NL shows that the eddies have modified the meridional profile of the zonal jet in addition to severely reducing its strength relative to the fixed point solution. The initially sinusoidal profile has been intensified at the channel center and weakened in the flanks of the center jet, with the appearance of an additional slight inflections of the jet profile near $y = 200$ km and $y = 800$ km. The QL upper layer jet has a meridional structure qualitatively different from that of the NL jet. The QL jet is intensified at the channel center and weakened on the flanks in agreement with the NL profile. However, additional flank jets are evident in the QL simulation near the locations of the minor inflections seen in the NL profile. These strong flank jets do not seem to form in the NL system. As the upper-level QL jet is stronger than the NL jet in some regions of the channel and weaker in other regions, the QL dynamics may track the behavior of NL more closely in the domain-average picture in which these differences in structure may be partially averaged out. The lower layer jets of NL and QL agree much more closely than the upper layer jets, although the QL dynamics still produces additional meridional structure on the flanks of the sharpened center jet. Additional work is in progress to identify

the physical explanation for these differences in structure, which we suspect are due to the development within the jet of critical layers for the equilibrated wave which are smoothed out by the turbulence of the NL model.

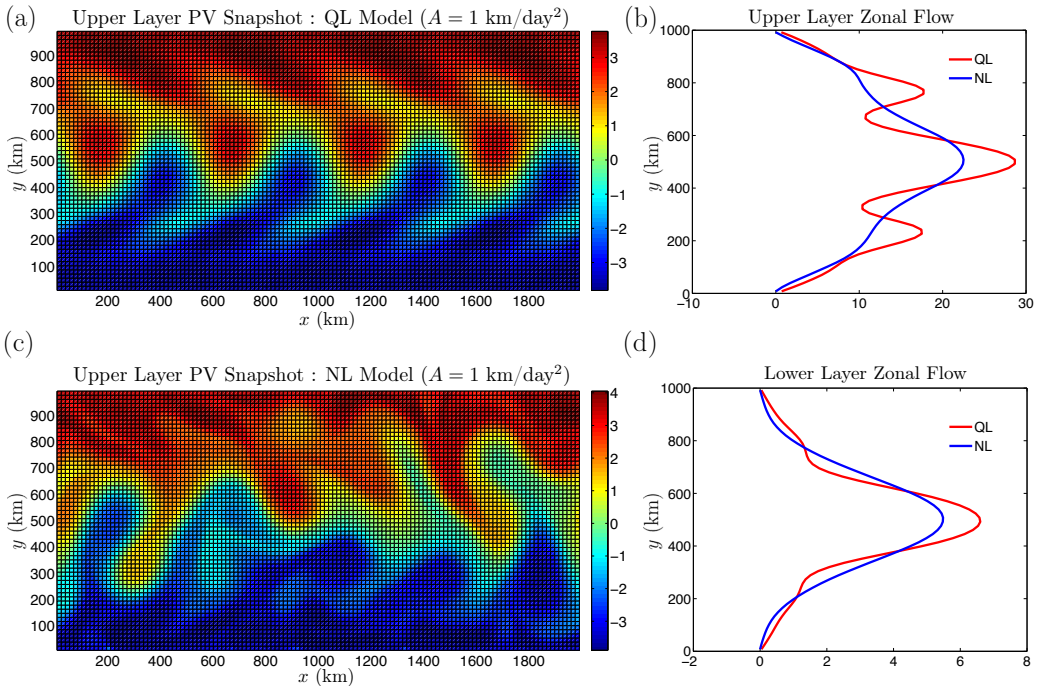


Figure 4: (a) Snapshot of upper layer PV in the QL model. (b) Upper level zonal mean flow in the QL (red) and NL (blue) models. (c) Snapshot of upper layer PV in the NL model. (d) Lower level zonal mean flow in the QL (red) and NL (blue) models. PV snapshots show qualitatively similar structures in both models, with NL showing a mixture of waves 4 and 5 as well as some small scale structure absent from the QL model. Zonal mean flows in QL and NL are close in amplitude but differ in meridional structure, with the QL dynamics producing a multiple jet structure that is not found in NL.

3.2 Model Response to Increasing Wind Stress

Motivated by the qualitative agreement between the QL and NL models in the reference configuration, we next vary the strength of the wind stress and compare the responses of NL and QL over a wide range of parameter space. Figure 5 shows the domain and time mean shear $U_1 - U_2$ as a function of the wind stress A for the NL and QL models. The abrupt change in the behavior of the QL curve at very weak wind stress is due to the onset of baroclinic instability, which sets in near $A = 0.25$ km/day in QL. The NL model is unstable for all plotted A values due to the absence of diffusion in NL system. Away from the instability boundary the shapes of the NL and QL curves are quite similar, with the behavior appearing to be loosely separated into two regimes. For relatively weak wind stress $A \lesssim 1$ km/day, the shear increases rapidly as A is increased. We emphasize that even in this relatively weak forcing regime the shear of the equilibrated state is much weaker than

that of the fixed point solution, and that the rate of increase of the shear with A is also much shallower than that which would result from the fixed point solution. The sharp increase of shear with A for the fixed point solution can be seen in the first few points of the QL curve before the flow becomes unstable. As A is increased beyond $A \approx 1$ km/day, the shear becomes less sensitive to changes in wind stress, indicating a degree of eddy saturation that occurs in both the QL and NL models. From Fig. 5 it does not appear that the shear ever becomes independent of A, which would correspond to complete eddy saturation, but we have not carried out model runs beyond $A = 7$ km/day.

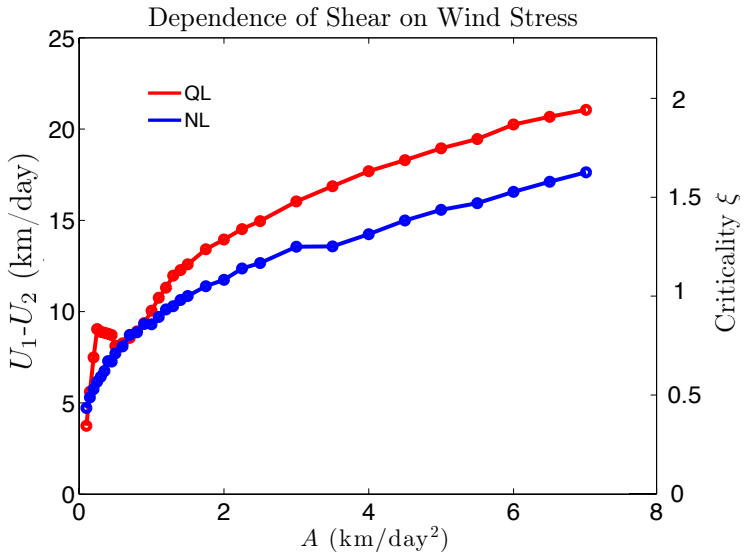


Figure 5: Domain and time average shear in statistical equilibrium in the QL (red) and NL (blue) models as functions of wind stress. Behavior in both models is characterized by increased sensitivity to wind stress at low values of A, transitioning to partial eddy saturation as the winds become stronger.

We also show the dependence of the criticality parameter ξ on wind stress in Fig. 5 (right axis). The criticality is defined as

$$\xi = \frac{f_0 s}{H_2 \beta} = \frac{F_2}{\beta}(U_1 - U_2) \tag{61}$$

and serves as a nondimensionalization of the shear. A commonly-invoked theoretical argument [18] suggests that baroclinic turbulence will adjust the mean state such that the lower layer PV gradient vanishes. This is motivated by the necessary condition for baroclinic instability that the PV gradient change sign somewhere in the domain: as the upper layer PV gradient is positive, zero PV gradient in the lower layer is the marginal state to which the eddies relax the mean state. Ignoring relative vorticity, the lower layer PV gradient is

$$\partial_y Q_2 = \beta - F_2(U_1 - U_2) \tag{62}$$

so that a vanishing gradient implies that $F_2(U_1 - U_2)/\beta = \xi = 1$. Values of the criticality $\xi > 1$ then correspond to interfacial slopes steeper than those expected from the marginal

instability criterion and a more strongly reversed lower layer PV gradient. Criticality values $\xi < 1$ indicate the opposite mean state properties. The model results shown in Fig. 5 indicate that $\xi = 1$ does not appear to be a preferred value for this system. In the NL dynamics, the turbulent state is subcritical for weak wind stress and passes through $\xi = 1$ to become supercritical for larger wind stress. The QL model shows similar behavior, although the subcritical range is obscured by the transition to baroclinic instability. These results are consistent with the recent idealized modeling study of Jansen and Ferrari [6], who showed that the turbulent state of the atmosphere can be moved from a subcritical to a supercritical regime by varying the external control parameters.

As we are interested in the role of the eddies in determining the equilibrated turbulent state, we next relate the mean states shown in Fig. 5 to the properties of the eddy field. Figure 6 (a) shows the eddy diffusivity K in the NL and QL models as a function of wind stress. Both models produce eddy diffusivities on the order of $\mathcal{O}(100 \text{ km}^2/\text{day}) \approx 1000 \text{ m}^2/\text{s}$, which is in the realistic range for the ocean at mid depths [10]. We estimate K using two independent methods. First, taking the interfacial slope s from our numerical simulations we can solve for K in (28) to obtain

$$K_{\text{theory}} = \frac{\tau}{f_0 \rho_1 s} = \frac{\tau/(\rho_1 H_1)}{F_1(U_1 - U_2)}. \tag{63}$$

We use the subscript 'theory' to indicate that this estimate of K is not based on observations of the PV flux and gradient from the numerical model, but is derived from the equilibrium state using the assumptions outlined in Section 2.2. We also estimate K directly from our numerical simulations using the definition

$$K_{\text{est}} = -\frac{\overline{v_1' q_1'}}{\partial_y Q_1} \tag{64}$$

which we calculate at every model timestep and average over the domain and in time to estimate K. Figure 6 (a) shows that the eddy diffusivity depends strongly on the wind stress in both the QL and NL dynamics, increasing by a factor of 4 as A is varied from 1 km/day^2 to 7 km/day^2. The diffusivity becomes somewhat less sensitive to A as A is increased. The QL model equilibrates with a larger shear than the NL model in Fig. 5, and correspondingly the QL eddy diffusivity is weaker than that of the NL model.

Mixing length arguments suggest that K is related to the EKE through a functional relationship of the form $K \sim \sqrt{EKE}$ [2]. Figure 6 (b) shows the domain-average total EKE as a function of wind stress in NL and QL, calculated using equation (60). The EKE behaves similarly in the two models as A is increased. For very small A values the QL EKE is exactly zero since the fixed point is stable. Beyond the instability threshold the EKE of both NL and QL increase close to linearly as wind stress is increased, with the QL EKE growing more rapidly with A than the NL EKE. For $A \lesssim 2$ km/day^2, the NL EKE exceeds the QL EKE as the NL model is further from the threshold for baroclinic instability. For larger values of the wind stress the eddies in the QL model contain more energy than those in the NL model in spite of the additional diffusion in the QL dynamics. This suggests that the EENL terms produce an effective eddy viscosity acting on the eddies themselves, scattering energy from the wavenumbers capable of extracting energy from the mean state to

small-scale damped waves where the energy is dissipated. In our numerical implementation of NL this small-scale damping may be accomplished by the flux limiter or by numerical diffusion. Although we do not quantitatively examine the scaling relationship between K and EKE in this work, the downward concavity of the diffusivity curve and approximate linearity of the EKE with A is suggests that a relationship similar to $K \sim \sqrt{EKE}$ may be applicable to our results. However, it appears that if such a functional relationship is valid it may not hold across both models, as for larger A values the QL EKE is greater than the NL EKE, while the diffusivity of the NL dynamics is always larger than that of the QL dynamics.

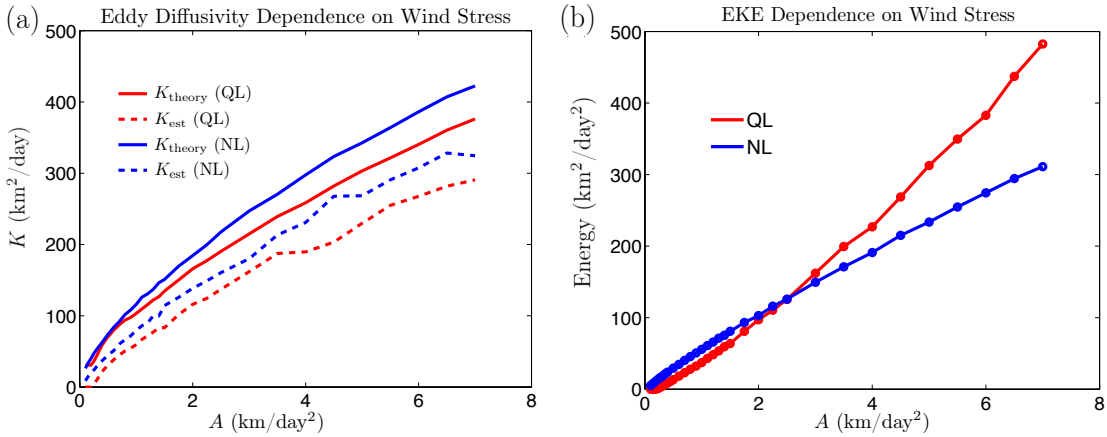


Figure 6: (a) Eddy diffusivity K in the QL (red) and NL (blue) models as functions of wind stress. K_{theory} is an estimate based on the equilibrium shear using equation (63), while K_{est} is an estimate based directly on the eddy flux and PV gradient in model simulations using equation (64). The results show that K increases substantially with increasing winds. (b) EKE as a function of wind stress in the QL (red) and NL (blue) models. In both models EKE increases approximately linearly with A.

3.3 Parameter Sensitivity

In Section 3.2 we presented the behavior of the NL model as the wind stress was increased and evaluated the ability of the mean-field dynamics to reproduce the observed NL behavior. In this section we vary the drag parameters d and r to test the robustness of the general agreement between the QL and NL models demonstrated in the previous section. Figure 7 shows the time and domain mean shear as a function of wind stress in the NL and QL models for four additional parameter cases. In panels (a) and (b) the bottom drag coefficient d is halved and doubled, respectively, relative to its reference value of $d = 500$ km^2/day. In panels (c) and (d) the Rayleigh drag coefficient r is halved and doubled with respect to its reference value of $r = 10^{-4}$ day^{-1}. For each case, all parameters are held constant except A and the drag parameter that was changed for that sensitivity test. In all cases the qualitative behavior of the NL model as A is increased is the same as discussed for the reference case. For weak wind stress the shear increases rapidly with A, with the sensitivity of the shear to the wind stress lessening as A is increased. For large wind stress the shear

appears to increase approximately linearly with A. It is clear from Fig. 7 that the success of the QL model in reproducing the NL shear scaling is robust to alternate choices of the the drag parameters.

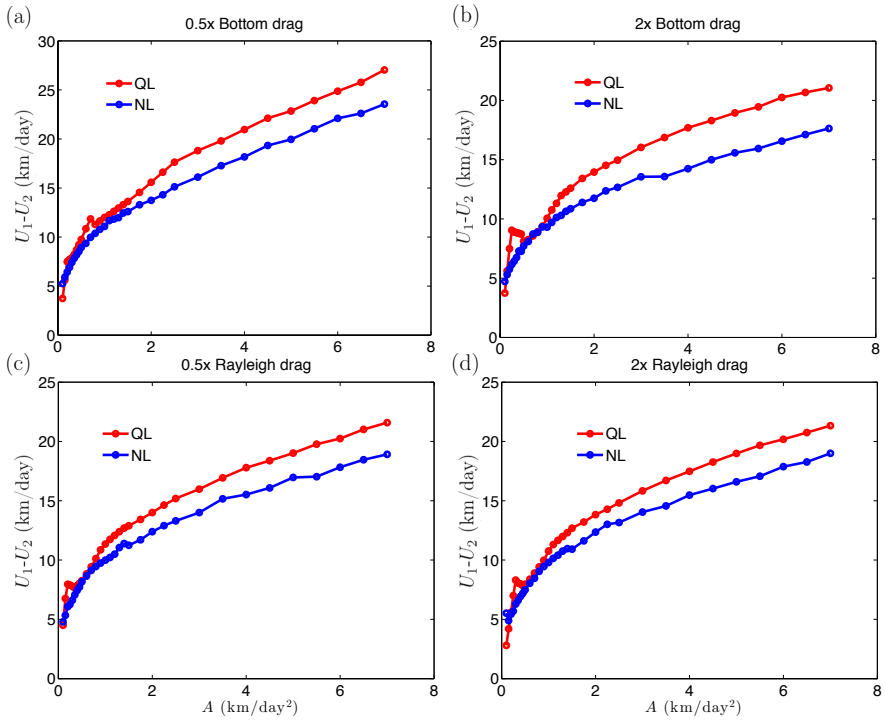


Figure 7: Testing the robustness of the agreement of QL and NL to changes in the drag parameters. All panels show the time and domain average shear in the QL (red) and NL (blue) models as functions of wind stress, for four different parameter cases. (a) $d = 250$ km^2/day, $r = 10^{-4}$ s^{-1}. (b) $d = 1000$ km^2/day, $r = 10^{-4}$ s^{-1}. (c) $d = 500$ km^2/day, $r = 0.5 \times 10^{-4}$ s^{-1}. (d) $d = 500$ km^2/day, $r = 2 \times 10^{-4}$ s^{-1}. The model behavior is similar in all cases.

Comparison of the NL curves in Fig. 7 (a) and (b) shows that the equilibrium shear in NL weakens as the bottom drag is increased for fixed A (note the difference in scale between (a) and (b)). This result is shown more clearly in Fig. 8 (b), which shows the equilibrium shear in QL and NL for the reference wind stress value $A = 1$ km/day^2 and three values of the bottom drag, corresponding to our reference drag $d = 500$ km^2/day and the halving and doubling cases shown in Fig. 7 (a) and (b). Both QL and NL show weakening shear with increasing bottom drag. Based on the scaling relation (28), this indicates that the eddy diffusivity K increases as the bottom drag is increased for constant winds. The physical explanation for this behavior is not obvious. From (29) increasing bottom drag results in a reduced lower level flow. However, this effect only changes the barotropic component of the flow, whereas the shear component is the one most relevant to the eddy physics. The shear associated with the fixed point solutions (35) and (36) also does not depend on d, in the limit $dF_2/r \gg 1$ relevant to this work. Understanding the physical mechanisms behind the parameter dependencies shown in Fig. 8 is an important future direction of the work.

The equilibrated shear also depends to some extent on the bulk Rayleigh drag as shown in Figs. 7 (c), (d) and 8 (a). As r is increased at constant wind, the shear decreases slightly in NL and somewhat more substantially in QL. The relation (28) implies that this decrease in shear results from an increase in the eddy diffusivity with increasing Rayleigh drag. This is somewhat surprising since the drag acts to remove energy from the eddy field, thus presumably weakening the eddy fluxes. However, the arguments in Section 2.2 were based on the assumption that the diabatic effects could be ignored in the mean budgets. In the complete dynamics, the drag r acts not only on the eddies but on the mean state as well, including directly damping the interface deflection through (18). The decrease in shear with increasing drag in Fig. 8 (a) likely results from a competition between the direct diabatic damping of the shear and the simultaneous reduction of the EKE and corresponding eddy diffusivity by the drag.

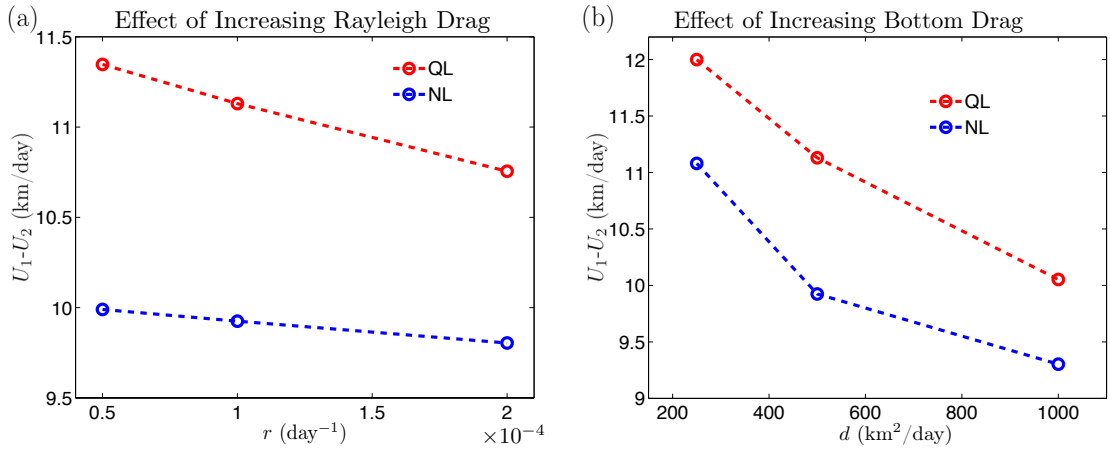


Figure 8: (a) Time and domain average shear in the QL (red) and NL (blue) models as functions of Rayleigh drag r. Wind stress is fixed at $A = 1$ km/day^2. (b) Same, but for varying bottom drag d. The equilibrium shear is seen to weaken with both forms of drag in both the QL and NL models.

In Fig. 9 we show the meridional structure of the shear for the reference wind stress $A = 1$ km/day^2 and the four alternate drag parameter choices previously discussed. As was found in the reference case (Fig. 4), important qualitative differences between the shear profiles of NL and QL are visible. The QL model forms multiple jets in all cases. Similar flank jets are also found in the NL model in the case of weak bottom drag (Panel (a)), but the additional jets are much weaker than their QL counterparts. Evidently, the wave-mean flow dynamics of a single zonal Fourier component has a strong tendency to produce multiple jets that is strongly suppressed in the NL dynamics. In the next section we will show the results of a preliminary attempt to produce a more realistic jet structure in the QL model by including additional zonal wavenumber components and modeling the effect of EENL interactions with stochastic forcing.

Based on our observations of the NL EKE and PV flux spectra in Fig. 3, we chose $k = 4$ as our single zonal wavenumber component in the QL dynamics. Figure 10 (a) shows the shear as a function of wind stress for reference drag values in the QL dynamics and in

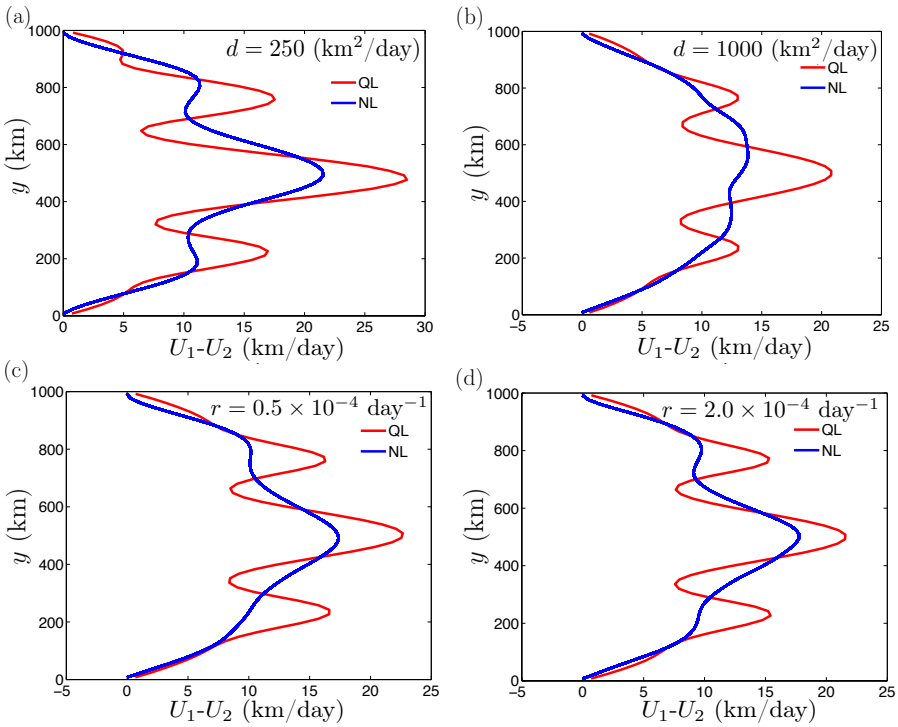


Figure 9: Comparison of the QL and NL profiles of time and zonal mean shear for different drag values. In all cases QL produces multiple jets that are not seen in NL. (a) $d = 250$ km^2/day, $r = 10^{-4}$ s^{-1}. (b) $d = 1000$ km^2/day, $r = 10^{-4}$ s^{-1}. (c) $d = 500$ km^2/day, $r = 0.5 \times 10^{-4}$ s^{-1}. (d) $d = 500$ km^2/day, $r = 2 \times 10^{-4}$ s^{-1}.

two alternate QL models in which $k = 3$ and $k = 5$ were chosen as the QL wavenumbers. Although the choice of wavenumber does have a quantitative effect on the shear, the shapes of the curves are quite similar for all three wavenumber choices. The $k = 3$ model maintains the largest shear, indicating that the eddy diffusivity is smallest for this choice of k. This is conceptually consistent with our observations of the NL spectrum, which showed that the EKE and PV flux associated with $k = 4$ and $k = 5$ were comparable while $k = 3$ was relatively weak. We take this as an indication that the $k = 3$ wave is less efficient at extracting energy from the equilibrated NL mean state than $k = 4$ and $k = 5$, consistent with the behavior shown in Fig. 10 (a) for the QL models.

Although we have chosen to work with the maximally-simplified QL dynamics in which only a single wavenumber is retained, it is also possible to formulate the QL dynamics with more than one wavenumber. In such a model each eddy component $\tilde{q}_{1,k}(y,t)$ has its own evolution equation of the form (56), (57). Each component interacts with the same mean field $Q_1(y,t)$, $Q_2(y,t)$, so that the waves are indirectly coupled through their mutual interactions with the mean flow. Figure 10 (b) shows an example of the results of a QL integration in which the first 8 wavenumber components $k \in (2\pi/L)\{1,\ldots,8\}$ were retained and reference parameter values were chosen. During the initial transient phase of instability growth, all unstable waves begin to grow. (See Fig. 2 (b) for the instability growth rates as a function of k for this case). Waves 3,4, and 5 grow to large amplitude, while all

other waves remain too weak to be seen on our axis scale. At equilibrium waves 4 and 5 persist. The structure of the resulting jets (not shown) is similar to the jets produced by the $k = 4$ single-wave model. In particular, the multi-wave QL dynamics continues to produce multiple jets in qualitative disagreement with NL solutions.

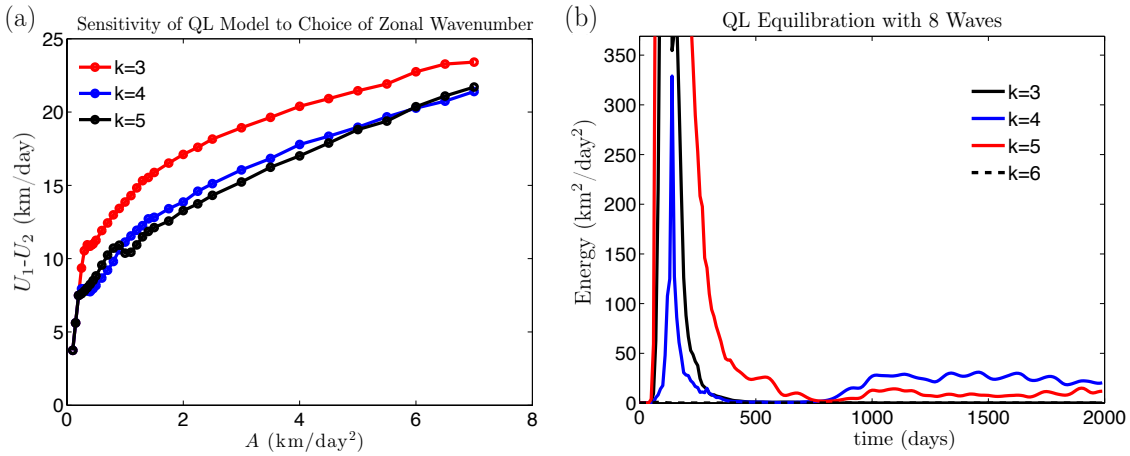


Figure 10: (a) Time and domain average shear as a function of wind stress for QL models in which different zonal wavenumbers were retained. All three cases show similar qualitative behavior, but the $k = 3$ QL model supports a larger equilibrium shear. (b) Time series of EKE for each wavenumber in a QL model run for reference parameter values in which the first 8 wavenumber components were retained. Initially all unstable waves grow, and some waves reach large amplitude before decaying to zero asymptotically. Two waves, $k = 4$ and $k = 5$, survive at finite amplitude in the statistical equilibrium state.

3.4 Stochastically-Driven QL Dynamics

In this section we present the results of a preliminary attempt to resolve the discrepancy between the jet structures of the QL and NL dynamics. Our approach is to introduce additional stochastic forcing into the QL model with multiple wavenumber components. The motivation for this model comes from several observations. First, it is clear from the NL energy spectrum (Fig. 3) that although the EKE is concentrated at $k = 4$ and $k = 5$ in the reference simulation, the wavenumber spectrum is populated up to about wave 10, with most of these waves producing non-negligible contributions to the PV flux. Second, the timeseries in Fig. 10 (b) shows that wavenumbers that are not present in the statistical equilibrium state still have the capacity to extract substantial energy from the mean flow as demonstrated by their rapid growth at early times. Although the equilibrium state has been adjusted so that waves other than $k = 4$ and $k = 5$ have negative growth rates on the mean state, the linear dynamics of these waves remains highly non-normal and thus can exhibit large transient growth if they are somehow excited. In the QL dynamics no such excitation occurs and these waves remain at zero amplitude forever. This is an unrealistic aspect of the QL model, since EENL interactions in the NL dynamics scatter energy into all wavenumbers. We make the simplifying assumptions that the EENL scattering is not

correlated in time and not dependent on the system state, and model the effect of the EENL terms by adding stochastic forcing to the evolution equations for all wavenumber components in the QL dynamics.

The stochastic QL model modifies the usual QL equations of motion to include an additional term on the RHS of the eddy PV equations:

$$\partial_t \tilde{q}_{1,k} = \left[-ikU_1 - r + \kappa(\partial_{yy} - k^2)\right] \tilde{q}_{1,k} + \left[-ik\partial_y Q_1\right] \tilde{\psi}_{1,k} + \varepsilon G_k(y,t) \tag{65}$$

$$\partial_t \tilde{q}_{2,k} = \left[-ikU_2 - r + \kappa(\partial_{yy} - k^2)\right] \tilde{q}_{2,k} + \left[-ik\partial_y Q_2 - dF_2(\partial_{yy} - k^2)\right] \tilde{\psi}_{2,k} + \varepsilon G_k(y,t). \tag{66}$$

The evolution equations for Q_1 and Q_2 remain the same, since EENL terms appear only in the eddy equations in the NL dynamics. In these equations $G_k(y,t)$ is a random forcing term that is applied equally in both layers for simplicity. We choose our random forcing to be δ-correlated in time and correlated in space with a structure $g_k(y,y')$ such that $\langle G_k(y,t) G_k^\star(y',t')\rangle = g_k(y,y')\delta(t-t')$ where angle brackets indicate the ensemble mean over realizations of the noise. We take the stochastic forcing to be equal for all wavenumbers so that $g_k(y,y') = g(y,y')$. We choose the spatial correlation structure of the noise to be given by

$$g(y,y') = \exp\left(-\frac{(y-y')^2}{\ell_c^2}\right) \exp\left(-\frac{(y-\frac{W}{2})^2}{2\Delta^2}\right) \exp\left(-\frac{(y'-\frac{W}{2})^2}{2\Delta^2}\right). \tag{67}$$

This correlation function produces noise with Gaussian correlation in space with correlation length ℓ_c and noise amplitude peaked in the channel center. We take $\ell_c = 78$ km and $\Delta = 156$ km. These choices ensure that the noise is locally smooth over a few model gridpoints and that the amplitude of the forcing is substantial over the middle third of the channel, tapering toward the boundaries. This choice of forcing structure was chosen for its simplicity. An important extension of this work will be to use observations of EENL scattering in the NL dynamics to understand what forcing structure is most realistic of the true EENL dynamics. The parameter ε controls the overall amplitude of the noise forcing and has units of day$^{-3/2}$ so that ε^2 has units of potential enstrophy injection rate. In the NL model the EENL interactions do not inject energy into the eddy field, but rather move energy from wavenumber to wavenumber while leaving the total EKE invariant. Our stochastic forcing is deficient in this sense because forcing of the form described above injects energy directly into the eddy field. The stochastic parameterization could thus be improved by augmenting the eddy dissipation such that the eddy energetics are consistent with those of NL [5]. For simplicity we ignore these concerns and proceed with the formulation described above as a first step. We refer to the stochastic dynamics introduced above as the stochastic quasilinear model (SQL).

Figure 11 summarizes several results from SQL. All SQL integrations were performed with reference values for the wind stress and drag parameters, and the first six zonal wavenumber components were retained in the calculations. Panel (b) shows the time evolution of the energy of each wavenumber component during a 10000 day model integration with stochastic forcing strength $\varepsilon = .05$ day$^{-3/2}$. It is clear that all waves are excited and fluctuate at finite amplitude. Wavenumbers 4 and 5 (cyan and green curves) are no longer dominant, with the largest contribution to the eddy energy being due to $k = 1$. This is clearly seen in panel (c) which shows the time average EKE spectrum. The spectrum is red

and in poor agreement with the NL spectrum shown in Fig. 3. Note, however, that our maximally simple forcing in which each component of the eddy PV is forced with equal amplitude leads to a preferential injection of energy into the large scales (enstrophy is injected into each scale equally). This forcing bias, together with the action of diffusion to dissipate higher wavenumbers, may underlie the complete lack of agreement between the SQL and NL spectra. Panel (d) shows the dependence of the total EKE on the forcing strength. As expected, the eddies are more energetic when they are forced more strongly, and we expect this stronger eddy field to suppress the equilibrium shear.

The SQL equilibrium shear profiles are shown in Fig. 11 (a) for three values of the stochastic forcing amplitude, with the results of the NL reference case reproduced for comparison. For the smallest value of ε the shear profile is close to that of the ordinary QL model, showing strong jets on the flanks of the center jet. As ε is increased both the amplitude and the structure of the jets undergo substantial changes. The amplitude of the shear decreases as ε is increased, consistent with our understanding that the stochastic forcing supports the transient growth of waves whose eddy fluxes barotropize the zonal flow. The flank jets also weaken substantially as ε increases. For $\varepsilon = .05$ day$^{-3/2}$ (red curve), the jet structure is quite similar to that of the NL model. The amplitude of the shear, however, is too weak for this choice of forcing. Keeping the correlation structure of the forcing fixed and varying the parameter ε only, it is not possible to accurately fit the NL jet structure: either the strength or the structure of the shear can be approximately matched, but matching both simultaneously is not possible. However, the reduction of the strength of the unrealistic flank jets when stochastic forcing is included is an encouraging result, and future work with the SQL model will involve experimenting with the forcing structure $g_k(y, y')$ to evaluate whether SQL is capable of producing fully-realistic jet structures and amplitudes.

4 Discussion and Conclusion

The goal of this work is to advance our theoretical understanding of the interactions between eddies and the large scale mean state in baroclinically unstable flow driven by surface stress. In particular, we are interested in understanding how eddy-mean flow interactions change in response to changes in the external forcing. The ACC provides a concrete example system for which the external forcing is changing and the eddy response must be understood if we are to understand the response of the mean state. Theoretical models of the ACC mean state typically rely on the assumption of eddy diffusion. The manner in which the eddy diffusivity is parameterized in terms of the large-scale variables in such models can be the determining factor in the model predictions of the ACC response to changing forcing. Although the sensitivity of model predictions to the eddy diffusion parameterization indicates that the eddy diffusivity is a quantity crucial to our understanding of the turbulent dynamics, no accepted theory exists to predict its value or how it depends on the external parameters of the system. In this work, we make progress on this problem by demonstrating that a reduced dynamics, the mean-field equations, reproduces many aspects of an idealized, but fully turbulent, model of the ACC, including the variation of important mean state properties as the external forcing is varied. Although the theoretical prediction of mean state variations with forcing remains an open problem, the identification of the mean-field model as a significantly simplified theoretical framework that contains the essence of the

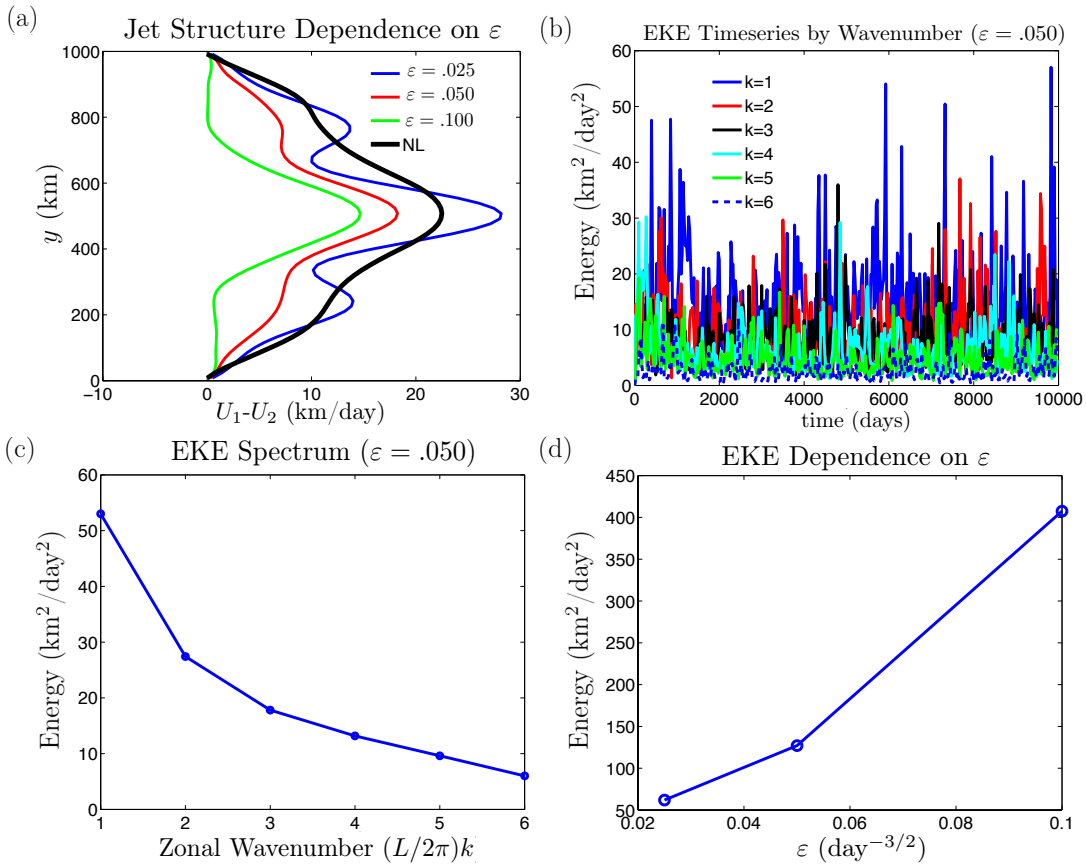


Figure 11: Summary of results of the stochastically-driven QL model. In all calculations reference parameter values were used and 6 zonal wavenumber components were retained. (a) Equilibrium shear profiles for three stochastic forcing strengths ε. The shear is seen to weaken overall as ε is increased, as well as shift toward a single-jet meridional structure. (b) Time series of EKE for each wavenumber in the stochastic model. All waves are sustained at finite energy in the equilibrium state. (c) Time average EKE spectrum for the stochastic model. In disagreement with the NL reference spectrum in Fig. 3, the eddy field is energetically dominated by the largest structures. (d) Total EKE as a function of stochastic forcing strength.

eddy-mean flow interaction and its dependence on the external model parameters provides a promising pathway for future efforts toward understanding the physics of the ACC response to forcing.

We used the idealized two-layer QG model in a flat-bottom channel configuration as our model ACC. We refer to this fully-nonlinear model formulation as the NL model. The flow was driven by wind stress at the surface, with the injected momentum being removed from the system in the lower layer by bottom friction. For very weak wind stress the flow is baroclinically stable, but the flow becomes unstable and turbulent for wind strengths that are realistic for the SO. By integrating the model equations numerically to a statistically-steady state for different values of the wind forcing, we assessed how the slope of the

interface, or equivalently the baroclinic component of the zonal mean flow, responded to changes in the winds. For weak winds the mean state was very sensitive to changes in forcing, with the shear increasing rapidly as the winds were increased. The shear responded less sensitively to further increases in the winds, demonstrating eddy saturation to some degree. However, the shear was not found to become entirely independent of the winds for any range of parameters tested, and so the model does not appear to exhibit complete eddy saturation.

We formulated the mean-field model by removing eddy-eddy advection terms from the equations of motion and retaining a single zonal wavenumber component in the eddy field. We refer to the mean-field equations by the abbreviation QL. In the domain mean picture, this highly simplified theoretical model reproduced many qualitative aspects of the NL dynamics. The dependence of the mean shear on the wind stress in QL mirrored that of NL, with differences occurring primarily at the quantitative level. The QL model also reproduced the sense of the dependence of the mean state on the imposed drag parameters. Our QL experiments demonstrated that the QL model robustly reproduces the domain-mean properties of the NL dynamics over a wide range of parameters in spite of the substantial simplifications made in deriving the model equations. Our results indicate that, in the domain mean sense, the strength and parameter-dependence of the eddy diffusivity are not fundamentally controlled by eddy-eddy processes.

We also compared the QL and NL models in the zonal mean, but not channel-mean, picture. When the channel-mean was not taken, important differences between the two models were evident. The equilibrated meridional structure of the shear flow in the NL takes the form of a broad single jet intensified at the channel center. In contrast, the QL equilibrium jet structure features multiple jets, with additional strong jets occurring on the flanks of the center jet. These qualitative differences in structure occur over all tested parameter ranges. This fundamental difference in structure points to qualitative differences in model physics that should be resolved if the QL dynamics is to be used to understand the NL model behavior.

In an attempt to improve the agreement between the meridional jet structures of the QL and NL model, we modified the QL dynamics to include additional wavenumber components and stochastic forcing as a parameterization of the EENL terms appearing in the NL dynamics. When stochastic forcing was included, the QL model attained an equilibrium state with a populated energy spectrum, and the flank jets were seen to weaken and disappear as the strength of the noise was increased. However, the energy spectrum of the stochastic model was in poor agreement with that of NL, and the QL jet amplitudes were too weak when compared with those of NL when the noise was strong enough to suppress the unrealistic multiple jets. We believe that these failures of the stochastic model may be due to bias introduced by the simplified structure chosen for the noise forcing as well as the diffusion introduced into the QL model for reasons of numerical stability.

The results presented here constitute a step toward understanding the relationship between the QL and NL dynamics but many aspects of the problem require additional research. At the level of the NL dynamics, a more detailed consideration of the eddy energy budget is required to understand the effects of the various forms of drag on the eddy field, and also how energy is passed between wavenumbers by eddy-eddy interactions. These results will be useful for informing the parameterization of the eddy-eddy terms in the QL dynamics.

Incorporating an improved noise parameterization is likely to improve the realism of the QL jet structure.

Beyond the problem of simply reproducing the NL simulation results with the QL model, it also remains an open problem to leverage the simplicity of the QL dynamics to develop a theoretical understanding of eddy-mean flow interactions in this problem. In the QL model without stochastic forcing, a possible approach is to identify and trace the unstable fixed point solution that exists, at least in some cases, beneath the fluctuating equilibrium state as a function of the wind stress. This fixed point consists of an equilibrated baroclinic wave and a steady mean flow, and becomes unstable in a supercritical Hopf bifurcation to an oscillatory state which becomes chaotic as the 'turbulent' QL state is reached. As the fluctuations in equilibrium appear to remain fairly close to this fixed point, the unstable fixed point solution is a relatively simple, time-independent mathematical object whose analysis may lead to a physical understanding of how eddy fluxes are determined in cooperation with the mean state. In the case of the stochastic QL model, a theoretical framework for analyzing the statistical mean equilibrium state has been developed by Farrell and Ioannou [4]. A promising future direction of this research is to use their covariance-matrix formalism to understand how the ensemble-mean eddy fluxes are determined by interaction with the mean state.

Our decision to model the ACC using the two-layer QG equations in a flat-bottomed reentrant channel was motivated by its simplicity. However, it is well-established that bottom topography plays an important role in the ACC dynamics. In a flat-bottom geometry the overall momentum balance must be between momentum injection by the winds and removal by bottom drag. This requires the development of a strong barotropic flow component capable of dissipating the required amount of momentum at the bottom. In the real ACC the momentum balance is between wind input and bottom form drag on topography, and the barotropic flow is weak. Flow over bottom topography also produces stationary waves that can have a significant effect on the ACC stratification by increasing the efficiency of eddy heat transport and correspondingly shallowing isopycnal slopes [1]. Our idealized model cannot account for these effects in its current formulation. However, topography could in principle be included in both the NL and QL models discussed in this work, and understanding the impact of topography in our framework is an important future direction.

5 Acknowledgments

I thank Glenn Flierl and Raf Ferrari for their help and guidance during my work on this project, and Brian Farrell for valuable comments on the manuscript. I also thank the organizers, participants, and Fellows of the 2014 WHOI GFD program for making the summer educational and enjoyable.

6 Appendix: Sensitivity of NL Model to Numerical Resolution

Figure 12 shows a comparison of nonlinear model runs at two different resolutions: the default resolution with 64×128 spatial grid points, and a doubled resolution run with

128×256 gridpoints. Reference parameter values are used for wind and drag. Panels (a) and (c) show snapshots of the upper layer PV fields. Panels (b) and (d) show the statistical mean upper and lower layer zonal jets. The PV fields show similar features and have similar characteristic magnitude in both cases. The higher resolution run shows some small-scale filamentation that is less prominent in the default resolution case. However, these small-scale features do not appear to dominate the flow. The time mean flow is also nearly identical in the two runs. The mean flow in the doubled-resolution run is not perfectly symmetric about the channel center. This is due to an insufficiently long period of averaging due to the increased computational cost to run the model at this resolution. Based on Fig. 12 we conclude that 64×128 gridpoints is sufficient to resolve the structures most important to the dynamics, at least in the parameter ranges considered here.

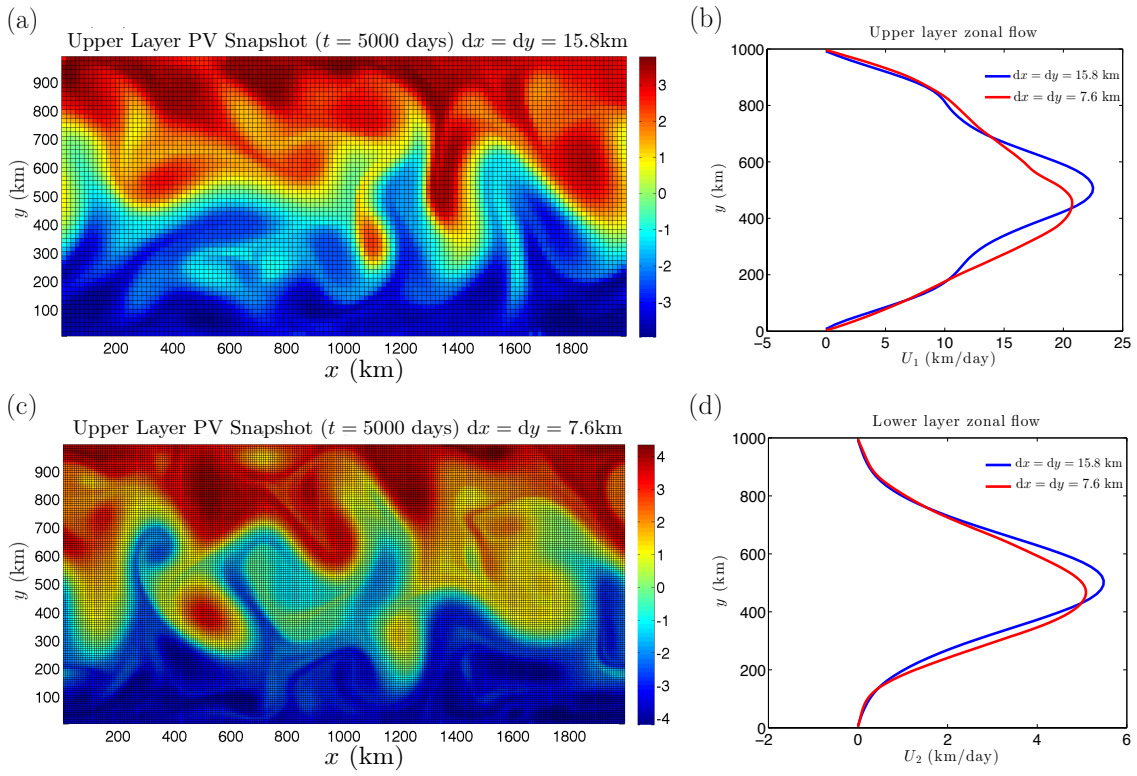


Figure 12: Test of spatial resolution in the NL model. Default resolution is 64×128 gridpoints, doubled resolution is 128×256 gridpoints. Panels (a) and (c) show snapshots of the upper layer PV fields. Although some smaller scale features are seen in the high resolution case that are absent from the default resolution case, the PV field still appears to be dominated by wave 4 and wave 5 structures. Panels (b) and (d) show the upper and lower level zonal mean flows in statistically steady state. High resolution jets (red) show some deviations from perfect cross-channel symmetry due to insufficiently long averaging periods. The high and default resolution flows are of nearly the same amplitude and similar meridional structure.

References

[1] R. Abernathey and P. Cessi, *Topographic Enhancement of Eddy Efficiency in Baroclinic Equilibration*, Journal of Physical Oceanography, 44 (2014), pp. 2107–2126.

[2] R. Abernathey, J. Marshall, and D. Ferreira, *The Dependence of Southern Ocean Meridional Overturning on Wind Stress*, Journal of Physical Oceanography, 41 (2011), pp. 2261–2278.

[3] C. W. Boening, A. Dispert, M. Visbeck, S. R. Rintoul, and F. U. Schwarzkopf, *The response of the Antarctic Circumpolar Current to recent climate change*, Nature Geoscience, 1 (2008), pp. 864–869.

[4] B. Farrell and P. Ioannou, *Structural stability of turbulent jets*, Journal of the Atmospheric Sciences, 60 (2003), pp. 2101–2118.

[5] B. F. Farrell and P. J. Ioannou, *A Theory of Baroclinic Turbulence*, Journal of the Atmospheric Sciences, 66 (2009), pp. 2444–2454.

[6] M. Jansen and R. Ferrari, *Macroturbulent Equilibration in a Thermally Forced Primitive Equation System*, Journal of the Atmospheric Sciences, 69 (2012), pp. 695–713.

[7] A. Kuo, R. Plumb, and J. Marshall, *Transformed Eulerian-mean theory. Part II: Potential vorticity homogenization and the equilibrium of a wind- and buoyancy-driven zonal flow*, Journal of Physical Oceanography, 35 (2005), pp. 175–187.

[8] C. Le Quere, C. Roedenbeck, E. T. Buitenhuis, T. J. Conway, R. Langenfelds, A. Gomez, C. Labuschagne, M. Ramonet, T. Nakazawa, N. Metzl, N. Gillett, and M. Heimann, *Saturation of the Southern Ocean CO2 sink due to recent climate change*, Science, 316 (2007), pp. 1735–1738.

[9] J. Marshall and T. Radko, *Residual-mean solutions for the Antarctic Circumpolar Current and its associated overturning circulation*, Journal of Physical Oceanography, 33 (2003), pp. 2341–2354.

[10] J. Marshall and K. Speer, *Closure of the meridional overturning circulation through Southern Ocean upwelling*, Nature Geoscience, 5 (2012), pp. 171–180.

[11] M. P. Meredith, A. C. N. Garabato, A. M. Hogg, and R. Farneti, *Sensitivity of the Overturning Circulation in the Southern Ocean to Decadal Changes in Wind Forcing*, Journal of Climate, 25 (2012), pp. 99–110.

[12] A. K. Morrison and A. M. Hogg, *On the Relationship between Southern Ocean Overturning and ACC Transport*, Journal of Physical Oceanography, 43 (2013), pp. 140–148.

[13] D. R. Munday, H. L. Johnson, and D. P. Marshall, *Eddy Saturation of Equilibrated Circumpolar Currents*, Journal of Physical Oceanography, 43 (2013), pp. 507–532.

[14] M. Nikurashin and G. Vallis, *A Theory of Deep Stratification and Overturning Circulation in the Ocean*, Journal of Physical Oceanography, 41 (2011), pp. 485–502.

[15] ——, *A Theory of the Interhemispheric Meridional Overturning Circulation and Associated Stratification*, Journal of Physical Oceanography, 42 (2012), pp. 1652–1667.

[16] P. A. O'Gorman and T. Schneider, *Recovery of atmospheric flow statistics in a general circulation model without nonlinear eddy-eddy interactions*, Geophysical Research Letters, 34 (2007).

[17] K. Srinivasan and W. R. Young, *Zonostrophic Instability*, Journal of the Atmospheric Sciences, 69 (2012), pp. 1633–1656.

[18] P. H. Stone, *Baroclinic Adjustment*, Journal of the Atmospheric Sciences, 35 (1978), pp. 561–571.

[19] D. Thompson and S. Solomon, *Interpretation of recent Southern Hemisphere climate change*, Science, 296 (2002), pp. 895–899.

[20] X.-Y. Yang, R. X. Huang, and D. X. Wang, *Decadal changes of wind stress over the Southern Ocean associated with Antarctic ozone depletion*, Journal of Climate, 20 (2007), pp. 3395–3410.

Models for Tropopause Height and Radiative-Convective Equilibrium

Shineng Hu

August 16, 2015

1 Introduction

Radiative-convective equilibrium (RCE) is the equilibrium state that a convective system reaches, where convective heating compensates radiative cooling. RCE has often been used to represent the planetary atmosphere on a global scale, and it is also useful for studying the tropical climates since tropical atmosphere is always believed to reside in a near RCE state.

There is a hierarchy of numerical models that have been developed to study the RCE system, ranging from radiative-convective models (RCMs) with parameterized convection to cloud resolving models (CRMs) with explicit convective processes. Those numerical models have led us to a much deeper understanding of the radiative-convective system, but the basic physics of RCE is still far from fully understood.

One important reason is that, although the RCE concept is already simplified from reality, the full RCE system (or model) is still too complicated at least in two aspects. First, radiative and convective processes are highly coupled and nonlinear. For example, convection transports water vapor from the boundary layer to the free troposphere, and it effectively influences the absorbed longwave radiation in the atmosphere, which in turn affects the local buoyancy (temperature) and thus convection. Second, current RCMs or CRMs are usually divided into tens of vertical layers. The exchange of mass and energy between different layers makes the system very difficult to diagnose.

To better understand the RCE system, one can try to simplify either aspect above, and this is one of the motivations of this study. Accordingly, the main goal of this study is twofold. The first part will stick with multiple layers, but try to reduce the coupling of the system by specifying some variables like lapse rate and optical depth. One main goal of this part is to understand the structure of earth's tropopause by taking advantage of the simplifications we made here. The second part will try to develop a maximally simplified RCE model with an interactive hydrological cycle to better understand the basic physics of RCE itself. This model contains two layers in troposphere, but the convective and radiative processes are still highly coupled. More details will be discussed below.

2 The height of tropopause

2.1 Background

Tropopause is usually referred to as the boundary between the troposphere and the stratosphere, even though it has various definitions based on the spatial structures of temperature, potential vorticity, and chemical concentrations. Understanding the height of tropopause is important to studying the column total ozone [1] and also the mass exchange between the troposphere and the stratosphere [2].

Despite its importance, many problems related to the observed features of the tropopause remain unsolved. From observations, the height of tropopause is higher in the tropics (16km) and lower in the polar regions (8km), and this transition sharply occurs in the mid-latitudes (Fig. 1a). What shapes this particular structure? Also, tropical tropopause, compared with the one in extratropics, is not only higher, but also colder and sharper (Fig. 1b). What can explain those three features simultaneously?

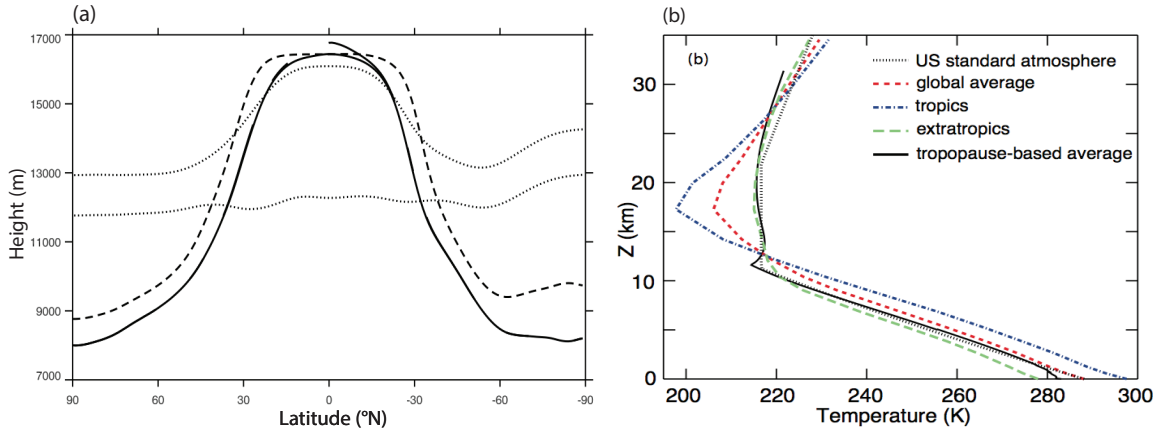


Figure 1: (a) Zonal-mean tropopause height as a function of latitude (solid). Figure adopted from [3]. (b) Zonal mean temperature profiles in different regions. Note the contrast between tropics (blue dash-dotted) and extratropics (green dashed). Figure adopted from [4].

To that end, we will use a extremely simplified tropopause model, and see whether this one-dimensional (1-D) model can reproduce the observed structure of tropopause. One thing noteworthy is that there will be no explicit dynamics in this model, although lateral transport and mid-latitude eddies are shown to be important in shaping the tropopause structure. Our following results will show that the observed features of tropopause can be surprisingly well captured by this simple 1-D model.

2.2 Numerical model

Let's start with a two-stream grey atmosphere, which is commonly used in early related studies [5, 6], and write the radiation transfer equation for longwave radiation as,

$$\frac{\partial D}{\partial \tau} = B - D, \qquad \frac{\partial U}{\partial \tau} = U - B, \tag{2.1}$$

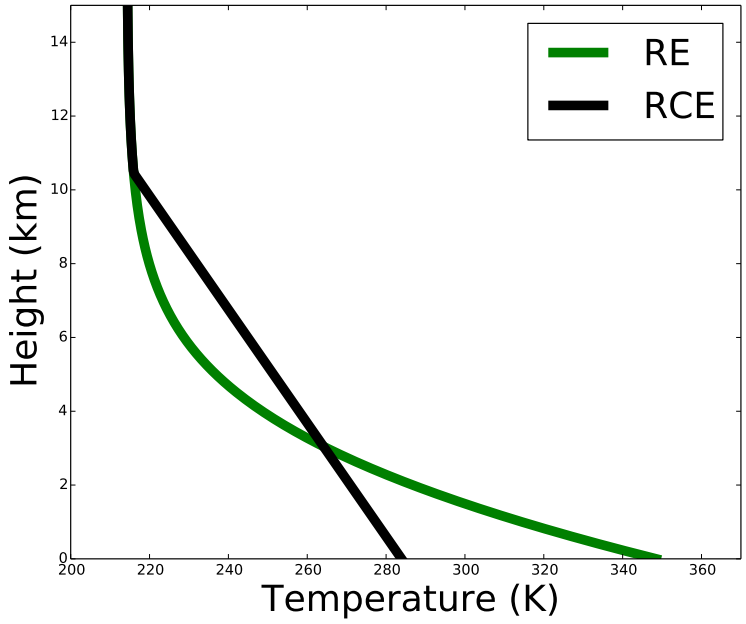


Figure 2: Temperature profiles for radiative equilibrium (green) and radiative-convective equilibrium (black).

where D and U are downward and upward longwave radiation, $B = \sigma T^4$ is Stefan-Boltzmann law, and τ is optical depth increasing downward. Then we can quickly rewrite Eq. (2.1) as

$$\frac{\partial I}{\partial \tau} = J - 2B, \qquad \frac{\partial J}{\partial \tau} = I \tag{2.2}$$

using

$$I = U - D, \qquad J = U + D. \tag{2.3}$$

Consider a radiative equilibrium (RE) state and assume that no shortwave radiation is absorbed by the atmosphere, then the convergence of net longwave radiation should vanish, which can be expressed as,

$$\frac{\partial I}{\partial \tau} = 0. \tag{2.4}$$

Furthermore, by using the boundary condition at the top of the atmosphere,

$$\tau = 0 : \qquad U = U_t = \sigma T_e^4, \quad D = 0 \tag{2.5}$$

we arrive at the solutions for radiative equilibrium,

$$D(\tau), U(\tau), B(\tau) = \left(\frac{\tau}{2}, \frac{\tau + 2}{2}, \frac{\tau + 1}{2} \right) U_t. \tag{2.6}$$

The temperature profile from this solution is shown as the green line in Fig. 2

Let's look at several features of RE solutions. When τ is very small, B/U is one half, and it physically means that near the top of the atmosphere, half of upward longwave radiation is coming from remote radiative transfer. Note that, until now, we have not used any boundary condition at the surface. So this observation will be valid for any cases with any lower boundary conditions, as long as the upper atmosphere is in the radiative equilibrium state; it will be used in the following analysis when we are deriving the analytical approximation of RCE solutions. Now consider at the surface, the difference between U and B is always non-zero. It indicates that there is always a temperature jump between ground temperature and near-surface air temperature in RE solutions, and this temperature jump is actually nontrivial, around $10K$. Another main feature is that the lapse rate near the surface is so high that convection can occur, which will result in a new equilibrium state called radiative-convective equilibrium (RCE) state.

Within this new equilibrium state, the system is in RE in the stratosphere, given by the solutions above, while in RCE in the troposphere with a uniform lapse rate. Mathematically, the temperature profile can be expressed as,

$$T(z) = \begin{cases} T_{re}, & z \geq H_T, \\ T_T + \Gamma(H_T - z), & H_T \geq z \geq 0, \end{cases} \tag{2.7}$$

where H_T is the tropopause height, and T_T is the tropopause temperature. The two vertical coordinates (τ and z) can be related by the formulation below,

$$\tau(z) = \tau_s \left[f_{H_2O} exp(-z/H_a) + (1 - f_{H_2O}) exp(-z/H_s) \right] \tag{2.8}$$

where τ_s is surface optical depth, f_{H_2O} is a linear parameter controlling the contribution of water vapor to optical depth compared with carbon dioxide, H_a is the scale height of water vapor (typically 2km), and H_s is the scale height of carbon dioxide (typically 8km). Additionally, we should add one more boundary condition at the surface,

$$\tau = \tau_s : \qquad U = B = \sigma T_s^4. \tag{2.9}$$

We can solve this set of equations numerically by integrating Eq. (2.1) from the top of the atmosphere to the surface given a certain H_T, and iterate the calculation to match both boundary conditions. The temperature profile associated with the RCE solutions is shown as the black curve in Fig. 2.

Note that the RCE solution above is a function of emission temperature T_e, surface optical depth τ_s, and tropospheric lapse rate Γ. The sensitivity of tropopause height to those three parameters are presented in Fig. 3. Tropopause height significantly increases when surface optical depth increases or tropospheric lapse rate decreases; tropopause height increases with outgoing longwave radiation but its impact is not as strong as the other two parameters.

2.3 Analytical approximation

To better understand the phase diagrams in Fig. 3, it is helpful to derive an analytical expression for tropopause height. However, finding the exact analytical solution is very difficult due to the multiple-layer model set-up and the nonlinearity of the system. Therefore,

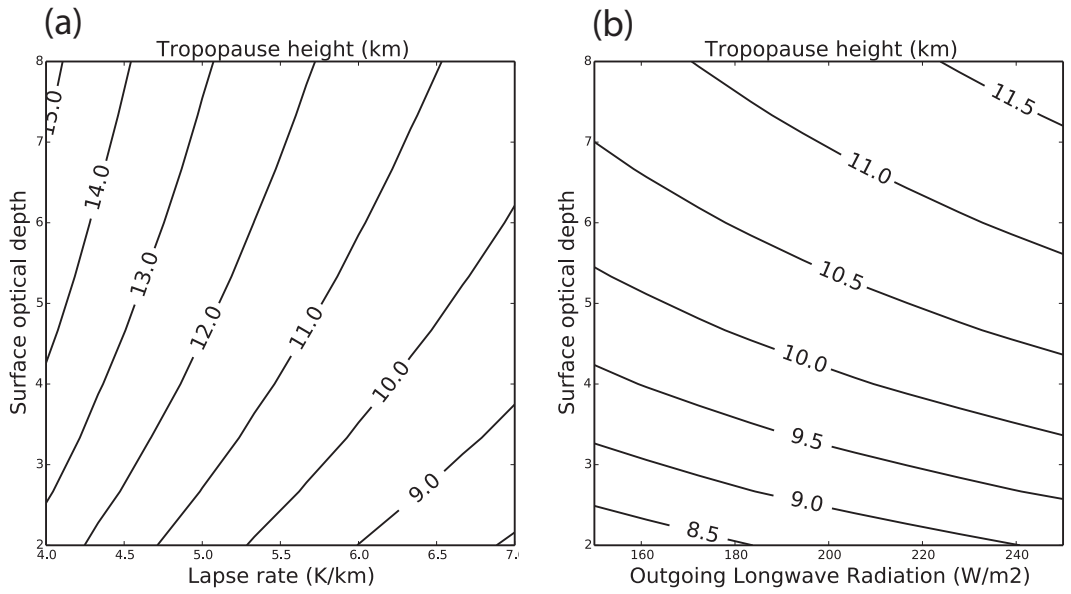


Figure 3: Tropopause height as a function of surface optical depth and (a) tropospheric lapse rate and (b) outgoing longwave radiation. Note that the parameter ranges are chosen as typical ranges of earth's atmosphere.

we make the following assumptions. First, temperatures above the tropopause are uniform and equal to $2^{-1/4}T_e$, which is obtained from Eq. (2.6) by simply setting τ to be zero. Second, as mentioned before, the lapse rate is held uniform throughout the whole troposphere below H_T. Third, the contribution to optical depth all comes from water vapor, that is, f_{H_2O} is set to 1. Fourth, B/U linearly increases from 1/2 (radiative equilibrium when τ is zero) at the tropopause to 1 (lower boundary condition) at the surface, which can be expressed as $B/U = 1 - z/2H_T$, and this assumption is fairly good for a wide range of surface optical depth and tropospheric lapse rate as shown in Fig. 4.

With those assumptions, we can eventually derive an analytical approximation for the height of tropopause as follows,

$$H_T = \frac{1}{16\Gamma}\left(CT_T + \sqrt{C^2T_T^2 + 32\Gamma\tau_s H_a T_T}\right), \tag{2.10}$$

where $C = ln4 \approx 1.38$. This expression agrees quite well with the sensitivity results presented in Fig. 3.

2.4 What shapes the observed tropopause structure from an RCE perspective? The role of Brewer-Dobson circulation

One question we can ask is could we reproduce the observed tropopause structure using this 1-D RCE model? To this end, we should first get the observed latitude dependence of lapse rate, optical depth, and outgoing longwave radiation, and then plug them into this simple model. We will go through those three parameters as inputs one by one, and only the zonal-mean annual-mean results will be discussed.

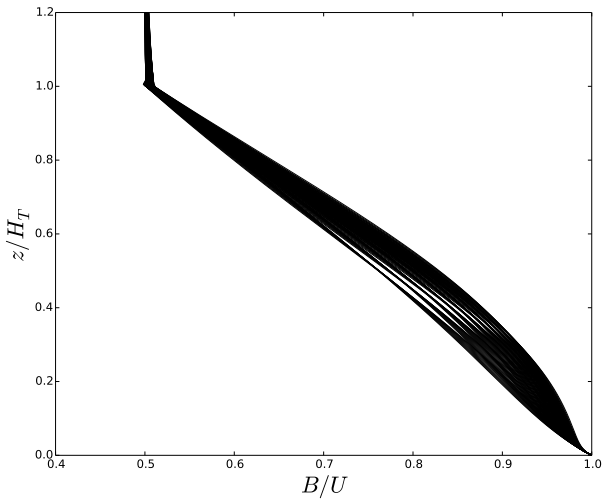


Figure 4: Validation of the assumption $B/U = 1 - z/2H_T$ for different combinations of surface optical depth and tropospheric lapse rate.

Tropospheric lapse rate decreases from $6.5K/km$ in the tropics to $4K/km$ in the polar regions [7, 8]. In the tropics, deep moist convection makes the lapse rate very close to the moist adiabatic lapse rate. However, in the polar regions, it is noteworthy that the lapse rate is not controlled by dry adiabatic lapse rate (about $9.8K/km$); instead, the value of $4K/km$ is determined by the critical lapse rate that satisfies the condition of baroclinic adjustments [7]. Surprisingly, the structure of lapse rate leads to an increase in the tropopause height from the tropics to polar regions by over 3km, which is opposite to the observed structure (blue line in Fig. 6).

As water vapor concentration decreases dramatically from tropics to polar regions, surface optical depth also decreases significantly. In this model, we are using a grey atmosphere; therefore, getting the observed equivalent surface optical depth is not very straightforward. An alternative way is to use the optical depth structure used in idealized GCMs. Following previous studies [9, 10], we will let the surface optical depth change sinusoidally from 7.2 in tropics to 1.8 in polar regions and set the parameter f_{H_2O} to 0.8. As a result, tropopause height decreases from tropics to polar regions by about 3km, which almost totally compensates the effect of optical depth (magenta line in Fig. 6).

Concerning the outgoing longwave radiation, it tends to yield a higher tropopause in the tropics, but as we mentioned above, its impact is very limited (red line in Fig. 6).

Combining all three effects, we will get an almost uniform tropopause height, about 10.5km, across different latitudes (black solid line in Fig. 6). Its pattern indeed mimics the observed pattern: higher in the tropics and lower in the polar regions with sharp changes occurring in the mid-latitudes. However, the tropical tropopause is only about 0.5km higher than the polar tropopause, which is much smaller than the observed value 8km. What might be missing?

Brewer-Dobson circulation in the stratosphere can cause the adiabatic heating (or cool-

ing) locally near the tropopause, therefore it can potentially change the height of tropopause. The heating rate is essentially determined by vertical velocity and thermal stratification,

$$\frac{\partial \theta}{\partial t} = -w\frac{\partial \theta}{\partial z} = Q_s. \tag{2.11}$$

From observations, the vertical gradient of potential temperature is positive all around the globe. So the ascents (descents) in the tropical (polar) stratosphere will induce a local adiabatic cooling (heating). Rough estimates from observations $w \approx 0.3mm/s, \partial\theta/\partial z \approx 40K/km$, give us a heating rate of $Q_s \approx 1K/day$. Due to this additional adiabatic heating, we need to rewrite the radiative equilibrium condition Eq. (2.4) as

$$\frac{\partial I}{\partial z} = \rho c_p Q_s, \tag{2.12}$$

where ρ is the density of air and c_p is the heat capacity of air.

After imposing the heating structure shown in the lower-right panel of Fig. 5, we can get a structure of tropopause height (thick green line in Fig. 6) very like observations (black solid line in Fig. 1a). The adiabatic cooling significantly raises the tropical tropopause by 3km, and the adiabatic heating also lowers the polar tropopause by 1km. Eventually we get a tropopause with very similar structure as observations in both the general pattern and the equator-pole contrast.

Brewer-Dobson circulation not only raises the tropical tropopause, but also makes it colder and sharper, which otherwise cannot be explained by any of the three parameters (Fig. 7a,b). Comparing our results (Fig. 7b) and the observations (Fig. 1b), we find that our model, in spite of its simplicity, well captures all three main features of tropical tropopause: higher, colder, and sharper than the extratropical tropopause. By adding ozone heating in the upper troposphere, we can get an even more realistic temperature profile (Fig. 7c) compared to observations (Fig. 1b).

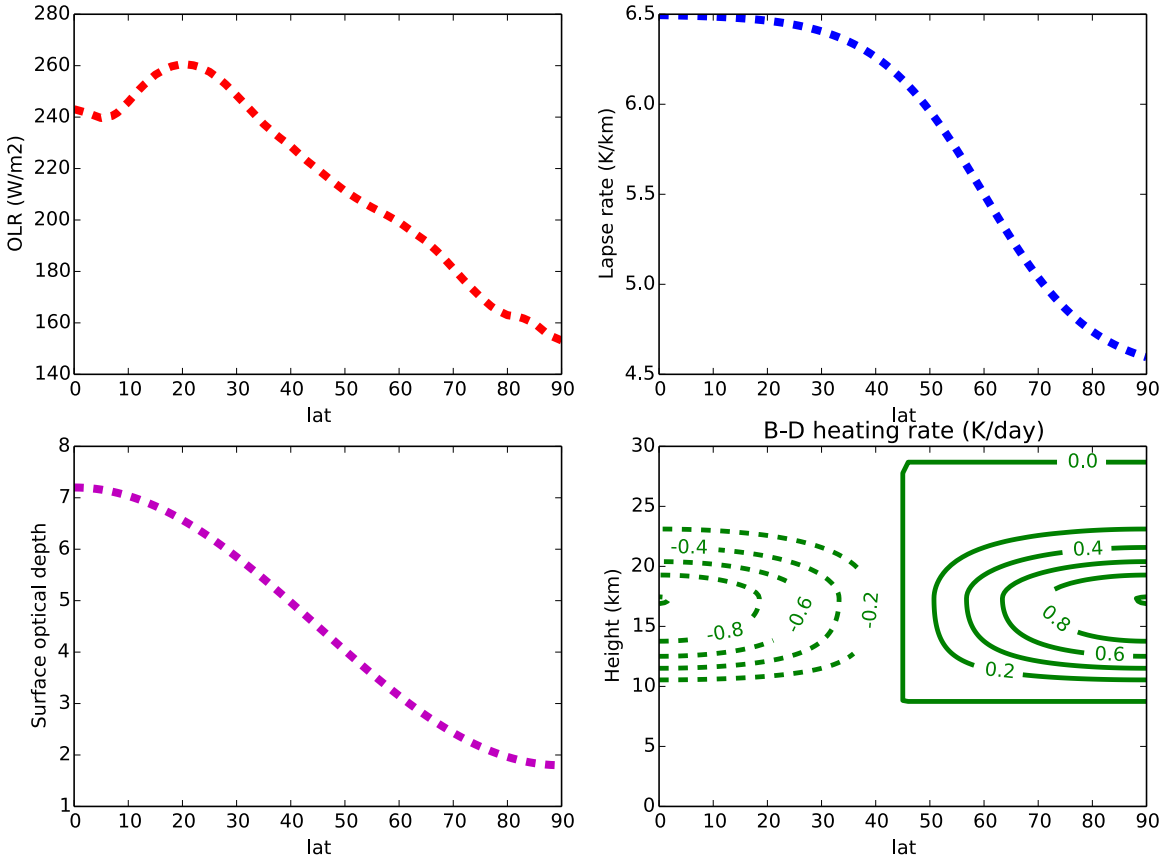


Figure 5: Meridional structure of outgoing longwave radiation, tropospheric lapse rate, surface optical depth, and adiabatic heating rate by Brewer-Dobson circulation, which have been used as inputs in the 1-D tropopause model. OLR is computed from NOAA OLR dataset. Lapse rate is a best-fit tangent curve to observed values. Surface optical depth is following the previous study with idealized GCMs. Adiabatic heating rates are artificially imposed just above the tropopause.

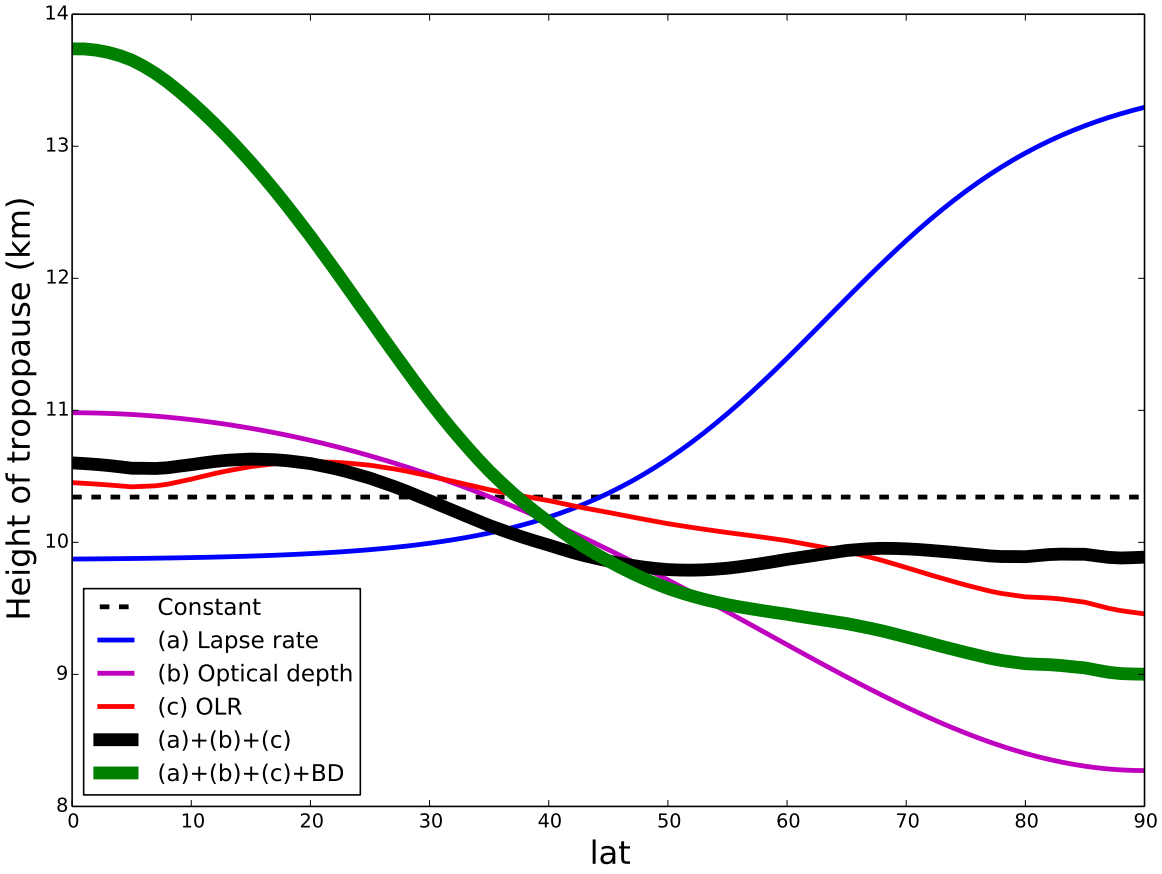


Figure 6: Meridional structure of tropopause height from the numerical solutions of 1-D tropopause model. The black dashed line is with all parameters constant. Three thin color lines are with only lapse rate (blue), optical depth (magenta), or OLR (red) changing with latitude. The thick black solid line is with all three parameters changing with latitude. The thick green solid line is with all three parameters changing with latitude and also an imposed adiabatic heating by Brewer-Dobson circulation.

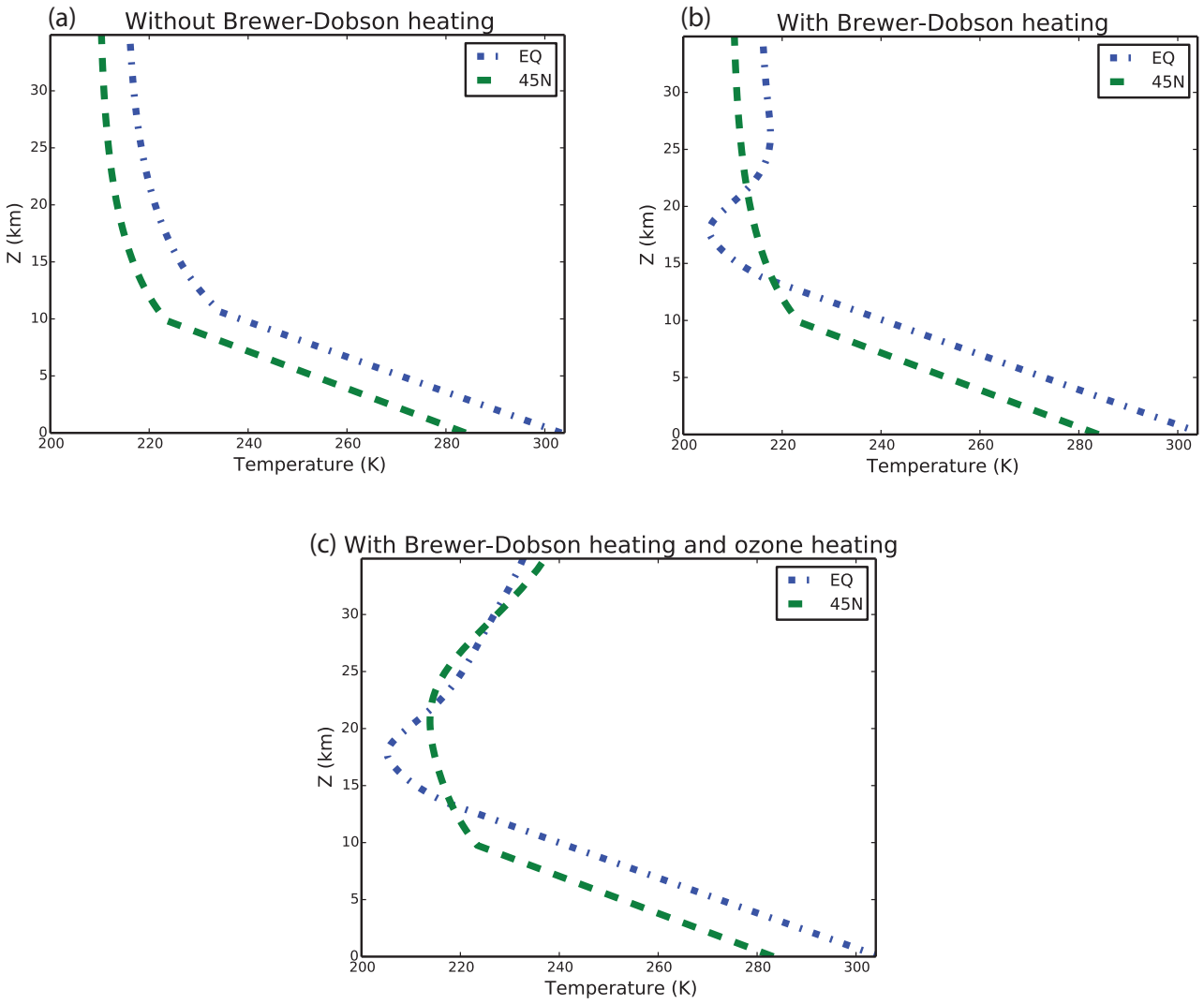


Figure 7: The contrast of temperature profiles at equator and 45^oN. (a) With all three parameters changing with latitude, (b) adding adiabatic heating (Brewer-Dobson circulation) onto case a, and (c) adding both adiabatic heating (Brewer-Dobson circulation) and diabetic heating (ozone) onto case a.

3 A two-layer RCE model

3.1 Background

In the tropopause model we just talked about, radiation and convection are two separate processes. Radiation accounts for the energy balance of the system, while convection acts to set the tropospheric lapse rate. In reality, the two processes are highly coupled. There are quite a few RCE studies with highly coupled models, such as radiative-convective models (RCMs) and cloud resolving models (CRMs), but they are too complicated to help us understand the principle physics of RCE states. For example, how does relative humidity change with climates in RCE states? It involves too many elements and layers. We will develop a maximally simplified radiative-convective equilibrium (RCE) model but still with an interactive hydrological cycle, and our main goal is to answer some basic questions related to RCE like the one about relative humidity we just mentioned.

3.2 Model set-up

Following a recent study [11], the RCE model has two equal-mass layers for troposphere, and its hydrological cycle is interactive. Here we only care about the equilibrium state, instead of the stability of the state. Effective emission temperature T_e is specified, and two other parameters are the precipitation efficiency ϵ_p and the surface wind V. All the other variables, including temperature, humidity, convective mass flux, tropopause height, and heating rate are calculated.

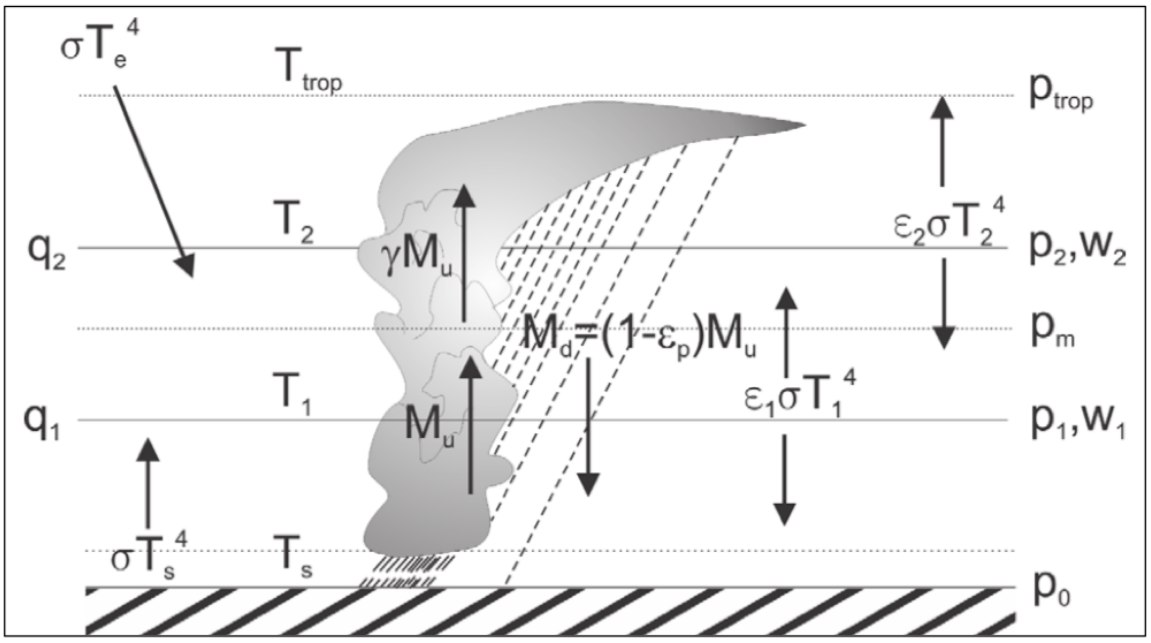


Figure 8: The schematic of the two-layer RCE model.

The basic model setup is shown in the schematic of Fig. 8. Surface pressure p_0 is specified, and the tropopause temperature is assumed to be $2^{-1/4}T_e$ as discussed in the

previous section. In the following discussions, the subscripts '1', '2', 's', 'b', 'm', and 'trop' stand for the lower layer, upper layer, surface, boundary layer, middle of troposphere, and tropopause, respectively. Note that 'middle' indicates the mid-level in terms of mass – in other words, in the pressure coordinate.

Let's start with the energy balance in the two layers. In the lower layer,

$$F_s - M_u(h_b - h_m) + Q_1 = 0, \tag{3.1}$$

where F_s is the surface turbulent enthalpy flux, M_u is the convective mass flux, h_b and h_m is the moist static energy (MSE) in the middle of troposphere and in the subcloud layer, and Q_1 is the radiative cooling. Surface turbulent enthalpy flux is essentially determined by surface wind speed, surface saturation specific humidity, and surface relative humidity,

$$F_s = \tilde{V}(h_s - h_b) \approx \tilde{V} L_v q_s^* (1 - \mathcal{H}_s), \tag{3.2}$$

where $\tilde{V}$ is defined as

$$\tilde{V} = \rho_b C_k V. \tag{3.3}$$

Similarly, in the upper layer,

$$M_u(h_b - h_m) + Q_2 = 0, \tag{3.4}$$

where we assume the MSE at the tropopause is the same as that in the subcloud layer, so there is no MSE exchange at the tropopause.

Energy balance should also hold in the clear air between clouds. In the lower layer, assuming convective downdraft as $(1 - \epsilon_p)M_u$, we can have that the large-scale subsidence in the clear air is $\epsilon_p M_u$. So the energy balance can be written as

$$\epsilon_p M_u \Delta S_1 + Q_1 = 0, \tag{3.5}$$

where we are assuming convection only occupies a very small portion of the domain area, and ΔS_1 is the contrast in dry static energy between the mid-troposphere and the surface,

$$\Delta S_1 = c_p(T_m - T_s) + g z_m. \tag{3.6}$$

Similarly, we can write the energy balance in the clear air of upper layer as,

$$\gamma M_u \Delta S_2 + Q_2 = 0, \tag{3.7}$$

where ΔS_2 is the contrast in dry static energy between the tropopause and the mid-troposphere,

$$\Delta S_2 = c_p(T_{trop} - T_m) + g(z_{trop} - z_m). \tag{3.8}$$

Note that here γ, not like ϵ_p, is calculated instead of specified, so it is not a parameter.

Finally let's consider the energy balance in the subcloud layer and at the surface. In the subcloud layer, surface turbulent enthalpy flux is balanced by the convective flux out of the layer and the flux due to shallow convection; that is,

$$M_u(h_b - h_m) + F_{shallow} = F_s. \tag{3.9}$$

If we parameterize $F_{shallow}$ as $(1-\alpha)F_s$, we can then get

$$M_u(h_b - h_m) = \alpha F_s. \tag{3.10}$$

Note that here α is calculated, not a parameter. The surface energy balance requires that

$$F_s - Q_s = 0. \tag{3.11}$$

The radiative heating rates in the lower layer, the upper layer, and at the surface can be defined as,

$$Q_1 = \sigma\varepsilon_1(T_s^4 - 2T_1^4 + \varepsilon_2 T_2^4), \tag{3.12}$$

$$Q_2 = \sigma\varepsilon_2[(1-\varepsilon_1)T_s^4 - 2T_2^4 + \varepsilon_1 T_1^4], \tag{3.13}$$

$$Q_s = \sigma[T_e^4 + \varepsilon_1 T_1^4 + \varepsilon_2(1-\varepsilon_1)T_2^4 - T_s^4]. \tag{3.14}$$

In addition to those three sets of energy balance equations, we also need to make assumptions of conserved saturation MSE and saturation entropy in the whole column, which are controlled by subcloud properties. Mathematically, it can be expressed as,

$$h_b = h_1^* = h_m^* = h_2^* = h_{trop}^* \tag{3.15}$$

and

$$s_b = s_1^* = s_m^* = s_2^* = s_{trop}^*. \tag{3.16}$$

Since this is a highly coupled and nonlinear system, we need to solve it numerically. Numerical solutions will be discussed in the following subsections.

3.3 State-dependent emissivity

Before we talk about the solutions, one thing we want to emphasize is the state-dependence of layer emissivities. Intuitively, we can imagine that when water vapor increases in a certain layer, more longwave radiation can be absorbed or emitted given a certain temperature; in other words, layer emissivity will increase. But, quantitively determining the relationship between emissivity and water vapor (or temperature) is not a trivial problem. As a first step, we can try to use an artificial exponential function to mimic the Clausius-Clapeyron relation, but the parameters we will choose are quite empirical. We notice that this relation is a key element coupling radiation and convection, so it will largely affect the model behavior. Thus, we decide to use more realistic formulation to get a more reasonable model behavior.

Following a very early study [12], we seek to find a relation between layer emissivity and water vapor path (WVP, sometimes also called total precipitable water, defined as the height of water if all water vapor in a certain layer condenses to liquid water). Carbon dioxide has been taken into account, but its concentration is held as constant around 400ppmv. The overlap of absorption spectrum between water vapor and carbon dioxide has also been corrected. Results are shown in Fig. 9. The WVP of current earth is about 27mm, ranging from 1mm in the polar regions to 42mm in the tropics, so the layer emissivity is roughly in the range of 0.6 to 0.8.

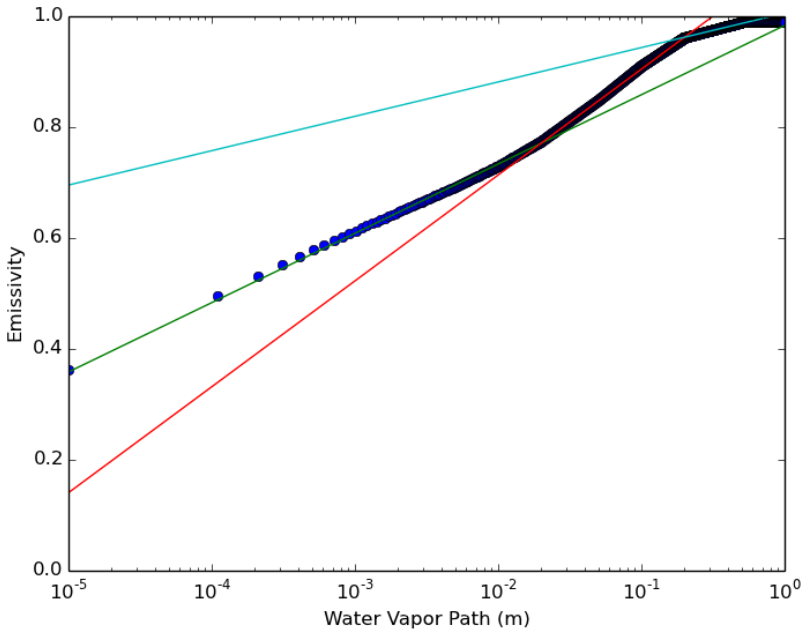


Figure 9: Relation between layer emissivity and water vapor path, with the best-fit lines in different ranges of water vapor path [12].

3.4 General behaviors of the two-layer RCE model

In general, our two-layer model well captures most important features of an RCE system, but here we will just discuss two examples about how the model behaves. The first thing we need to check is the lapse rate. By definition, the lapse rate can be computed as the temperature contrast between the top and the bottom of a layer divided by the layer thickness. Then we can compare this value with the theoretical value of moist adiabatic lapse rate, which can be computed with the average layer temperature. In principle, those two values should be approximately the same, and this feature is well captured by our two-layer model as expected (Fig. 10).

Another example we want to show is how convection strength changes with climates. Recall that Eq. (3.4) basically states that in the second layer, radiative cooling balances the convective heating, which is determined by both convective mass flux and the MSE contrast between the mid-troposphere and the tropopause (note that $h_b = h^*_{trop} = h_{trop}$ because the tropopause is extremely dry). When SST increases, on one hand, radiative cooling will increase because both temperature and emissivity increase; on the other hand, the MSE contrast also increases because the satiation specific humidity increases significantly with temperature following Clausius-Clapeyron relation. In comparison, the MSE contrast increases much faster than the radiative cooling, so the convective mass flux will decrease, as is seen from the full cloud resolving simulations (Fig. 11b). This feature is also captured by our two-layer model (Fig. 11a). Note that here we are focused on RCE states, and things will be different if one considers weak temperature gradient (WTG) simulations.

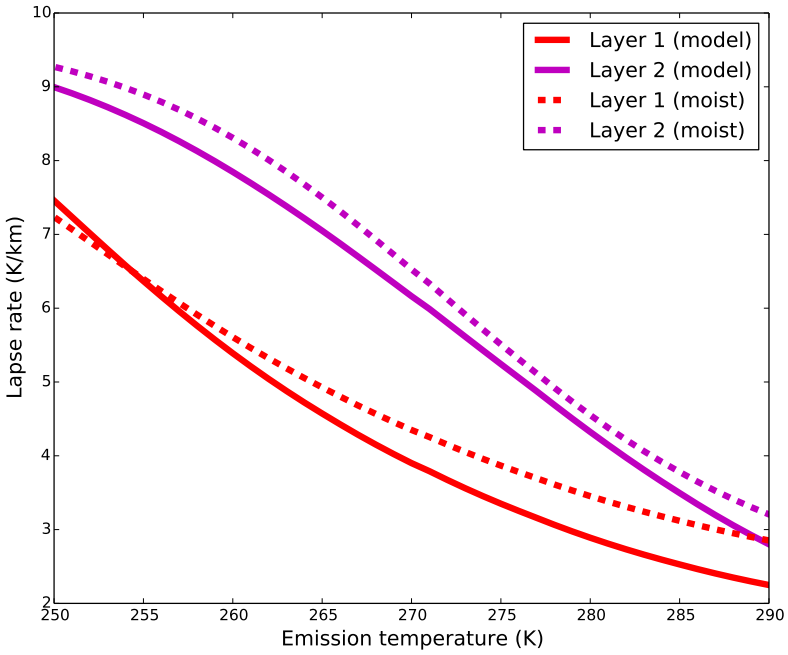


Figure 10: Comparison of the model lapse rate and the theoretical moist adiabatic lapse rate in both layers for different emission temperatures.

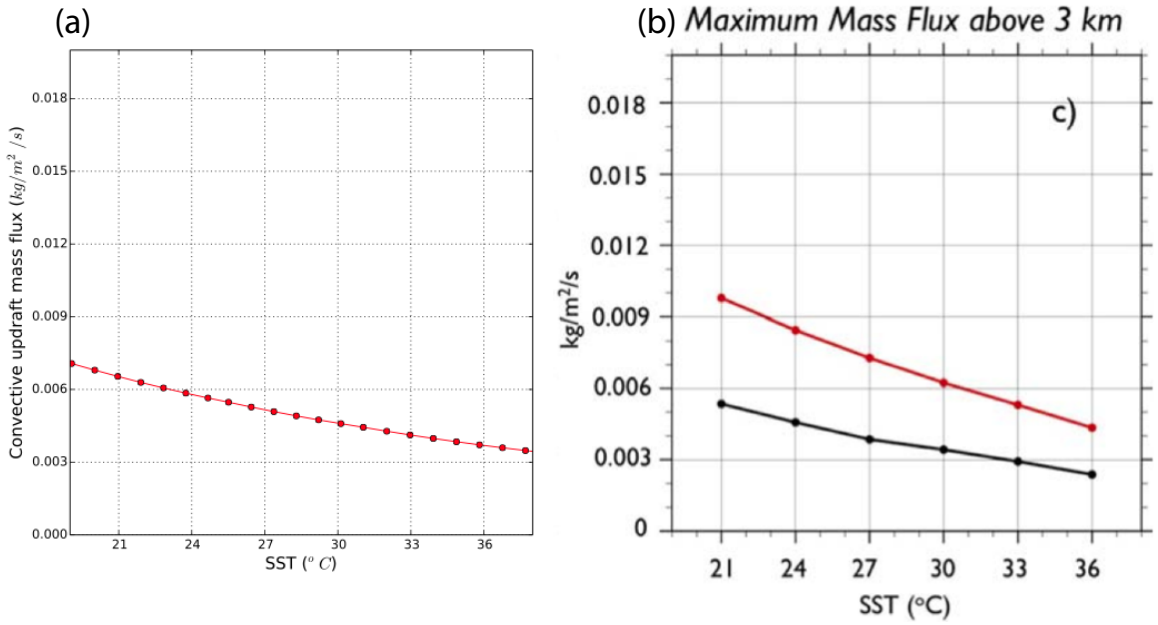


Figure 11: Comparison of (a) the convective mass flux (in the lower layer) from our two-layer model and (b) the maximum convective mass flux (red line) from a cloud resolving model [13].

3.5 How does relative humidity change with climates from an RCE perspective?

How relative humidity changes with climates is a fundamental and important question, but we still lack a complete answer to that. Some studies regard it as a constant variable in climate models, but this assumption still needs to be justified. Among all arguments related to this question, large-scale circulation argument might be known and accepted by most people in the community. It basically states that large-scale circulation transports a saturated air parcel to a warmer place, which makes it under-saturated, and the relative humidity is simply determined by the temperature difference between two places; therefore as long as the large-scale circulation does't change too much, global (or tropical) average relative humidity in the troposphere might not be able to change too much. Here we will focus on RCE states without any large-scale circulation, and our results suggest that even in RCE states, tropospheric relative humidity can not change much and a key element is the radiation-convection coupling.

Seen from our two-layer model results, tropospheric relative humidity hardly changes within a broad range of SST. When SST increases from 10^oC to 40^oC, surface relative humidity slowly increases from 0.7 to 0.9; however, the relative humidity in the free troposphere changes little and sits at a value of 0.5. To better understand this behavior, we will make use of the energy balance equations mentioned above and conduct energetics analysis.

Let's start with the surface relative humidity. From Eq. (3.2), we can easily get

$$\frac{1}{F_s}\frac{\partial F_s}{\partial T_s} = \frac{1}{q_s^*}\frac{\partial q_s^*}{\partial T_s} - \left(\frac{\mathcal{H}_s}{1-\mathcal{H}_s}\right)\frac{1}{\mathcal{H}_s}\frac{\partial \mathcal{H}_s}{\partial T_s}, \tag{3.17}$$

Similarly, from Eq. (3.14), we can get

$$\frac{1}{Q_s}\frac{\partial Q_s}{\partial T_s} = \frac{A+B}{C}, \tag{3.18}$$

where

$$\begin{aligned} A &= 4[T_e^3\frac{\partial T_e}{\partial T_s} + \varepsilon_1 T_1^3\frac{\partial T_1}{\partial T_s} + \varepsilon_2(1-\varepsilon_1)T_2^3\frac{\partial T_2}{\partial T_s} - T_s^3], \\ B &= \frac{\partial \varepsilon_1}{\partial T_s}T_1^4 + [(1-\varepsilon_1)\frac{\partial \varepsilon_2}{\partial T_s} - \varepsilon_2\frac{\partial \varepsilon_1}{\partial T_s}]T_2^4, \\ C &= T_e^4 + \varepsilon_1 T_1^4 + \varepsilon_2(1-\varepsilon_1)T_2^4 - T_s^4. \end{aligned} \tag{3.19}$$

Since layer emissivity is close to 1, the second term B/C is relatively small compared with A/C and can be neglected. Then we can simplify Eq. (3.18) as

$$\frac{1}{Q_s}\frac{\partial Q_s}{\partial T_s} \approx \frac{4}{\bar{T}}\frac{\partial \bar{T}}{\partial T_s} \approx \frac{8}{\bar{T}} \tag{3.20}$$

where $\bar{T}$ is the average tropospheric temperature, and it is roughly $280K$ when surface temperature is $300K$. The second approximation is because the value of $\partial\bar{T}/\partial T_s$ can be assumed to be 2, which indicates that, given a certain increase in surface temperature, tropospheric temperature is amplified by a factor of 2 due to the lapse rate feedback. Then

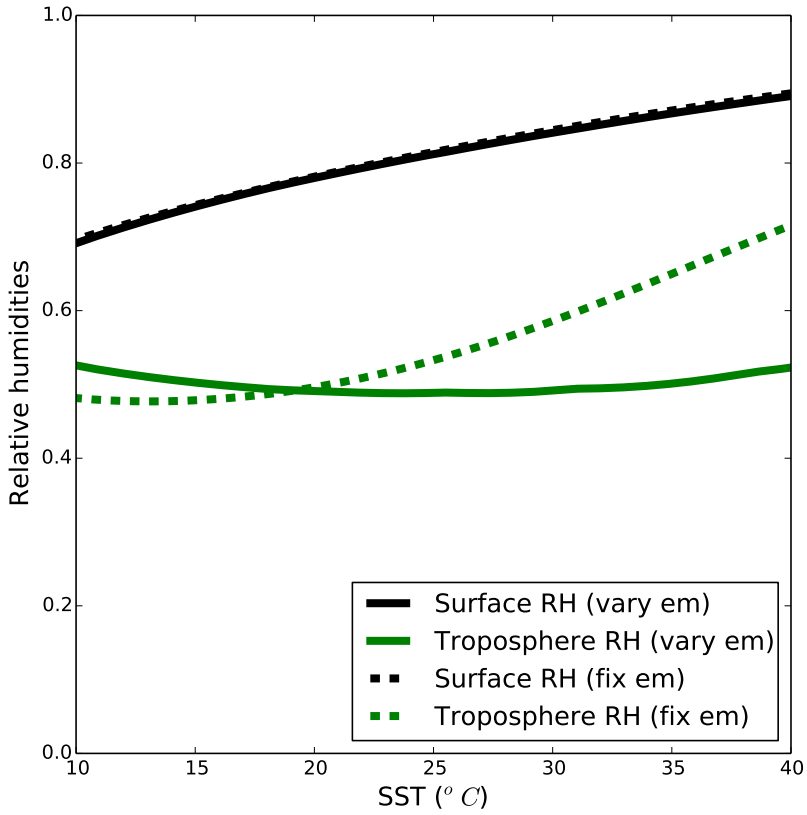


Figure 12: Surface relative humidity (black) and tropospheric relative humidity (green) as a function of sea surface temperature for the cases with varying emissivity (solid) and fixed emissivity (dashed). Note the unchanged tropospheric relative humidity for the case of varying emissivity.

combining Eq. (3.17) and Eq. (3.20), and recalling the surface energy balance Eq. (3.11), we can eventually arrive at,

$$\begin{aligned} \frac{1}{\mathcal{H}_s}\frac{\partial \mathcal{H}_s}{\partial T_s} &\approx \left(\frac{1-\mathcal{H}_s}{\mathcal{H}_s}\right)\left(\frac{1}{q_s^*}\frac{\partial q_s^*}{\partial T_s} - \frac{8}{\bar{T}}\right) \\ &\approx \left(\frac{0.18}{0.82}\right)\left(\frac{6.2\%}{1K} - \frac{8}{280K}\right) \\ &\approx 0.7\%/K, \end{aligned} \tag{3.21}$$

which is a good estimate to the black solid line in Fig. 12.

With a similar method, we can get how tropospheric relative humidity changes with SST. Combining Eqs. (3.4)(3.5), we can get

$$h_b - h_m = \epsilon_p \Delta S_1 Q_2 / Q_1, \tag{3.22}$$

and thus

$$\frac{1}{h_b - h_m}\frac{\partial(h_b - h_m)}{\partial T_s} = \frac{1}{Q_2}\frac{\partial Q_2}{\partial T_s} - \frac{1}{Q_1}\frac{\partial Q_1}{\partial T_s} + \frac{1}{\Delta S_1}\frac{\partial \Delta S_1}{\partial T_s}. \tag{3.23}$$

Then we rewrite $(h_b - h_m)$ as

$$h_b - h_m = h_m^* - h_m = L_v q_m^*(1 - \mathcal{H}_m). \tag{3.24}$$

Also, we rewrite ΔS_1 as

$$\Delta S_1 = (h_m^* - L_v q_m^*) - (h_b - L_v q_s) = L_v(q_s^* \mathcal{H}_s - q_m^*). \tag{3.25}$$

Substituting Eqs.(3.24)(3.25) into Eq.(3.23), we get

$$\frac{1}{\mathcal{H}_m}\frac{\partial \mathcal{H}_m}{\partial T_s} = \left(\frac{1 - \mathcal{H}_m}{\mathcal{H}_m}\right)\left[\frac{1}{q_m^*}\frac{\partial q_m^*}{\partial T_s} - \frac{1}{q_s^*\mathcal{H}_s - q_m^*}\frac{\partial(q_s^*\mathcal{H}_s - q_m^*)}{\partial T_s} + \frac{1}{Q_1}\frac{\partial Q_1}{\partial T_s} - \frac{1}{Q_2}\frac{\partial Q_2}{\partial T_s}\right] \tag{3.26}$$

When layer emissivity varies with SST (or WVP), the three terms in Eq. (3.26) are

$$\frac{1}{q_m^*}\frac{\partial q_m^*}{\partial T_s} \approx 9.4\%/K, \tag{3.27}$$

$$-\frac{1}{q_s^*\mathcal{H}_s - q_m^*}\frac{\partial(q_s^*\mathcal{H}_s - q_m^*)}{\partial T_s} \approx -4.4\%/K, \tag{3.28}$$

$$\frac{1}{Q_1}\frac{\partial Q_1}{\partial T_s} - \frac{1}{Q_2}\frac{\partial Q_2}{\partial T_s} \approx -5.0\%/K, \tag{3.29}$$

which are the values when SST is 300K. Substituting Eqs.(3.27)(3.28)(3.29) into Eq.(3.26), we get

$$\frac{1}{\mathcal{H}_m}\frac{\partial \mathcal{H}_m}{\partial T_s} \approx 0.0\%/K. \tag{3.30}$$

That is why tropospheric relative humidity changes little with a broad range of climate with SST varying from 10^oC to 40^oC.

Since this RCE system is highly coupled, it is hard to say which single element in this coupled system causes the almost constant tropospheric relative humidity. Actually, it is due to the radiation-convection coupling itself, in other words, the feedback between radiation and convection. To test this hypothesis, we can fix the layer emissivity to a constant value, which kills the coupling between convection and radiation, and see how the relative humidity changes with SST.

Interestingly, we find that the case with fixed emissivity experiences a significant increase in the tropospheric relative humidity with SST. This can be partly explained by the same energetics analysis we conducted above, and the new values of the three terms in Eq. (3.26) are

$$\frac{1}{q_m^*}\frac{\partial q_m^*}{\partial T_s} \approx 9.4\%/K, \tag{3.31}$$

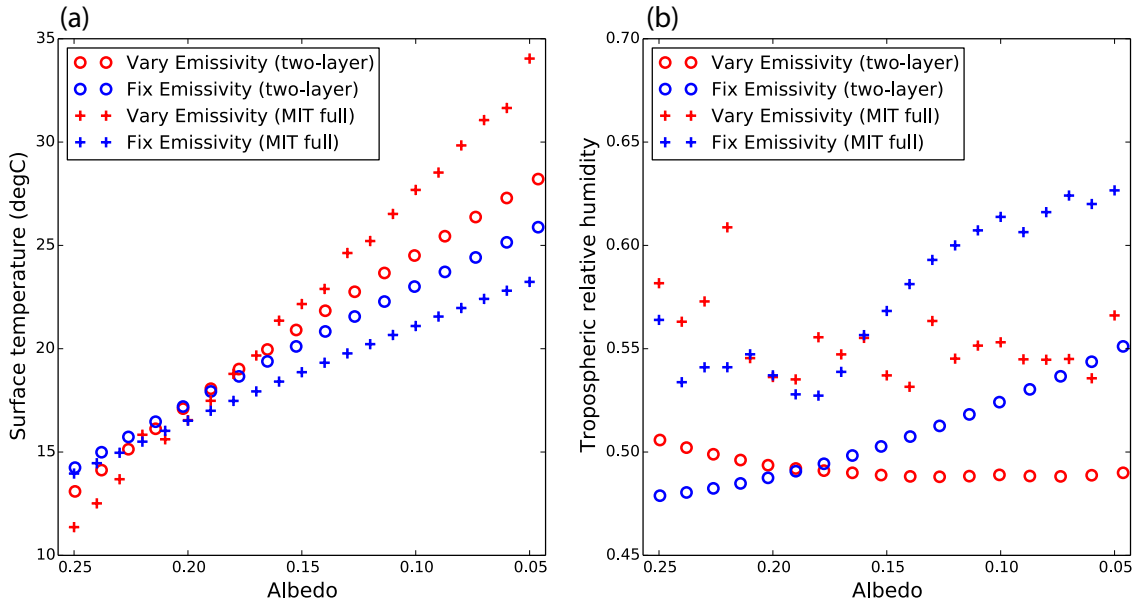


Figure 13: Comparison of (a) surface temperature and (b) tropospheric relative humidity between our two-layer model (circle) and the MIT full radiative-convective model (cross). Horizontal axis has been changed to albedo for the ease of comparison.

$$-\frac{1}{q_s^*\mathcal{H}_s - q_m^*}\frac{\partial(q_s^*\mathcal{H}_s - q_m^*)}{\partial T_s} \approx -4.1\%/K, \tag{3.32}$$

$$\frac{1}{Q_1}\frac{\partial Q_1}{\partial T_s} - \frac{1}{Q_2}\frac{\partial Q_2}{\partial T_s} \approx -2.9\%/K. \tag{3.33}$$

Again, they are the values when SST is 300K. Substituting Eqs.(3.31)(3.32)(3.33) into Eq.(3.26) and taking $\mathcal{H}_m$ as 0.55, we get

$$\frac{1}{\mathcal{H}_m}\frac{\partial \mathcal{H}_m}{\partial T_s} \approx 2.0\%/K, \tag{3.34}$$

which is a good estimate for the green dashed line in Fig. 12.

One might have noticed that fixing the layer emissivity will not change the behavior of surface relative humidity too much (black dashed line in Fig. 12). That is because the state dependency of emissivity only enters the term B in Eq. (3.19), and as we have mentioned, B is a small term and it will not change the results too much in two cases.

To further test our hypothesis about the stabilizing effect of radiation-convection coupling on the tropospheric relative humidity, we have also used a full radiative-convective model (MIT column RCM) and the results are shown in Fig. 13. In the MIT full RCM, we can modify surface albedo to change the external forcing, so for the ease of comparison, we use surface albedo as horizontal axes in Fig. 13. For the cases with varying emissivity (red), both models exhibit an almost unchanged relative humidity; for the cases with fixed emissivity (blue), both models exhibit increasing relative humidity when the system gets

warmer. Those trends agree especially well when albedo is in the range of 0.05 to 0.20, and surface temperature is in the range of 15^oC to 30^oC, which is exactly the temperature range of current tropics.

4 Conclusions and discussions

This study develops two idealized models with different simplifications from the full radiative-convective equilibrium system. The first model is a one-dimensional tropopause model with multiple layers but less coupling. We find that this 1-D model, even without any explicit dynamics, can well reproduce the observed meridional structure of tropopause height, and can also explain why the tropical tropopause is higher, colder, and sharper than the extratropical tropopause. Brewer-Dobson circulation is a key element in explaining the observed features of the tropopause. We will use an idealized GCM to further test this hypothesis.

The second model is a two-layer RCE model with interactive hydrological cycle. This extremely simplified model captures the main features of RCE states very well, and its simplicity helps us better understand the basic physics of RCE. With this model, we find that tropospheric relative humidity hardly changes climate in a broad range of SST from 15^oC to 30^oC, which is exactly the temperature range of current tropics. We want to emphasize that it is an alternative argument, aside from the conventional argument with large-scale circulation, on explaining the unchanged relative humidity with climates people usually assume. The key element of this argument is the coupling between radiation and convection. Once we kill this coupling or use an unrealistic relation between emissivity and water vapor path, we might fail to observe this phenomenon.

For future investigation, we are hoping to combine the two simplified models to better understand how the height of tropopause changes with climate (temperature, humidity, etc.), and going further, large-scale circulation such as lateral transport and mid-latitude eddies will be incorporated in the model as well to form a full picture.

Acknowledgements

I would like to thank Geoff Vallis and Kerry Emanuel for suggesting this wonderful project, for their invaluable guidance, and continuous supervision throughout the summer. I also would like to thank Andy Ingersoll for a lot of insightful discussions. Thanks also to the directors for organizing a smooth program. Last, but certainly not least, I would like to thank all Fellows for a memorable summer.

References

[1] G Vaughan and JD Price. On the relation between total ozone and meteorology. *Quarterly Journal of the Royal Meteorological Society*, 117(502):1281–1298, 1991.

[2] James R Holton, Peter H Haynes, Michael E McIntyre, Anne R Douglass, Richard B Rood, and Leonhard Pfister. Stratosphere-troposphere exchange. *Reviews of Geophysics*, 33(4):403–439, 1995.

[3] LJ Wilcox, BJ Hoskins, and KP Shine. A global blended tropopause based on era data. part i: Climatology. *Quarterly Journal of the Royal Meteorological Society*, 138(664):561–575, 2012.

[4] Geoffrey K Vallis. *Atmospheric and oceanic fluid dynamics: fundamentals and large-scale circulation*. Cambridge University Press, 2006.

[5] Issac M Held. On the height of the tropopause and the static stability of the troposphere. *Journal of the Atmospheric Sciences*, 39(2):412–417, 1982.

[6] Pablo Zurita-Gotor and Geoffrey K Vallis. Determination of extratropical tropopause height in an idealized gray radiation model. *Journal of the Atmospheric Sciences*, 70(7):2272–2292, 2013.

[7] Peter H Stone and John H Carlson. Atmospheric lapse rate regimes and their parameterization. *Journal of the Atmospheric Sciences*, 36(3):415–423, 1979.

[8] II Mokhov and MG Akperov. Tropospheric lapse rate and its relation to surface temperature from reanalysis data. *Izvestiya, Atmospheric and Oceanic Physics*, 42(4):430–438, 2006.

[9] Dargan MW Frierson, Isaac M Held, and Pablo Zurita-Gotor. A gray-radiation aquaplanet moist gcm. part i: Static stability and eddy scale. *Journal of the atmospheric sciences*, 63(10):2548–2566, 2006.

[10] Paul A O'Gorman and Tapio Schneider. The hydrological cycle over a wide range of climates simulated with an idealized gcm. *Journal of Climate*, 21(15):3815–3832, 2008.

[11] Kerry Emanuel, Allison A Wing, and Emmanuel M Vincent. Radiative-convective instability. *Journal of Advances in Modeling Earth Systems*, 6(1):75–90, 2014.

[12] DO Staley and GM Jurica. Flux emissivity tables for water vapor, carbon dioxide and ozone. *Journal of Applied Meteorology*, 9(3):365–372, 1970.

[13] Marat Khairoutdinov and Kerry Emanuel. Rotating radiative-convective equilibrium simulated by a cloud-resolving model. *Journal of Advances in Modeling Earth Systems*, 5(4):816–825, 2013.

An Experimental Investigation of the Rossby Two-slit Problem

Alexis Kaminski

August 2014

1 Introduction

Rossby waves arise in the oceans as a response to forcing, either by buoyancy or by the actions of winds at the sea surface [3]. They have been observed in satellite altimeter data (for example, Kelly and Thompson (2002)[1] show excellent agreement between observed sea surface height and a Rossby wave model) as well as in moorings.

The case of Rossby waves in closed basins (i.e. Rossby basin modes) is a classical problem in geophysical fluid dynamics [2, 3, 4]. There is observational evidence from moorings for these modes; for example, Warren et al. (2002)[9] observed a large signal in moorings in the Mascarene Basin (off the coast of Madagascar) which they attributed to a barotropic Rossby mode.

When examining ocean circulation in basins with incomplete barriers, Pedlosky et al. (1997)[7] and Pedlosky and Spall (1999)[8] found that barriers extending through most of the ocean basin were surprisingly inefficient at blocking the transmission of Rossby wave energy from one subbasin to the next. Pedlosky (2000)[5] developed the linear theory further for the case of a long thin island extending nearly the entire meridional length of the basin, with only small gaps between the north and south ends of the island and the basin boundary, and found that for certain forcing symmetries, waves forced in the eastern subbasin were able to easily slip around the island into the western subbasin. Following this, Pedlosky (2001)[6] looked at the amplitude of reflected and transmitted waves through a barrier with two or three small gaps. However, the theory derived in Pedlosky (2000)[5] and Pedlosky (2001)[6] neglects nonlinear effects and friction in the main basin interiors, and it is unclear what effect these neglected processes will have. As such, investigation of this problem in a laboratory setting might be able to shed some light as to how well the linear theory captures the physics of Rossby waves impinging on a barrier with small gaps, such as ocean ridges or island chains with small gaps between neighbouring islands.

The remainder of this report will proceed as follows. In section 2, the details of the linear theory for the problem of Rossby modes interacting with a barrier with two gaps in it will be outlined, as well as giving a scaling for the likely impact of nonlinear effects. In section 3, the laboratory setup and the associated troubleshooting of the apparatus will be described. Then, in section 4, the measured flows will be described and compared with the linear theory. The quantitative details of the integral constraint derived in the linear theory will be examined using the laboratory data, and some effects of nonlinearity will be examined. We find that while the linear theory captures the large-scale structures of the flow as the Rossby modes encounter the barrier, viscosity and nonlinearity appear to significantly affect

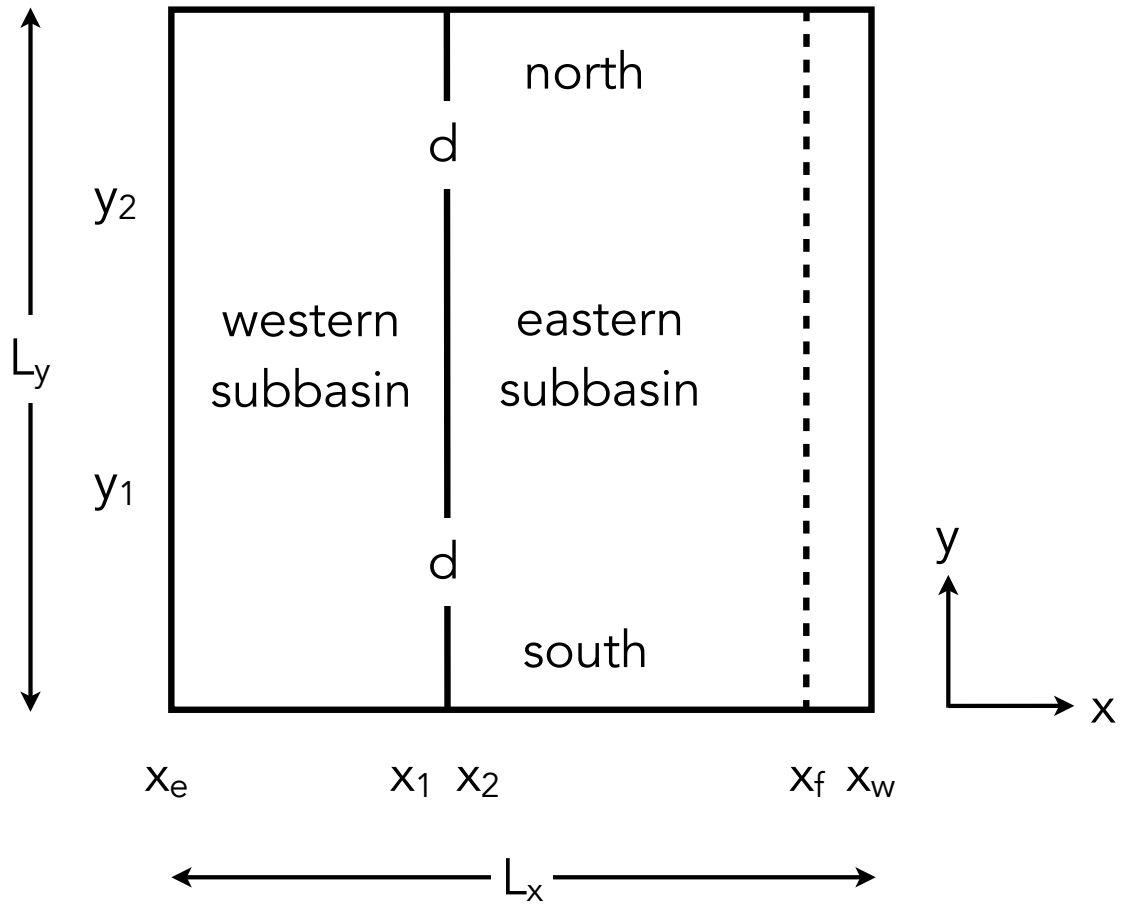


Figure 1: Geometry of basin, barrier, and forcing.

the flow along the boundaries and the barrier through strong boundary currents, and in the gaps through the formation of vortices. Additionally, there is no indication of resonance in the experimental results, despite the predicted resonance at the basin and subbasin normal mode frequencies in the linear theory, indicating additional missing physics. Finally, in section 5, future directions for the problem based on the laboratory results observed here will be discussed.

2 Linear theory

The geometry of the system being considered, shown in figure 1, is similar to that of Pedlosky (2001)[6]: a barrier from x_1 to x_2 extends the length of a square basin in the meridional direction with only two small gaps of length $d \ll L$, and is symmetric with respect to the y-coordinate. The flow is forced at the location $x = x_f$, similar to the forcing in Pedlosky (2000)[5].

As in Pedlosky (2000)[5], the governing equation for the flow is the linearized quasi-geostrophic potential vorticity equation for the β-plane. The fluid is homogeneous and the

flow is barotropic. In nondimensional form, with lengths scaled by the basin length L and times scaled by the characteristic Rossby wave period $(\beta L)^{-1}$, the equation of motion is

$$\nabla^2\Psi_t + \Psi_x = -r\nabla^2\Psi + A\nabla^4\Psi + W(x,y,t)\,. \tag{1}$$

The first term on the right-hand side of (1), proportional to r, is a linear drag term representing the effect of bottom friction, and may be thought of as a ratio between Stommel's boundary layer thickness to L. The second term, proportional to the nondimensional viscosity A, represents the effect of lateral friction. It can be thought of as a cubed ratio of the Munk scale δ_M to L [5]. Both A and r are assumed to be small parameters. Finally, the third term is a forcing term.

As in Pedlosky (2000)[5], we assume that the forcing is harmonic with frequency ω_0,

$$W = \mathrm{Re}[\mathrm{e}^{i\omega_0 t} w(x,y)]\,. \tag{2}$$

We then search for solutions $\Psi(x,y,t)$ of the form

$$\Psi = \mathrm{Re}[\mathrm{e}^{i\omega_0 t}\psi(x,y)]\,. \tag{3}$$

Substituting the above expressions for W and Ψ into (1) thus leads to the following partial differential equation for the spatial structure of the stream function $\psi(x,y)$:

$$i\omega\nabla^2\psi + \psi_x = A\nabla^4\psi + w(x,y) \tag{4}$$

where $\omega = \omega_0 - ir$.

On the basin boundaries and the peninsulas, we set $\psi = 0$, while the island is assumed to have constant value Ψ_I, to be determined.

The vertical structure of the forcing may be represented as a Fourier sine series

$$w = \sum_{n=1} w_n(x)\sin n\pi y\,. \tag{5}$$

Additionally, if the forcing is localized in x, for instance at some location $x = x_f$, then the x-dependent coefficients $w_n(x)$ may be represented using a Dirac delta as

$$w_n = W_n\delta(x - x_f), \qquad x_2 < x_f < x\,. \tag{6}$$

2.1 Gaps

In the gaps, the characteristic length scales of the flow are assumed to be much smaller in the y-direction than in the x-direction. Correspondingly, x-derivatives are assumed to be negligible when compared with y-derivatives, and (4) becomes

$$i\omega\psi_{yy} - A\psi_{yyyy} = 0 \tag{7}$$

in the gaps.

Defining $y' = y_2 + d - y$ in the northern gap and $y' = y - y_1 + d$ in the southern gap, the solution to (7) is

$$\psi = A_1 + B_1\frac{y'}{d} + C_1\exp\left(\frac{(1+i)y'}{\delta}\right) + C_2\exp\left(\frac{-(1+i)y'}{\delta}\right)\,, \tag{8}$$

where $\delta = \sqrt{2A/\omega}$. The corresponding boundary conditions are that the streamfunction ψ be continuous across the gaps, i.e.

$$\begin{aligned} \psi = 0, \quad & y' = 0\,, \\ \psi = \Psi_I, \quad & y' = d\,, \\ \psi_y = 0, \quad & y' = 0, d\,. \end{aligned} \tag{9}$$

Applying (9) allows for the determination of the coefficients in (8) as

$$A_1 = -\Psi_I \frac{\rho(1-q)}{(1+i)[1+q-\rho(1-i)(1-q)]} \tag{10}$$

$$B_1 = \Psi_I \frac{(1+q)}{[1+q-\rho(1-i)(1-q)]} \tag{11}$$

$$C_1 = -\Psi_I \frac{q\rho}{(1+i)[1+q-\rho(1-i)(1-q)]} \tag{12}$$

$$C_2 = \Psi_I \frac{\rho}{(1+i)[1+q-\rho(1-i)(1-q)]} \tag{13}$$

with $\rho = \delta/d$ and $q = \exp(-(1+i)/\rho)$. The values of A_1, B_1, C_2, C_2 are proportional to Ψ_I which is yet to be determined.

Thus, the streamfunction along the longitudes of the barrier is

$$\psi = \begin{cases} 0, & 0 < y < y_1 - d \\ \psi_{gap}, & y_1 - d < y < y_1 \\ \Psi_I, & y_1 < y < y_2 \\ \psi_{gap}, & y_2 < y < y_2 + d \\ 0, & y_2 + d < y < 1 \end{cases} \tag{14}$$

which may be written as $\psi = \Psi_I g(y)$.

2.2 Basins

In the basin interior, lateral friction is neglected and (4) becomes

$$i\omega\nabla^2\psi + \psi_x = w(x,y)\,. \tag{15}$$

The solution may be represented as

$$\psi = \mathrm{e}^{ikx} \sum_{n=1} \phi_n(x) \sin n\pi y\,, \tag{16}$$

so that solving for ψ amounts to solving

$$\frac{\mathrm{d}^2\phi_n}{\mathrm{d}x^2} + a_n^2\phi_n = \frac{w_n(x)}{i\omega}\mathrm{e}^{-ikx}\,, \tag{17}$$

in which $a_n^2 = k^2 - n^2\pi^2$ and $k = 1/(2\omega)$, for ϕ_n.

Solving (17) with the boundary conditions for ψ and continuity of the streamfunction at $x = x_f$,

$$\left.\frac{\mathrm{d}\phi_n}{\mathrm{d}x}\right|_{x=x_f+} - \left.\frac{\mathrm{d}\phi_n}{\mathrm{d}x}\right|_{x=x_f-} = -\frac{i}{\omega} W_n \mathrm{e}^{-ikx_f} \tag{18}$$

gives

$$\phi_n = \begin{cases} A_n^+ \frac{\sin a_n(x-x_e)}{\sin a_n(x_2-x_e)}\,, & x_f \le x \le x_e \\ A_n^- \frac{\sin a_n(x-x_e)}{\sin a_n(x_2-x_e)} + B_n \frac{\cos a_n(x-x_e)}{\sin a_n(x_2-x_e)}\,, & x_2 \le x \le x_f \\ D_n \frac{\sin a_n(x-x_w)}{\sin a_n(x_1-x_w)}\,, & x_w \le x \le x_1 \end{cases} \tag{19}$$

in which

$$B_n = \frac{W_n \mathrm{e}^{-ikx_f}}{i\omega a_n} \sin a_n(x_2 - x_e) \sin a_n(x_f - x_e) \tag{20}$$

$$A_n^- = \Psi_I g_n \mathrm{e}^{-ikx_2} - \frac{W_n \mathrm{e}^{-ikx_f}}{i\omega a_n} \cos a_n(x_2 - x_e) \sin a_n(x_f - x_e) \tag{21}$$

$$A_n^+ = A_n^- + B_n \cot a_n(x_f - x_e) \tag{22}$$

$$D_n = \Psi_I g_n \mathrm{e}^{-ikx_1} \tag{23}$$

In the above, g_n refers to the Fourier sine transform of $g(y)$,

$$g_n = 2\int_0^y g(y) \sin n\pi y \,\mathrm{d}y\,. \tag{24}$$

2.3 Integral constraint

To determine the value of Ψ_I, we may apply Kelvin's circulation theorem by integrating around a contour, C_I, bordering the island. This gives

$$\oint_{C_I} \mathbf{u} \cdot \,\mathrm{d}\mathbf{s} = 0 \tag{25}$$

for time-periodic motion with frequency ω. This may be expressed as

$$i\omega \oint_{C_I} \nabla\psi \cdot \mathrm{n}\,\mathrm{d}l - A \oint_{C_I} \nabla\nabla^2\psi \cdot \mathbf{n}\,\mathrm{d}l = 0\,. \tag{26}$$

When the above expressions for ψ are substituted into the above, the resulting algebraic expression is for this integral constraint is

$$\Psi_I \left[\sum_{n=1} \frac{\mu_n g_n a_n \cos n\pi y_1}{n\pi} \frac{\sin a_n[L_x - l_x]}{\sin a_n(x_2 - x_e) \sin a_n(x_1 - x_w)} - \frac{2B_1 l_x}{d}\right] = -i \sum_{n=1} \frac{\mu_n W_n \cos n\pi y_1}{\omega n\pi} \exp ik(x_2 - x_f) \frac{\sin a_n(x_f - x_e)}{\sin a_n(x_2 - x_e)} \tag{27}$$

where $\mu_n = 1 - (-1)^n$. This may then be solved for the island constant Ψ_I.

It should be noted that when n is even, the resulting value of Ψ_I is identically zero. This may be understood by considering that for the symmetric island geometry described

here, the integral around the island may be satisfied on the eastern side of the island alone: the meridional velocities along the eastern side integrate to zero.

However, when n is odd, $\Psi_I \neq 0$ in general. In contrast to the even-n case, the velocities along the eastern side of the island do not integrate to zero. This requires some response on the western side of the island in order to satisfy the integral constraint.

Finally, it should be noted that within the theoretical framework considered here, the result for the integral constraint is also applicable to the nonlinear equations of motion, i.e. integration around the island would still imply some response on the western side of the island to satisfy the integral constraint.

2.4 Scaling for nonlinearity

While the theory presented above is a linearized theory for the problem being considered, in a real fluid it is expected that nonlinear effects may become important in certain flow regimes. By considering the relative importance of the nonlinear terms in the full nonlinear quasigeostrophic potential vorticity equation, we can predict under what circumstances nonlinear effects may become important.

We begin by defining a characteristic forcing velocity $U_e = \omega_0 A_{\text{forcing}}$. Then, the (dimensional) stream function in the basin is expected to scale with this forcing velocity as $\psi_e \sim U_e L$. In addition, x distances are expected to scale roughly with L everywhere. In the basin, y distances scale approximately with L. We can then define a parameter NL_{basin} as a scaling of the relative importance of the nonlinear $J(\psi, \nabla^2\psi)$ terms compared with the linear $\beta\psi_x$ term, i.e.

$$\frac{J(\psi, \nabla^2\psi)}{\beta\psi_x} \sim \frac{U_e}{\beta L^2} = NL_{\text{basin}}\,. \tag{28}$$

However, in the gaps the y distances do not scale with L, but rather with d. A similar parameter NL_{gap} can thus be defined as

$$\frac{J(\psi, \nabla^2\psi)}{\beta\psi_x} \sim \frac{U_e L}{\beta d^3} = NL_{\text{gap}}\,. \tag{29}$$

For gaps which are small compared with the length of the barrier, $d \ll L$, this implies that $NL_{\text{gap}} \gg NL_{\text{basin}}$, i.e. nonlinear effects are expected to be more significant in the gaps than in the basin interior.

3 Laboratory setup

The general experimental setup is shown in figure 2, with the corresponding parameters listed in table 1. The resulting laboratory apparatus is shown in figure 3, and consists of a square tank with a sloping bottom on a rotating table to create a laboratory analogue to the β-effect. A meridional barrier, constructed from $1/8''$ acrylic, is placed in the tank; the geometry of the barrier is symmetric with respect to y. The 45 cm-long forcing paddle is mechanically forced by a scotch yoke-type mechanism which allows for different forcing frequencies ($\mathcal{O}(0.10)$ rad/s) and amplitudes (up to 3 cm). A rigid lid is used in order to reduce surface gravity-capillary modes in the system.

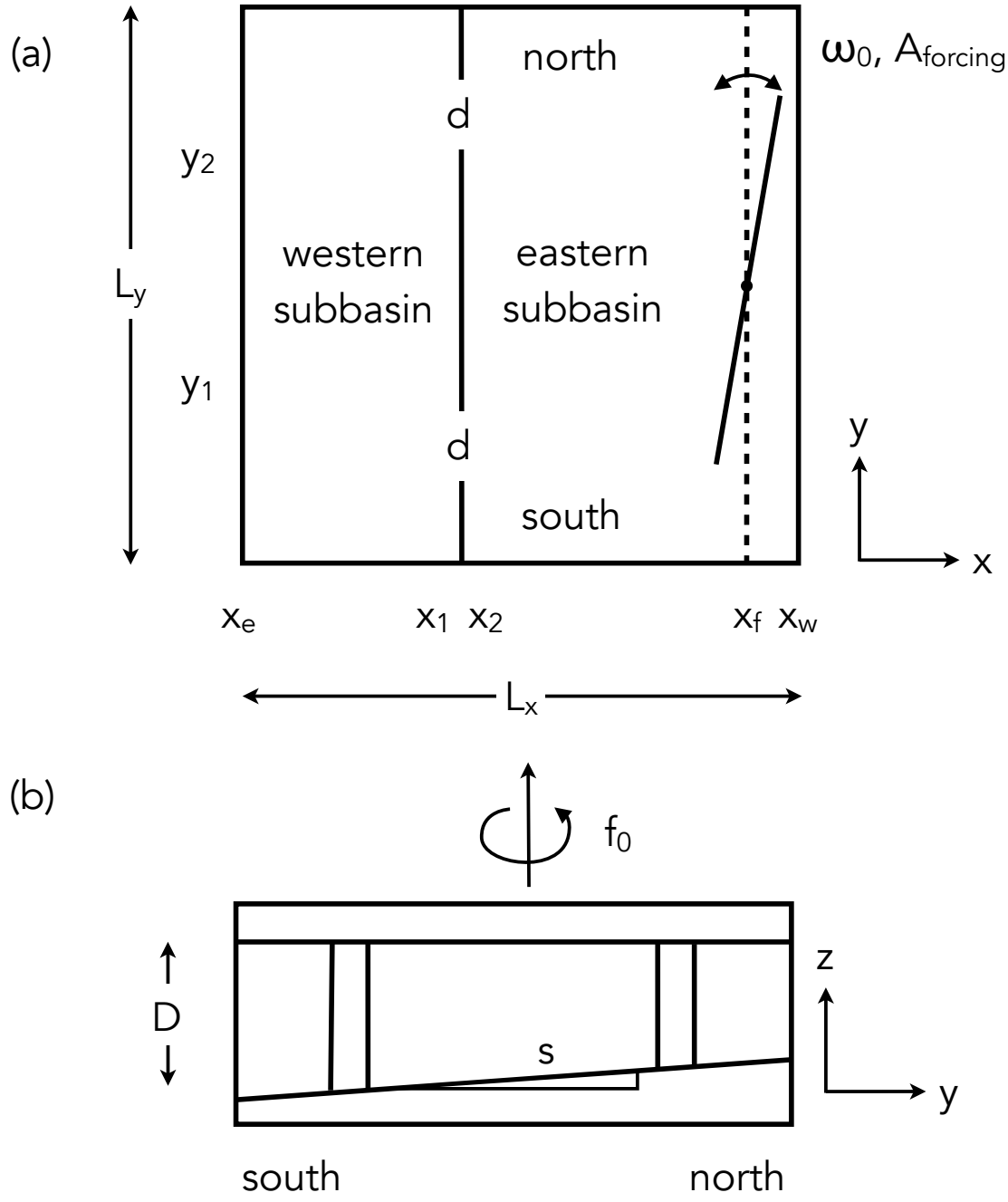


Figure 2: Geometry of basin, barrier, and forcing for laboratory experiments. (a) xy-plane. (b) yz-plane.

Parameter	Value	Parameter	Value
L_x	60 cm	y_1	15 cm
L_y	60 cm	y_2	45 cm
d	4 cm	D	20 cm
x_w	0 cm	s	2/15
x_e	60 cm	A_{forcing}	0.7-3.0 cm
x_1	22 cm	ω_0	0.05-0.15 rad/s
$x_2 - x_1$	$1/8''$	f_0	2.0-3.5 rad/s
x_f	57 cm		

Table 1: Experimental parameters used in laboratory experiments, corresponding to figure 2.

Two types of experiments are performed: dye visualization and particle image velocimetry (PIV). For the dye experiments, a syringe pump capable of low flow rates ($\mathcal{O}(1\,\text{ml/hr})$) is used to inject dye at various locations within the basin. For the PIV experiments, saltwater with density $\rho \approx 1200\,\text{kg/m}^{-3}$ is used as the working fluid. The fluid is seeded with $50\,\mu$ particles with density $\rho \approx 1160\,\text{kg/m}^{-3}$. A green laser (532 nm/1064 nm) is used to illuminate the flow, with a pulse rate of 1-10 Hz. The software used to compute the flow velocities is LaVision's DaVis software.

The Ekman layer depth, relative to the change in depth due to the sloping bottom, is $\delta_{Ek}/\Delta D \sim 0.01 - 0.02$. The Munk scale, relative to the gap width, is $\delta_M/d < 0.25$ (where the gap needs to be at least twice the Munk scale, to prevent the boundary layers from blocking flow through the gap). Given that $d \ll L$, this confirms the assumption in § 2 that A is a small parameter.

We anticipate a resonant response at approximately the normal-mode frequencies associated with the full basin in the absence of the barrier as well as those for each individual subbasin [8, 5]. These frequencies are computed in nondimensional form (relative to $(\beta L) \sim 1$) by

$$\omega_{nm} = \frac{1}{2\pi\sqrt{m^2(L_y/L_x)^2 + n^2}}\,, \tag{30}$$

in which the integers m and n refer to the mode number in the x and y directions, respectively [3]. These frequencies are $\omega_F = 0.1075$, $\omega_E = 0.0802$, and $\omega_W = 0.0548$ for the full, eastern, and western basins respectively.

4 Results

4.1 Initial results and troubleshooting

Initial dye visualization experiments (not shown) showed evidence of oscillations at approximately the forcing frequency. However, one limitation of the dye visualization technique is that it only gives a Lagrangian description of particle paths, and not a full picture of the overall flow field, which motivated the move to full PIV measurements of the flow.

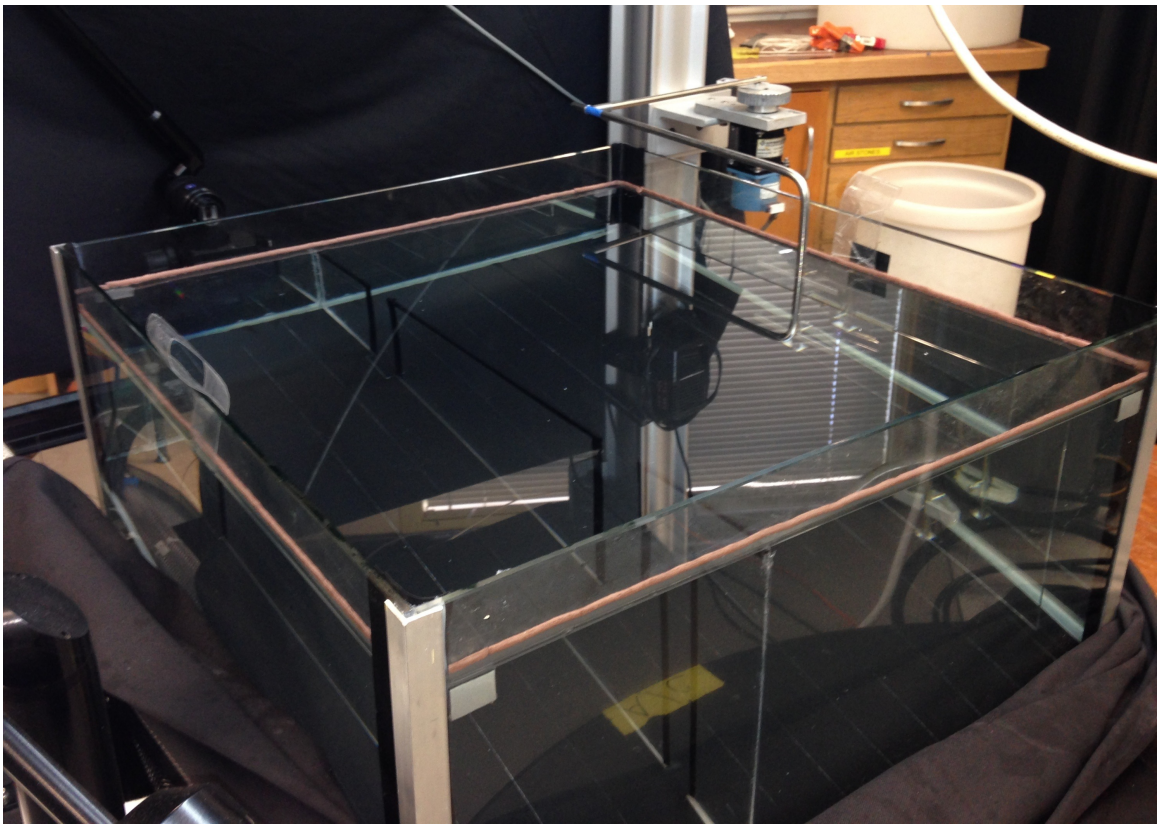


Figure 3: Laboratory apparatus corresponding to figure 2.

However, upon taking initial PIV measurements of the flow, a high-frequency oscillation of $\mathcal{O}(1\,\mathrm{Hz})$ was observed with amplitude comparable to that of the signal being forced. The high-frequency signal persisted even when the forcing was turned off entirely, as shown in figure 4; as such, it was necessary to determine the cause of the oscillations in order to reduce or eliminate them.

Several possible sources of oscillations were examined:

1. Surface gravity waves.
   Although a rigid lid for the tank is included in the experimental apparatus, it seemed possible that the lid may have still been able to wobble, thus forcing a flow in the interior of the tank. However, sealing down the lid did not appear to reduce the observed oscillations.

2. Table off-balance.
   Given the addition of the camera, camera mount, and laser to the table, it was thought that perhaps the weight may have unbalanced the table surface. However, adding weight to the opposite side of the table to offset this potential imbalance did not appear to reduce the oscillations.

3. Table not level.
   It was found that the table was not properly level, i.e. the axis of rotation was not parallel to gravity. Re-levelling the table appeared to help slightly but did not substantially reduce the oscillations.

4. Additional vibration introduced by control system.
   Due to the setup of the control system used to maintain the table's rotation speed, a peak in frequency of $\mathcal{O}(1\,Hz)$ could be observed in the associated frequency spectrum. By changing the manner in which the table was operated, the oscillations observed within the tank were reduced significantly.

With the oscillations observed in figure 4 substantially reduced, the PIV measurements of the forced flow could then be measured.

4.2 Comparison of PIV results with linear theory

PIV experiments were carried out at five forcing frequencies ($\omega_0 = 0.0690$, 0.0882, 0.1010, 0.1134, $0.1355\,\mathrm{rad/s}$) and three forcing amplitudes ($A_{\mathrm{forcing}} = 0.7, 2.0, 2.7\,\mathrm{cm}$) with a table rotation rate of $15\,\mathrm{rpm}$ ($f_0 = 3.1\,\mathrm{rad/s}$).

Figures 5 and 6 show the stream functions and velocities, respectively, from the laboratory measurements and the corresponding linear theory for forcing with $\omega_0 = \omega_F = 0.1355\,\mathrm{rad/s}$ (at $f_0 = 3.1\,\mathrm{rad/s}$) and $A_{\mathrm{forcing}} = 2.0\,\mathrm{cm}$. As the two figures show, despite the noise apparent in the experimental data, the linear theory does capture many of the large-scale features of the observed flow from the experiments. In particular, the computed streamfunction shown in figure 5 shows the Rossby wave propagating through the barrier with the correct frequency, as predicted by the linear theory.

Using the computed streamfunction from the experimental data, the experimental island constant can be computed. This is done by taking the average value of the streamfunction

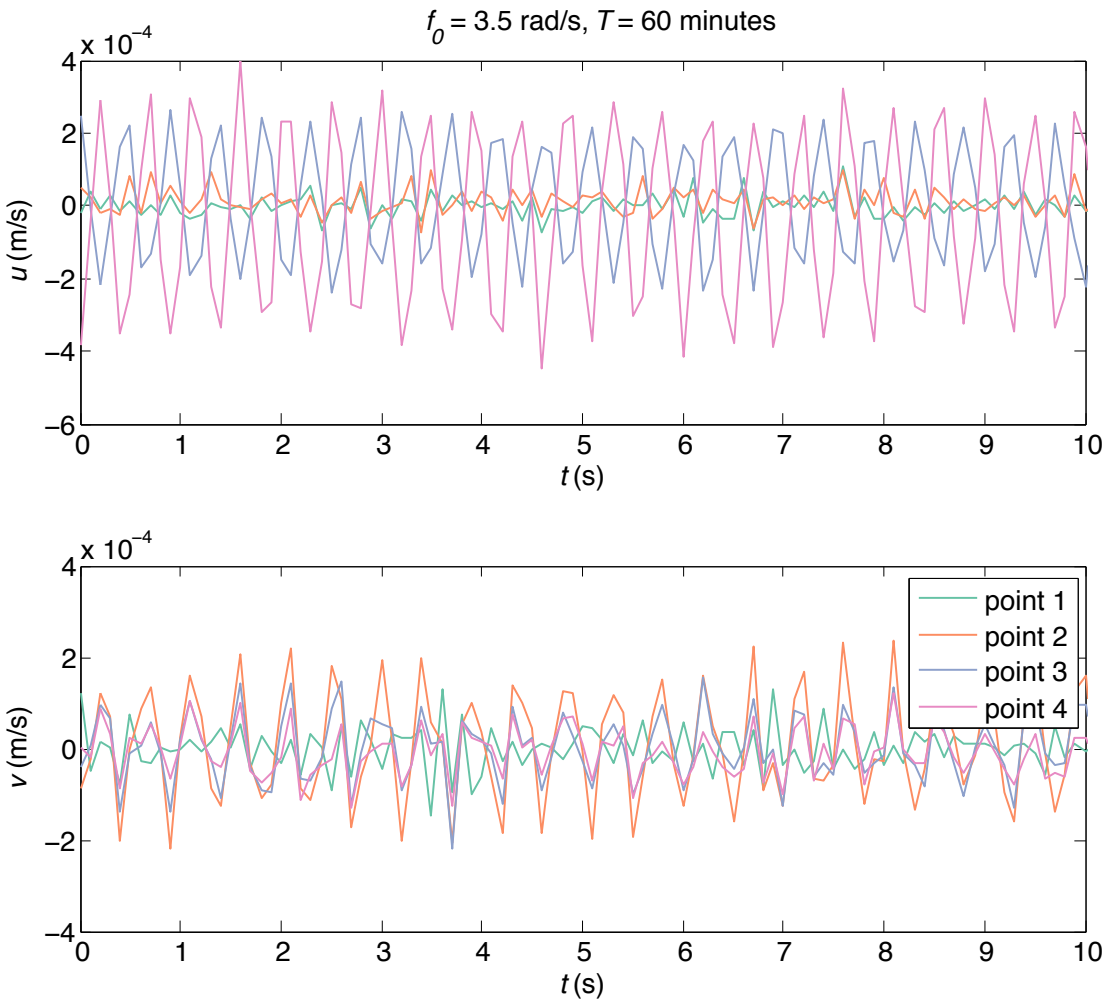


Figure 4: Horizontal (top) and vertical (bottom) velocities at four points in basin. Points 1, 2, 3, and 4 correspond to the northern gap, southern gap, centre of the western subbasin, and centre of the eastern subbasin, respectively.

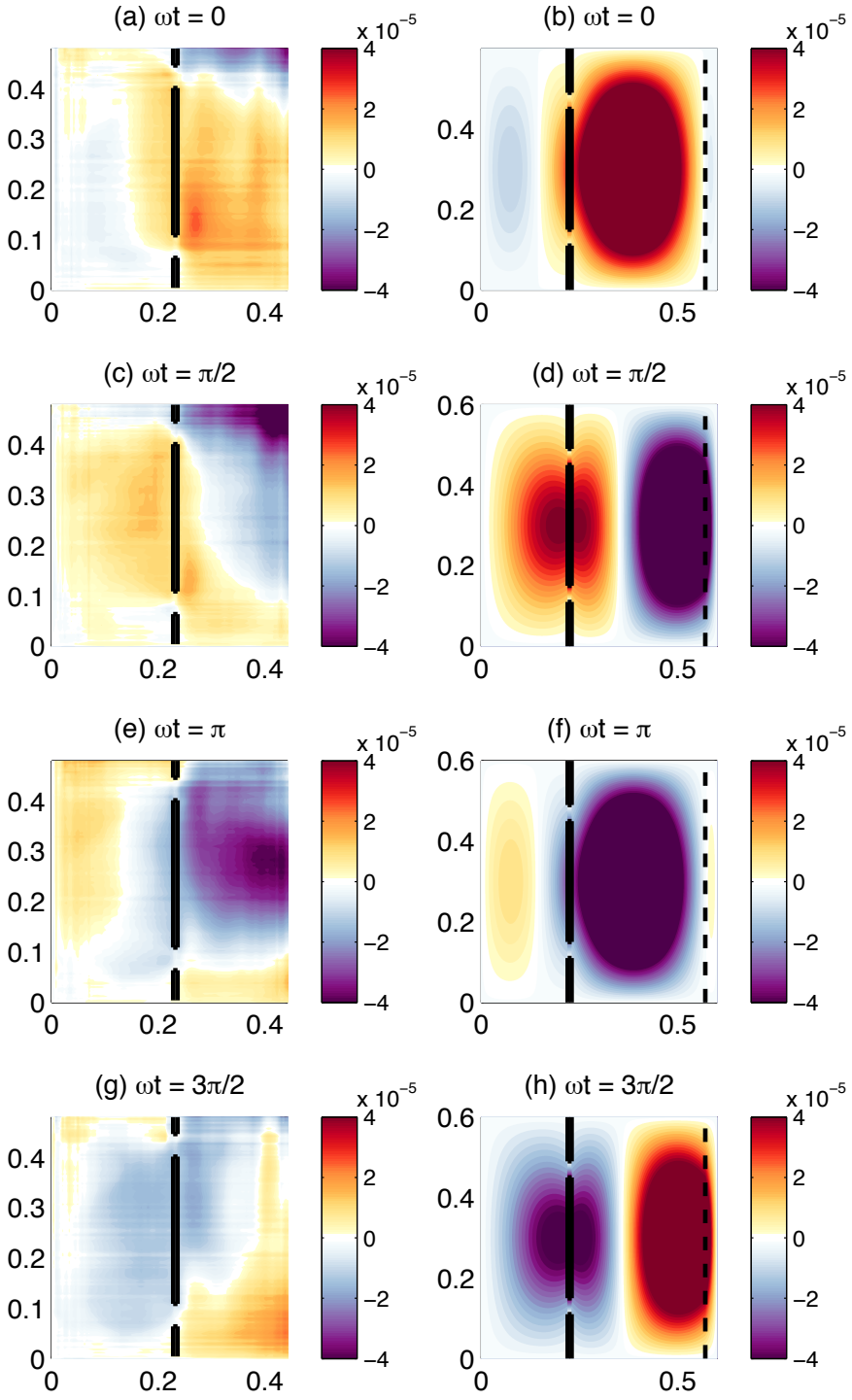


Figure 5: Comparison of streamfunction in laboratory results (left) and theoretical model (right) for $\omega_0 = \omega_F = 0.1355\,\mathrm{rad/s}$ at $f_0 = 3.1\,\mathrm{rad/s}$ and $A_{\mathrm{forcing}} = 2.0\,\mathrm{cm}$.

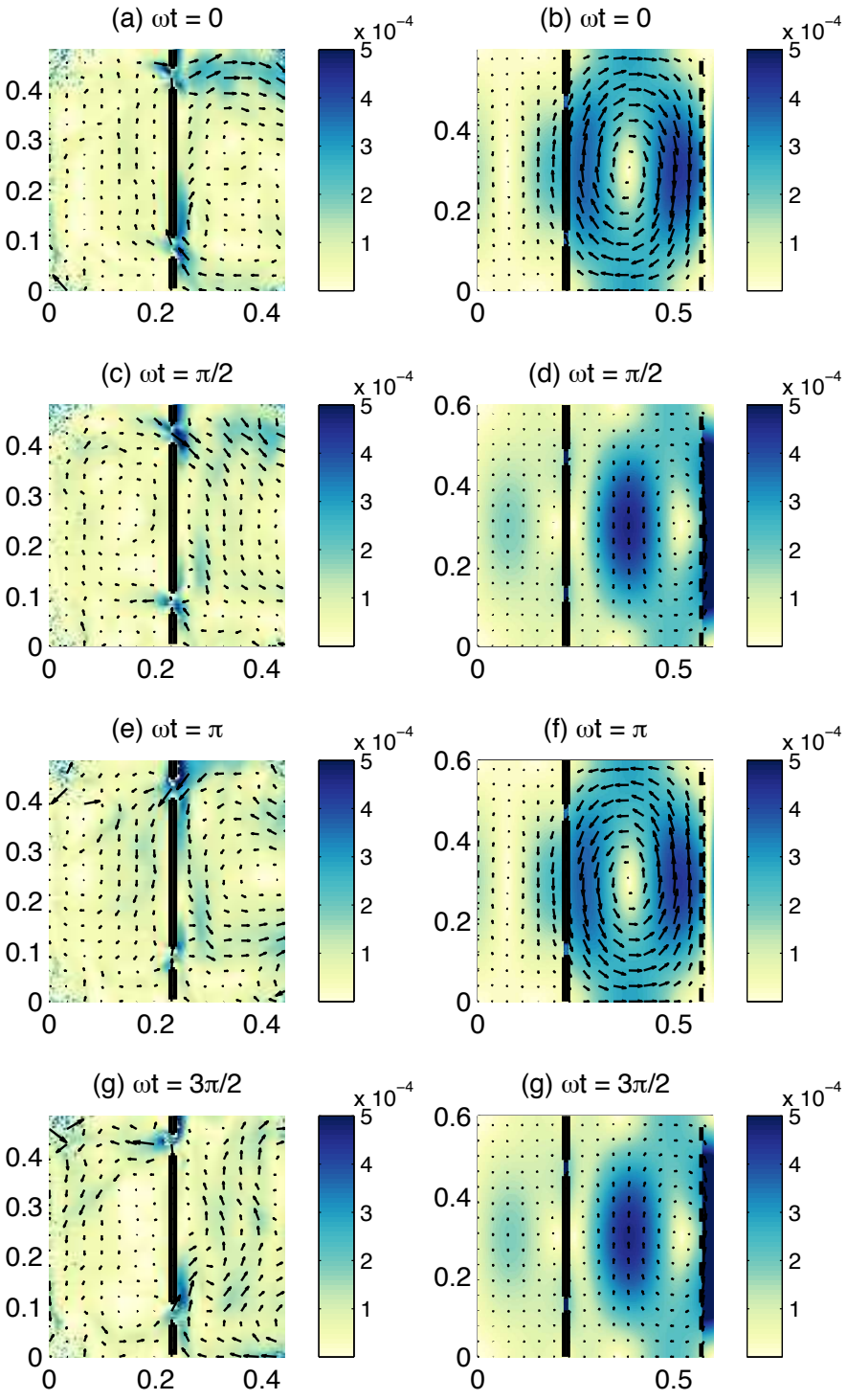


Figure 6: Comparison of velocities (arrows) and speeds (colours) in laboratory results (left) and theoretical model (right) for $\omega_0 = \omega_F = 0.1355\,\mathrm{rad/s}$ at $f_0 = 3.1\,\mathrm{rad/s}$ and $A_{\mathrm{forcing}} = 2.0\,\mathrm{cm}$.

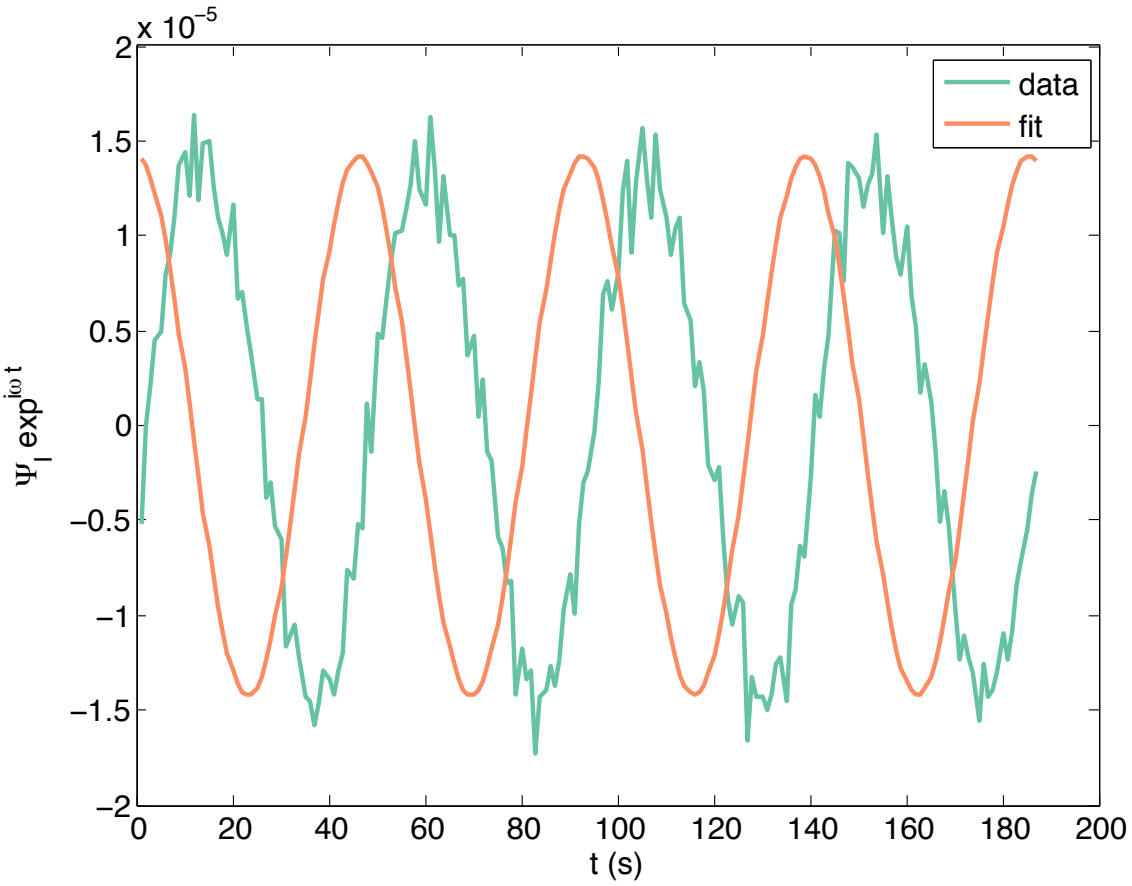


Figure 7: Mean value of streamfunction around island and corresponding computed fit, $\Psi_I \cos \omega t$.

around the island at each point in time, which oscillates with frequency ω, and computing the root-mean-square velocity corresponding to this timeseries. An example of the data and corresponding fit for $\omega_0 = 0.1355\,\text{rad/s}$ and $A_{\text{forcing}} = 2\,\text{cm}$ is shown in figure 7. Excellent agreement is seen between the amplitude of the oscillations in the experimental data and the computed amplitude Ψ_I.

A summary of the computed values of the island constant for all forcing frequencies and forcing amplitudes is shown in figure 8. Of note is the apparent lack of resonance, despite three of the forcing frequencies ($\omega_0 = 0.0690, 0.1010, 0.1355\,\text{rad/s}$) corresponding to the western subbasin, eastern subbasin, and full basin normal mode frequencies. This is particularly surprising given that the theoretical predictions of Pedlosky and Spall (1999)[8] and Pedlosky (2000)[5] indicate that resonant peaks should be observed in the value of Ψ_I (see, for instance, figure 4 of [5]).

An additional way in which the experimental results may be compared with the linear theory outlined above is by further examining the integral constraint defined in (25). We

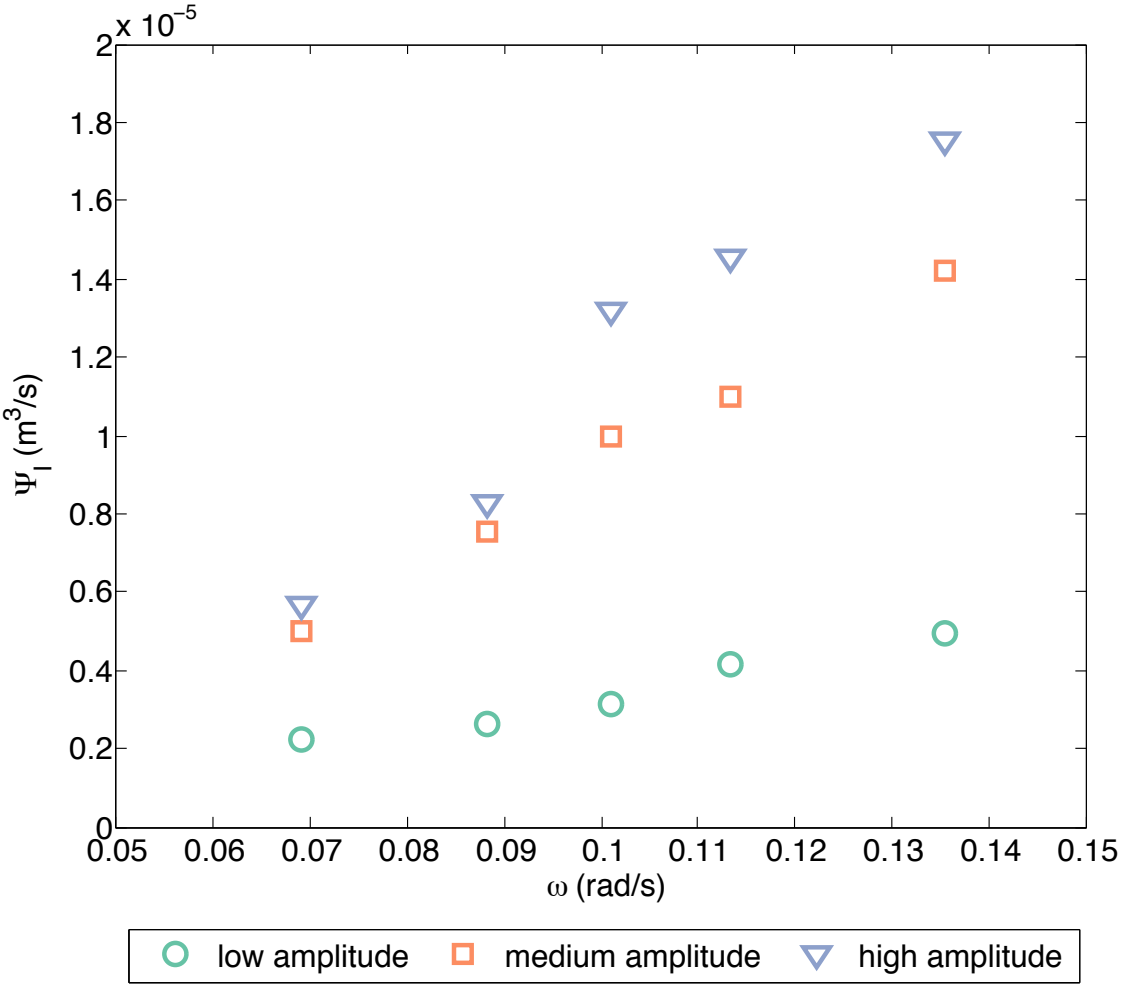


Figure 8: Computed values of island constant Ψ_I for different forcing frequencies and amplitudes.

redefine this expression as

$$\oint_{C_I} \mathbf{u} \cdot \mathrm{d}\mathbf{s} = \text{residual}\,. \tag{31}$$

By computing the individual components of the contour integral, the residual is determined.

The mean value of the computed residual is shown in figure 9, and the rms amplitudes of oscillation of each of the individual components of the contour integral are shown in figure 10. The amplitudes of the mean values are smaller than the rms amplitudes; however, it is clear from figure 10 that the amplitude of the residual value is comparable to that of some of the individual terms in the contour integral. This points to some missing physics in the contour integral. As the contour integral should be valid for nonlinear flows, the missing terms are likely related to missing viscous effects not accounted for in the linear theory (e.g. the assumption that lateral friction is negligible in the basin). This is, perhaps, unsurprising given that the laboratory's smaller length scales; viscous effects might be expected to be more dominant in the laboratory than at, say, geophysical scales.

4.3 Nonlinear effects

Figure 11 shows a zoomed-in view of the northern gap for $\omega_0 = 0.1134\,\mathrm{rad/s}$ and $A_{\mathrm{forcing}} = 2.7\,\mathrm{cm}$. For these parameters, the scalings for nonlinearity in the basin and in the gaps are $NL_{\mathrm{basin}} = 0.004$ and $NL_{\mathrm{gap}} = 13.7$, respectively. As such, it is expected that significant nonlinear effects will be observed in the gap regions for these parameters.

Examination of figure 11 does show evidence of nonlinear effects. In particular, we can see the formation of a vortex on the western side of the barrier, and a strong boundary current on the eastern side of the barrier. Both the vortex and the strong east-west asymmetry are characteristic of strong nonlinear effects which are not well captured by the linear theory.

5 Discussion and conclusions

Here we have presented laboratory experiments corresponding to the theory of Pedlosky (2000)[5] and Pedlosky (2001)[6], which predicts that barriers with relatively small gaps may be quite inefficient in preventing the transmission of Rossby waves through the barriers. Comparisons between the linear theory and the experimental results indicate that the theory does capture the large-scale structure of the flow; however, preliminary analysis of the results points to the importance of additional physics not captured by the linear theory in its current form.

First, additional forcing frequencies would be of great interest, particularly to determine whether resonant peaks are seen in the laboratory experiments (as the linear theory suggests) or not. Given that only five frequencies were considered above, it may be the case that the resonant peaks were "missed" in some sense. On the other hand, if there is no evidence of resonance, it may be indicative of some additional physics not fully captured by the linear theory.

Modification of the forcing symmetry and/or the barrier symmetry and number of gaps would be of interest. Blocking off one of the gaps to confirm that only smaller-scale structures are able to pass through the gap would be of interest, as would adjusting the forcing

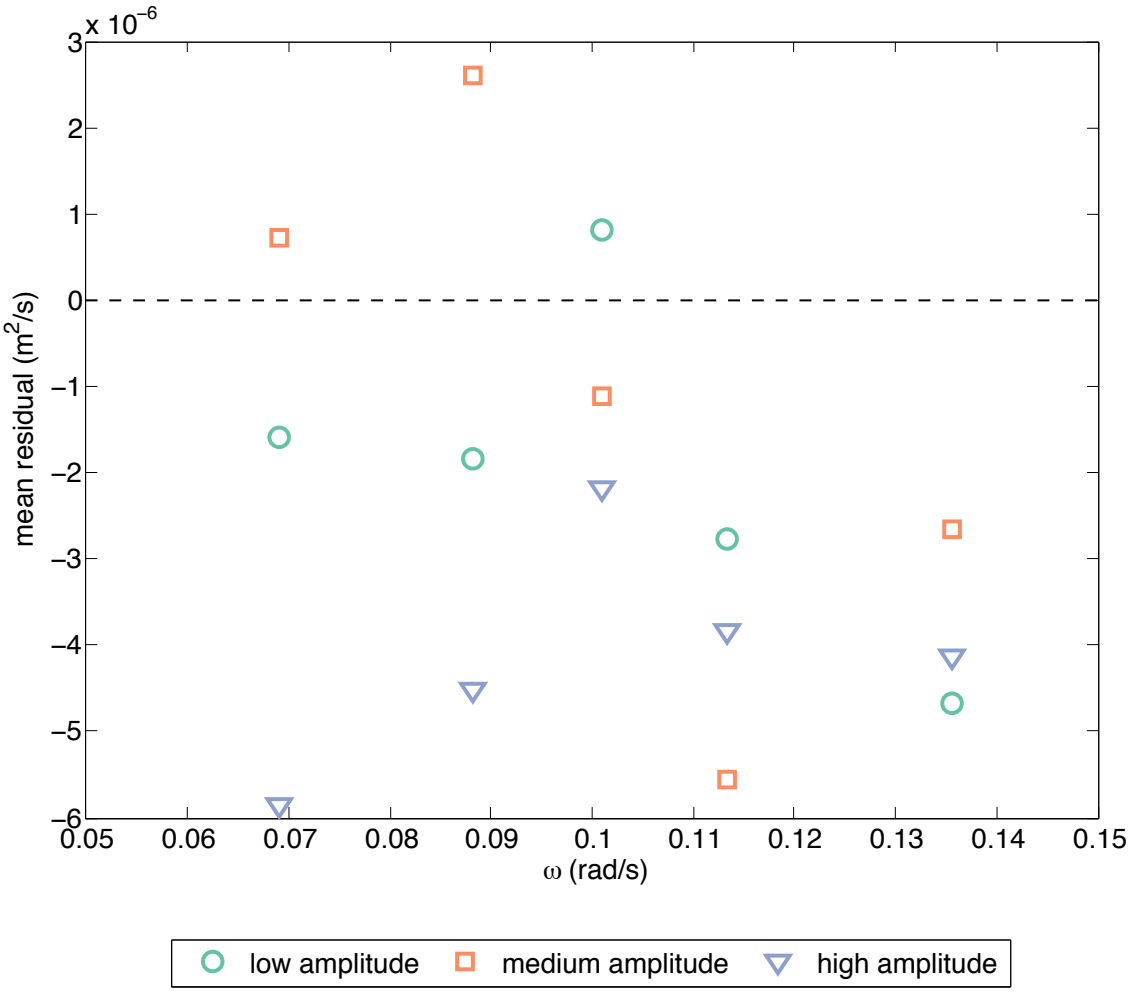


Figure 9: Mean values of the computed residual from (31).

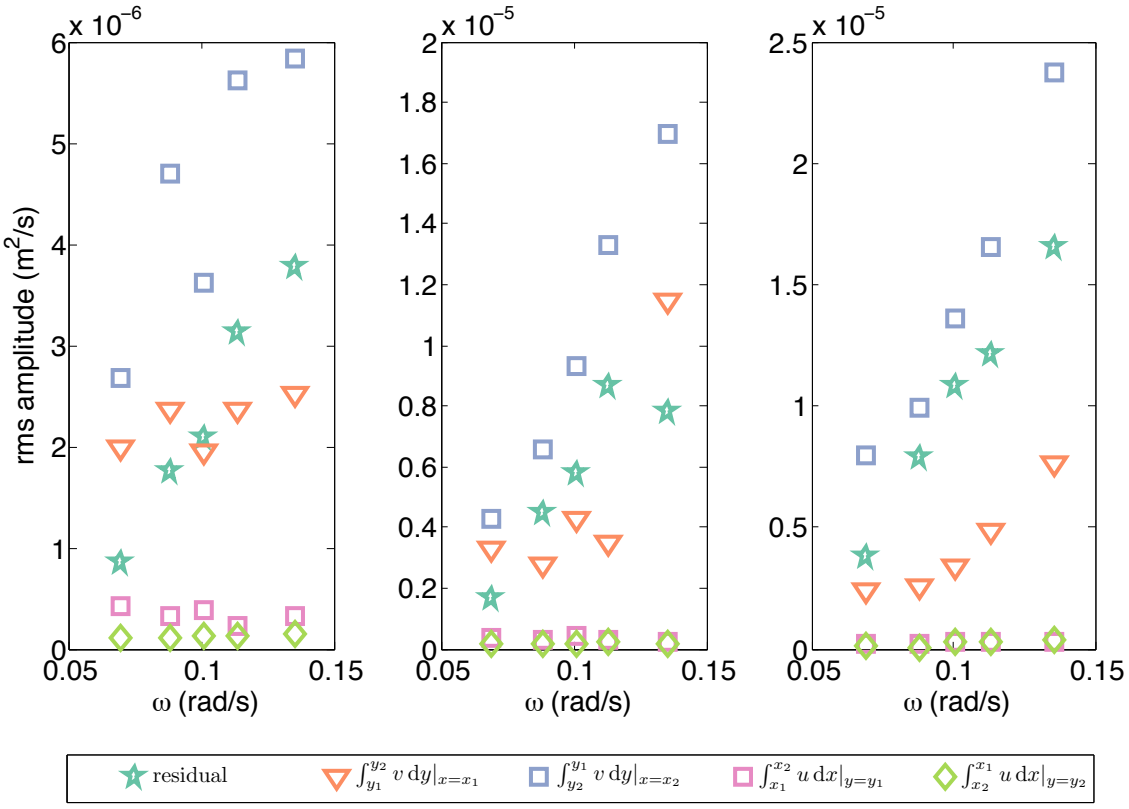


Figure 10: Root-mean-square amplitudes of each term in (31).

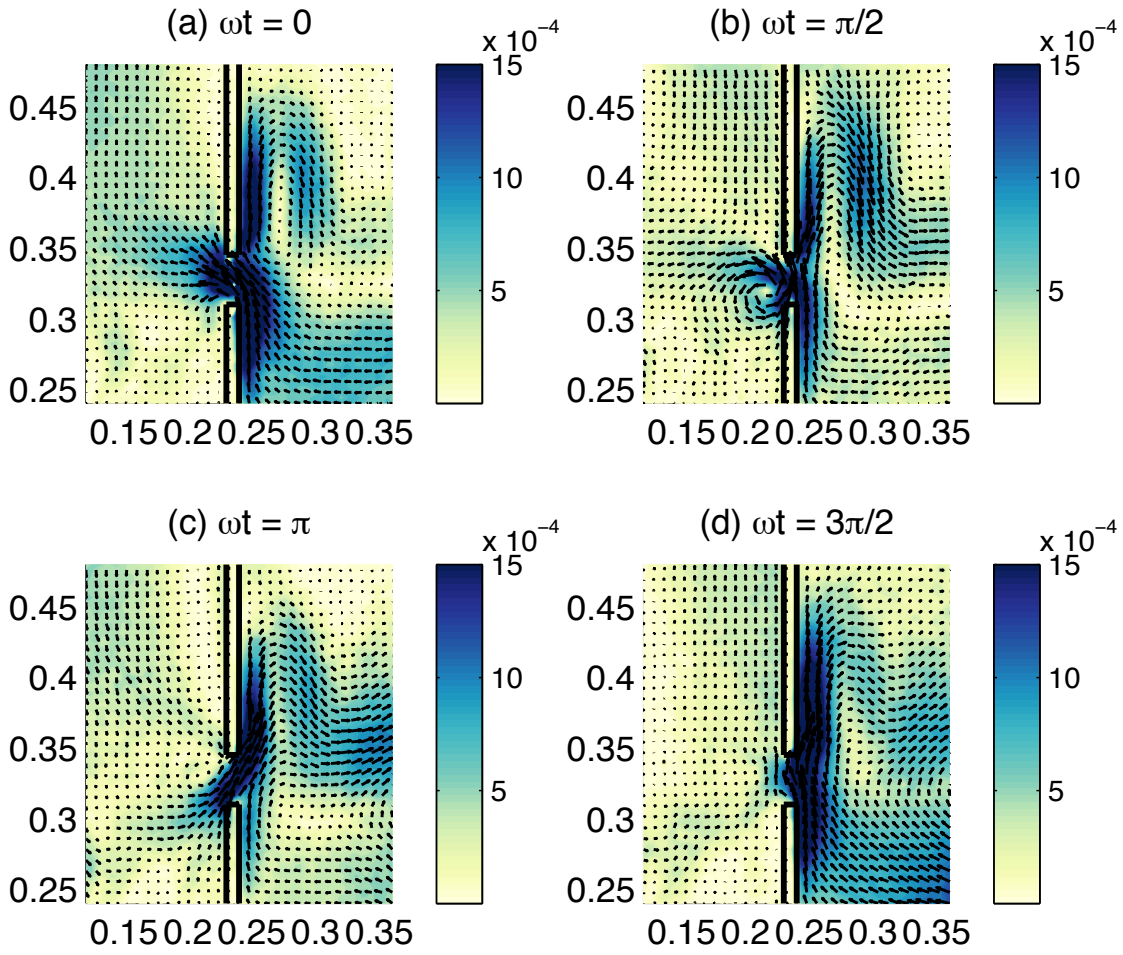


Figure 11: Velocities (arrows) and speeds (colours) in laboratory results, zoomed in around the northern gap, for $\omega = 0.1134\,\mathrm{rad/s}$ at $f_0 = 3.1\,\mathrm{rad/s}$ and $A_{\mathrm{forcing}} = 2.7\,\mathrm{cm}$.

to one in which n is even. As Pedlosky and Spall (1999)[8] describe (see their figure 7), additional gaps in the barrier allow for different ways in which the waves may be transmitted through the barrier; comparing these scenarios with laboratory results could reveal further interesting features of the flow.

Additional forcing amplitudes could also be considered. In particular, it would be interesting to see if increasing the forcing could lead to separation of the vortex forming at the western side of the gap, shown in figure 11. Additionally, varying the forcing may help to quantify under what conditions these vortices form. As such vortices are an additional means by which wave energy may be transmitted into the western subbasin, and are a nonlinear effect not described by the linear theory described above, understanding their behaviour is of great interest.

Finally, numerical simulations corresponding to the laboratory length scales would be of great interest. In particular, while it is difficult to compute contributions due to viscosity from the measured velocities, owing to the noisy nature of the data and the higher derivatives that appear in viscous terms, these terms could be more easily computed from numerical data. This may help in quantifying the effects of nonlinearity and viscosity suggested to be of some importance by the experimental results.

6 Acknowledgments

I would like to thank Joe Pedlosky and Karl Helfrich, my advisors for this project. I am very grateful for their patience and guidance in getting set up in the lab and working through the theory. Thanks for proposing such an interesting problem to look at for the summer! I would also like to thank Anders Jensen for helping to design and build the experimental apparatus that was used for this project – his help was very much appreciated!

References

[1] K. A. Kelly and L. Thompson, *Scatterometer winds explain damped Rossby waves*, Geophys. Res. Lett., 29 (2002), pp. 52–1–4.

[2] M. S. Longuet-Higgins, *Planetary waves on a rotating sphere*, Proc. Roy. Soc. A, 279 (1964), pp. 446–473.

[3] J. Pedlosky, *A study of the time dependent ocean circulation*, J. Atmos. Sci., 22 (1965), pp. 267–272.

[4] ——, *Fluctuating winds and the ocean circulation*, Tellus, 19 (1967), pp. 250–257.

[5] ——, *The transmission of Rossby waves through basin barriers*, J. Phys. Oceanogr., 30 (2000), pp. 495–511.

[6] ——, *The transparency of ocean barriers to Rossby waves: the Rossby slit problem*, J. Phys. Oceanogr., 31 (2001), pp. 336–352.

[7] J. Pedlosky, L. J. Pratt, M. A. Spall, and K. R. Helfrich, *Circulation around islands and ridges*, J. Mar. Res., 55 (1997), pp. 1199–1251.

[8] J. Pedlosky and M. A. Spall, *Rossby normal modes in basins with barriers*, J. Phys. Oceanogr., 29 (1999), pp. 2332–2349.

[9] B. A. Warren, T. Whitworth III, and J. H. LaCasce, *Forced resonant undulation in the deep Mascarene Basin*, Deep-Sea Res. II, 49 (2002), pp. 1513–1526.

A Study of Heat Transport and the Runaway Greenhouse Effect using an Idealized Model

Paige Martin

1 Introduction

The Earth, in its current state, is in a delicate balance of incoming radiation from the sun, and outgoing radiation emitted by our planet. For the Earth to maintain a certain climate, the amount of incoming and outgoing radiation must be equal, and thus we say that the Earth is in radiative equilibrium. What would happen to the temperature on Earth if something were to perturb our current equilibrium? This work looks into the possibility of a runaway greenhouse effect and how an increase in the incoming solar radiation would affect the temperature of the Earth's surface. We use a one- and two-column model, first in pure radiative equilibrium, then with added convection, and allow the atmosphere's optical depth to vary with temperature and allow for lateral heat transport between the two columns. Our goal is to investigate whether the lateral heat transport from the equator toward the poles might mitigate the impact of the runaway greenhouse effect.

2 Brief Background

The majority of the Earth's energy comes from the incoming solar radiation, which passes through most of the components of the atmosphere, straight to the ground. This light is reflected back upwards through the atmosphere, largely in the form of infrared radiation, some of which gets absorbed by the atmosphere. The absorbed radiation thus heats up the atmosphere, and this is why the Earth's surface temperature is warm enough for human habitation [9] [3]. This atmospheric warming is called a greenhouse effect.

In this work, we will investigate the runaway greenhouse effect. This phenomenon occurs when the Earth absorbs more radiation than it can emit, and thus is no longer in equilibrium. If the surface temperature of the planet increases, more water vapor is formed in the atmosphere, which then traps more of the upward infrared radiation, leading to an even higher concentration of water vapor in the atmosphere, and so the cycle continues, and hence the term "runaway" [4] [5] [9]. In Ingersoll (1969), it is proposed that Venus underwent such a runaway greenhouse effect that eventually caused its oceans to boil away [4]. Here, we take a preliminary look into how close the Earth is to experiencing the same fate as Venus.

Symbol	Meaning	Value
S_0	solar constant	1366 W/m^2
α	albedo	0.3
σ	Stefan-Boltzmann constant	$5.67 * 10^{-8}$
T_g	ground temperature	
T_a	atmospheric temperature	

Table 1: Table of symbols used in this section.

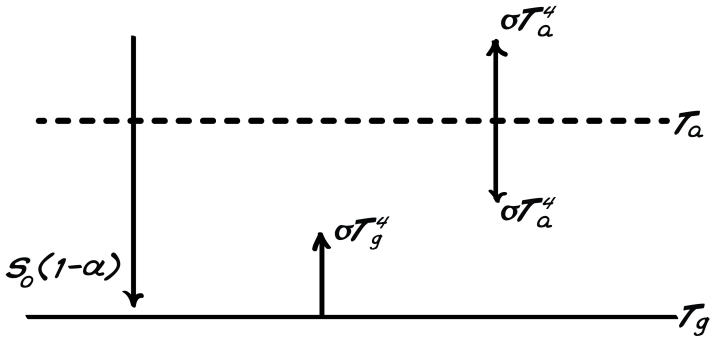


Figure 1: A one-column, two-layer model of radiative balance.

3 The Model

3.1 One-Column Radiative Equilibrium

(Note: the following derivations are obtained by following the steps in Geoffrey Vallis' 2014 GFD Lecture Notes [9]). We first consider a very simple radiative equilibrium model of the atmosphere, starting with a single column, two-layer setup. As illustrated in figure 1, the only radiative fluxes present are the incoming solar radiation, $S_0(1-\alpha)$, the upward heat flux from the ground, given by the Stefan-Boltzmann Law, σT_g^4, and the upward and downward heat fluxes emitted from the atmosphere layer, σT_a^4 [7] [1] [8] [6]. (See table 1 for the meanings and values of the symbols used in this and following sections.) To keep the notation consistent, it is common to represent the incoming solar radiation in the form of the Stefan-Boltzmann Law in terms of an effective emitting temperature of the Earth, which we call T_e. In equation form, this gives [6] [2]

$$\sigma T_e^4 = S_0(1-\alpha). \tag{1}$$

The equations for the model illustrated in figure 1 can be written assuming radiative equilibrium at each layer - the ground and the atmosphere.

At the surface of the Earth:

$$\sigma T_e^4 + \sigma T_a^4 = \sigma T_g^4 \tag{2a}$$

At the top of the atmosphere:

$$\sigma T_e^4 = \sigma T_a^4 \tag{2b}$$

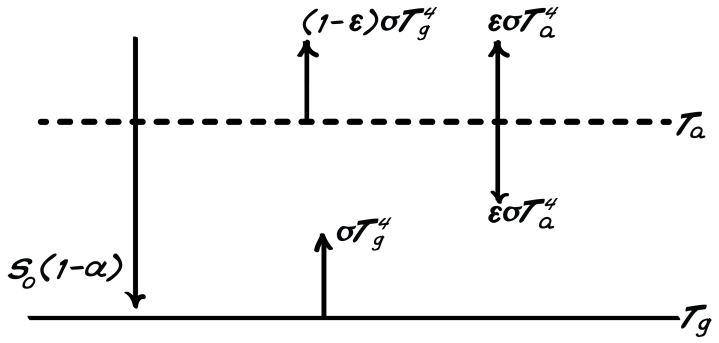


Figure 2: A one-column, two-layer model of radiative balance with an absorbing atmosphere.

This clearly gives $T_a = T_e$, and thus $T_g = 2^{\frac{1}{4}} T_e$. For an emitting temperature $T_e = 255K$, this implies a ground temperature of about $T_g = 303K$, which overestimates the actual ground temperature of around $288K$, so we choose to improve on this model.

The model assumes that the Earth is a perfect black body and thus absorbs no radiation. A small tweak to this model yields slightly more realistic results; we assume that the atmosphere boundary emits only a fraction ϵ (called the "emissivity") of the incoming radiation flux (see figure 2) [2]. Again, balancing incoming and outgoing radiation at each of the two layers in the model gives the following equations.

At the surface

$$\sigma T_e^4 + \epsilon\sigma T_a^4 = \sigma T_g^4 \tag{3a}$$

At the top

$$\sigma T_e^4 = \epsilon\sigma T_a^4 + (1-\epsilon)\sigma T_g^4. \tag{3b}$$

We now have a set of equations that depend on the parameter ϵ, which varies between 0 and 1. However, this restriction on the value of ϵ is too strict to allow for a runaway greenhouse effect , and so we chose to upgrade this model to one that is continuous in the vertical direction. The new setup is still in one-column form, but instead of having layers and setting up radiative balance at the interfaces for specific values of the emissivity, we have a continuous atmosphere written in terms of the optical depth τ. The emissivity and optical depth are essentially measuring the same quantity - how much of the incoming radiation gets transmitted, i.e. a measure of the atmosphere's opacity to long wave radiation. The equations in this case are more complicated, and cannot be simply read directly from the diagram, but are derived from the radiative transfer equations.

3.1.1 Radiative Transfer Equations

Let's begin by considering a beam of radiating particles traveling through a medium that is also emitting radiation. The intensity I of the beam of particles changes according to the equation

$$dI = (B - I)d\tau, \tag{4}$$

where B is the radiation emitted by the medium, given by the Stefan-Boltzmann Law (σT^4), and τ is the optical depth. Applying this equation to the atmosphere, and assuming a two-stream atmosphere, we allow for only upward (U) and downward (D) radiation. We define τ in the standard way, i.e. decreasing with atmospheric height such that $\tau = 0$ at the top of the atmosphere. The equations are thus

$$dU = -(B - U)d\tau \tag{5a}$$

$$dB = (B - D)d\tau. \tag{5b}$$

Notice the negative sign in the first equation, since the upward radiation U increases with decreasing τ. In the atmosphere, B refers to σT^4 for the T at a certain height, dictated by the value of τ. Dividing both sides by $d\tau$, we are left with the differential equations

$$\frac{dU}{d\tau} = U - B \tag{6a}$$

$$\frac{dB}{d\tau} = B - D. \tag{6b}$$

It turns out to be convenient to change variables to $(U + D)$ and $(U - D)$, which gives

$$\frac{d(U + D)}{d\tau} = U - D \tag{7a}$$

$$\frac{d(U - D)}{d\tau} = U + D - 2B. \tag{7b}$$

$$\frac{d(U - D)}{d\tau} = 0 \tag{8}$$

These are the equations that we would like to find solutions for, imposing the boundary conditions

$$D = 0 \tag{9a}$$

$$U = \sigma T_e^4 \tag{9b}$$

for $\tau = 0$.

The solution for the given boundary condition is the following:

$$D = \sigma T_e^4 \left(\frac{\tau}{2}\right) \tag{10a}$$

$$U = \sigma T_e^4 \left(1 + \frac{\tau}{2}\right) \tag{10b}$$

$$B = \sigma T_e \left(\frac{1 + \tau}{2}\right). \tag{10c}$$

The final equation is the one we are interested in - the one that relates the temperature of the atmosphere (in B) to the emitting temperature (T_e).

We are interested in the ground temperature of the Earth, but so far we have not accounted for a black surface at the ground level when z=0. From equation 10, we know

Symbol	Value
a	1.12
b	.14
c	5413

Table 2: The values of the constants used in the expansion of optical depth in terms of temperature.

the expression for the upward radiation. This is true at all levels, and so at z=0 with temperature T_g, we have

$$T_g^4 = \left(\frac{2+\tau}{1+\tau}\right) T_s^4, \tag{11}$$

where I use T_s to denote the temperature of the surface, i.e. just above the actual ground. Below we plug this equation in to replace T_s with T_g, the ground temperature which is the desired value.

The next step is to expand the optical depth in terms of temperature. We write the optical depth first in terms of z:

$$\tau = \tau_0 e^{-z/H_a}, \tag{12}$$

where z is the vertical height (in km) and the constant H_a corresponds to the characteristic height of water vapor in the atmosphere, which is taken to be 2 km. Because this work is studying the temperature only at the surface of the Earth, we take $z = 0$. However, we expand τ_0 as a function of temperature, defined as

$$\tau_0(T_g) = a + b\left(e^{c\left(\frac{1}{288} - \frac{1}{T_g}\right)}\right). \tag{13}$$

The constants a,b, and c (see Table 2) are determined by model tuning in conjunction with discussions with Andy Ingersoll.

With $z = 0$, and plugging in for τ and T_g, the final equation is

$$\sigma T_g^4 = \sigma T_e^4 \left(1 + \frac{a + b\left(e^{c\left(\frac{1}{288} - \frac{1}{T_g}\right)}\right)}{2}\right). \tag{14}$$

The solutions to this equation for various emitting temperatures, as well as other analysis, are shown in the Results Section.

Before adding convection to the model, we first add a term to the equation that allows for additional heat loss, and that will lead into the two-column model that will be discussed later. The new term takes the form $-kT$ and yields the slightly altered equation:

$$\sigma T_g^4 = (\sigma T_e^4 - kT_g)\left(1 + \frac{a + b\left(e^{c\left(\frac{1}{288} - \frac{1}{T_g}\right)}\right)}{2}\right). \tag{15}$$

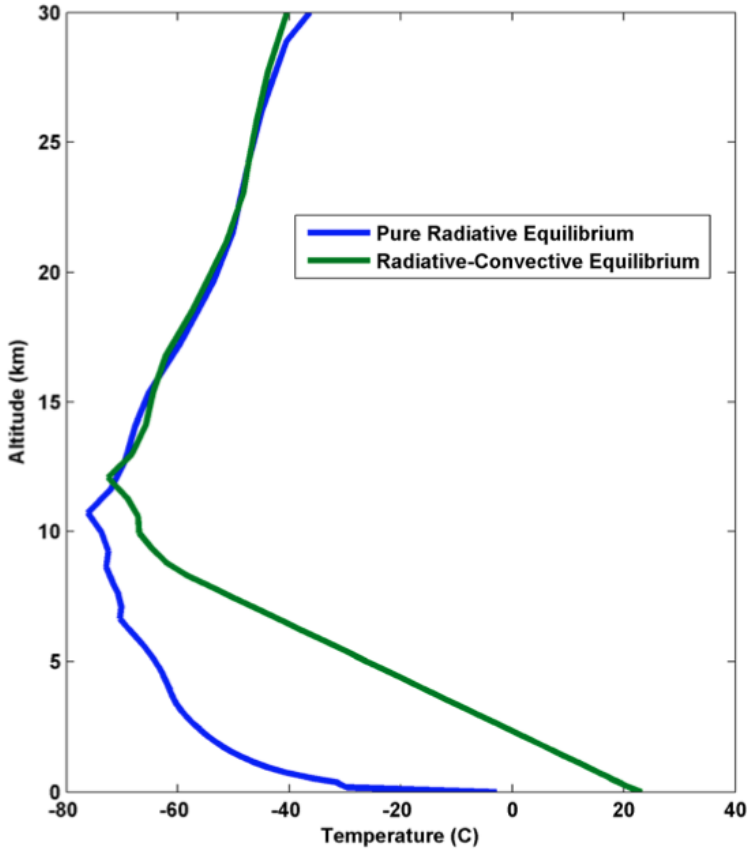


Figure 3: The temperature profiles of radiative and radiative-convective equilibrium. (Graph by Kerry Emanuel [2])

In the next section we will allow for atmospheric convection to occur, in addition to radiation.

3.2 One-Column Radiative-Convective Equilibrium

Convection plays an important role in the atmosphere, particularly in the troposphere, and greatly influences the heating in the lower atmosphere, as shown in figure 3. From this graph, the temperature profile in the lower atmosphere clearly changes with the addition of convection; instead of a large increase in temperature very close to the ground in the purely radiative case, the temperature change is roughly linear with both radiative and convective effects [2]. Convection, therefore, is an important process to include in our model.

The approach differs from the purely radiative case outlined above. We can no longer assume radiative balance as in the previous section, as we have extra terms in that come into play from the convection. We will begin with the same radiative transfer equations

$$\frac{dU}{d\tau} = U - B \tag{16a}$$

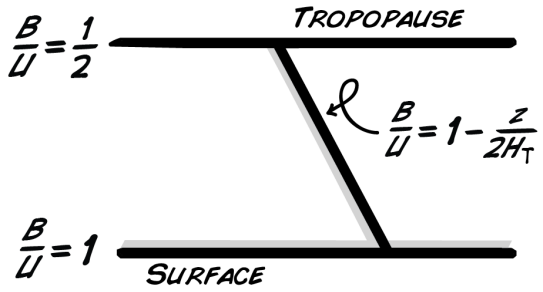


Figure 4: A depiction of the approximation used in the RCE model.

$$\frac{dB}{d\tau} = B - D. \tag{16b}$$

From here, we solve the equations in a different manner, and derive two equations with two unknowns: T_g, which still refers to the ground temperature and H_T, the height of the tropopause. The tropopause is the boundary between the troposphere and the stratosphere.

We will only consider the first of the radiative transfer equations, and rearrange to give

$$\frac{1}{U}\frac{dU}{d\tau} = 1 - \frac{B}{U}, \tag{17}$$

so that we can write it as

$$\frac{d(ln(U))}{d\tau} = 1 - \frac{B}{U}. \tag{18}$$

Next we make an assumption that has been shown to be a good match to data. As depicted in figure 4, we assume that the quantity B/U varies linearly from the tropopause down to the surface, according to the equation

$$\frac{B}{U} = 1 - \frac{z}{2H_T}, \tag{19}$$

where z is the height in the vertical direction (the altitude). This expression can be substituted into equation (18), which yields

$$\frac{d(ln(U))}{d\tau} = \frac{z}{2H_T}. \tag{20}$$

Since the right-hand side of the equation is now in terms of z, we make the substitution from τ to z in the derivative term, according to

$$\tau(z) = \tau_s e^{\frac{-z}{H_a}}. \tag{21}$$

Thus,

$$d\tau = \frac{-1}{H_a}\tau_s e^{\frac{-z}{H_a}} dz. \tag{22}$$

Note that here we use the symbol τ_s instead of τ_0 as in the RE case. This is because we use a different value of the optical depth when accounting for convection, in order to keep the ground temperature at 288K. The dependence on temperature still takes the same form, except for a scaling factor of 6.4 in the RCE case, compared to the RE constants. A higher value of the optical depth for convection than radiation makes sense, if we look at figure 3 and notice the rapid temperature increase very close to the ground in the radiative equilibrium temperature profile in blue. For convection to play its role and heat the ground to the same temperature, the optical depth must be greater.

The differential equation then becomes

$$\frac{d(ln(U))}{dz} = -\frac{z}{2H_T}\frac{\tau_s}{H_a}e^{\frac{-z}{H_a}}. \tag{23}$$

This equation can be solved using integration by parts, which gives

$$ln\left(\frac{U(z=H_T)}{U(z=0)}\right) = \int_0^{H_T} -\tau_s H_a e^{\frac{-z}{H_a}}\frac{1}{2H_T H_a}dz \tag{24}$$

$$ln\left(\frac{2\sigma T_T^4}{\sigma T_g^4}\right) = \frac{-\tau_s H_a}{2H_T}, \tag{25}$$

where T_T is the temperature of the tropopause, given by

$$T_T = \frac{T_e}{2^{\frac{1}{4}}}. \tag{26}$$

Before we can evaluate the left-hand side of the equation, we introduce an equation that will be used in the remainder of this section. The following equation relates the temperature of the tropopause to the temperature of the ground, which is the quantity we are truly interested in. To do this, we use the lapse rate (denoted Γ), which is a measure of how much the temperature in the atmosphere decreases with increasing altitude. i.e. $\frac{dT}{dz}$. For simplicity, we are assuming a constant lapse rate, and so the temperature of the tropopause is related to the temperature of the ground in a linear fashion:

$$T_T = -\Gamma H_T + T_g. \tag{27}$$

So Γ is the slope of the line in a temperature vs. height graph that connects the ground temperature with the tropopause temperature. In our model, we use a global average value of 6.5 K/km for the lapse rate [9].

Going back to equation (25), the left-hand side can be approximated as

$$ln\left(\frac{2T_T^4}{T_g^4}\right) = ln(2) + 4ln\left(\frac{1}{1+\frac{\Gamma H_T}{T_T}}\right) \tag{28}$$

$$= ln(2) - 4\left(\frac{\Gamma H_T}{T_T}\right), \tag{29}$$

where we plug in for T_g according to equation (27) and approximate the last term in the first equation above by using a Taylor series to obtain equation (29). If we put all of these approximations together, we get a quadratic equation in H_T:

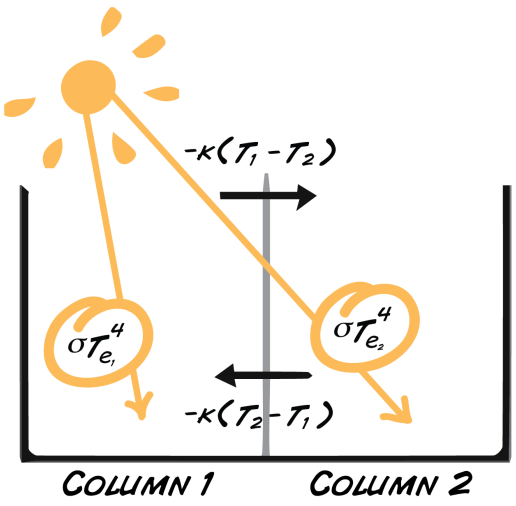


Figure 5: The two-column model of radiative balance with an absorbing atmosphere.

$$-8\Gamma H_T^2 + 2ln(2)(H_T T_T) + \tau_s H_a T_T = 0. \tag{30}$$

Solving for H_T gives one of the two equations used in the one-column RCE model, and the other is simply equation (25):

$$H_T = \frac{1}{16\Gamma}\left(CT_T + \sqrt{C^2T_T^2 + 32\Gamma\tau_s H_a T_T}\right) \tag{31a}$$

$$T_g = T_T + \Gamma H_T, \tag{31b}$$

with $C = 2ln(2)$. Equations 31 are the two equations with two unknowns, the height of the tropopause H_T and the temperature of the ground T_g, that we solve to find the radiative-convective equilibrium solutions in a one-column model.

3.3 Two-Column Radiative Equilibrium

Now we take a step backwards and disregard convection, but create a two-column model that allows for lateral heat transport between the two columns. The setup is identical to the one-column RE model (see figure 5), with different values of the emitting temperature (T_e) and ground temperature (T_g), and thus also the optical depth (τ), as well as an added term defined as $k(T_{g_i} - T_{g_j})$ to represent the exchange of heat between the columns (where i and j denote either column 1 or column 2). Here, k is considered to be a diffusion coefficient, and this diffusion term is an approximation to the standard diffusion term $k\frac{d^2T}{dy^2}$, where y denotes the latitude. We consider column 1 to be a rough analogy to the tropics, and column 2 to represent the midlatitudes. This is taken into account by allowing for the emitting temperatures to be those corresponding to the specific regions.

The final equations for the two-column case are simply extensions from the one-column case, and therefore I will not go through the derivations a second time:

$$\sigma T_{g_1}^4 = (\sigma T_{e_1}^4 - k(T_{g_1} - T_{g_2}))(1 + \frac{\tau(T_{g_1})}{2}) \tag{32a}$$

$$\sigma T_{g_2}^4 = (\sigma T_{e_2}^4 - k(T_{g_2} - T_{g_1}))(1 + \frac{1 + \tau(T_{g_2})}{2}). \tag{32b}$$

As before, the optical depth τ is written as a function of temperature according to the equations

$$\tau(T_{g_1}) = a + b\left(e^{c\left(\frac{1}{288} - \frac{1}{T_{g_1}}\right)}\right) \tag{33a}$$

$$\tau(T_{g_2}) = a + b\left(e^{c\left(\frac{1}{288} - \frac{1}{T_{g_2}}\right)}\right). \tag{33b}$$

4 Preliminary Results

4.1 One-Column RE

We begin by showing the solutions for the ground temperature T_g of the RE equations (equations 14) as a function of the emitting temperature T_e, as shown in figure 6. There are a couple features worth noting; first, the shape of the solution resembles a backward C-curve. This implies that for temperatures below about 268K, there are two equilibrium solutions. The bottom branch is the branch we are currently on (for our current emitting temperature of 255K, the ground temperature is the expected 288K), and the top branch shows a much higher ground temperature for a given emitting temperature.

Another important feature is that there is a critical temperature (in this graph at 268K) above which there are no solutions to the equations. That is, there are no equilibrium points in the system, and we thus deduce that the system is in a runaway greenhouse regime, where the planet is endlessly warming.

We are also interested in the stability of these equilibrium points. If we start by considering a time-dependent equation:

$$C\frac{dT_g}{dt} = (1 + \frac{\tau}{2})\sigma T_e^4 - \sigma T_g^4, \tag{34}$$

and linearizing, we find that the bottom branch (the branch we are currently on) is stable, whereas the top branch is unstable.

In the one-column RE model, we also added in a term that allowed for additional heat loss (see equation 15), and the solutions in this case are shown in figure 7. Allowing for additional heat loss creates an added upper branch to the C-curve diagram, but one that curves upward rather steeply. This branch arches into an extremely high ground temperature range, and so is likely not of great importance to our current Earth (thankfully), but this does reveal an interesting aspect of the system that we are studying. If we take another look at equation 15, we notice that for high ground temperature, the kT term will dominate. The different shape of the curve, then, at high ground temperatures is not a surprise.

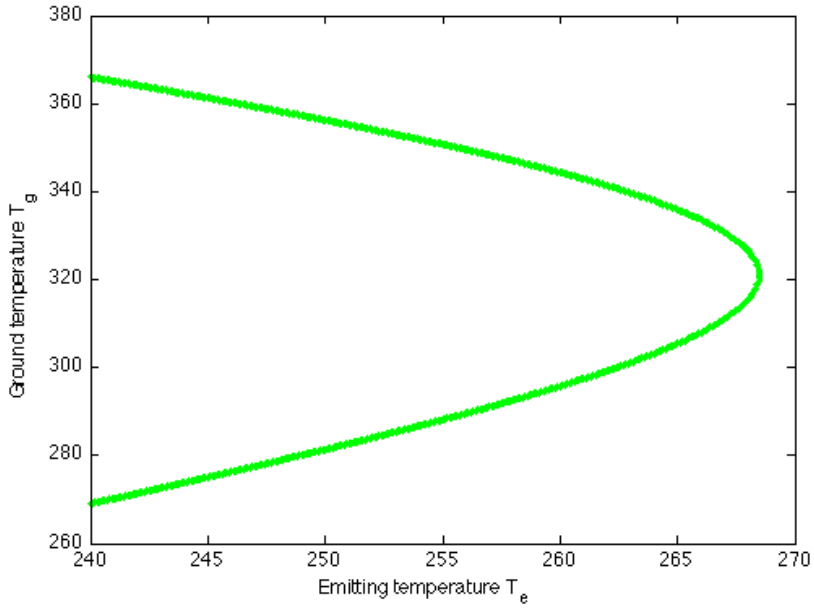


Figure 6: Emitting temperature plotted vs. ground temperature for the RE model.

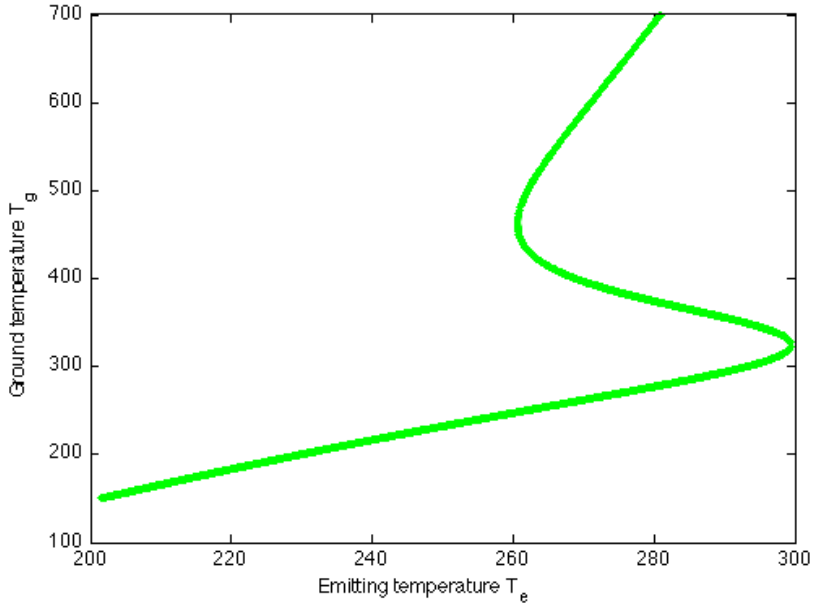


Figure 7: Emitting temperature plotted vs. ground temperature for the RE model with additional heat loss term.

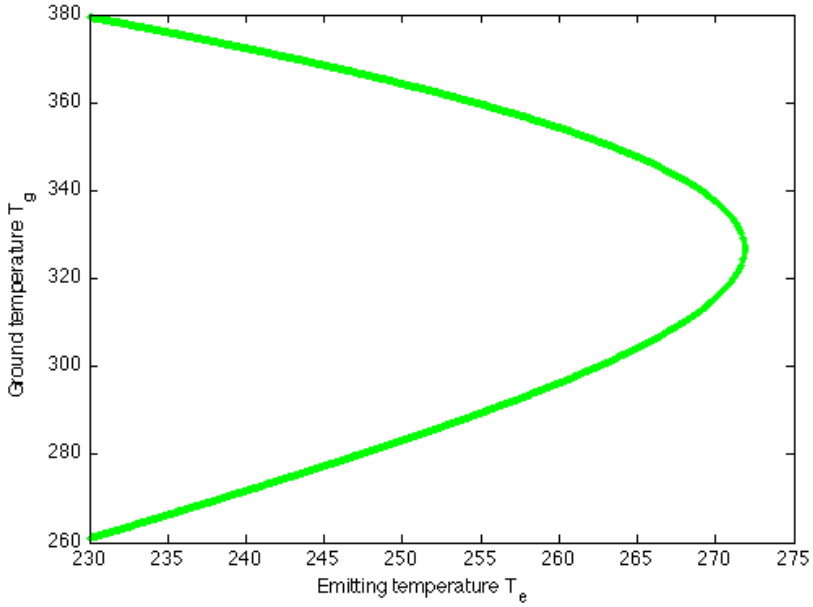


Figure 8: Emitting temperature plotted vs. ground temperature for the RCE model.

It is also worth noting that the tip of the C-curve is at a much higher critical temperature of nearly 300K. This also makes intuitive sense, as the extra term is allowing for heat to be taken out of the system, allowing for an equilibrium solution for greater amounts of incoming solar radiation.

4.2 One-Column RCE

In the RCE model, we observe the same C-shaped curve (see figure 8) as we did in the RE case. This time, however, the critical emitting temperature before the onset of a runaway greenhouse effect is a couple of degrees higher, around 272K. This is a bit misleading, however, because we have increased the constants in the equation for optical depth, and so the comparison must take this into account. In figure 9, we show the RE C-curve using the values for the optical depth in the RCE model. We see here that the critical emitting temperature has dropped to under 200K. When taking this into account, we see that adding convection to the model greatly mitigates the runaway greenhouse effect.

Since we had a second variable, H_T in the RCE model, we show the solution of the tropopause height also for varying emitting temperatures (see figure 10). The shape of the curve is still C-like, though slightly asymmetric. Still, there exist two solutions for a range of emitting temperatures, and above a critical value there exist no solutions.

This work is meant to be a study of the system described in previous sections, and so we present here a couple of parameter studies, meant to shed light on the behavior of our model. Figure 11 shows the model's dependence on the lapse rate Γ. The three curves plotted are the full solution for the ground temperature for 3 different values of the lapse rate, shown in the plot's legend. Similarly, figure 12 depicts the height of the tropopause

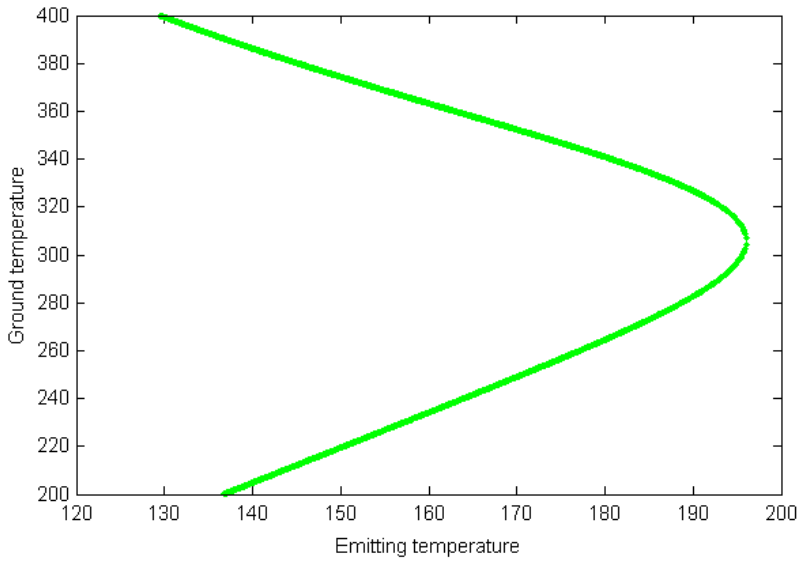


Figure 9: Emitting temperature plotted vs. ground temperature for the RE model when using the values for the RCE optical depth.

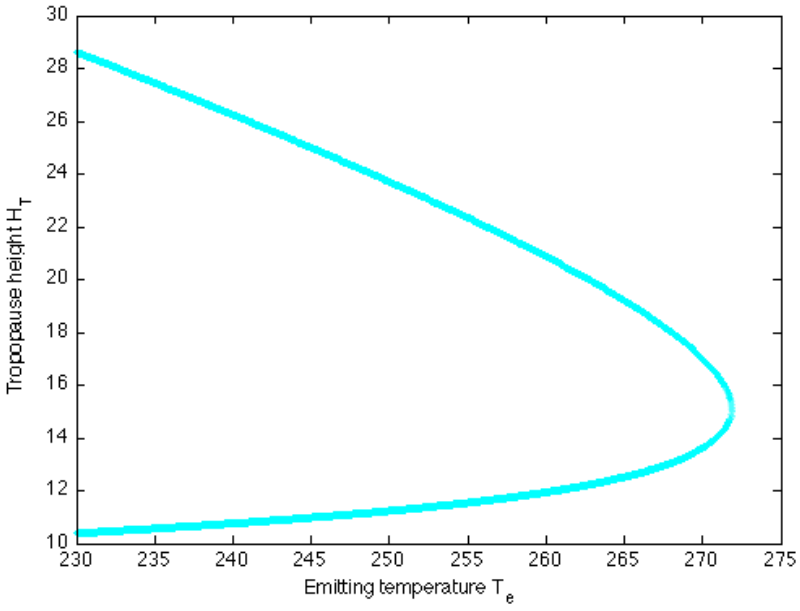


Figure 10: Emitting temperature plotted vs. ground temperature for the RCE model.

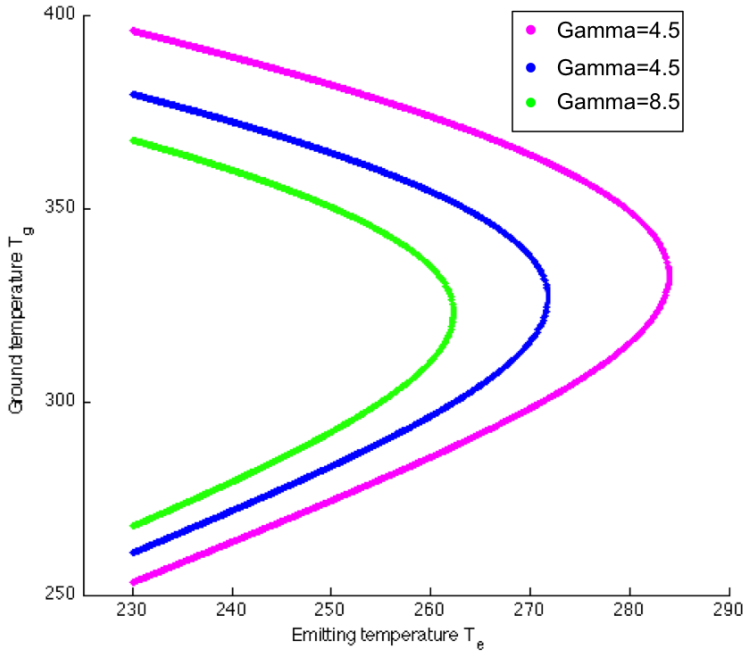


Figure 11: Emitting temperature plotted vs. ground temperature for the RCE model with three values of gamma.

for the same three values of the lapse rate. In both cases, as the lapse rate decreases, the critical emitting temperature increases and the range of ground temperatures increases, as well. In the previous plots, the lapse rate used was a constant 6.5 K/km. However, in reality the lapse rate also depends on temperature, and taking this into account would be a logical next step in this study.

Figure 13 illustrates the relationship between the emitting temperature and the lapse rate. As mentioned in the previous paragraph that the critical temperature increases with decreasing lapse rate, this figure shows the whole spectrum. According to the plot, today's emitting temperature of 255K would be critical if the lapse rate were a constant value of 10 K/km across the globe. We portray the same information in a different way in figure 14. The vertical axis still shows the lapse rate, but the horizontal axis shows the fractional distance to the sun (according to the Earth-sun distance) needed to obtain the critical emitting temperature. For example, the Earth is located at 1 Earth-sun distance and would reach critical emitting temperature at a lapse rate of 10 K/km. Venus, located at about 0.72 Earth-sun distance away from the sun requires a lapse rate of just over 2 K/km to achieve a runaway greenhouse effect. For the global average lapse rate of around 6.5 K/km [9], as used in this work, the Earth would need to be at roughly 0.9 Earth-sun distance to be at a critical emitting temperature, assuming all other elements of the system remain the same.

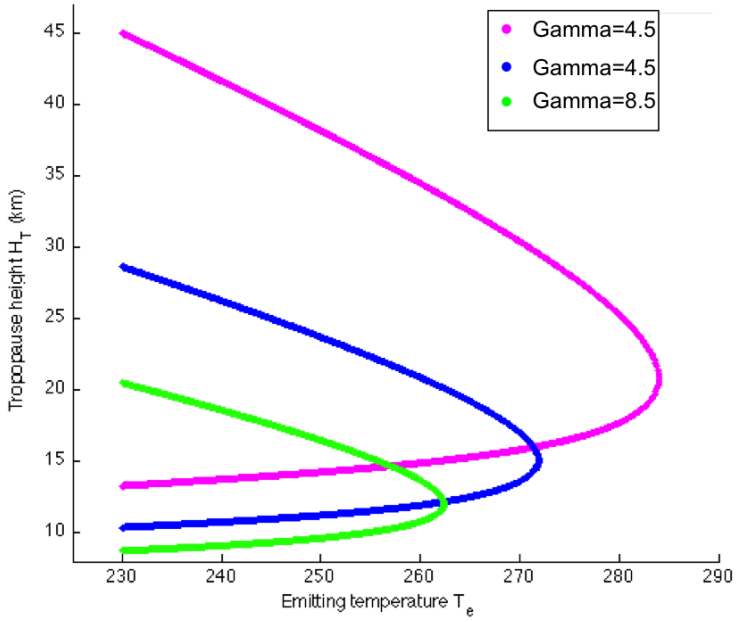


Figure 12: Emitting temperature plotted vs. tropopause height for the RCE model with three values of gamma.

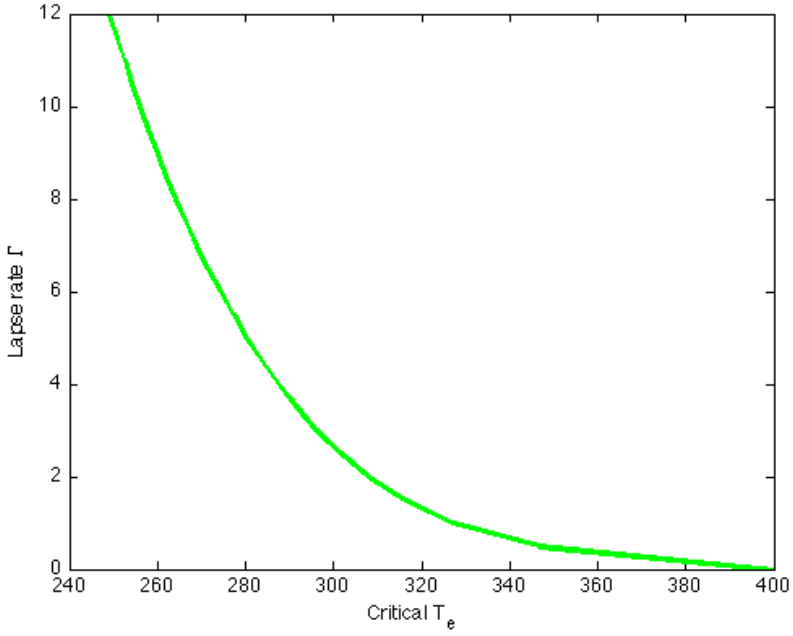


Figure 13: Critical emitting temperature plotted vs. lapse rate for the RCE model.

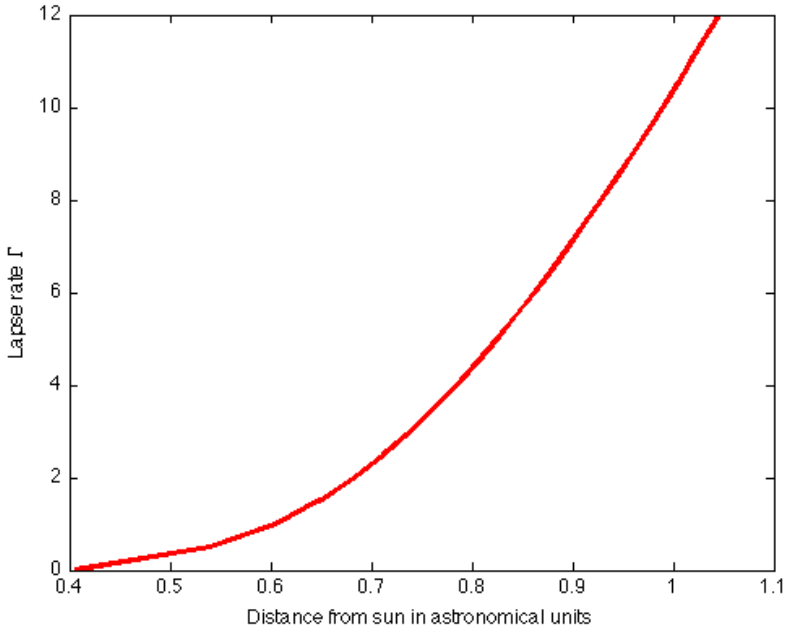


Figure 14: Distance to the sun vs. lapse rate for the RCE model.

4.3 Two-Column RE

Finally, we take a look at the two-column radiative equilibrium case. Figure 15 shows two plots, one for each of the two columns. Once again, we see the familiar C-curves, though this time there appear to be more than one C-curve in each column. Even though the plots do not show two full C-shaped curves, they both indicate that the full solution may actually be made up of two C-curves, as we postulated. This is more obvious in the top graph, for column 1 (the warmer temperature, mimicking the tropics). The outer C-curve is visible, and the upper branch of what appears to be a second C-curve is visible. We postulate that there is a missing lower branch to the inner C-curve, which may be missing due to numerical difficulties. Having two C-curves makes intuitive sense, since each column displays a C-curve separately as we showed in the one-column RE case, and with the exchange of heat between the columns, each column now portrays two C-curves in the solution. More work is certainly required to confirm our speculations.

Regarding the runaway greenhouse effect and its onset, the two-column model does not show a great improvement; i.e. the critical emitting temperature is only slightly higher than the 268K that we obtained from the single column case. The first column has a very similar value of around 269K, while the second column appears to have a slightly lower critical emitting temperature, which we infer to be in the neighborhood of 265K. However, the system is quite different with two columns, and the emitting temperature T_e is defined as 20 Kelvin lower in the second column than in the first. So it is possible that the two-column RE approach mitigates the runaway greenhouse effect very slightly, but the system is different enough and the numbers are not robust enough at this time to draw any further

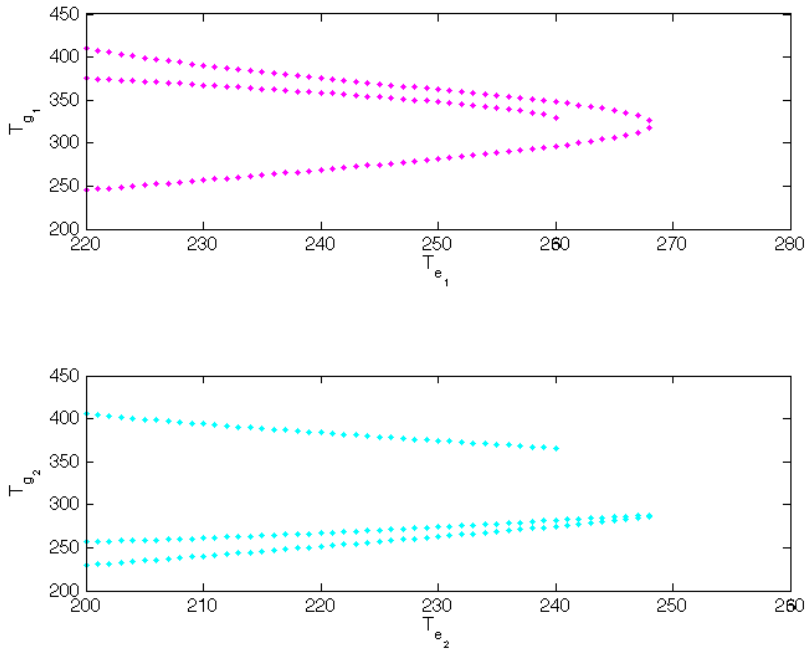


Figure 15: Emitting temperature plotted vs. ground temperature for both columns in the two-column RE model. Upper graph: column 1 (the tropics); Lower graph: column 2 (the midlatitudes).

conclusions.

5 Conclusions and Future Work

These results are preliminary, as this is but the starting framework for understanding the Earth system's behavior regarding the runaway greenhouse effect and lateral heat transport. However, there are a couple key features that are worth pointing out. We find that in both the RE and RCE models, there exists a critical emitting temperature, above which there are no equilibrium solutions, and the system experiences a runaway greenhouse effect.

In the one-column RCE model, we find a critical emitting temperature of around 268K. In the one-column RCE model, it is a couple degrees higher, around 272K. However, because we use a different value of the optical depth for the RCE than the RE model, these numbers are not easily comparable. In fact, if using the more realistic RCE value of the optical depth, we see that our planet would already be in a runaway greenhouse regime in the RE model. So, convection mitigates the runaway greenhouse effect quite a bit, raising the critical emitting temperature from just under 200K in the RE model to 272K in the RCE model.

Adding a second column hardly alters the critical emitting temperature, contrary to

what we expected. With a warmer column pumping heat into a cooler column, we anticipated a higher critical emitting temperature in the first column, and thus a higher threshold for the onset of the runaway greenhouse effect. However, the two-column model is still in need of tuning and we will thus continue developing the system and hopefully better understand it in a future work.

Beyond a more in-depth analysis of the two-column RE model in the future, we would like to add an analysis of a two-column RCE model as well. We would also like to express the lapse rate as a function of temperature, in addition to the temperature-dependence of the optical depth, as we have done here. Finally, we would like to perform more parameter studies in order to have a more thorough understanding of the Earth system's behavior.

6 Acknowledgements

I would like to thank Geoff Vallis for his patience and guidance while advising me on this project. I would also like to acknowledge Andy Ingersoll and Steve Meacham for helpful discussions. Finally, I thank the directors and participants in the 2014 Woods Hole Oceanographic Institute's GFD Summer Program for this wonderful research opportunity.

References

[1] M. Budyko, *The effect of solar radiation variations on the climate of the earth*, Tellus, 21 (1969), pp. 611–619.

[2] K. Emanuel, *Gfd 2014 lecture notes.*

[3] R. Goody and Y. Yung, *Atmospheric Radiation: Theoretical Basis*, Oxford University Press, 1989.

[4] A. P. Ingersoll, *The runaway greenhouse: A history of water on venus*, Journal of the Atmospheric Sciences, 26 (1969), pp. 1191–1198.

[5] W. Ingram, *A very simple model for the water vapour feedback on climate change*, Q. J. R. Meteorol. Soc., 136 (2010), pp. 30–40.

[6] R. A. Lindzen, *Dynamics in Atmospheric Physics*, Cambridge University Press, 1993.

[7] G. R. North, *Theory of energy-balance climate models*, Journal of Atmospheric Sciences, 32 (1975).

[8] W. D. Sellers, *A global climatic model based on the energy balance of the earth-atmosphere system*, Journal of Applied Meteorology, 8 (1969).

[9] G. K. Vallis, *Six easy pieces concerning the structure of the atmosphere and ocean*, 2014. Walsh Cottage Lectures.

A Numerical Study of the Downstream Development of a Baroclinic Instability

Daniel Mukiibi

August 16, 2015

Abstract

The downstream development of baroclinic instability is studied in a 2-layer non-linear Quasi-Geostrophic (QG) model with a semi-infinite downstream extense and rigid meridional walls. Starting with a baroclinic current in a channel, a perturbation is invoked at the entrance of the channel upstream and its spatial and temporal downstream development is studied. For matters of simplicity, the study considers 2 y-modes which offers two advantages: it's simple enough to follow the 2 modes easily, and it also gives insight into the more complicated scenario of having more than a mode leading to interaction of the different modes and hence modifying the dynamics of the flow. The boundary condition at the channel entrance upstream is a temporally oscillating perturbation at $x = 0$; downstream, the potential vorticity is zero at $x = \infty$. In the $y-$direction, the derivatives of the stream function (i.e velocity) at the meridional walls are zero.

It has been found by Pedlosky (2011) that in a simple finite amplitude model of a spatially developing baroclinic instability, there's a regime during which the spatial and temporal evolution of the instability amplitude along x, t characteristics exhibits chaotic behaviour. This chaotic behaviour resulting from periodic initial conditions at the channel entrance leads to sharp and abrupt spatial variations downstream. In the current study, the non-linear development of a baroclinic instability is studied numerically in a slightly more complicated but compelling manner. In the current study, persistence of these features and the circumstances over which this behaviour persists in a more realistic oceanic model forms the main motivation for the study.

The Model

Consider the Quasi-Geostrophic(QG) equations for a two-layer model similar to the one studied in (Pedlosky, 1970) and (Pedlosky, 2011). Here, the formulation will be given in terms of the potential vorticity, q. The equations of motion in the two layers are given as below:

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_1\frac{\partial}{\partial x}\right)q^{(1)}+Q_{y1}\frac{\partial\psi^{(1)}}{\partial x}+\mathrm{J}(\psi^{(1)},q^{(1)})=-r\nabla^2\psi^{(1)}. \tag{1}$$

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_2\frac{\partial}{\partial x}\right)q^{(2)}+Q_{y2}\frac{\partial\psi^{(2)}}{\partial x}+\mathrm{J}(\psi^{(2)},q^{(2)})=-r\nabla^2\psi^{(2)}. \tag{2}$$

Where the jacobian, J, is defined as $J(a,b)=\dfrac{\partial a}{\partial x}\dfrac{\partial b}{\partial y}-\dfrac{\partial b}{\partial x}\dfrac{\partial a}{\partial y}$ and for a channel with no topography as is the case in this study,

$$Q_1=\nabla^2\psi_1-\mathrm{F}_1(\psi_1-\psi_2)+\beta y.$$
$$Q_2=\nabla^2\psi_1+\mathrm{F}_2(\psi_1-\psi_2)+\beta y.$$

Proposing a truncated Fourier series solution to the equations (1) and (2) of the form

$$\begin{pmatrix} q^{(1)} \\ \psi^{(1)} \end{pmatrix}=\begin{pmatrix} q_1^{(1)}(x,t)\sin\pi y+q_2^{(1)}(x,t)\sin 2\pi y+\cdots \\ \psi_1^{(1)}(x,t)\sin\pi y+\psi_2^{(1)}(x,t)\sin 2\pi y+\cdots \end{pmatrix}. \tag{3}$$

$$\begin{pmatrix} q^{(2)} \\ \psi^{(2)} \end{pmatrix}=\begin{pmatrix} q_1^{(2)}(x,t)\sin\pi y+q_2^{(2)}(x,t)\sin 2\pi y+\cdots \\ \psi_1^{(2)}(x,t)\sin\pi y+\psi_2^{(2)}(x,t)\sin 2\pi y+\cdots \end{pmatrix}. \tag{4}$$

Substituting equation (3) into equation equation (1) and projecting onto $\sin\pi y$ and $\sin 2\pi y$ yields equations (5) and (6) respectively

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_1\frac{\partial}{\partial x}\right)q_1^{(1)}+Q_{y1}\frac{\partial\psi_1^{(1)}}{\partial x}-\frac{\pi^2}{2}q_2^{(1)}\frac{\partial\psi_1^{(1)}}{\partial x}+\frac{\pi^2}{4}q_1^{(1)}\frac{\partial\psi_2^{(1)}}{\partial x}\quad-$$
$$\frac{\pi^2}{4}\psi_1^{(1)}\frac{\partial q_2^{(1)}}{\partial x}+\frac{\pi^2}{2}\psi_2^{(1)}\frac{\partial q_1^{(1)}}{\partial x}=-rq_1^{(1)}. \tag{5}$$

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_1\frac{\partial}{\partial x}\right)q_2^{(1)}+Q_{y1}\frac{\partial\psi_2^{(1)}}{\partial x}+\frac{\pi^2}{4}q_1^{(1)}\frac{\partial\psi_1^{(1)}}{\partial x}-\frac{\pi^2}{4}\psi_1^{(1)}\frac{\partial q_1^{(1)}}{\partial x}=-rq_2^{(1)}. \tag{6}$$

In the reduced model, the equations (5)and (6) are the equations of motion in layer 1. In a similar way, the equations of motion in layer 2 are obtained as :

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_2\frac{\partial}{\partial x}\right)q_1^{(2)}+Q_{y2}\frac{\partial\psi_1^{(2)}}{\partial x}-\frac{\pi^2}{2}q_2^{(2)}\frac{\partial\psi_1^{(2)}}{\partial x}+\frac{\pi^2}{4}q_1^{(2)}\frac{\partial\psi_2^{(2)}}{\partial x}\quad-$$
$$\frac{\pi^2}{4}\psi_1^{(2)}\frac{\partial q_2^{(2)}}{\partial x}+\frac{\pi^2}{2}\psi_2^{(2)}\frac{\partial q_1^{(2)}}{\partial x}=-rq_1^{(2)}. \tag{7}$$

$$\left(\frac{\partial}{\partial t}+\mathrm{U}_2\frac{\partial}{\partial x}\right)q_2^{(2)}+Q_{y2}\frac{\partial\psi_2^{(2)}}{\partial x}+\frac{\pi^2}{4}q_1^{(2)}\frac{\partial\psi_1^{(2)}}{\partial x}-\frac{\pi^2}{4}\psi_1^{(2)}\frac{\partial q_1^{(2)}}{\partial x}=-rq_2^{(2)}. \tag{8}$$

$q \equiv$ potential vorticity, $\psi \equiv$ stream function, $r \equiv$ dissipation coefficient.

Linearising the equations (5),(6),(7) and (8) yields a slightly modified set of equations from which the linear stability of the equations can be investigated. Proposing potential vorticities and stream functions of the form below:

$$q_j^{(\ell)}=\hat{q}_j^{(\ell)}\mathrm{e}^{ik(x-ct)} \quad \text{and} \quad \psi_j^{(\ell)}=\hat{\psi}_j^{(\ell)}\mathrm{e}^{ik(x-ct)} \quad \text{for} \quad \ell=1,2; \quad j=1,2. \tag{9}$$

where x is the downstream coordinate and k is the x-direction wave number.
The superscripts in equation (9) denote layers while the subscripts denote the wave mode under consideration. The linearised forms of equations (5) and (7) can be presented in matrix form as

$$\begin{pmatrix} r+ik(\mathrm{U}_1-c)-\dfrac{ikQ_{y1}(\mathrm{K}_1^2+\mathrm{F}_2)}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)} & \dfrac{-ikQ_{y1}\mathrm{F}_1}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)} \\ \dfrac{-ikQ_{y2}\mathrm{F}_2}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)} & r+ik(\mathrm{U}_2-c)-\dfrac{ikQ_{y2}(\mathrm{K}_1^2+\mathrm{F}_1)}{\mathrm{K}_1^2(K_1^2+\mathrm{F}_1+\mathrm{F}_2)} \end{pmatrix}\begin{pmatrix}\hat{q}_1^{(1)}\\ \hat{q}_1^{(2)}\end{pmatrix}=0 \tag{10}$$

where $\mathrm{K}_1^2=k^2+\pi^2$ and k is the x-direction wave number.

The linearised forms of equations (6) and (8) take a similar form and can be presented in matrix form in the same form as equation (10) but with K_1 swapped for $\mathrm{K}_2=k^2+2\pi^2$.

For non-trivial solutions, the determinant of the matrix in equation (10) must vanish, thus,

$$\left(r+ik(\mathrm{U}_1-c)-\frac{ikQ_{y1}(\mathrm{K}_1^2+\mathrm{F}_2)}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)}\right)\left(r+ik(\mathrm{U}_2-c)-\frac{ikQ_{y2}(\mathrm{K}_1^2+\mathrm{F}_1)}{\mathrm{K}_1^2(K_1^2+\mathrm{F}_1+\mathrm{F}_2)}\right)- \left(\frac{-ikQ_{y1}\mathrm{F}_1}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)}\right)\left(\frac{-ikQ_{y2}\mathrm{F}_2}{\mathrm{K}_1^2(\mathrm{K}_1^2+\mathrm{F}_1+\mathrm{F}_2)}\right)=0 \tag{11}$$

For matters of simplicity, discarding the β-effect from our consideration and also defining the shear as $\mathrm{U}_s=\mathrm{U}_1-\mathrm{U}_2$. Also considering that $\mathrm{F}_1=\mathrm{F}_2=\mathrm{F}$, then

$$\begin{aligned} Q_{y1}=\mathrm{F}\mathrm{U}_s, \quad Q_{y2}=-\mathrm{F}\mathrm{U}_s \quad \text{thus} \\ Q_{y1}+Q_{y2}=0, \quad Q_{y1}\cdot Q_{y2}=-\mathrm{F}^2\mathrm{U}_s^2 \end{aligned} \tag{12}$$

Equation (11) generates a quadratic equation in c whose solution is obtained to be

$$c=\left(\frac{-ir}{k}+\frac{1}{2}(\mathrm{U}_1+\mathrm{U}_2)\right)\pm\frac{\mathrm{U}_s}{2}\left(1-4\mathrm{F}\left(\frac{\mathrm{F}^3}{y^2}-\frac{\mathrm{F}x^2}{y^2}+\frac{x}{y}\right)\right)^{1/2} \tag{13}$$

$$\text{where} \qquad x=\mathrm{K}_1^2+\mathrm{F}, \qquad y=\mathrm{K}_1^2(\mathrm{K}_1^2+2\mathrm{F}) \quad \text{and thus}$$

$$c-\mathrm{U}_B=\frac{-ir}{k}\pm\frac{\mathrm{U}_s}{2}\left(\frac{\mathrm{K}_1^2-2\mathrm{F}}{\mathrm{K}_1^2+2\mathrm{F}}\right)^{1/2}. \tag{14}$$

Thus c is generally complex and can be written as $c = c_r + ic_i$ where c_r and c_i are the real and imaginary parts of c respectively. In case the term in the last parentheses vanishes, then the dispersion relation reduces to

$$c = \left(\frac{-ir}{k} + \mathrm{U}_B\right) \qquad \text{where} \quad \mathrm{U}_B = \frac{1}{2}(\mathrm{U}_1 + \mathrm{U}_2), \quad \text{the barotropic flow.} \tag{15}$$

It is also observed that the decay of the perturbation is proportional to the dissipation in the system and is higher for lower wave numbers. In this case, the shear does drop out of the dispersion relation rendering the dissipation, r, and the magnitude of the barotropic flow as the parameters governing the flow dynamics in the channel under study.

However if $\mathrm{K}_1 < (2\mathrm{F})^{1/2}$, then the last parentheses in equation (14) also contribute to the complex part of c and in turn the shear, U_s and the Froude number, F become effective parameters of the system.

In the following, simulations will be conducted in two main categories; the first being the case when the dissipation, $r = O(1)$ and the second being the case when the dissipation is almost zero in the system i.e $r = O(\Delta)$ for very small Δ.

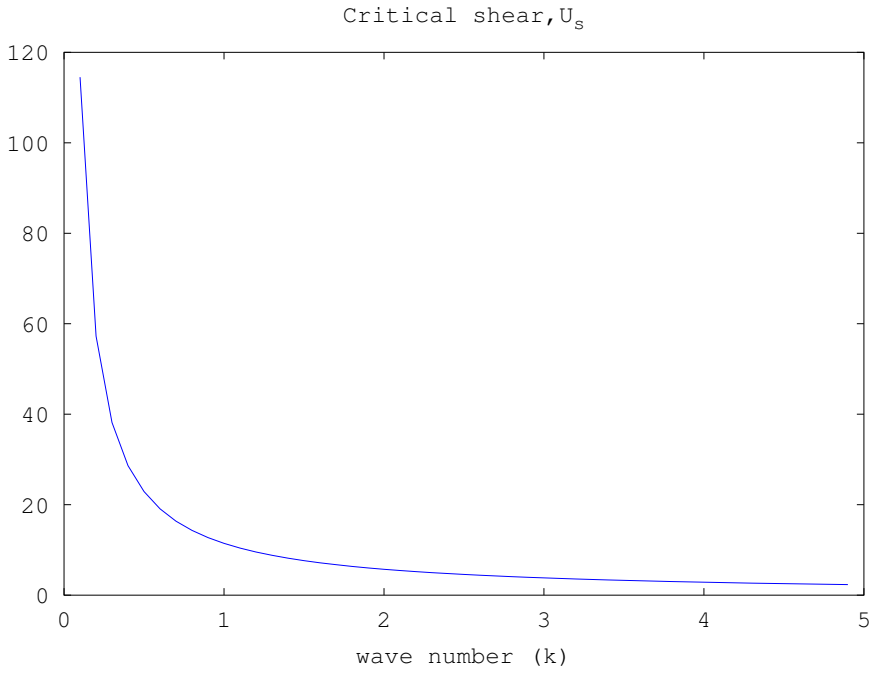
(i) $r = O(1)$
In this case, the only way of having a growing instability is that $2\mathrm{F} > \mathrm{K}_1^2$ and the whole term in the parentheses must be large enough to outweigh the decay term, $\frac{-ir}{k}$. This yields the relation for the marginal condition on U_s in order to have a growing instability. With values selected as $r = 4.6, \mathrm{F} = 40, l = \pi \quad \text{and} \quad 0 < k < 5.0$

$$\mathrm{U}_s = \frac{2r}{k}\left(\frac{2\mathrm{F} + \mathrm{K}_1^2}{2\mathrm{F} - \mathrm{K}_1^2}\right)^{1/2} \tag{16}$$

(ii) $r = O(\Delta)$
With r small, the marginal curve is given in terms of the parameter, F and takes the form

$$\mathrm{F} = \frac{\mathrm{K}_1^2}{2} = \frac{k^2 + l^2}{2} \tag{17}$$

F thus has a parabolic form with a minimum at $k = 0$ and takes the form below:



Numerical Simulations and Results

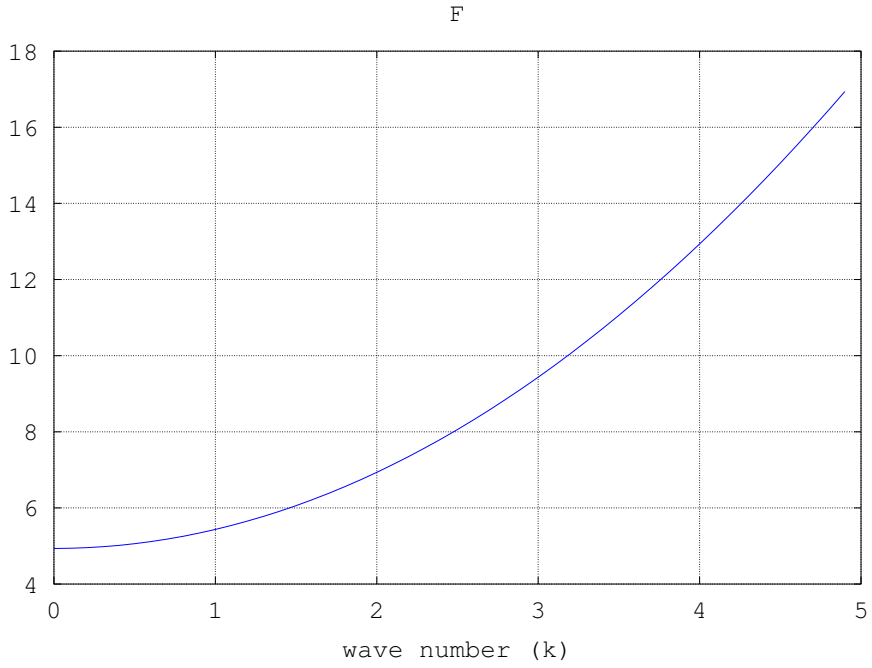
The case $r = O(1)$

The model set up is such that the system is slightly above its neutral criticality. The barotropic velocity in the channel is set at $U_B = 13.125$ and the most unstable mode is to be k = 4.34 . The parameter values of the model when neutrally critical and those used in the numerical simulation are given in table 1.

Table 1: Model parameters used for the case $r = O(1)$

Parameter	Symbol	Critical values	Simulation value
Shear	U_s	3.00	3.25
Froude number	F	40.0	40.0
Dissipation	r	4.60	4.60

With the shear slightly increased above the neutrally critical value of the shear, an initial perturbation is set up at the entrance of the channel (i.e at $x = 0$) and it's downstream development with time is studied. The boundary conditions are such that at the channel entrance, a perturbation, $A_o \sin(\omega t)$ with frequency $\omega = kU_B$ is seeded. At the end of the channel, spongy boundaries are considered such that whatever is incident on the wall, goes in and none of it is reflected back into the channel. The meridional walls are rigid and the potential vorticities there are zero*.



Discussion

It's observed that to first order, the solution has both the barotropic and baroclinic modes on $\sin(\pi y)$ as the leading terms of the solution to the nonlinear set of equations (5,6,7 and 8). The leading terms of the solution exhibit the oscillations of the initial perturbation imposed at the entrance of the channel (see figures 1 and 3) all the way downstream. The amplitude of the perturbation grows initially until it reaches a finite amplitude and thereafter momentarily stabilises before eventually decaying off to zero. The stabilisation in growth of the perturbation at finite amplitude is longer downstream.

In all cases and for all the modes considered, it is observed that the part of the growing perturbation behind the front reaches finite amplitude before saturation such that in the regions ahead of the front, the amplitude of the perturbation remains constant in the vicinity of the front and decays quite quickly away from the front downstream. The slightly unique cases amongst the modes considered in this study are those of the baroclinic and barotropic modes on $\sin(2\pi y)$ which appear to be relatively smooth compared to the leading order terms. With the exception of a few oscillations whose amplitudes are still small near the channel entrance, any information about the oscillatory nature of the perturbation is lost downstream and the resulting correction to the mean flow is of a largely smooth structure.

Behind the front, wiggles are observed in the spatial and temporal structure of the perturbation before finally reaching finite amplitude in the vicinity of the front. Ahead of the front, the perturbation has already reached finite amplitude and therefore remains constant (for longer time scales) or immediately decays (for shorter time scales) ahead of the front. This is in agreement with Pedlosky (2011) who highlighted that the correction to the mean flow carries the oscillatory information of the perturbation only behind the front

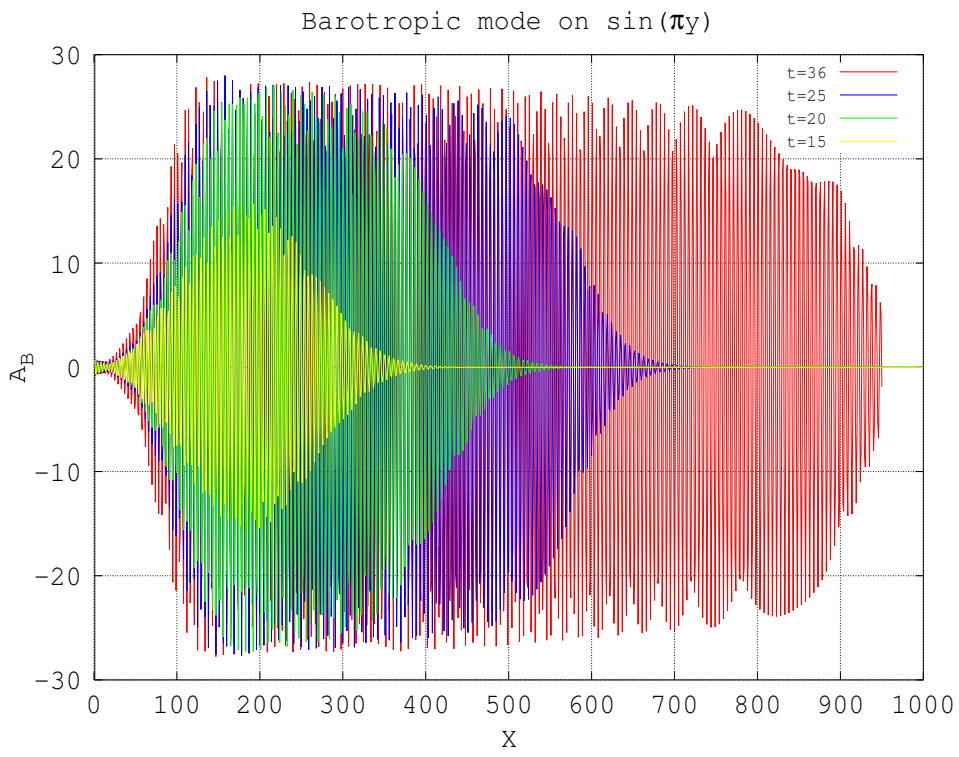


Figure 1: Snapshots of the barotropic potential vorticity on the sin(πy) mode with $U_s = U_{so} + 0.25$ and $r = r_o$

during which time the perturbation also attains finite amplitude. In that study, they also observe that ahead of the front, the structure is basically a smooth one with the perturbation having reached finite amplitude.

However, it is also observed that the baroclinic mode on sin(2πy) is also the largest of the projections on the sin(2πy) modes. This is consistent with the results obtained by (Pedlosky, 2011) which highlighted that the first order correction to the mean flow was fully baroclinic.

However, the non-linear simulations carried out in this study reveal that there's a small contribution to the mean flow correction by the barotropic component carried by the barotropic component on the sin(2πy) mode. It's the smallest of all the components but it is worth mentioning that although the asymptotic approach adopted using the finite amplitude model in Pedlosky (2011) fails to capture this contribution, it is not necessarily zero as observed in figure 2.

A probable explanation as to why the asymptotic approach shows that the sin($2\pi y$) barotropic mode does not contribute to the mean flow correction could be that as observed from figure (2), the average of this mode over a period is zero. So, it could be asserted that the reason for the failure to capture to this mode in the theory is not because it's so small in magnitude but it's because it vanishes on average. More interestingly for all time, apart

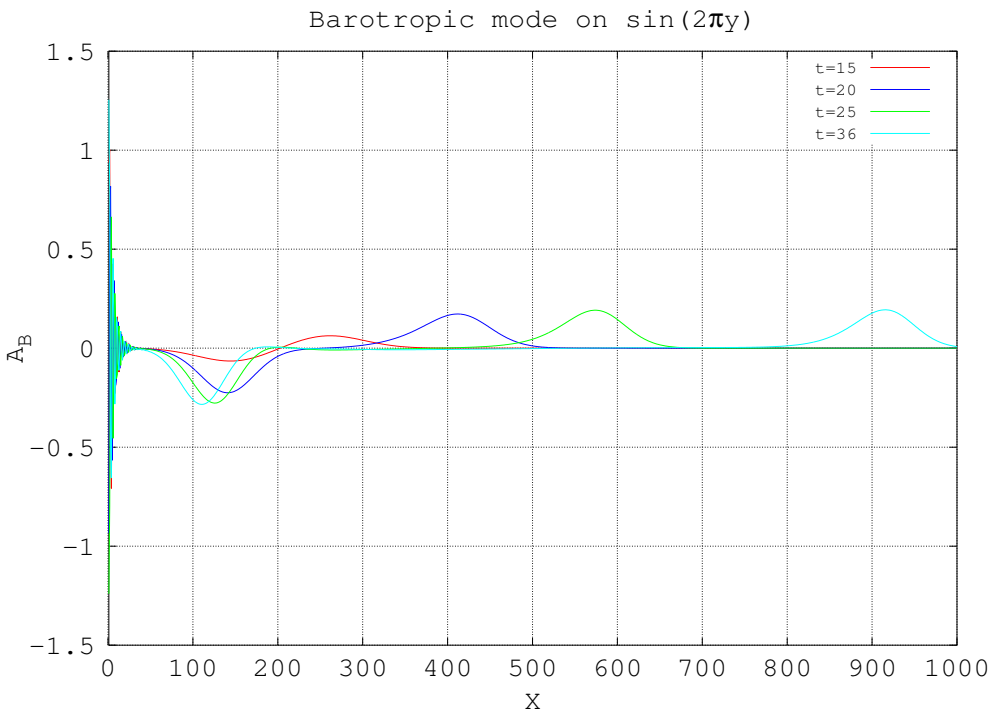


Figure 2: Snapshots of the baroclinic potential vorticity on the sin(2πy) mode with $U_s = U_{so} + 0.25$ and $r = r_o$

from the initial transients, the $\sin(2\pi y)$ barotropic mode manifests as a periodic oscillation whose average over a period vanishes.

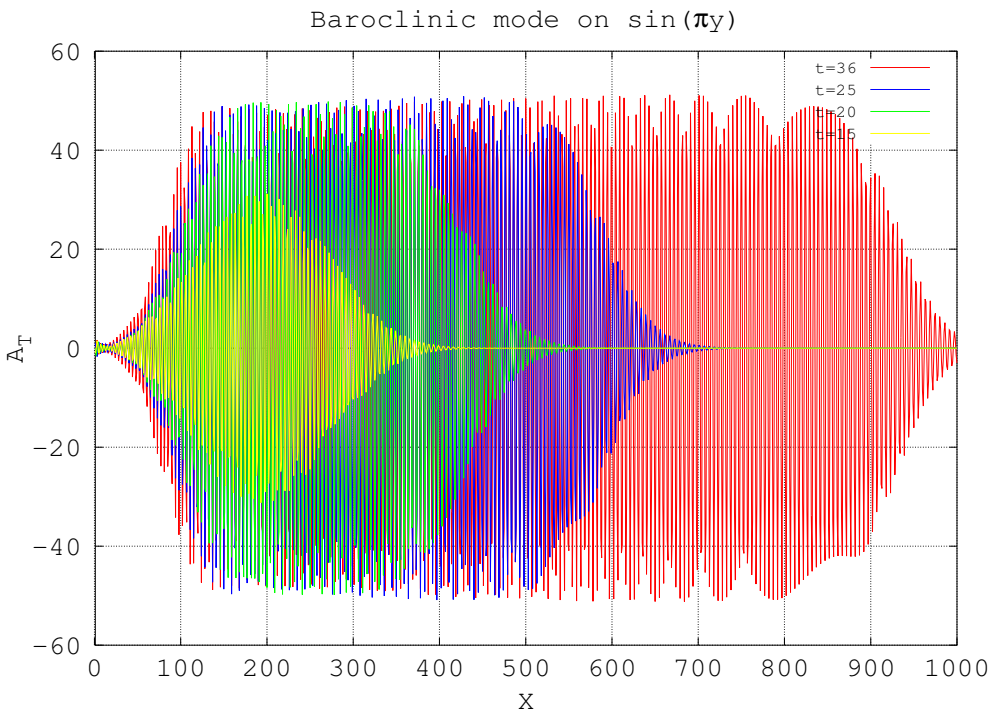


Figure 3: Snapshots of the baroclinic potential vorticity on the sin(πy) mode with $U_s = U_{so} + 0.25$ and $r = r_o$

The case r=O(Δ)

Table 2: Model parameters used for the case $r = O(\Delta)$

Parameter	Symbol	Critical values	Simulation value
Shear	U_s	1.30	1.30
Froude number	F	4.9348	4.9348 + 0.02
Dissipation	r	0.001	0.001

The barotropic flow in this case is reduced to $U_B = 1.65$ and the most unstable mode corresponds to the wave number $k = 0.15$.

Several simulations are carried out with various degrees of super criticalities (i.e for increasing values of Δ) and the results are presented in the following figures. For most values of Δ, the flow does not seem to change significantly but it happens that as Δ increases, more features emerge ahead of the front for longer times.

As the theory predicts, the largest component of the correction to the leading order solution is baroclinic (the sin$2\pi y$ baroclinic mode). However, the fully non-linear solution shows that a barotropic contribution is also present. The latter is initially small (≈ 0) but develops with time until it is one order of magnitude lower than the baroclinic correction.

Discussion

It's observed that at the leading order, the dominant part of the flow is the $\sin\pi y$ barotropic mode. The $\sin\pi y$ baroclinic mode is lower than the former but it's significantly appreciable. This is in agreement with the findings of Pedlosky (2011) that showed that in this regime, the leading order solution is the $\sin\pi y$ mode and that its baroclinic counterpart is an order of magnitude lower. Although not as much as an order of magnitude, the fully nonlinear solutions strongly yield similar results.

At the next order, the major correction component to the mean flow is found to be fully baroclinic (i.e the $\sin 2\pi y$ baroclinic mode). The $\sin 2\pi y$ barotropic mode is consistently zero throughout the time of the simulation. This is also in agreement with the findings from the multi-scale asymptotics which yielded that the correction to the mean flow is fully baroclinic.

Increasing the degree of super criticality leads to a complete break down of the predictions of the linear and weakly non-linear theory. In this case, at leading order, the dominant term is the $\sin\pi y$ baroclinic mode as opposed to the $\sin\pi y$ barotropic mode predicted by theory. Also, at the next order, the barotropic correction to the mean flow becomes appreciable which is, of course, another difference from the case considered when the dynamics are slightly super critical. The other remarkable feature that emerges with increasing levels of supercriticalities is that the features formed ahead of the front become more apparent and highly variable downstream as one would expect when the non-linearirities in the system are at full operation.

In conclusion, the findings from this study qualitatively show that the degree of dissipation in the system is a major determinant of the dynamics of the flow. When the system is substantively dissipative, the marginal curve is given in terms of the shear and the dominant correction component to the mean flow is largely baroclinic. In the case when the dissipation is so small, the marginal curve is expressed in terms of the parameter, F - the Froude number. In this case, the lowest order compnent is found to be barotropic and the correction is fully baroclinic.

For further study, it would be meaningful to consider using a periodic channel so that the flow statistics can be obtained with a good degree of accuracy to enable giving a quantitative account of the dynamics of the flow and how the different components exchange the energy in both spatial and temporal considerations. Of course, inclusion of the $\beta-$ effect would also serve the purpose of comparing the results obtained to what happens in a real oceanographic scenario.

Acknowledgement

My sincere thanks go to the directors of the Geophysical Fluid Dynamics (GFD) summer school of 2014 for giving me the opportunity to be one of the fellows this year. It has been such a remarkable experience with talks and academic discussions with scientists from several institutions. From the principle lectures offered by Professors Kerry Emanuel and Geoffery Vallis to the research project, every moment has been remarkable and unforgettable.

Sincere gratitude also to my advisor Prof. Joseph Pedlosky of Woods Hole Oceanographic Institution who gave the academic direction of the project. The prompt email replies which had to be made at times late in the night or very early in the morning depicted tremendous commitment and willingness to help me accomplish the project. Thanks a lot, Joe.

I would also like to acknowledge the much welcomed assistance and support offered by Professor Glenn Flierl of Massachussetts Institute of Technology who not only provided regular guidance during the course of the project but also helped in providing the numerical code used in the study. Glenn's regular explanations made such a huge impact toward a better understanding of the problem studied and in turn shaped the direction of the study. Equally important was the support and academic guidance given by Stephen Meacham throughout the study. Thanks so much for all your efforts in seeing this project come to its completion.

References

Pedlosky, J. (1970). Finite-amplitude baroclinic waves. *Journal of Atmospheric Sciences*, 27:15–30.

Pedlosky, J. (2011). The nonlinear downstream development of baroclinic instability. *Journal of Marine Research*, 69:705–722.

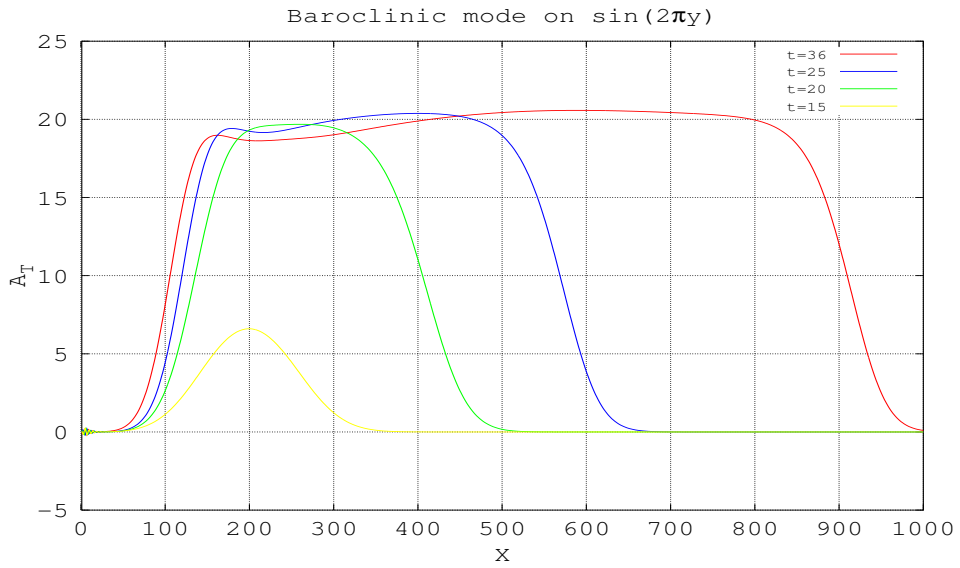


Figure 4: Snapshots of the baroclinic potential vorticity on the sin(2πy) mode with $U_s = U_{so} + 0.25$ and $r = r_o$

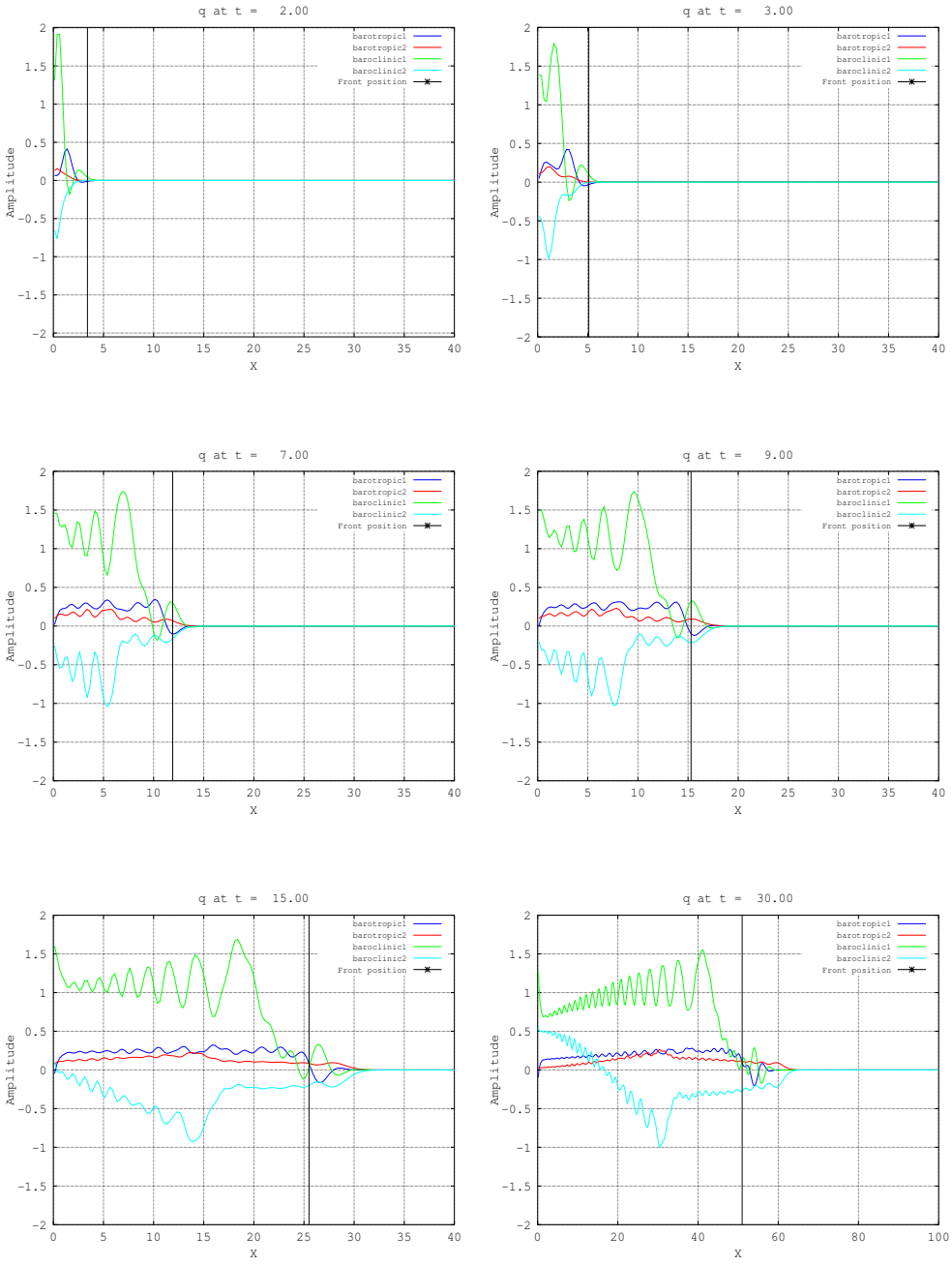


Figure 5: Snapshots of the barotropic and baroclinic potential vorticities. $\Delta = 0.02$

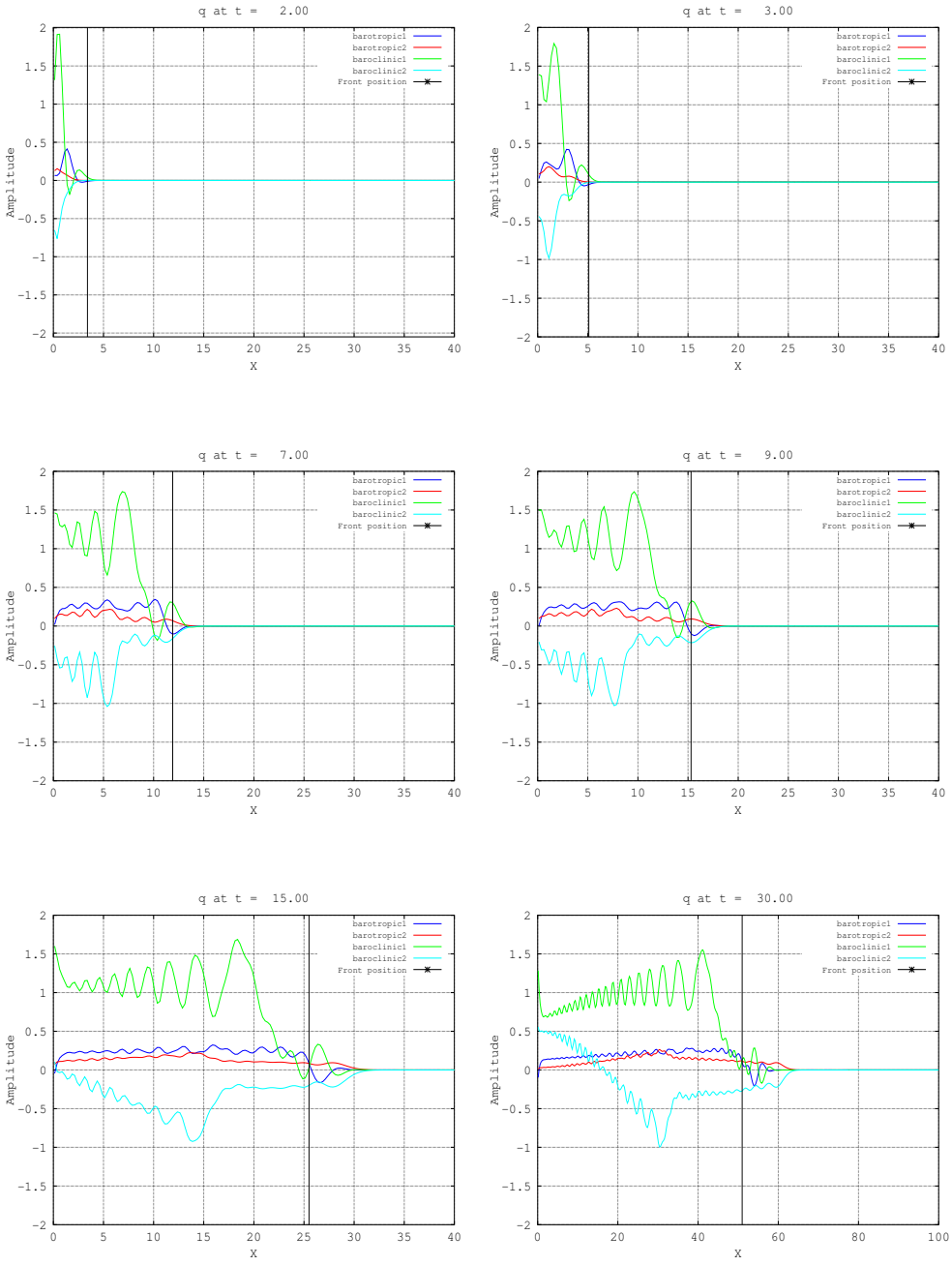


Figure 6: Snapshots of the barotropic and baroclinic potential vorticities. $\Delta = 0.05$

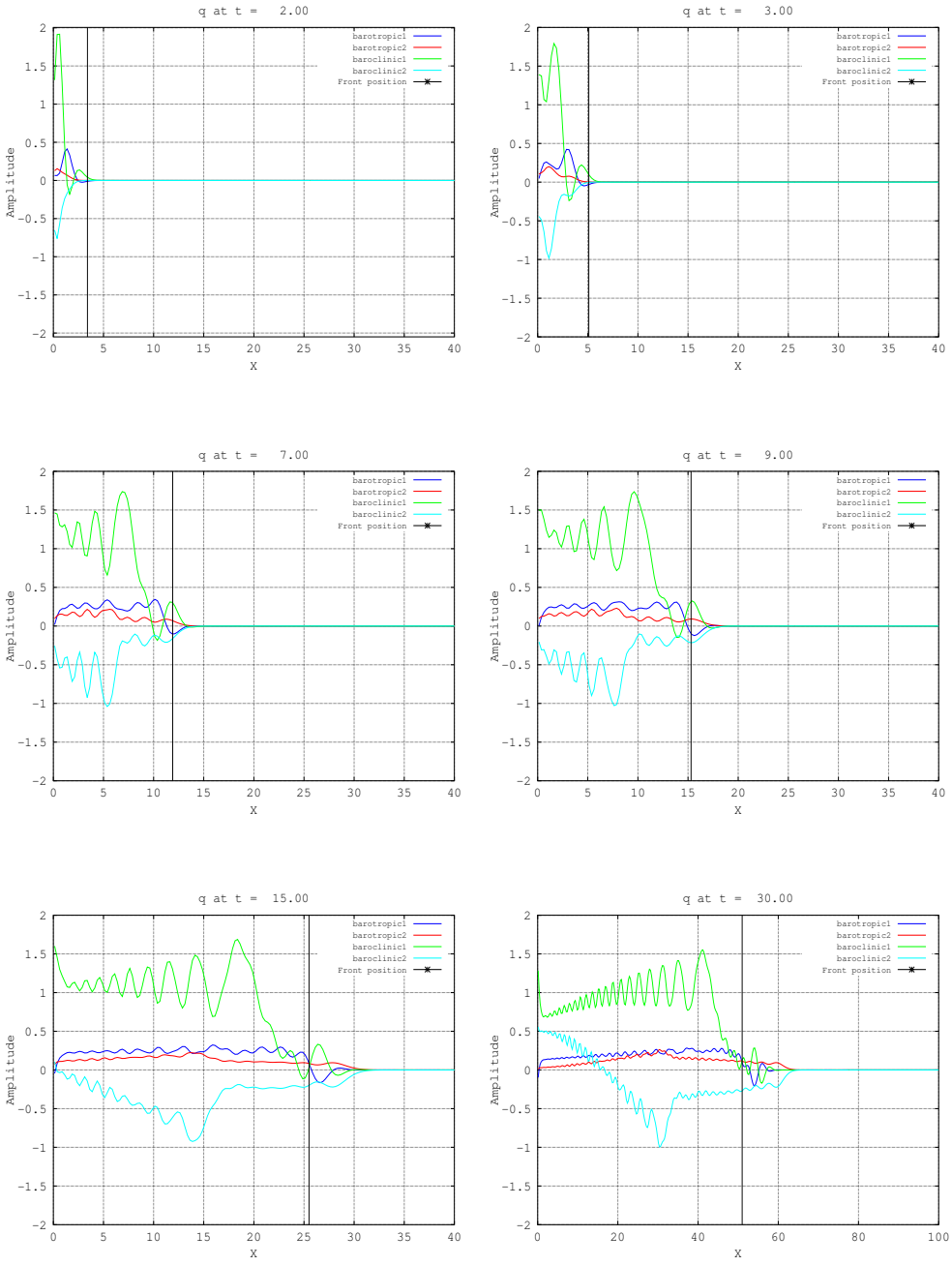


Figure 7: Snapshots of the barotropic and baroclinic potential vorticities. $\Delta = 0.1$

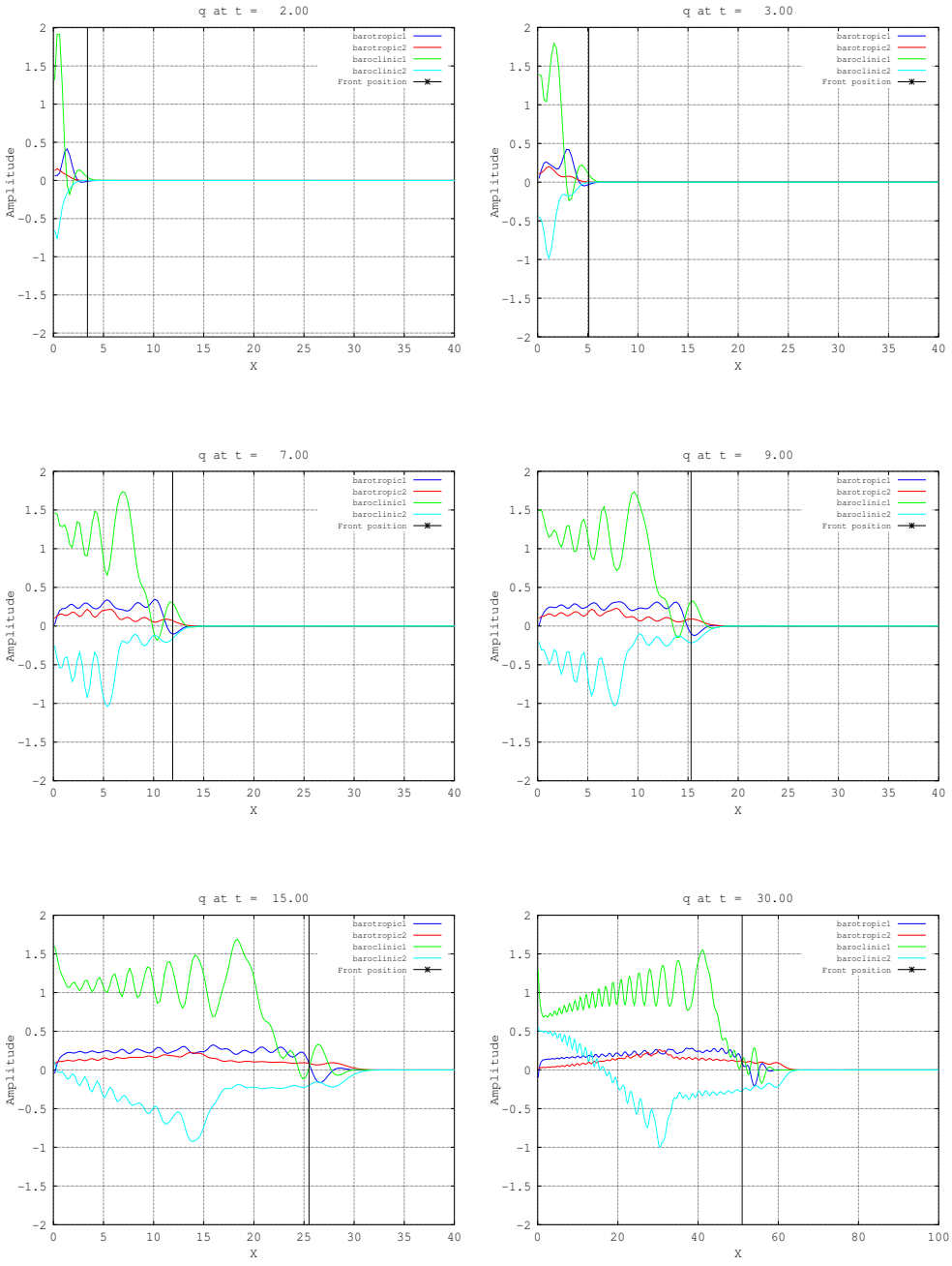


Figure 8: Snapshots of the barotropic and baroclinic potential vorticities. $\Delta = 0.7$

Conceptual Model Analysis of the Influence of Temperature Feedbacks on Polar Amplification

Ashley E. Payne

August 16, 2015

1 Introduction

Polar amplification, the increased response of surface warming at high latitudes to radiative forcing relative to the global mean, is found in idealized models, paleoclimate records and in observations of the Arctic [3, 4, 7]. However, attribution of its causes is a complex problem because of the high degree of interaction between the different mechanisms at work.

Recent research has focused on the role of longwave feedbacks in polar amplification. Winton [2006] used twelve climate models from the fourth IPCC assessment report to compare the relative magnitudes of various feedbacks in the Arctic and globally for CO_2 doubling in 1%/year CO_2 increase experiments. He found that a large portion of the enhanced warming in the Arctic is attributable to the effect of longwave feedbacks, which include cloud, water vapor and temperature effects. Neutralizing the surface albedo feedback at high latitudes by replacing it by its global mean value in the feedback calculation still resulted in some degree of Arctic amplification relative to the global mean, indicating that the surface albedo feedback is important, but not dominant, in explaining polar amplification.

The temperature feedback can be broken into the Planck feedback, the background equilibrium response of the climate system, and the lapse rate feedback, the non-uniform change in temperature vertically. Both have been argued to be important to polar amplification in recent work. Pithan and Mauritsen [2014] performed a feedback analysis on data from phase 5 of the Coupled Model Intercomparison Project (CMIP5) and found that the temperature feedback dominates Arctic amplification, attributing this dominance to the interaction between the Planck feedback and the lapse rate feedback. Pithan and Mauritsen [2014] attribute a large part of the polar amplification to the Planck response itself, arguing that colder temperatures, as found at high latitudes relative to the low latitudes, require greater temperature increases for a given radiative forcing. The role of the lapse rate feedback in relation to the surface albedo feedback was investigated in Graversen et al. [2014] . By suppressing the lapse rate feedback in Community Climate System Model, version 4 (CCSM4) simulations, they found that a portion of polar amplification can be explained by its effect at high latitudes. However, the high level of interaction with the surface albedo feedback makes it difficult to consider the lapse rate independently.

The complexity of the problem highlights the importance of a simplified approach to gain insight into the roles of the Planck feedback and the lapse rate feedback on polar amplification. Here, we use a hierarchy of simple conceptual models to isolate the two different feedbacks and investigate whether (and if so, how much) polar amplification can be attributed to each.

As discussed in Pithan and Mauritsen [2014], the Planck feedback is indeed much weaker in the colder high-latitudes. However, we argue that this by itself cannot cause polar amplification. Instead, it is the qualitative difference in the lapse rate response between the low- and high-latitudes which is expected to contribute to polar amplification. While the low-latitude lapse rate is set by moist convection and is expected to decrease with increasing temperatures, leading to a negative lapse rate feedback, in the high-latitudes, models generally suggest a destabilization of the vertical temperature profile with warming [6], leading to positive lapse rate feedback.

The paper is organized as follows. In section 2, we describe our model setup, data, and method of quantifying feedback parameters. In section 3 we describe the radiative equilibrium response of surface temperature to increases in emissivity in a simple two-layer energy balance model. Against this equilibrium baseline, we investigate the Planck feedback and the lapse rate feedback in sections 4 and 5, respectively, using a hierarchy of models. In section 6, we look at the impact of these feedbacks on polar amplification, with and without the water vapor feedback. We summarize our results and future directions in section 7.

2 Data and Methods

2.1 Models

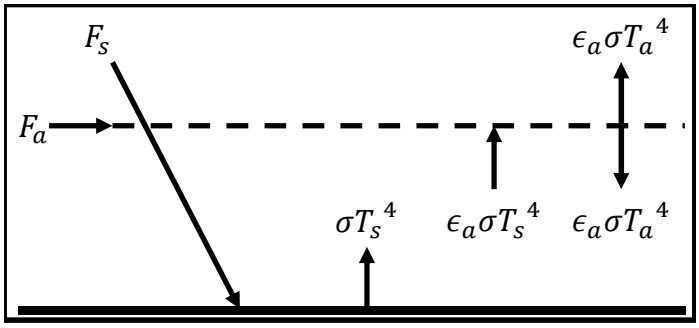


Figure 1: Illustration of a simple two-layer energy balance model with a surface layer and an atmospheric layer. F_s is the incoming solar radiation, F_a is atmospheric heat transport, T_s is the temperature of the surface layer, T_a is the temperature of the atmospheric layer, ϵ_a is the atmospheric emissivity and σ is the Stefan-Boltzmann constant.

We investigate the role of the longwave temperature feedback using a hierarchy of simple models. The simplest is a two-layer one-dimensional energy balance model (EBM), in which the system is represented by two layers; the bottom layer represents the planetary boundary layer and the top layer represents the free troposphere (Fig. 1). Energy input to the system

is in the form of solar insolation, F_s and meridional atmospheric heat transport, F_a, which can be turned on or off. The temperatures of the surface layer and the atmospheric layer are represented by T_s and T_a, respectively. The influence of increasing CO_2 concentration on warming is translated as increases in atmospheric emissivity, ϵ_a, which is assumed to be less than one.

For this simple model, we can write a system of equations that describe the balance of radiation in the two layers:

$$C_s \frac{dT_s}{dt} = F_s + \epsilon_a \sigma T_a^4 - \sigma T_s^4, \quad (1)$$

$$C_a \frac{dT_a}{dt} = F_a + \epsilon_a \sigma T_s^4 - 2\epsilon_a \sigma T_a^4, \quad (2)$$

where, C_s and C_a are the heat capacities of the surface and atmospheric layers, respectively.

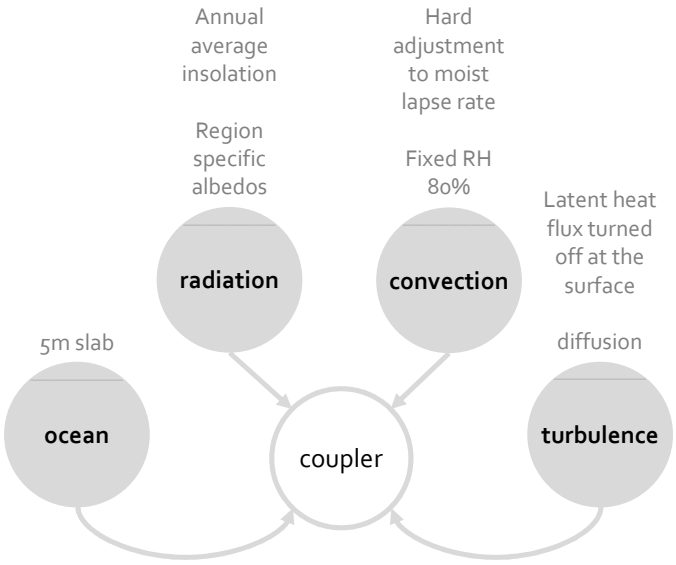


Figure 2: An illustration of the column model composed using the CliMT package, in which various components are linked through a coupler.

To step up complexity, we use a 26-layer radiative-convective column model. The model is based on the Climate Modelling and Diagnostics Toolkit (CliMT), available as a Python wrapper to various existing Fortran components [1]. Figure 2 describes the package set-up, in which ocean, radiation, convection and turbulence components interact using a coupler.

With the CliMT package we set up two different configurations. The first is a column model equivalent of the simple two-layer EBM described previously, in which a grey-gas radiative scheme is coupled to a slab ocean and a simple turbulence scheme. For this model, as well as for the simple two-layer EBM described previously, the role of water vapor is

not incorporated. The second model has an added level of complexity, with a multi-band radiative scheme based on the Column Radiation Model by the NCAR Community Climate Model (CCM3). In order to maintain the energy budget of the column model, but also to constrain the role of moisture, we turn off latent heat flux at the surface and fix relative humidity to 80% in the model troposphere and to zero in the model stratosphere.

When considered, convection is parameterized by hard adjustment to the moist adiabatic lapse rate. For each timestep, the temperature profile of the moist adiabatic lapse rate is calculated from the surface temperature. The intersection of this profile and the model-calculated temperature profile is found and all temperatures below that level are adjusted to the moist adiabat. The additional heat to the column due to this adjustment is pulled from the slab ocean and assumes efficient redistribution of heat from the surface throughout the column.

For the column model, in order to provide some level of comparison to the simple two-layer model, we use the average of the temperature profile between 500 and 600 hPa to represent the atmospheric temperature, T_a.

2.2 Data

We use ERA-Interim reanalysis to constrain parameter values in the following analysis, calculated as the climatological mean of the zonal average. In the following analysis, we will consider a "high-latitude" (HL) column and a "low-latitude" (LL) column. Parameter values for each of these columns are estimated based on ERA-Interim reanalysis data at 80°N and 20°N, respectively. An average insolation value of 367 W m^{-2} is used for non-region specific calculations.

To approximate the role of meridional atmospheric heat transport, we use the zonally averaged vertically integrated convergence of total energy flux (W m^{-2}) (Fig. 3a). In a given location, for the column models, we represent the atmospheric heating profile as a normal distribution that is capped at the tropopause and constrained to integrate to the value of total energy flux corresponding to its latitude, converted to units of K day^{-1} (Fig. 3b).

2.3 Feedback analysis

In order to quantify the contribution of the Planck and lapse rate feedbacks to changes in the surface temperature, we use the partial radiative perturbation method [9]. By this method, a feedback, λ_x, is represented as:

$$\lambda_x = \left(\frac{\partial R}{\partial x}\right)\left(\frac{\partial x}{\partial T_s}\right), \tag{3}$$

where R is the radiation at the top of the atmosphere (TOA) and T_s is the surface temperature. As our focus is on temperature feedbacks, Eqn. 3 becomes:

$$\lambda_T = \left(\frac{\partial R}{\partial T}\right)\left(\frac{\partial T}{\partial T_s}\right), \tag{4}$$

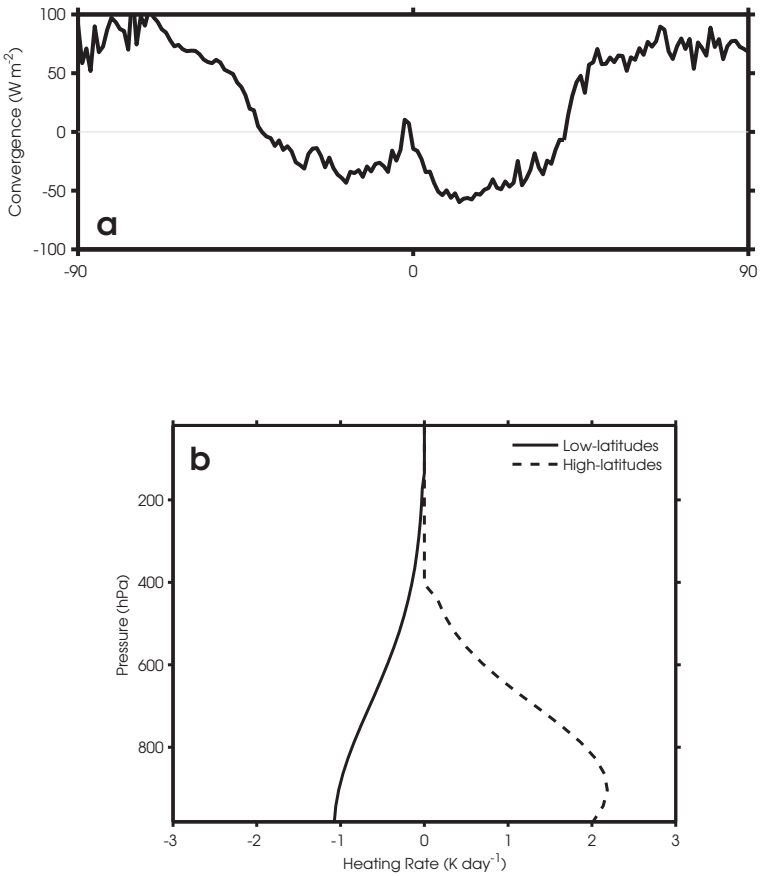


Figure 3: (a) Meridional variation of zonally averaged total energy flux calculated from ERA-Interim reanalysis (climatological mean over 1979 to 2013). (b) Heating rate profiles for the (solid line) low-latitudes and (dashed-line) high-latitudes in the column model.

and,

$$\lambda_T = \lambda_P + \lambda_{LR} \tag{5}$$

where λ_P is the Planck feedback and λ_{LR} is the lapse rate feedback.

In order to calculate the lapse rate feedback, we solve for the temperature feedback, λ_T and the Planck feedback, λ_P. The Planck feedback is the response of the TOA radiation budget in the absence of all other feedbacks and is based on the Stefan-Boltzmann law, so that the feedback is larger for higher temperatures. It is calculated as the derivative of the Stefan-Boltzmann equation and is negative as it dampens the temperature response to

perturbations in R by increasing the emission of longwave radiation[1]:

$$\lambda_P = -4\sigma T_e^3, \tag{6}$$

where, T_e is the effective emission temperature:

$$T_e = \frac{1}{\sigma}\left(F_s + F_a\right)^{1/4} \tag{7}$$

The temperature feedback, λ_T, is calculated as the TOA radiative response associated with the atmospheric and surface temperature changes, per degree of surface warming:

$$\lambda_T = \left(\frac{\partial R}{\partial T_a}\right)\left(\frac{\partial T_a}{\partial T_s}\right) + \left(\frac{\partial R}{\partial T_s}\right)\left(\frac{\partial T_s}{\partial T_s}\right), \tag{8}$$

where, from Fig. 1, the energy balance at the top of the atmosphere for OLR, R, can be written as:

$$R = -\epsilon_a \sigma T_a^4 - (1-\epsilon_a)\sigma T_s^4 \tag{9}$$

Taking the partial derivatives of Eqn. 9 with respect to T_a and to T_s and fitting them back into Eqn. 8, the temperature feedback is calculated as:

$$\lambda_T = -\left(4\epsilon_a \sigma T_a^3\right)\left(\frac{\partial T_a}{\partial T_s}\right) - 4\left(1-\epsilon_a\right)\sigma T_s^3 \tag{10}$$

so that:

$$\lambda_{LR} = -\left(4\epsilon_a \sigma T_a^3\right)\left(\frac{\partial T_a}{\partial T_s}\right) - 4\left(1-\epsilon_a\right)\sigma T_s^3 - \left(-4\sigma T_e^3\right) \tag{11}$$

3 Radiative equilibrium

Here, we consider the system in radiative equilibrium and the response of surface temperature to changes in emissivity. From Fig. 1, at equilibrium and assuming no atmospheric heat transport ($F_a = 0$), we can write the balance of radiation at the TOA as,

$$F_s = (1-\epsilon_a)\sigma T_s^4 + \epsilon_a \sigma T_a^4, \tag{12}$$

where, solving for the T_a term in Eqn. 1, we can simplify Eqn. 12 to:

$$2F_s = 2\sigma T_s^4 - \epsilon_a \sigma T_s^4 \tag{13}$$

We use $F_s = \sigma T_e^4$, to rewrite Eqn. 13 as:

$$T_e^4 = T_s^4 - \frac{1}{2}\epsilon_a T_s^4, \tag{14}$$

[1]The Planck response could alternatively be defined as the radiative response to vertically uniform warming. While not formally identical, the two definitions yield virtually the same results.

and substitute $\epsilon_a + \delta\epsilon_a$ and $T_s + \delta T_s$ for ϵ_a and T_s, respectively, into Eqn. 14,

$$T_e^4 = (T_s + \delta T_s)^4 - \frac{1}{2}(\epsilon_a + \delta\epsilon_a)(T_s + \delta T_s)^4 \tag{15}$$

To find the response of surface temperature to changes in emissivity, we solve Eqn. 15 for δT_s,

$$\delta T_s = \frac{T_s}{4(2 - \epsilon_a)}\delta\epsilon_a \tag{16}$$

At equilibrium, we can see from Eqn. 16 that the most basic state of the system shows that the change in surface temperature is proportional to the background state of surface temperature, T_s. For a given change in emissivity, assumed to be globally constant, we can compare the response of the high-latitudes (HL) and low-latitudes (LL):

$$\frac{\delta T_s^{HL}}{T_s^{HL}} = \frac{\delta T_s^{LL}}{T_s^{LL}} \tag{17}$$

We, therefore, expect a somewhat larger change in low-latitude surface temperature, δT_s^{LL}, relative to the change in high-latitude surface temperature, δT_s^{HL}, because of the greater initial temperatures (in general) in the low-latitudes relative to the high-latitudes. Thus, for this radiative equilibrium state, amplification of surface temperatures occurs in the tropics rather than at high-latitudes. Tropical amplification is evident in solutions to Eqns. 1 and 2 at equilibirum. Figure 4 shows the difference in surface temperatures for a (blue) high- and (orange) low-latitude column.

4 The Planck feedback

The radiative equilibrium response discussed above implies a change in the lapse rate. To isolate the effect of the Planck feedback on the surface temperature response to changing emissivity, we consider a configuration of the two-layer EBM with a fixed lapse rate, in which the temperature of the atmospheric layer is dependent on the surface temperature. We present a perturbation analysis of surface temperature to changes in emissivity in section 4.1 and in section 4.2, we compare the solution for the two-layer experiment to its equivalent in the column grey-gas model. For this section, we use an average insolation value of 367 W m^{-2}.

4.1 Fixed lapse rate

Modifying the setup in section 3, we fix the lapse rate, so that the change in temperature with height is a constant, ΔT_c. As in section 3, we perform a perturbation analysis using the simple two-layer EBM, but with the added constraint:

$$T_a = T_s - \Delta T_c, \tag{18}$$

which we fit into Eqn. 12 and, using the definition of F_s, get:

$$T_e^4 = (1 - \epsilon_a)T_s^4 + \epsilon_a(T_s - \Delta T_c)^4 \tag{19}$$

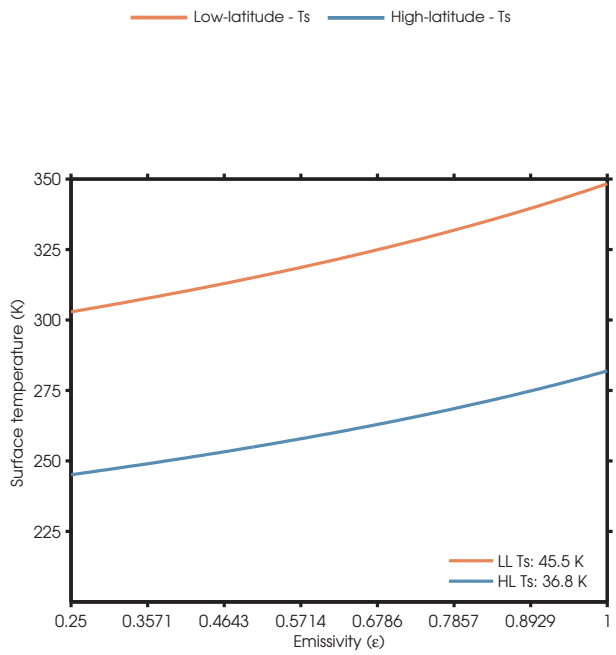


Figure 4: Comparison between solutions for the two-layer model between (orange) low-latitude surface temperature and (blue) high-latitude surface temperature response to increasing emissivity, where the legend indicates the total change in T_s for each region.

As in section 3, we are interested in the surface temperature response to changing emissivity. Thus, we substitute $\epsilon_a + \delta\epsilon_a$ and $T_s + \delta T_s$ for ϵ_a and T_s, respectively, in Eqn. 19:

$$T_e^4 = (1 - (\epsilon_a + \delta\epsilon_a))\,(T_s + \delta T_s)^4 + (\epsilon_a + \delta\epsilon_a)\,((T_s + \delta T_s) - \Delta T_c)^4 \tag{20}$$

Linearizing and ignoring higher order terms (where we assume T_s is much greater than the perturbation terms), we are left with:

$$4\epsilon_a T_s^3 \delta T_s - 4\epsilon_a (T_s - \Delta T_c)^3 \delta T_s - 4T_s^3 \delta T_s = \delta\epsilon_a (T_s - \Delta T_c)^4 - \delta\epsilon_a T_s^4 \tag{21}$$

Solving Eqn. 21 for δT_s:

$$\delta T_s = \frac{(T_s - \Delta T_c)^4 - T_s^4}{4\epsilon_a (T_s^3 - (T_s - \Delta T_c)^3) - 4T_s^3} \delta\epsilon_a \tag{22}$$

Assuming that ΔT_c is much less than T_s we can simplify Eqn. 22 to:

$$\delta T_s \approx \Delta T_c \delta\epsilon_a \tag{23}$$

Equation 23 shows that the change in surface temperature is proportional to the magnitude of the lapse rate, ΔT_c. Thus, for a given lapse rate we do not expect amplified surface

warming at high latitudes based on the differential Planck response alone. Instead, we may expect reduced warming in the high-latitudes, due to the typically small lapse rate found there.

4.2 Comparison to column model and feedback analysis

We extend the analysis in section 4.1 to the column model with a grey-gas radiative scheme, as described in section 2.1. Results for the column model are qualitatively similar to those of the two-layer model. Figures 5a and 5b show a comparison between the response of (grey line) atmospheric temperature to increasing atmospheric emissivity relative to (black line) surface temperature, in (solid) radiative equilibrium and with (dashed) a fixed lapse rate, for the two-layer and column grey gas models, respectively. For the fixed lapse rate solution, ΔT_c is adjusted so that the atmospheric temperature at an emissivity of 0.3 is equal to the temperature of the radiative-equilibrium solution in each model and approximates the dry adiabatic lapse rate.

The results discussed above may seem to be at odds with the notion that the Planck feedback ought to be weaker at high-latitudes. The reason for the lack of enhanced warming lies in the similarly weaker radiative forcing, associated with a given increase in the optical thickness of the atmosphere. The effect of the Planck feedback in modifying radiative forcing is seen in Fig. 6a, which shows the radiative forcing for the (solid) radiative equilibrium state and (dashed) fixed lapse rate case for the two-layer EBM. Due to our use of a constant Planck feedback calculation, the slight residual difference between the radiative forcing and Planck feedback is manifested as a small lapse rate feedback, even for the fixed lapse rate case. For comparison, Fig. 6b shows the (black) fixed lapse rate radiative forcing seen in Fig. 6a plotted with the radiative forcing for fixed lapse rate in the (blue) high- and (orange) low-latitudes. At higher emissivity, the larger lapse rate for the radiative equilibrium state shows the larger radiative forcing.

5 Lapse rate changes and feedbacks

Here, we include the role of meridional atmospheric heat transport to investigate its role in setting the sensitivity of the lapse rate feedback to changes in atmospheric emissivity. We consider two cases: (1) a system in radiative-dynamic equilibrium, that is with the lapse rate controlled by the interaction of radiation and a prescribed atmospheric heat transport, and (2) a system with a convective lapse rate in the low-latitudes. As atmospheric heat transport changes sign in the meridional direction, we now consider two columns for each experiment: (1) a high-latitude column and (2) a low-latitude column.

5.1 Two-layer EBM

5.1.1 Radiative equilibrium

For a small change in atmospheric emissivity, $\delta\epsilon_a$, we assume a small change in the surface temperature, δT_s, and the atmospheric temperature, δT_a. How amplification will affect meridional heat transport is an active area of research, however, to isolate out the effects

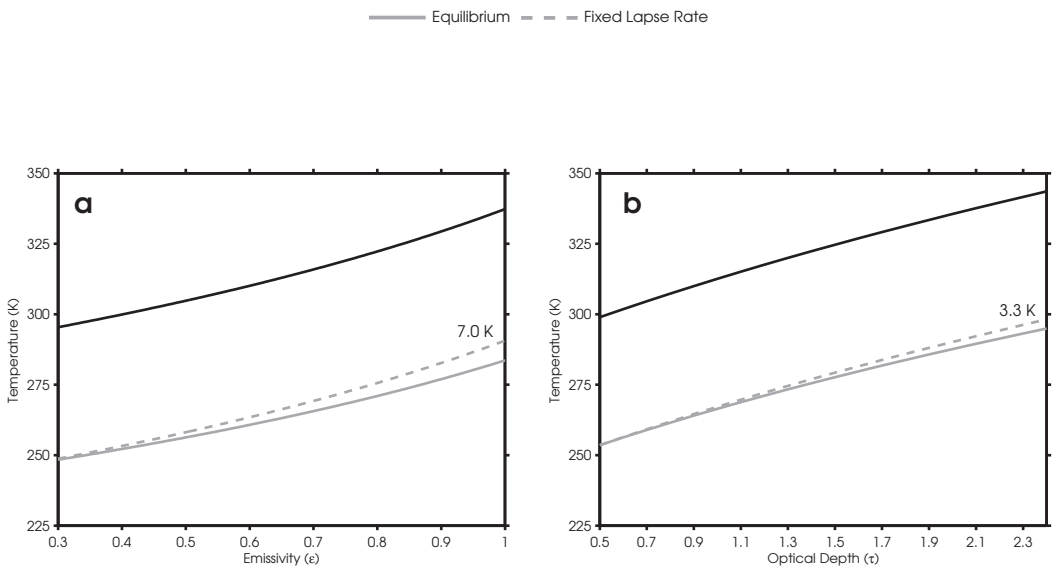


Figure 5: Comparison between solutions for the (a) two-layer model and (b) column grey-gas model showing the trend in (black line) surface and (grey line) atmospheric temperatures for the (solid line) radiative equilibrium and (dashed line) fixed lapse rate experiments.

of local feedbacks versus changes in the heat transport, we will assume a constant heat transport which is unaffected by changes in emissivity. At equilibrium, with these small changes, Eqn. 2 becomes,

$$F_a + (\epsilon_a + \delta\epsilon_a)\sigma(T_s + \delta T_s)^4 = 2(\epsilon_a + \delta\epsilon_a)\sigma(T_a + \delta T_a)^4 \tag{24}$$

Linearizing Eqn. 24 and ignoring higher order terms, leaves us with an expression for the response of the lapse rate to a change in emissivity,

$$8\epsilon_a\sigma T_a^4\left(\frac{\delta T_a}{T_a} - \frac{\delta T_s}{T_s}\right) = -F_a\left(\frac{\delta\epsilon_a}{\epsilon_a} + 4\frac{\delta T_s}{T_s}\right) \tag{25}$$

For a positive atmospheric heat flux convergence, where $F_a > 0$, the L.H.S. of Eqn. 25 must be negative, and thus $\delta T_a/T_a < \delta T_s/T_s$. Since generally $T_s > T_a$, this implies that $\delta T_a < \delta T_s$ and, thus, an increase in the lapse rate. If the atmospheric heat flux convergence is negative (i.e. $F_a < 0$), we may instead expect a decrease in the lapse rate.

Due to the meridional differences in the energy balance between incoming solar insolation and outgoing longwave radiation, the low-latitudes are characterized by a net export of heat and the high-latitudes are characterized by a net import of heat. This can be seen in Fig. 3. Based on Eqn. 25, we, therefore, expect to see a greater increase in surface temperatures relative to atmospheric temperatures in the high-latitudes (an increase in the lapse

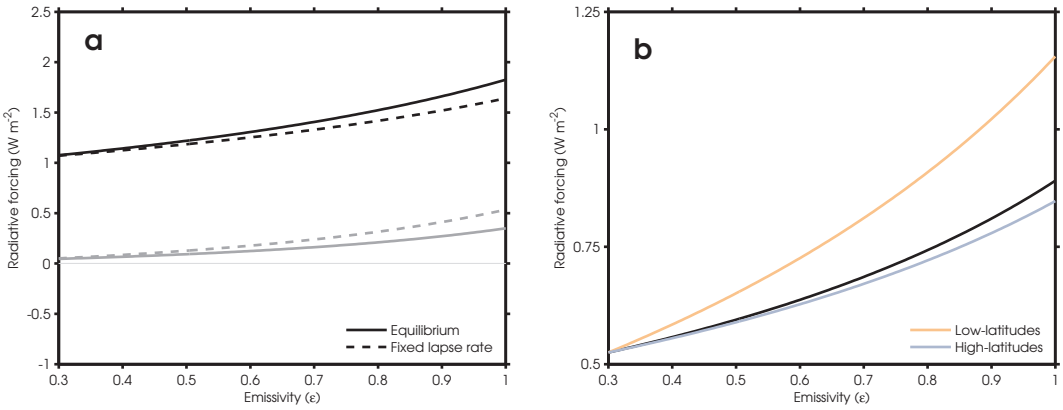


Figure 6: For the two-layer EBM, (a) comparison between the (black) radiative forcing and the (grey) lapse rate feedback for the (solid line) radiative equilibrium and (dashed line) fixed lapse rate experiments. (b) Comparison of the radiative forcing between the (blue) high- and (orange) low-latitudes with the (black) radiative forcing for the fixed lapse rate experiment.

rate). For the tropics, the radiative-dynamic equilibrium solution in Eqn. 25, would instead suggest a decrease in the lapse rate. However, the radiative-dynamic equilibrium solution is also statically unstable in the tropics, which would trigger convection. In practice, Eqn. 25 is thus not applicable in the tropics, where the lapse rate is set by convection.

We compare solutions of T_s and T_a for the high-latitude column to a convective low-latitude column in the next section. The system of equations for the high-latitudes are then (from Eqns. 1 and 2, at equilibrium):

$$T_s = \left[\frac{F_a + 2F_s}{2\sigma - \epsilon_a \sigma}\right]^{1/4}, \tag{26a}$$

$$T_a = \left[\frac{F_a + \epsilon_a F_s}{2\epsilon_a \sigma - \epsilon_a^2 \sigma}\right]^{1/4}, \tag{26b}$$

5.1.2 Convective lapse rate

We add a level of realism to our two-layer model by allowing the lapse rate in the low-latitude scenario to be set by convection. In the low-latitudes, where the lapse rate is set by convection, the two-layer system is described by the energy balance at the TOA and by a temperature dependent lapse rate:

$$F_s + F_a = (1 - \epsilon_a)\sigma T_s^4 + \epsilon_a \sigma T_a^4, \tag{27a}$$

$$T_a = T_s - \Delta T \tag{27b}$$

In contrast to section 4.1, ΔT is not constant, but is dependent on the moist adiabatic lapse rate, Γ_w,

$$\Gamma_w = g\frac{1+{}^{L_v r}/_{R_d T}}{C_{pd}+{}^{L_v^2 r}/_{R_v T^2}}, \tag{28}$$

where L_v is the latent heat of vaporization, R_d is the specific gas constant of dry air, C_{pd} is the heat capacity of dry air at constant pressure, and R_v is the specific gas constant of water vapor. Since the amount of moisture is dependent on temperature, the mixing ratio of the mass of water vapor to that of dry air, r changes with temperature and pressure:

$$r = {}^{0.622 e_s}\,/_{(P-e_s)}, \tag{29}$$

where e_s is the saturation vapor pressure:

$$e_s = 6.11 \exp \frac{L_v}{R_v}\left(\frac{1}{273}-\frac{1}{T}\right) \tag{30}$$

The atmospheric temperature is assumed to be at 500 hPa and is calculated by integrating the change in temperature, T, along the moist adiabatic lapse rate profile with height.

5.2 Comparison to column model and feedback analysis

Here, we compare the solutions of the two-layer EBM to those from the column grey-gas model. The column grey-gas model is setup for the same two cases as the two-layer model: (1) a radiative-dynamic equilibrium case, and (2) a case with a radiative-convective column in the low-latitudes.

5.2.1 Role of atmospheric heat transport

In this section, we want to examine the effect of atmospheric heat transport on the radiative-dynamic equilibrium solution. We are thus assuming a radiative-dynamic equilibrium solution also for the low-latitude column, ignoring the effects of convection. The more realistic case with a convective lapse rate in the tropical column is discussed in section 5.2.2.

The interpretation of Eqn. 25 in section 5.1.1 is supported by solutions for the two-layer EBM and the column grey-gas model. Figure 7 shows a comparison of the temperature response to changes in optical thickness for the (solid lines) radiative equilibrium and (dashed lines) radiative-dynamic equilibrium for the (a,c) two-layer model and (b,d) column grey-gas model. Temperature trends for the two models are very similar and consistently show an increase in the lapse rate with the introduction of heat transport, evident in the flattening of the atmospheric temperature trend in Figs. 7c and 7d. In the high-latitudes, the lapse rate of the radiative-dynamic equilibrium state increases with increasing optical thickness. The opposite occurs at low-latitudes.

While there is a large response in the lapse rate to atmospheric heat transport, the surface temperature response still shows a slightly higher increase in low-latitude surface temperatures relative to high-latitude temperatures. Introduction of heat transport to the models

slightly reduces the amount of surface heating in the low-latitude scenario and slightly increases the surface heating in the high-latitudes. However, the models still show amplification of low-latitude surface temperature relative to the high-latitudes.

Focusing on the high-latitudes where there is a large response to atmospheric heating, Fig. 8 shows the radiative response (the product of the total surface temperature change with the feedback parameter) of the Planck and lapse rate feedbacks to changes in emissivity at high latitudes for the two experiments (a) without and (b) with atmospheric heat transport. As inferred from Fig. 7, the magnitude of the lapse rate feedback is larger in the presence of atmospheric heat transport. However, the dynamic heat transport also strongly changes the radiative forcing by changing the lapse rate of the background state, as discussed in section 4.1. The effect of the dynamic heat transport on the base-state lapse rate, and thus the radiative forcing, here approximately cancels the effect of the lapse rate feedback. For the high-latitude atmospheric temperature response, near $\epsilon_a = 0.3$ in Fig. 8, the radiative forcing changes sign from negative (cooling) to positive (warming). This change in sign is associated with a change in sign of the lapse rate, which for a small epsilon is very efficiently stabilized by the atmospheric heat transport.

5.2.2 Effect of convection on low-latitude temperature response

In the previous section, no polar amplification was evident due to the amplified response of the low-latitudes to increases in atmospheric emissivity (or optical depth). In order to investigate the effect of convection on the comparative role of the low-latitudes to the high-latitudes and polar amplification, we compare the temperature response at low-latitudes, with active heat transport, with- and without- convection in the two models (Fig. 10). Two processes must be taken into account to adequately explain the figure: (1) the role of convection on the lapse rate of the background state and (2) the negative lapse rate feedback set by moist convection.

Figure 9 shows the same as Fig. 8, but for the (a) non-convective and (b) convective low-latitudes. Comparison of the two panels shows a large decrease in the amount of radiative forcing for a given change in emissivity. The role of convection in the tropics acts to decrease the lapse rate of the background state. This reduction results in the large decrease of radiative forcing, as discussed in section 4.1. Superimposed on its influence on the background state lapse rate, convection also imposes a negative lapse-rate feedback in the tropics, as the moist adiabatic lapse rate decreases with increasing temperature. Figure 9b shows an increasingly negative lapse rate feedback with increasing emissivity, associated with the increase in the lapse rate feedback with rising temperature (causing an exponential rise in the saturation specific humidity).

The reduction of the radiative forcing and the negative lapse rate feedback, resulting from the addition of moist convection, have strong implications for the sensitivity of the surface temperature with increasing optical thickness. Figure 10 shows the weakening of the tropical surface temperature response when convection is taken into account. Compared to the almost constant lapse rate in Fig. 7, there is a large decrease in the lapse rate, which is

reflected in the feedback analysis.

6 Impact on polar amplification

6.1 Dry longwave perspective

Here, we bring together the results from previous sections to compare the high- and low-latitude columns in the two-layer EBM and column grey-gas models in order to investigate their impact on polar amplification. We use a low-latitude scenario with a convective lapse rate and negative heat transport and a high-latitude scenario with a lapse rate set by radiation and positive heat transport.

Figures 11a and 11b show temperature trends for the two-layer model and column grey-gas model, respectively. Both models show an amplification of surface temperatures in the high-latitudes relative to the low-latitudes. This amplification is associated with a large increase in the high-latitude lapse rate, as opposed to a decrease in the low-latitude lapse rate, which can be seen in Fig. 11c and 11d, where the response of lapse rate in each region is plotted against emissivity. These results are in line with the perturbation analysis in section 5 and provide evidence for the role of the lapse rate feedback in modifying the sensitivity of surface temperature to changes in emissivity (or optical thickness).

The amplification of polar surface temperatures relative to the tropics implies a weakening of the surface temperature gradient. In the free troposphere, the opposite case is found, where polar atmospheric temperatures increase at a smaller rate than tropical atmospheric temperatures, in line with the sign of the lapse rate feedback response in each region. The difference in the warming pattern between the surface and atmosphere in the two regions is a robust feature found in previous research [8] and has implications on mid-latitude dynamics and changes in the atmospheric heat transport with warming.

6.2 The role of water vapor and heat transport

The analysis in the previous sections focused on the roles of the Planck and lapse rate feedbacks in a dry atmosphere. Here, we use a multi-band radiative transfer scheme based on CCM3 in the column model, and allow for changes in specific humidity. For simplicity, the relative humidity is held fixed at 80%. We will compare temperature trends between a convective low-latitude scenario and a high-latitude scenario, both with atmospheric heat transport.

Figure 12a shows the temperature trends for the column model, with the absolute temperature shifted such that the atmospheric temperature is the same as the surface temperature at 280 ppmv for better comparison. The results indicate that the more realistic column model is in agreement with the lapse rate results from the simpler models, where lapse rate increases in the high-latitudes and decreases in the low-latitudes. However, amplification of polar surface temperatures is no longer evident relative to low-latitude surface temperatures. Surface warming in the low-latitudes is almost double that at high-latitudes.

The difference between the results in section 6.1 and those in Fig. 12a is directly attributable to the the presence of the water vapor feedback. To isolate this feedback, we fix the vertical specific humidity profile to their value at 380 ppmv. We then compare the low-latitude surface temperatures for this fixed scenario to one in which water vapor instead responds to increasing CO_2 concentrations. Figure 12b shows the difference in low-latitude surface temperatures for the two experiments. Suppressing the WV feedback leads to a suppression of surface temperature warming in the convective tropics and the re-emergence of polar amplification. Without the water vapor feedback, the model behaves similarly to the results described in section 6.1, using a grey-gas radiative transfer scheme or the EBM. However, with an approximately constant relative humidity in the models, the water vapor feedback results in a strongly enhanced tropical warming, while having little effect at high latitudes. This cancels the differential effect of the lapse rate feedback, and instead leads to low-latitude amplification. Additional feedbacks (such as ice-albedo or clouds) and /or changes in the atmospheric heat transport are, thus, likely to be necessary to explain polar amplification in Earth's atmosphere and in GCMs.

7 Summary and conclusions

In this project, we used a hierarchy of conceptual models of increasing complexity to test the roles of the Planck feedback and the lapse rate feedback in polar amplification. Using a simple column model with a fixed lapse rate, we argue that the Planck feedback alone cannot explain polar amplification. Instead, the surface warming expected for a given change in atmospheric optical thickness depends primarily on the prescribed lapse rate itself (but not on the temperature of the base state). While the strength of the Planck feedback does increase with the temperature of the base state, the effective radiative forcing, associated with a given change in atmospheric optical thickness, increases approximately similarly with the base state temperature. As a result, the net warming for a given change in atmospheric optical thickness is approximately independent of the base state temperature. The differing response of the lapse rate at high- and low-latitudes instead does cause polar amplification. At high latitudes, the lapse rate is sensitive to atmospheric heat transport and increases strongly in response to increasing atmospheric optical thickness.

When we consider the role of water vapor in our column model by using a multi-band radiative transfer scheme, and holding atmospheric relative humidity constant, low-latitude surface temperatures changes in response to CO_2 increases are amplified relative to high-latitude surface temperature changes. This indicates that the stronger water vapor feedback in the tropics cancels the polar amplification associated with the lapse rate feedback. To isolate the role of water vapor, we fix water vapor to 380 ppmv for a range of CO_2 concentrations. With the water vapor feedback essentially removed, high latitude surface temperatures are again amplified relative to low-latitude surface temperatures. These results suggest that, while the lapse rate feedback does play a role in Arctic amplification of surface temperatures, the effect is masked in comparison by the large water vapor feedback at low latitudes.

Future work for this project includes:

- The calculation of feedback parameters for the column models for better comparison to the results of the two-layer model experiments. This task is important for quantifying the role of the lapse rate feedback in the higher order models. In addition, we can use feedback analysis in the realistic radiative scheme column model to investigate the role of the water vapor feedback in relation to the Planck and lapse rate feedbacks.

- The parameterization of a dynamic heat transport between the high- and low-latitudes in the two-layer EBM and column models. In the current setup, the two regions are completely decoupled and used only for comparison. Our results suggest a decrease in the meridional temperature gradient at the surface and an increase in the gradient in the free troposphere. This has implications on dynamic heat transport in the mid-latitudes.

8 Acknowledgements

I would like to thank Malte Jansen for giving me the opportunity to work on this interesting problem for the summer. He was an excellent advisor and I am very grateful for his incredible patience with me. I truly enjoyed my experience in the program and would like to thank the directors, main lecturers and all others I had the opportunity to talk to over the course of the summer for their insight.

References

[1] R. Caballero and M. Steder, *Climt: An object-oriented climate modelling and diagnostics toolkit* `http://maths.ucd.ie/~rca/climt/`, 2008.

[2] R. G. Graversen, P. L. Langen, and T. Mauritsen, *Polar amplification in CCSM4: Contributions from the lapse rate and surface albedo feedbacks*, Journal of Climate, (2014), pp. 4433 – 4450.

[3] S. Manabe and R. J. Stouffer, *Sensitivity of a global climate model to an increase of CO_2 concentration in the atmosphere*, Journal of Geophysical Research, 85 (1980), pp. 5529-5554.

[4] G. H. Miller, R. B. Alley, J. Brigham-Grette, J. J. Fitzpatrick, L. Polyak, M. C. Serreze, and J. W. White, *Arctic amplification: can the past constrain the future?*, Quaternary Science Reviews, 29 (2010), pp. 1779 – 1790.

[5] F. Pithan and T. Mauritsen, *Arctic amplification dominated by temperature feedbacks in contemporary climate models*, Nature Geoscience Letters, 7 (2014), pp. 181-184.

[6] M. Serreze, A. Barrett, J. Stroeve, D. Kindig, and M. Holland, *The emergence of surface-based Arctic amplification*, The Cryosphere, 3 (2009), pp. 11 – 19.

[7] S. Solomon et al., eds., *Climate Change 2007: The Physical Science Basis*, Cambridge University Press, 2007.

[8] ——, *Climate Change 2007: The Physical Science Basis, Ch. 10*, Cambridge University Press, 2007.

[9] R. Wetherald and S. Manabe, *Cloud feedback processes in a general circulation model*, Journal of Atmospheric Sciences, 45 (1988), pp. 1397 – 1415.

[10] M. Winton, *Amplified arctic climate change: What does surface albedo feedback have to do with it?*, Geophysical Research Letters, 33 (2006).

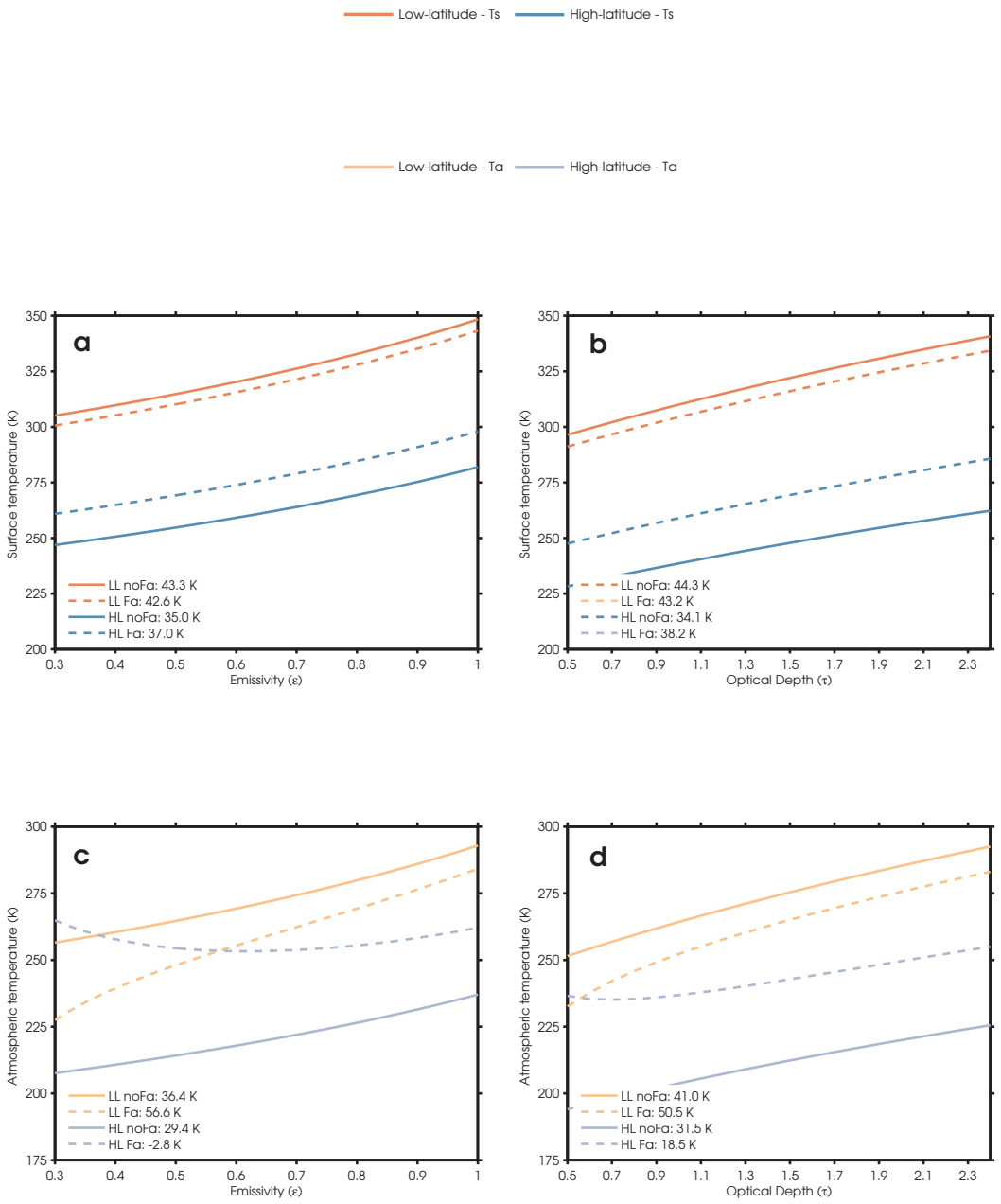


Figure 7: (a,b) Surface temperature and (c,d) atmospheric temperature response to increasing emissivity for the (orange) low-latitudes and (blue) high-latitudes, (solid line) without and (dashed line) with atmospheric heat transport for the (a,c) two-layer EBM and (b,d) column grey-gas model. The numbers in the legends of each panel show the temperature change between the thickest atmospheric limit and the thinnest limit for each scenario.

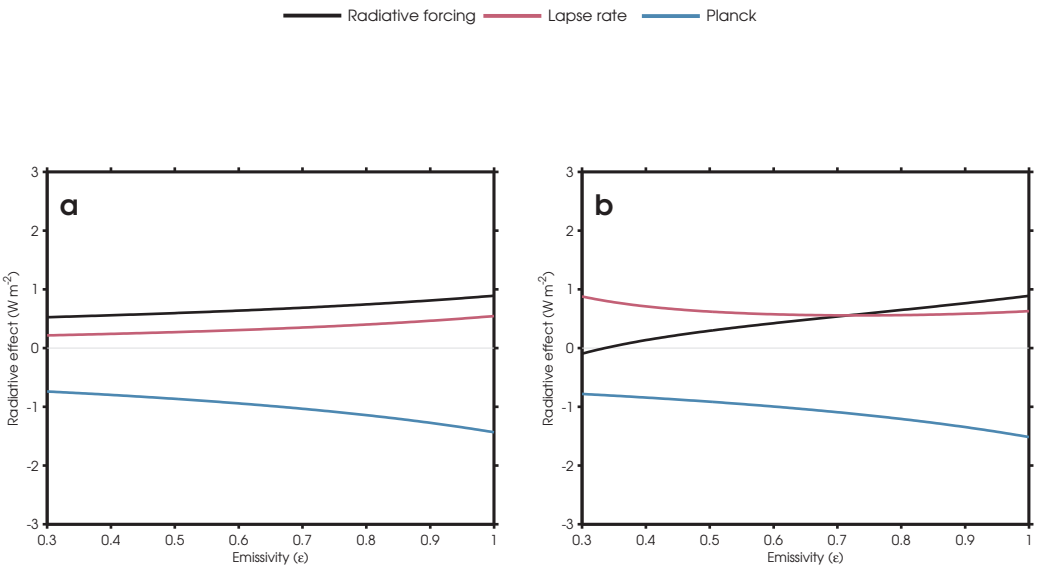


Figure 8: At high-latitudes, the radiative response of the (blue) Planck feedback and the (pink) lapse rate feedback to (black) radiative forcing (a) without atmospheric heat transport and (b) with atmospheric heat transport.

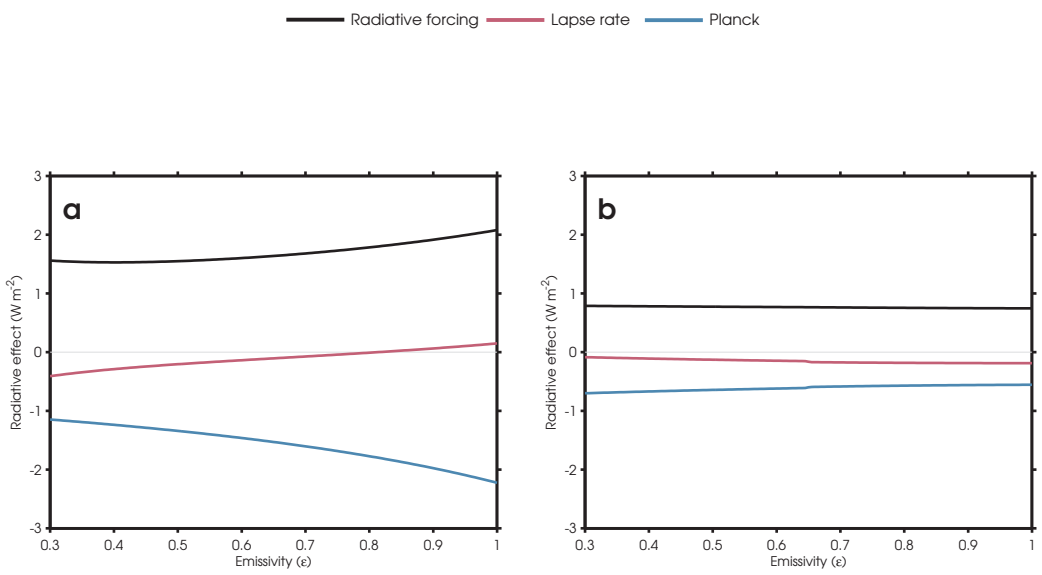


Figure 9: At low-latitudes, the radiative response of the (blue) Planck feedback and the (pink) lapse rate feedback to (black) radiative forcing (a) without and (b) with convection.

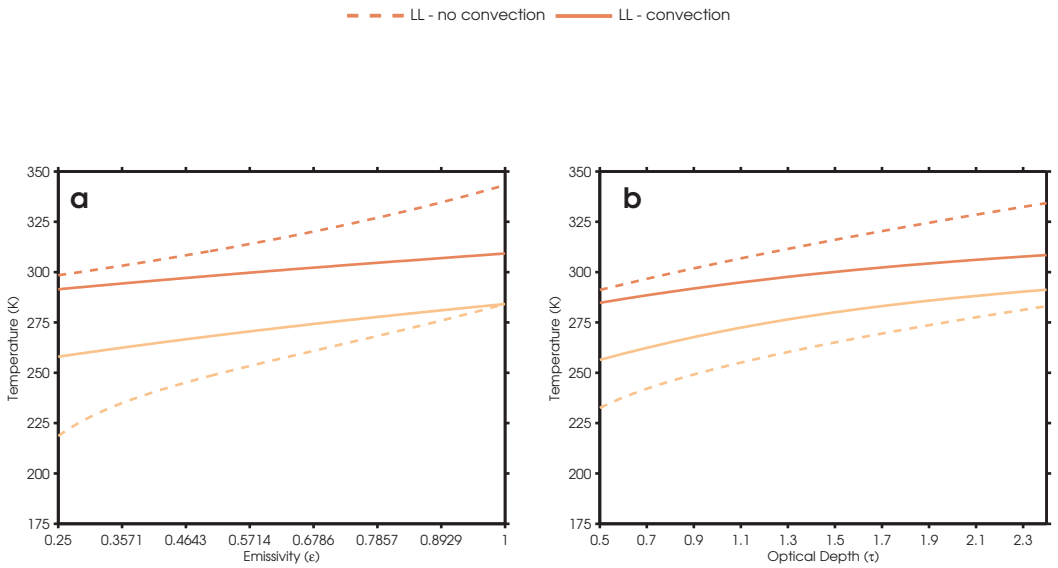


Figure 10: Comparison between (dark orange) surface temperature and (light orange) atmospheric temperature (solid line) without convection and (dashed line) with convection for the (a) two-layer EBM and (b) column grey-gas model.

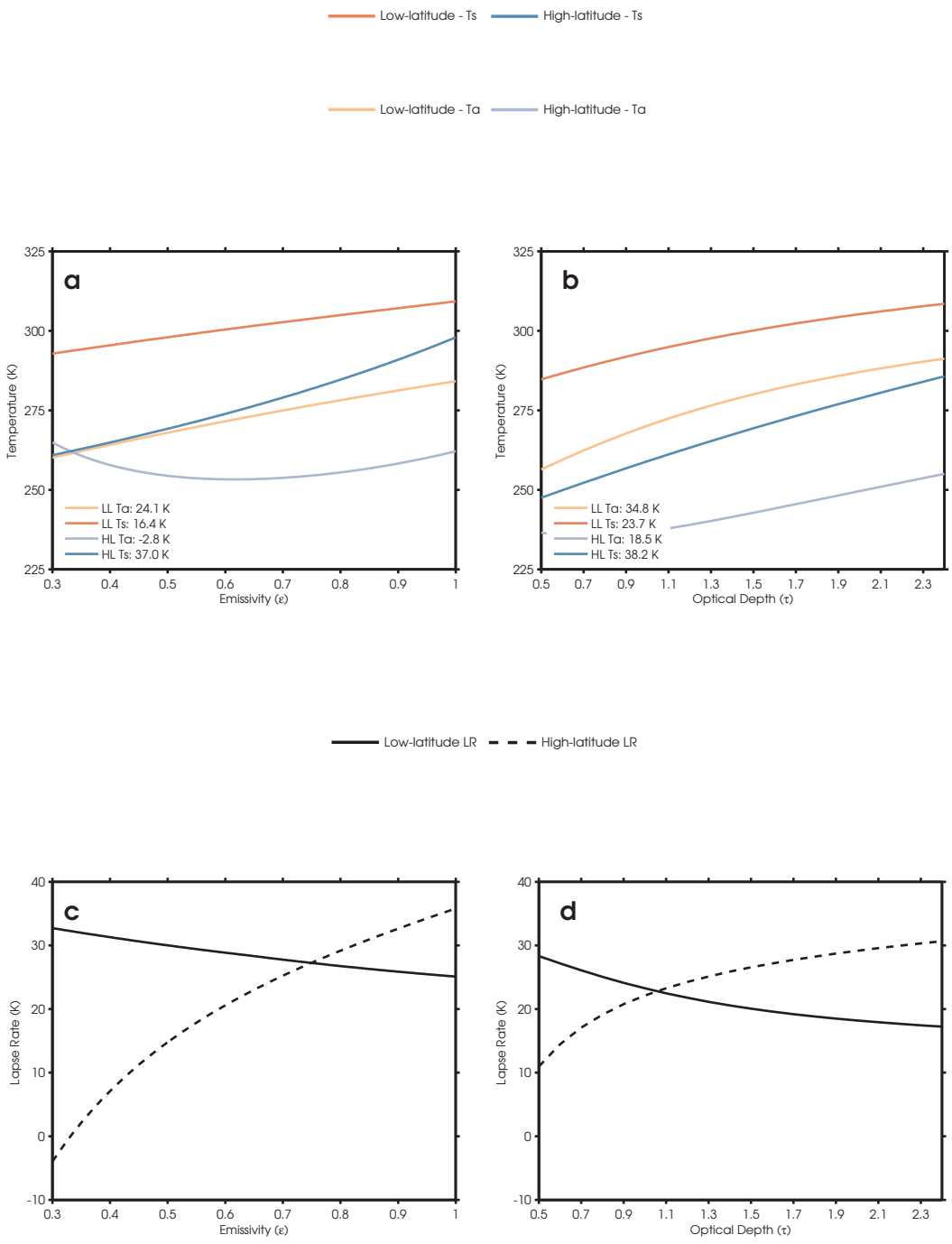


Figure 11: For the (a,c) two-layer EBM and the (b,d) column grey-gas model, (a,b) Surface temperature and atmospheric temperature response to increasing emissivity (optical depth) and (c,d) lapse rate changes for the (blue, black dashed) high- and (orange, black solid) low-latitude scenarios. The numbers in the legends of panels (a) and (b) are the same as in Fig. 7

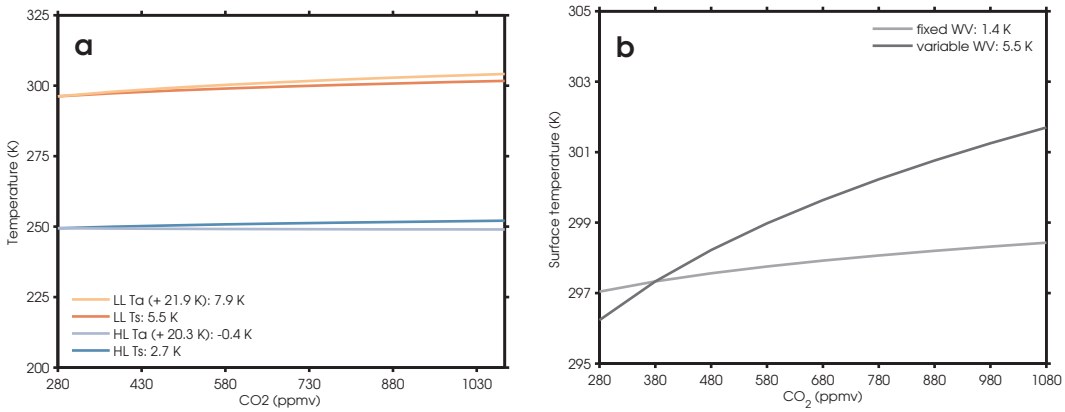


Figure 12: (a) (dark colors) Surface temperature and (light colors) atmospheric temperatures for the (orange) low- and (blue) high-latitudes in the column model with a multi-band radiative transfer scheme. (b) Comparison between the convective tropical surface temperature response to increasing CO_2 for a realistic radiative scheme with (light grey) moisture fixed to 380 ppmv levels and (dark grey) freely varying atmospheric moisture.

Thermobaric Effects on Double-Diffusive Staircases

Erica Rosenblum

August 16, 2015

1 Introduction

Under the sea ice of the polar oceans, warm salty water flows below the cold, fresh surface waters. Meanwhile, deep below the surface, geothermal heating warms the cold, salty waters of the deep Arctic. In each of these systems, an unstable temperature gradient competes with a stabilizing salinity gradient, creating conditions susceptible to two instabilities. One of these instabilities, called double-diffusive convection, is due to the fact that heat diffuses 100 times faster than salt. This instability can lead to the formation of thermohaline staircases in which a series of well-mixed, convective layers are separated by sharp interfaces in both temperature and salinity. A second instability that exists in this system, called the thermobaric instability, arises because cold water is more compressible than warm water and therefore becomes denser with pressure

In the Arctic, waters from the North Atlantic provide the warm, salty water in the aforementioned system, leading to the formation of double-diffusive staircases (Figure 1). Contained within this North Atlantic water is enough heat to melt all of the Arctic sea ice, were it transported to the surface [10],[13]. Processes controlling vertical heat flux in this area of the world are not well understood, but double-diffusive convection has been identified as an important mechanism [9], and may therefore play an important role in the rapid disappearance of the Arctic sea ice.

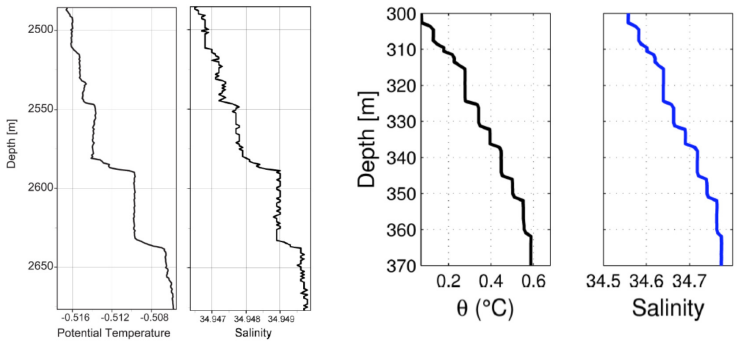


Figure 1: Examples of double-diffusive staircases found in the deep Arctic (left) and the shallow Arctic (right). Profiles taken from [13], [12]

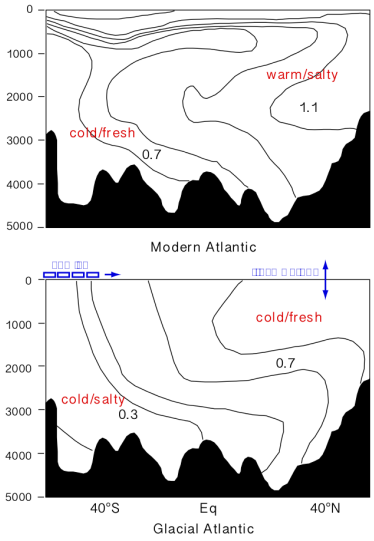


Figure 2: Cartoon description showing that the Atlantic was stably stratified by a salt gradient during the last glacial maximum, taken from [1]

Several thousand meters deeper, quietly sits the strongly salt-stratified waters of the deep Arctic. This water is slowly being heated from below by geothermal vents, again leading to a salt-stratified system that is heated from below. Double-diffusive staircases exist in this context as well, however, these staircases are notably thicker than than those found in the shallow Arctic (Figure 1).

This system is under a large amount of pressure and is potentially thermobarically unstable despite its strong salinity gradient. Interestingly, because this system has a strong salinity stratification and is being heated geothermally, it bears resemblance to the Atlantic Ocean during the last glacial maximum (Figure 2). It has been proposed that the bottom of this glacial Atlantic Ocean slowly warmed until an intense overturn occurred due to the thermobaric instability, leading to the rapid warming events observed in the paleoclimate record [1].

On the other side of the planet, cool, fresh surface water flows above warm, salty water near the coasts of Antarctica. Because the salinity is weakly stratified, it is susceptible to both double-diffusive convection and thermobaric instability, making it surprisingly similar to the deep Arctic, rather than its shallow Arctic counterpart. Susceptibility to these two instabilities has been observed in the Weddell sea. Here, the thermobaric instability has been invoked to explain periods of persistent deep convection and the formation of polynas [7], which have important implications for general ocean circulation as well as sea ice formation. Secondly, a recent study has suggested that double-diffusive convection may also have an important impact on the vertical heat transport in this area [11].

Despite their coexistence in these important and exciting areas of the ocean, neither the impact of thermobaricity on double-diffusive staircases, nor the influence of double-diffusion on the thermobaric instability has been examined. Here, we take the first step in studying how these two effects may interact by recreating a simple 1-D model of double-

diffusive staircases [4] and, for the first time, including thermobaric effects. Specifically, we examine how large changes in the thermal expansion coefficient affect the structure of the double-diffusive staircases found in the polar oceans. In the following section, these two instabilities are described in greater detail. In Section 3, we present the theory and numerical model which form the basis of our work, followed by our additions that allow us to include thermobaric effects in Section 4. Preliminary results are discussed in Section 5 and lastly, possible implications and future work is discussed in Section 6.

2 One system, two instabilities

When a pot of water is being heated from below, regular thermal convection can occur because cold water is denser than warm water. However, when a stable salinity gradient is added, the stability of the system depends on the ratio of the two background gradients. This quantity is well described by the density ratio: $R_0 = \frac{\beta \overline{S_z}}{\alpha \overline{T_z}}$, where $\overline{T_z}$ and $\overline{S_z}$ are the background temperature and salinity gradients, α is the thermal expansion coefficient, and β is the saline contraction coefficient. Thus, a larger density ratio describes a more stable system and vice versa (see Figure 3).

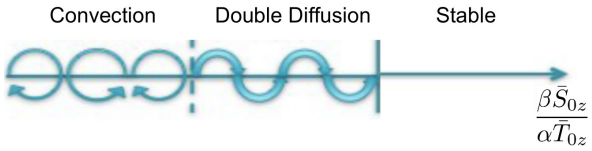


Figure 3: Cartoon describing the various stages of stability as function of the density ratio (see text) in a system with a negative temperature and salinity gradient.

2.1 Thermobaric instability

The thermal expansion coefficient is a negative value that is defined as

$$\alpha = \frac{1}{\rho_0}\left(\frac{\partial \rho}{\partial T} + \frac{\partial^2 \rho}{\partial T \partial P} + \frac{\partial^2 \rho}{\partial T^2} + ...\right), \tag{1}$$

where ρ_0 is a mean density, p is pressure, and T is temperature. The first term is negative and implies that the density of a parcel decreases as its temperature increases. The two nonlinear terms describe cabbeling and thermobaricity, respectively, which imply that α varies as both a function of pressure and temperature (Figure 4). This implies that by simply moving our system to a deeper pressure, the density ratio (and the stability) will decrease.

Moreover, as described by the thermobaric term, cold water is more compressible than warm water. Therefore, the thermal expansion coefficient increases with depth more rapidly for cold water. This allows for the possibility of a local instability in which a small perturbation to the system could cause a cool, fresh water parcel to fall towards the warmer, saltier part of the system. The colder parcel compresses enough (equivalently, α increases enough) to make the parcel denser than its salty surroundings causing it to continue to fall

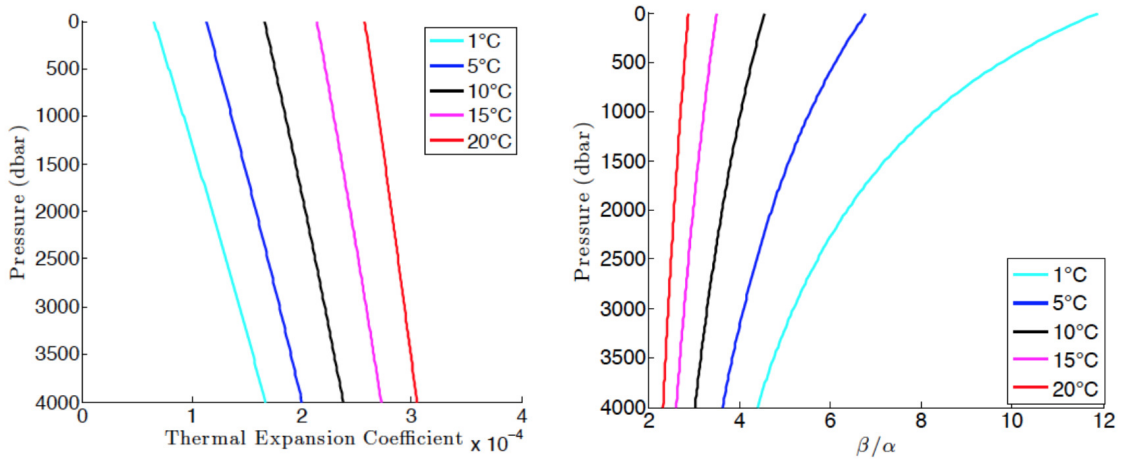


Figure 4: The thermal expansion coefficient (left) and the ratio of the salinity contraction coefficient to the thermal expansion coefficient (right) as a function pressure for several temperatures. The figures demonstrate that the thermal expansion coefficient increases with depth more rapidly in cold water compared to warm water.

towards the warm, salty part of the system [3],[6]). This instability is notably different and often stronger compared to regular thermal convection because it is internally driven by this thermobaric effect rather than continually forced at a boundary [2], [7].

2.2 Double-diffusive convection

Between the stable and fully convective states lives the double-diffusive instability (Figure 3). The initial instability of double-diffusive convection is well-described by a linear stability analysis [17]. Conceptually, it can be understood by imagining a small perturbation that causes a cool, fresh fluid parcel to fall toward the warm, saltier region of the system. The parcel will quickly diffuse in the surrounding heat but its salinity will remain nearly constant, causing the parcel to become warmer but remain fresher than its surroundings. Now much lighter than its environment, the parcel will float to a point somewhat higher than its initial position, where it will quickly diffuse its heat, maintaining its salinity to become denser than its surroundings and fall into the warm, salty area below. This fluid motion continues to oscillate, with the amplitude growing with each oscillation. Eventually, turbulent effects become dominant and a coherent structure emerges. This is called a thermohaline staircase, where each 'step' is made up of a cool, fresh layer over a warmer, saltier layer separated by a well-mixed interface.

3 Previous work

Turner and Stommel[14] performed the first laboratory experiments in which a stable salt gradient was heated from below. They found that within a few minutes, the bottom layer would overturn and form a well mixed layer. Within the hour, a series of well-mixed layers

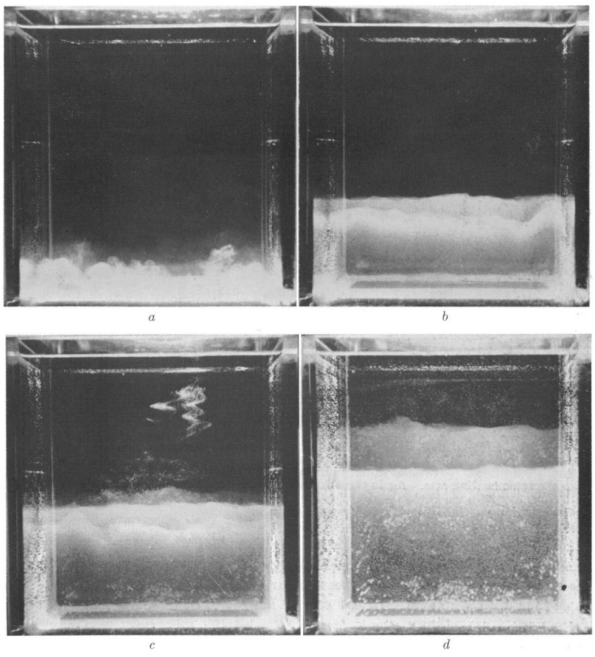


Figure 5: Image from Stommel and Turner 1964, showing results from a laboratory experiment where a stable salinity gradient is heated from below. (a) shows that one mixed layer has formed after 10 minutes. (b)-(d) shows that subsequent layers form after 25, 60 and 90 minutes, respectively.

(i.e. a double-diffusive staircase) formed in a similar fashion (Figure 5). In this section, we discuss a theory and numerical model that was developed to describe what had been seen in these laboratory experiments and form the foundation of our work.

3.1 Theory

Turner[15] describes the evolution of a system with an initially linear, negative salt gradient, $\overline{S_z}$, and a uniform background temperature, $\overline{T}$, that is heated from below by a constant heat flux, H (Figure 6). In particular, the evolution of the first layer as well as the growth and eventual formation of the second layer are discussed.

3.1.1 Evolution of the first layer

By implementing the conservation of heat and salt, the salinity (S_1), temperature (T_1) and thickness (h_1) of the first well mixed layer can be described by

$$\Delta S_1 = \frac{-1}{2} h \overline{S_z}, \qquad Ht = \rho c h_1 \Delta T_1, \tag{2}$$

where ΔS_1 and ΔT_1 are the change in salt and temperature across the top of the 1st layer (and are positive), t is time, ρ is the mean density of the system and c is the specific heat. To close this system of equations, we invoke

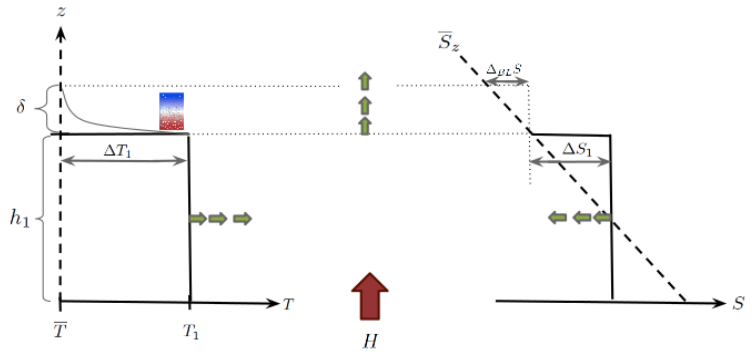


Figure 6: Cartoon depiction of the theory describing the formation of the first two mixed layers. The green arrows indicate that the layer warms, freshens and grows with time.

$$\alpha \Delta T_1 + \beta \Delta S_1 = 0, \tag{3}$$

which relates the two interfaces and implies that the density contributions from the heat and salt combine to create a continuous density profile (Figure 7).

By algebraically solving these equations, the expressions for the evolution of ΔT_1, ΔS_1 and h_1 are

$$h_1 = \sqrt{\left(\frac{\widetilde{H}_* t}{\widetilde{S}_*}\right)}, \qquad -\alpha g \Delta T_1 = \beta g \Delta S_1 = \sqrt{\widetilde{H}_* \widetilde{S}_* t}, \tag{4}$$

where $\widetilde{H}_*$ and $\widetilde{S}_*$ are the buoyancy fluxes, defined by

$$\widetilde{H}_* = \frac{-\alpha g H}{\rho c}, \qquad \widetilde{S}_* = \frac{-\beta g \overline{S_z}}{2}. \tag{5}$$

Thus, this layer, which is heated from below, warms and grows taller with time. As the layer grows, it overtakes the less salty water above and becomes fresher (Figure 6).

3.1.2 Evolution of boundary layer, formation of layer 2

While this layer grows, heat and salt are diffused through the top of the interface, causing a boundary layer to form. Because heat diffuses 100 times faster than salt, the diffusion of salt through the interface is neglected and a simple heat equation is used to describe its thermal evolution,

$$\frac{\partial \theta}{\partial t} = \kappa \frac{\partial^2 \theta}{\partial z^2}, \qquad \begin{aligned} \theta(z = h_1) &= T_1 \\ \theta(z \to \infty) &= \overline{T}, \end{aligned} \tag{6}$$

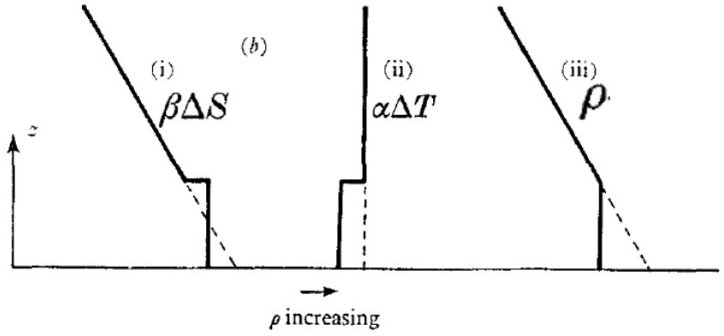


Figure 7: Cartoon depiction of equation (3)

where κ is the thermal diffusivity, θ is the temperature of the boundary layer and $\overline{T}$ is the initial uniform background temperature.

As heat diffuses through the boundary layer, cool, fresh water lies above increasingly warmer, salty water. These are the conditions necessary for the double-diffusive instability to occur. Turner suggests that this boundary layer will overturn and consequently form a second convective mixed layer when the temperature gradient grows strong enough such that it is linearly unstable to double-diffusive convection. The criterion for overturn is then given by Veronis[16], who found that the onset for double-diffusive convection between two freely moving boundaries is given by

$$R = R_T - \frac{\sigma}{\sigma + 1} R_S \geq \frac{27\pi^4}{4}, \tag{7}$$

where $R_T = \frac{-\alpha g \Delta_{BL} T \delta^3}{\kappa\nu}$, $R_S = \frac{\beta g \Delta_{BL} S \delta^3}{\kappa\nu}$, δ is the scaling length of the boundary layer, ν is the kinematic viscosity, and $\sigma = \frac{\nu}{\kappa}$ is the Prandtl number. Δ_{BL} is a positive value that signifies the change in temperature and salinity above the interface. Thus,

$$\Delta_{BL} T = T_1 - \overline{T}, \qquad \Delta_{BL} S = -\overline{S_z}\delta,$$

because we have assumed that the diffusion of salt through the top of the layer can be neglected (Figure 6).

3.2 Numerical model

Using the framework put forth in [15], Huppert[4] developed a one-dimensional numerical model that could describe this system with multiple (N) layers. However, some additional assumptions are included in this model:

- Loss of heat through the top of the staircase can be neglected in the heat budget.
- Only the top layer grows. Interior layers have a constant thickness unless they merge together to form a thicker layer.
- If two layers create a neutral density profile, the two layers merge.

- heat and salt fluxes across the interface are given by experimentally determined flux laws [5]:

$$\begin{aligned} \phi_i &= 0.32\kappa \left(\frac{-\alpha g}{\kappa\nu}\right)^{1/3} (\Delta T_i)^{4/3} \left(\frac{\alpha \Delta T_i}{\beta \Delta S_i}\right)^2 \\ \psi_i &= \begin{cases} \frac{-\alpha}{\beta}\phi_i \left(1.85 + 0.85\left(\frac{\beta\Delta S_i}{\alpha \Delta T_i}\right)\right) & , 1 < \frac{\beta\Delta S}{-\alpha\Delta T} \leq 2 \\ -0.15\phi_i \left(\frac{\alpha}{\beta}\right) & , 2 < \frac{\beta\Delta S}{-\alpha\Delta T}. \end{cases} \end{aligned} \tag{8}$$

(Equations (8) appear slightly differently in [5] because we have followed the formalism of [15] and [4] and defined α as a negative value).

3.2.1 Evolution of initial layer

The equations describing the evolution of the initial layer are given by equation 4. When only one layer is present in the simulation, the magnitude of the interfaces are given by

$$\Delta T_1 = T_1 - \overline{T_t} \qquad \Delta S_1 = S_1 - \overline{S_t}. \tag{9}$$

$\overline{T_t}$ and $\overline{S_t}$ are the background temperature and salinity at the top of the staircase, which can be expressed as

$$\overline{T_t} = \overline{T}, \qquad \overline{S_t} = \overline{S_z} h_1 + \overline{S_0}, \tag{10}$$

where $\overline{S_0}$ is the initial salinity at the bottom of the system (Figure 8).

3.2.2 Evolution of multiple layers

Once two or more layers are present in the staircase, flux conservation laws are used to describe the evolution of the layers. The general form of these equations are

$$\rho c \frac{\mathrm{d}}{\mathrm{d}t}[h_i(T_i - \overline{T_i}] = \rho c(\phi_{i-1} - \phi_i), \qquad \frac{\mathrm{d}}{\mathrm{d}t}[h_i(S_i - \overline{S_i})] = \psi_i - \psi i - 1,$$

where ϕ_i and ψ_i are the fluxes through the ith layer.

$$\overline{T_i} = \overline{T}, \qquad \overline{S_i} = \overline{S_z}(d_i - h_i/2) + \overline{S_0}$$

are the background temperature and salinity imposed by the initial conditions, where $d_i = \sum_{j=1}^{i} h_j$.

3.2.3 Evolution of interior layers

Because $\overline{T_i}$,$\overline{S_i}$, and h_i are constant for all interior layers ($i < N$), equation (11) becomes

$$\rho c h_1 \frac{\mathrm{d}T_1}{\mathrm{d}t} = H - \rho c \phi_1, \qquad h_1 \frac{\mathrm{d}S_1}{\mathrm{d}t} = -\psi_1 \tag{11}$$

for the first layer, and

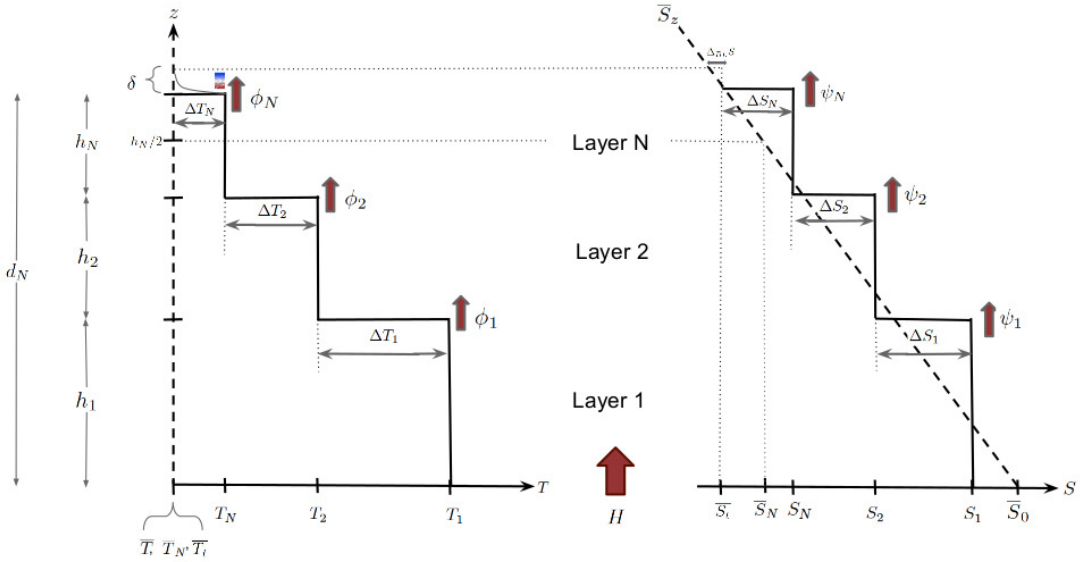


Figure 8: Cartoon depiction of the N-layer model.

$$h_r \frac{\mathrm{d}T_r}{\mathrm{d}t} = \phi_{r-1} - \phi_r, \qquad h_r \frac{\mathrm{d}S_r}{\mathrm{d}t} = \psi_{r-1} - \psi_r, \tag{12}$$

for all other layers, where $1 < r \leq N-1$.

3.2.4 Evolution of top layer

The thickness of the top layer grows with time, and the heat and salt flux exiting the top of the staircase are neglected, thus equation (11) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(h_N(T_N - \overline{T_N})\right) = \phi_{N-1}, \qquad \frac{\mathrm{d}}{\mathrm{d}t}\left(h_N(S_N - \overline{S_N})\right) = \psi_{N-1}. \tag{13}$$

Finally, by satisfying equation (3), the third equation describing the top layer is

$$\alpha\left(T_N - \overline{T}_t\right) + \beta\left(S_N - \overline{S}_t\right) = 0.$$

Again, $\overline{T}_t = \overline{T}$, but now $\overline{S}_t = \overline{S}_z d_N + \overline{S}_0$.

3.2.5 Boundary layer

Equation (6) describes the evolution of the initial layer and boundary layer, with the adjusted boundary condition, $\theta(z = d_N) = T_N$. The criterion for overturn is given by equation (7) where, by assuming that the temperature profile of the boundary layer becomes relatively linear at the time of overturn, we let

$$\delta = \frac{2}{T_N - \overline{T_N}} \int_{d_N}^{\infty} \theta \mathrm{d}z.$$

3.2.6 Formation of new, N+1 layer

Once a boundary layer has overturned to form a new layer, the characteristics of this new layer are determined by conserving heat and salt and by satisfying equation (3):

$$\begin{aligned} \left(T_{N+1} - \overline{T}\right) h_{N+1} = \int_{d_N}^{\infty} \theta \mathrm{d}z, \qquad S_{N+1} &= \overline{S}_z \left(d_N + \frac{h_{N+1}}{2} \right) + \overline{S}_0 \\ \alpha \left(T_{N+1} - \overline{T}_t\right) + \beta \left(S_{N+1} - \overline{S}_t\right) &= 0, \end{aligned}$$

where the salinity at the top of the staircase has become $\overline{S}_t = \overline{S}_z \left(d_N + h_{N+1}\right) + \overline{S}_0$.

3.2.7 Merging layers

Lastly, if the density is neutral across an interface (i.e. $\alpha \Delta T + \beta \Delta S = 0$), the two layers will merge while conserving heat and salt. So the characteristics of the new layer are given by

$$\begin{aligned} T_{new} &= \frac{h_i T_i + h_{i+1} T_{i+1}}{h_i + h_{i+1}}, \qquad S_{new} = \frac{h_i S_i + h_{i+1} S_{i+1}}{h_i + h_{i+1}}, \\ h_{new} &= h_i + h_{i+1}. \end{aligned}$$

4 Adding thermobaric effects

The equations presented in the previous section differ slightly from those found in [4]. This is because they assume that $\overline{T} = \overline{S}_0 = 0$ and also because their equations are written using the following non-dimensional scaling where hatted variables indicate dimensional quantities:

$$t = \sqrt{S_*}\widehat{t}, \qquad z = \frac{S_*^{3/4}}{H_*^(1/2)} \widehat{z}, \qquad T = \frac{-\alpha_0 g \widehat{T}}{H_*^{1/2} S_*^{1/4}}$$

$$S = \frac{\beta g \widehat{S}}{H_*^{1/2} S_*^{1/4}}, \qquad \phi = \frac{-\alpha_0 g \widehat{\phi}}{H_*}, \qquad \psi = \frac{\beta g \widehat{\psi}}{H_*}.$$

The scaling for the buoyancy fluxes are defined as

$$H_* = \frac{-\alpha_0 g H}{\rho c}, \qquad S_* = \frac{-\beta g \overline{S_z}}{2}, \tag{14}$$

where α_0 is a constant, mean thermal expansion coefficient (note that this differs from $\widetilde{H}_*$).

In order to include thermobaric effects, we simply introduce a new, non-dimensional variable,

$$\alpha = \frac{\widehat{\alpha}}{\alpha_0}.$$

We set $\widehat{\alpha} = \alpha_0 + c\widehat{z}$, where c is determined by linearizing alpha near the local pressures of each experiment. The resulting equations using this non-dimensionalization are given below.

4.1 Heat and salt fluxes through interfaces

Non-dimensionalizing equations (8) leads to

$$\begin{aligned}
\phi &= 0.32\left(\frac{\alpha Q}{\sigma}\right)^{1/3}(\Delta T)^{4/3}\left(\frac{\alpha\Delta T}{\Delta S}\right)^{2} \\
\psi &= \begin{cases} \alpha\phi\left(1.85 - 0.85\left(\frac{\Delta S}{\alpha\Delta T}\right)\right), & 1 < \frac{\Delta S}{\alpha\Delta T} \le 2 \\ 0.15\alpha\phi, & 2 < \frac{\Delta S}{\alpha\Delta T}. \end{cases}
\end{aligned}$$

4.2 Growing 1st layer

Non-dimensionalizing equations (4) leads to

$$h_1 = \sqrt{\alpha t}, \qquad T_1 = \overline{T} + \sqrt{\frac{t}{\alpha}}, \qquad S_1 = \overline{S_0} - \sqrt{\alpha t}.$$

4.3 Evolution of interior layers

Non-dimensionalizing equations (11) and (12) leads to

$$\begin{aligned}
h_1\frac{\mathrm{d}T_1}{\mathrm{d}t} &= 1 - \phi_1, \qquad & h_1\frac{\mathrm{d}S_1}{\mathrm{d}t} &= \psi_1 \\
h_r\frac{\mathrm{d}T_r}{\mathrm{d}t} &= \phi_{r-1} - \phi_r, \qquad & h_r\frac{\mathrm{d}S_r}{\mathrm{d}t} &= \psi_{r-1} - \psi_r.
\end{aligned}$$

4.4 Evolution of top layer

Non-dimensionalizing equations (13) leads to

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\left(h_N(T_N - \overline{T_0})\right) &= \phi_{N-1}, \qquad & \frac{\mathrm{d}}{\mathrm{d}t}\left(h_N\right)) &= \psi_{N-1} - \frac{\mathrm{d}h_N}{\mathrm{d}t}(2d_N - \overline{S_0}) \\
\alpha(T_N - \bar{T}) &= S_N + 2d_N - \bar{S}_0.
\end{aligned}$$

After much algebra, these equations can be rewritten as

$$\begin{aligned}
\frac{\mathrm{d}h_N}{\mathrm{d}t} &= \frac{1}{2h_N}(\alpha\phi - \psi), \qquad \frac{\mathrm{d}T_N}{\mathrm{d}t} = \frac{1}{h_N}\left(\phi - \frac{T_N - \overline{T}}{2h}(\alpha\phi - \psi)\right) \\
\frac{\mathrm{d}S_N}{\mathrm{d}t} &= \frac{1}{h_N}\left(\psi - \frac{\alpha\phi - \psi}{2h_N}(S_N + 2d_N - \overline{S_0})\right).
\end{aligned}$$

4.5 Evolution of boundary layer

Non-dimensionalizing equations (6) and (7) leads to

$$\frac{\partial\theta}{\partial t} = Q\frac{\partial^2\theta}{\partial z^2}, \qquad R = \frac{\delta^3}{\sigma Q^2}\left(\alpha(T_N - \overline{T}) - \frac{2\sigma\delta}{\sigma+1}\right).$$

This is semi-analytically solved using the solution for heat through a finite, one-dimensional system with temperatures fixed on each boundary, where the length of the system, x_f, is chosen such that it is effectively infinity. Following [8], that semi-analytical solution is given by

$$\theta(x,t) = v + w, \qquad v(x) = T_N + (\overline{T} - T_N)\frac{x}{x_f} \tag{15}$$

$$w(x,t) = \sum_{n=1}^{N_{max}} b_n \sin\lambda_n x \exp -\lambda_n^2 kt, \qquad \lambda_n = \frac{n\pi}{x_f} \tag{16}$$

$$b_n = \frac{2}{x_f}\int_0^{x_f} -v(x)\sin\frac{n\pi x}{x_f}\mathrm{d}x. \tag{17}$$

4.6 Formation of new layer

Non-dimensionalizing equations (14) leads to

$$\begin{aligned}
\left(T_{N+1} - \overline{T}\right)h_{N+1} = \int_{d_N}^{\infty}\theta\mathrm{d}z \equiv \gamma, \qquad S_{N+1} &= -2\left(d_N + \frac{h_{N+1}}{2}\right) + \overline{S}_0 \\
\alpha\left(T_{N+1} - \overline{T}_t\right) + \left(S_{N+1} + 2(d_N + h_{N_1}) - \overline{S_0}\right) &= 0.
\end{aligned}$$

After much algebra, we find that this leads to

$$\begin{aligned}
h_{N+1} &= \sqrt{\alpha\gamma}, \qquad T_{N+1} = \overline{T} + \sqrt{\gamma/\alpha} \\
S_{N+1} &= \overline{S_0} - 2d_N - \sqrt{\gamma\alpha}.
\end{aligned}$$

4.7 Merge layers

Finally, while the criterion for merging becomes $\alpha\Delta T = \Delta S$, the equations which describe the new layer that results from merging (equations (14)), remain unchanged.

5 Preliminary results

To begin exploring how the thermobaric instability and double-diffusive convection interact, we first address the question of how the structure of the staircases may change as a function of α. One could imagine two possible outcomes of such an experiment. First, increasing α may have the same effect as increasing the heating rate or increasing the background temperature gradient. This would cause the density ratio to decrease as α increases, leading to a less stable system that transports heat and salt more efficiently. Alternatively, one could imagine that the steps might reorganize in such a way as to maintain its density ratio. In this case, the salinity interface would increase or the temperature interface would decrease, causing $\overline{S_z}/\overline{T_z}$ to increase and the density ratio to remain constant. Finally, in either case, we would expect the overall thickness of the layers to increase with α in order to sustain the larger interfaces or fluxes that would result in either scenario.

To try to answer this question, we ran our model using the same parameters that were used by [4] for a variety of αs. Specifically, we chose αs such that they were equivalent to systems where double-diffusive staircases are found in the deep and shallow Arctic, a laboratory and finally, one that would correspond to the Marianas trench, as a way to gauge how the staircases would behave for very large, yet realistic expansion coefficients (see Tables 1 and 2 for full description of parameters used). Although α is computed as a function of depth within each experiment, the staircases presented are sufficiently small that we may consider α to be constant.

Our results are presented in Figures 9-14. Figure 9 shows how the top of each layer evolves in time for each experiment and demonstrates that layers appear to form and merge more rapidly for larger values of α. This can be rationalized upon inspection of equation (7), the expression for R, which must reach a critical value in order for a new layer to form. One can see that R increases as a function of α and thus a smaller temperature gradient is required for overturn.

The rapid formation of layers also appears to correspond to rapid merging events. This might be because a layer that takes a short time to form will be smaller, less resilient and will then quickly have a similar density to the layer below compared to a layer that takes longer to form.

Secondly, the fourth subplot of Figures 12-14, which show the evolution of the layer heights for the first three layers and compares them to each experiment, demonstrate that once a quasi-steady layer forms (i.e. persists without merging immediately), it is likely to be larger when formed under conditions with a larger α. This result is consistent with what is observed in the shallow Arctic compared to the deep Arctic (see Figure staircase).

This result is further supported by Figures 10 and 11, which show the final form of the staircase at the end of each experiment. Moreover, one can see that the interfaces in salinity seem to increase as a function of α, while the temperature interfaces behave in the opposite manner. This can be explored further through the first and second subplots of Figures 12-14, which show how the two interfaces evolve for the first three layers. Although the result is far from indisputable, the same trend seems to exist, particularly for the more stable bottom two layers.

The size of the temperature interfaces, ΔT, are closely related to the fluxes across the interfaces (equation (8) but are also functions of α. These two quantities appear to have

opposite effects on both fluxes. Interestingly, the fifth and sixth subplots of Figures 12-14 clearly demonstrate that the heat flux decreases with α while the salt flux (even more strongly) increases with α. This implies that for the heat flux, the impact of ΔT outweighed that of α while the opposite was true for the salt flux. Again, these results are most clear for the more stable bottom two layers.

Lastly, we test our hypothesis from the beginning of this section and analyze how the density ratio varies for large changes in α. Upon inspection of the third and sixth subplot of Figures 12-14, it appears that the ratio of the two interfaces, $\Delta S/\Delta T$, increases with α, leaving the density ratio fairly constant by comparison (although there may be a small decrease as a function of α). This result suggests that increasing α has consequences which may not include destabilizing the water column but are reflected in the structure of the staircase.

Table 1: Simulation Parameters

κ	σ	H_*	S_*	T_{BG} (°C)	S_{BG} (psu)
$1.4 \cdot 10^{-7}$	7	$-1.48\alpha_0 \cdot 10^{-2}$	.167	.1	35

Table 2: Experiments

Description	α_0 (10^{-4} °C^{-1})	c (10^{-8} °C^{-1}m^{-1})	depth range (m)
lab	-.53	-2.88	0-.85
shallow Arctic	-.61	-2.86	292-316
deep Arctic	-1.24	-2.6	2611-2814
Marianas Trench	-2.92	-1.67	10800-10860

6 Implications, speculations and future work

Although we have not yet directly tested the effects of thermobaricity on double-diffusive staircases, these results help us to understand how the staircases change as a function of α. From our results, we might expect that in a system where thermobaric effects are important (i.e. α increases quickly with depth), the staircase will have certain features. First, we would expect that the layer thickness and salinity interfaces would increase with depth while the the temperature interfaces would decrease with depth (or increase less with depth , as this seems to be a necessary characteristic of all staircases). This in turn would lead to a divergent heat flux and a convergent salinity flux.

These results may explain why and how double-diffusive steps appear to increase with depth in nature. Furthermore, one could speculate that these thermobaric effects on the staircases may be important for understanding thermobaric convection. For example, between thermobaric overturning events, one might suspect that the system could be double-diffusively unstable (see Figure 3) and would therefore contain double-diffusive staircases. If the staircase supports converging and diverging vertical fluxes, they may impact the rate

at which the system reaches the density ratio required for the system to be thermobarically unstable.

However, these are merely speculations and require further scrutiny. Specifically, in the future we aim to determine how and under what conditions we might expect thermobaric effects to be important. This will be done by comparing experiments in which α is held constant to those in which α is allowed to vary appreciably within a staircase, based on the temperature gradient. We will explore a parameter range in which we will vary the heating rate, the salinity gradient and α. These experiments will include simulations with parameters that correspond to those found in the Antarctic as well as the shallow and deep Arctic, where both double-diffusion and thermobaricity may be important. This will hopefully lead us to a nice set of predictions that will allow us to begin to compare our results to observations.

7 Acknowledgements

A big thank you to the staff of the 2014 GFD summer school, particularly the organizers Glenn Flierl, Raffaele Farrari and Antonello Provenzale. The principal lecturers, Geoff Vallis and Kerry Emanuel are also acknowledged for impressively packing the cottage and packing our heads with knowledge. George Veronis is thanked for patiently waiting for our softball skills to arrive all summer. The friendship of the fellows and all of those who spent their summer on the porch helped make this summer especially fun. Finally, Mary-Louise Timmermans, George Veronis and Andy Ingersoll are warmly thanked for their time and supervision of this project.

References

[1] J. F. Adkins, A. P. Ingersoll, and C. Pasquero, *Rapid climate change and conditional instability of the glacial deep ocean from the thermobaric effect and geothermal heating*, Quaternary Science Reviews, 24 (2005), pp. 581–594.

[2] K. Akitomo, T. Awaji, and N. Imasato, *Open-ocean deep convection in the Weddell Sea: two-dimensional numerical experiments with a nonhydrostatic model*, Deep Sea Research Part I: Oceanographic Research Papers, 42 (1995), pp. 53–73.

[3] A. Gill, *Circulation and bottom water production in the Weddell Sea*, Deep Sea Research and Oceanographic Abstracts, 20 (1973), pp. 111–140.

[4] H. E. Huppert and P. F. Linden, *On heating a stable salinity gradient from below*, Journal of Fluid Mechanics, 95 (1979), p. 431.

[5] H. E. Huppert and J. S. Turner, *Double-diffusive convection and its implications for the temperature and salinity structure of the ocean and Lake Vanda*, Journal of Physical Oceanography, 2 (1972), pp. 456–461.

[6] T. B. Lø yning and J. E. Weber, *Thermobaric effect on buoyancy-driven convection in cold seawater*, Journal of Geophysical Research, 102 (1997), p. 27875.

[7] M. G. McPhee, *Marginal Thermobaric Stability in the ice-covered upper ocean over Maud Rise*, Journal of Physical Oceanography, 30 (2000), pp. 2710–2722.

[8] D. Powers, *Boundary Value Problems, Sixth Edition: and Partial Differential Equations*, Elsevier, 2006.

[9] T. Radko, *Double-Diffusive Convection*, Cambridge University Press, Cambridge, 2013.

[10] B. Rudels, E. P. Jones, U. Schauer, and P. Eriksson, *Atlantic sources of the Arctic Ocean surface and halocline waters*, Polar Research, 23 (2004), pp. 181–208.

[11] W. J. Shaw and T. P. Stanton, *Dynamic and double-diffusive instabilities in a weak pycnocline. Part I: Observations of heat flux and diffusivity in the vicinity of Maud Rise, Weddell Sea*, Journal of Physical Oceanography, 44 (2014), pp. 1973–1991.

[12] M.-L. Timmermans, L. Rainville, L. Thomas, and a. Proshutinsky, *Moored observations of bottom-intensified motions in the deep Canada Basin, Arctic Ocean*, Journal of Marine Research, 68 (2010), pp. 625–641.

[13] M.-L. Timmermans, J. Toole, R. Krishfield, and P. Winsor, *Ice-Tethered Profiler observations of the double-diffusive staircase in the Canada Basin thermocline*, Journal of Geophysical Research, 113 (2008), p. C00A02.

[14] J. Turner and H. Stommel, *A new case of convection in the presence of combined vertical salinity and temperature gradients*, Proceedings of the National Academy of Sciences, 52 (1964), p. 49.

[15] J. S. Turner, *The behaviour of a stable salinity gradient heated from below*, Journal of Fluid Mechanics, 33 (1968), pp. 183–200.

[16] G. Veronis, *On finite amplitude instability in thermohaline convection*, Journal of Marine Research, (1965), p. 1.

[17] G. WALIN, *Note on the stability of water stratified by both salt and heat*, Tellus, 16 (1964), pp. 389–393.

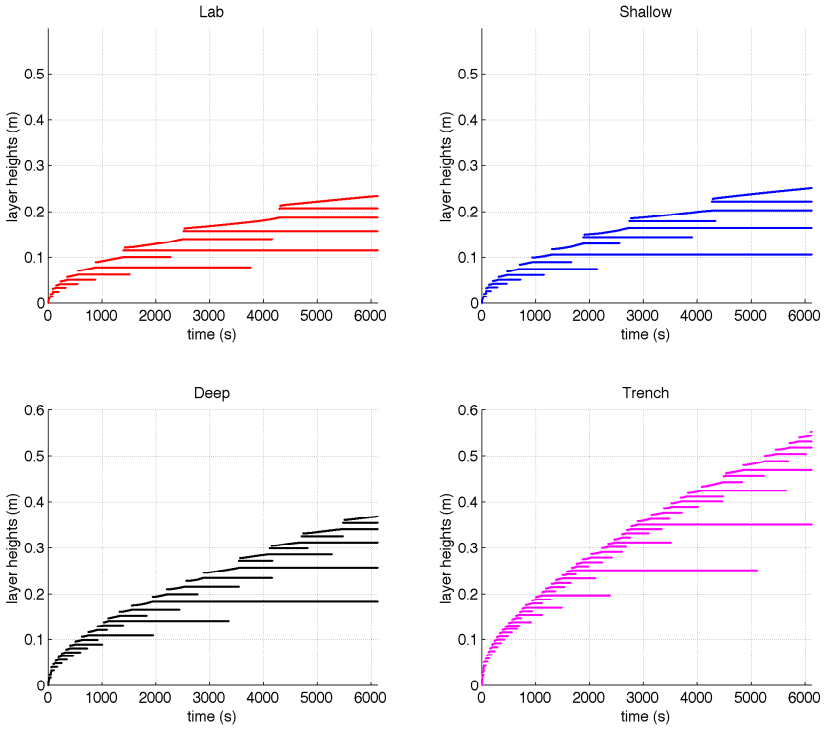


Figure 9: The evolution of the layer heights for the four experiments.

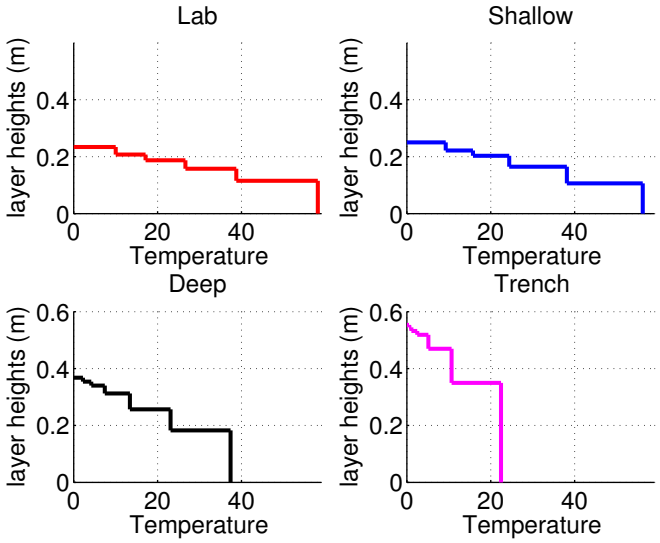


Figure 10: Final form of the staircases in temperature-depth space for the four experiments.

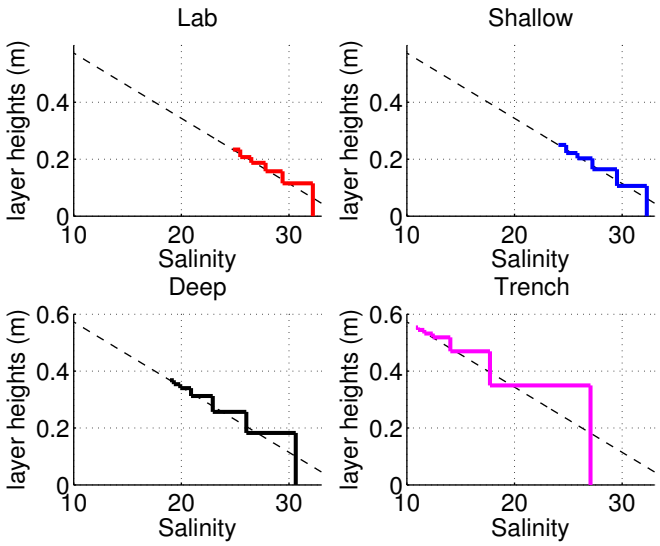


Figure 11: Final form of the staircases in salinity-depth space for the four experiments.

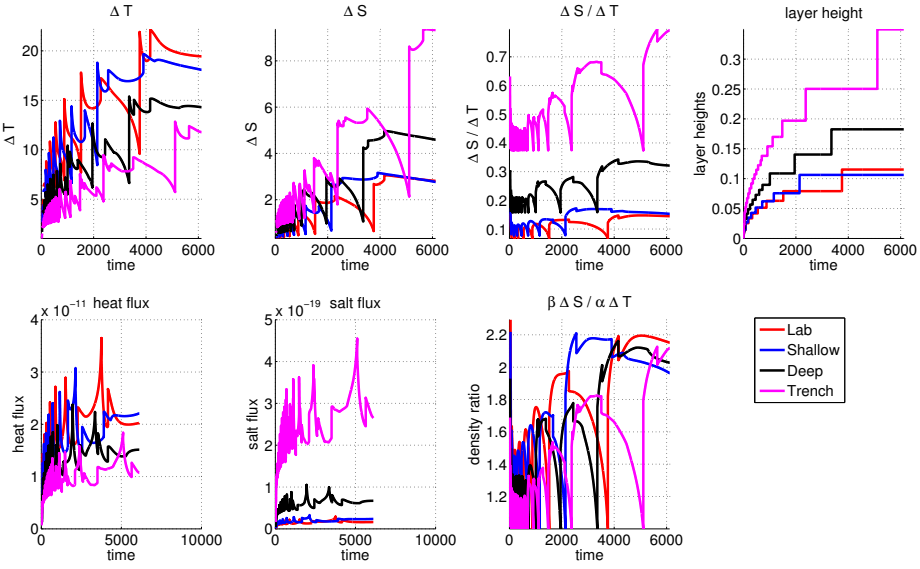


Figure 12: Comparison of the bottom layer between the four experiments. Evolution of the layer heights, magnitude and ratio of the interfaces, heat and salt fluxes, and density ratio is plotted.

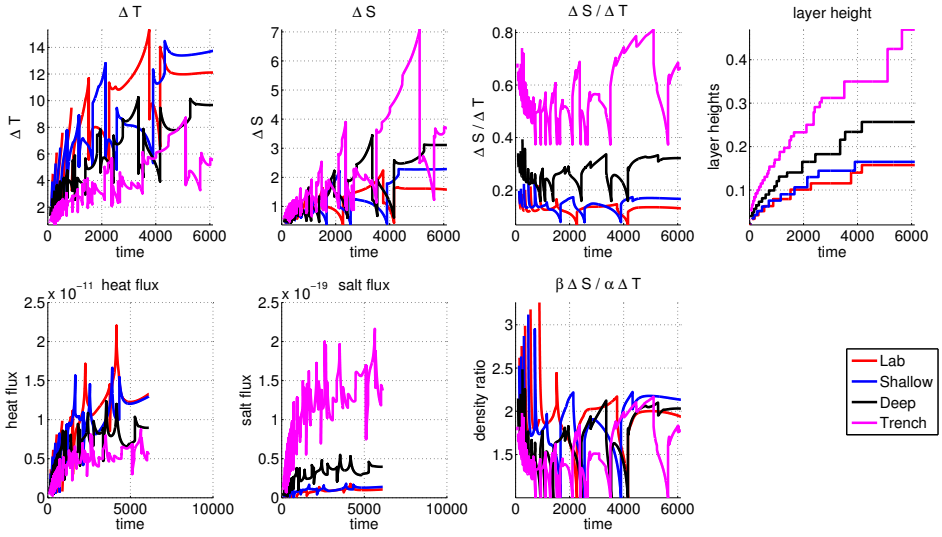


Figure 13: Comparison of the 2nd layer between the four experiments. Evolution of the layer heights, magnitude and ratio of the interfaces, heat and salt fluxes, and density ratio is plotted.

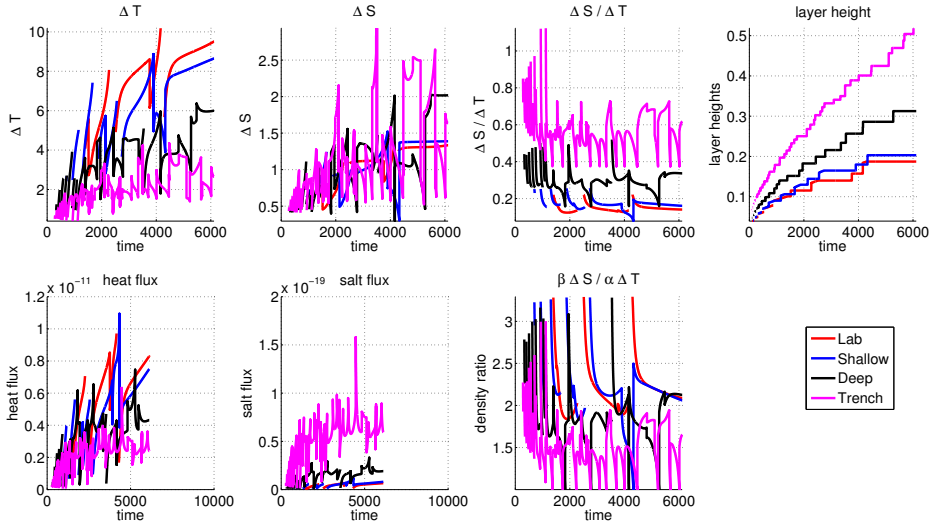


Figure 14: Comparison of the third layer between the four experiments. Evolution of the layer heights, magnitude and ratio of the interfaces, heat and salt fluxes, and density ratio is plotted.

The Most Minimal Seed for Transition to Turbulence in Shear Flow

Geoff Stanley

September 30, 2014

Abstract

A key question to develop our understanding of turbulence in shear flows is: what is the smallest perturbation to the laminar flow that causes a transition to turbulence, and how does this change with the Reynolds number R? Finding this so-called "minimal seed" is as yet unachievable in direct numerical simulations of the Navier-Stokes equations, but there exist low-dimensional dynamical systems that model those aspects of the full flow which are considered essential to turbulence. We search for the minimal seed in one such model, owing to Waleffe (1997). We employ an optimization technique, reviewed by Kerswell *et al.* (2014), to find non-linear optimal perturbations, with some modifications. In particular we apply the technique to a new regime in which the edge that is sought is an internal boundary to the basin of attraction of the laminar flow state. We find such a boundary in Waleffe's model and calculate the minimal seed on this edge.

1 Introduction

1.1 Shear Turbulence

It is well known that the Navier-Stokes equations for fluid flow present one of the greatest challenges of classical physics. The equations in general have not been solved analytically, and even numerical solutions have been unobtainable until recent decades owing to the vast range of scales that must be captured in a turbulent flow. One aim of study has been to consider the simplified case of a constant density, incompressible fluid with simple boundary conditions such as for flow through a pipe first studied by Reynolds, or between two concentric cylinders as first studied by Couette. Yet fluid flow even in these idealized settings remains far from understood. One major problem is how a laminar flow transitions to a turbulent flow.

For instance, plane Poiseuille flow[1] (PPF) admits a laminar flow that is linearly stable for all $R < R_c = 5772.22$, where $R = LU/\nu$ is the Reynolds number, R_c is the critical Reynolds number, L is the half-width of the channel, U is the maximum flow speed, and ν is

[1] This is flow between two infinite parallel plane sheets with a no-slip boundary condition and forced by a pressure gradient also parallel to the sheets.

the kinematic viscosity (Orszag, 1971). Yet laboratory experiments of such flows typically observe the laminar flow transitioning to turbulence at much lower Reynolds numbers, around $R \approx 1000$. A similar scenario exists for plane Couette flow[2] (PCF): the laminar flow is linearly stable for all R out to ∞ but turbulence is typically observed with Reynolds numbers as low as $R \approx 350$. This disagreement between experiment and linear stability theory ultimately owes to noise in the experimental setup: finite perturbations to the linear flow can, in dynamical systems parlance, move the system out of the basin of attraction of the stable laminar flow and into the basin of attraction of some other, turbulent state. Since the basin of attraction of the laminar flow diminishes in size with increasing R, a given experimental setup subject to a certain amount of noise will observe transition to turbulence at some sub-critical R.

Reversing this problem presents a key question, as posed in an illuminating paper by Trefethen *et al.* (1993): for a given R, what is the minimal perturbation to the laminar flow that causes the flow to transition to turbulence? This perturbation is called the "*minimal seed*", although sometimes this term simply refers to the magnitude of the actual perturbation. These authors conjectured that the magnitude of the minimal seed should be of order R^γ for some value of γ which will depend on the problem under consideration. For shear flows with stable eigenvalues for the laminar flow, they state $\gamma \leq -1$ citing a technical report that was later published (Kreiss *et al.*, 1994).

Following this conjecture there has been significant work to determine γ for different flows, using theory, laboratory experiments, direct numerical simulation, and low-dimensional model analogues. Many studies focus on a particular form of perturbation that is thought to be especially effective at generating turbulence—effective in the sense that a small energy injection to this mode then excites other modes in such a way as to amplify the total energy. By careful study of the asymptotics of flows with initial conditions in these particular modes,Chapman (2002) derived a R^{-1} scaling for PCF and $R^{-3/2}$ for PPF.

In the laboratory, Hof *et al.* (2003) searched for the minimal seed in pipe Poiseuille flow by injecting fluid through six small drill holes equally spaced around the circumference of the pipe downstream of where the flow was deemed to be laminar. The magnitude and total time of the injections was controlled and downstream observations made to determine whether the flow became turbulent. They found extremely good agreement with a R^{-1} scaling for the minimal seed, in this case the energy of their injections.

However, there is a common problem with the above studies: they begin with a specified form of initial perturbation, then determine the magnitude of that perturbation which is necessary for transition to turbulence. This is even true in the laboratory where the drill holes are specified and therefore allow only certain forms of perturbations. A complete solution to the minimal seed problem must determine the *form* of the perturbation as well as its magnitude. This is certainly a very difficult problem, and the above studies are important steps towards a full understanding.

Another approach, and the one taken here, is to study low-dimensional analogues of the full Navier-Stokes equations for shear turbulence. In these simpler systems it is possible (though still non-trivial) to find the exact form of the minimal seed. Many such models for

[2]This flow is the same as PPF but is forced instead by a constant velocity difference between the sheets.

various flow geometries were developed in the 1990's; Baggett & Trefethen (1997) compared them and found the particular γ for each. We focus in particular on the model by Waleffe (1997, hereafter W97). Waleffe formed by Galerkin truncation a four-dimensional model of sinusoidal flow[3] and speculated an R^{-1} scaling for the minimal seed of this flow, based on the fact that one component of the lower branch fixed point in the 4D model scales as R^{-1}. However, the minimal seed will involve non-zero values for the other three components; the possibility exists that the minimal seed could scale as R^{γ} with $\gamma < -1$. Waleffe's model also elucidated a "Self-Sustaining Process" that is supposedly a crucial building block of shear turbulence. This will be further discussed in Section 3.1; for now note that streamwise vortices are a critical flow structure in initiating turbulence: a small amount of energy in a streamwise vortex mode can, by advecting the mean shear, lead to a large transient growth of energy. Indeed these or their close cousins constitute the particular form studied intensively, as in the aforementioned studies.

The W97 model is a development of an earlier model (Waleffe, 1995,a) model that was included in the study by Baggett & Trefethen (1997). However, the 1995 and 1997 models are significantly different, and only much later was the first attempt made to find the minimal seed and the particular γ for the W97 model (Cossu, 2005). Cossu employed a three-dimensional search algorithm (having eliminated one dimension as described in Section 3.4). The particular "minimal seeds" found by Cossu are those minimal perturbations (for each R) which lead to *permanent* turbulence. However, initial conditions for the W97 model can be quickly found which demonstrate *transient* turbulence—they behave turbulently for a time but ultimately decay to the laminar state. It is these perturbations which, we argue, constitute candidates for the minimal seed.

To find these minimal seeds (over a range of Reynolds numbers), we will use a new optimization technique, recently reviewed by Kerswell *et al.* (2014) and discussed in Section 2. With some modifications, we apply it to a new class of problems wherein the turbulence is transient, versus those problems illustrated by Kerswell *et al.* or even the region of the W97 model that was studied by Cossu (2005) in which the turbulence is persistent. The W97 model itself and results from the search for its minimal seed are given in Section 3, before concluding in Section 4.

1.2 Definitions and Examples: Optimal Perturbations, The Edge, and the Minimal Seed

We now formulate the problem in dynamical systems terms and give a few brief definitions.

The Navier-Stokes equations are shifted so that the laminar state is at the origin. They may then be written in the following generic form:

$$\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} = F(\mathbf{x}) = L\mathbf{x} + N(\mathbf{x}), \tag{1}$$

where L is a linear operator and N a non-linear function. This dynamical system defines a new function $\mathbf{X}(t, \mathbf{x}_0)$ that is the solution to (1) which also satisfies $\mathbf{X}(0, \mathbf{x}_0) = \mathbf{x}_0$.

The standard dot product is used together with the L_2 norm $\|\mathbf{x}\|_2$, written simply as $|\mathbf{x}|$. For Navier-Stokes, the state $\mathbf{x}$ represents the discretized velocity field $\mathbf{u}$ and hence the

[3]This is PCF but with free-slip boundary conditions, leading to a sinusoidal crosswise profile of the laminar streamwise flow.

kinetic energy $E = \frac{1}{2}\int \mathbf{u}\cdot\mathbf{u}\ \mathrm{d}V$ is technically given by the quadratic form $\frac{1}{2}\mathbf{x}\cdot\mathbf{x}$. Since this maps one-to-one onto $|\mathbf{x}| = (2\cdot E)^{1/2}$, we often refer to $|\mathbf{x}|$ as simply the energy and sometimes write it as $\mathcal{E}(\mathbf{x})$.

The *orbit* starting from an initial condition $\mathbf{x}_0$ is the set of points $\{\mathbf{X}(t,\mathbf{x}_0)\ :\ t\in[0,\infty)\}$.

A *fixed point* of a dynamical system is an $\mathbf{x}_{FP}$ for which $F(\mathbf{x}_{FP})=0$. The stability of the fixed point is found by computing the eigenvalues and eigenvectors of F linearized about $\mathbf{x}_{FP}$. If all eigenvalues have negative real part, $\mathbf{x}_{FP}$ is stable. If at least one eigenvalue has a positive real part, $\mathbf{x}_{FP}$ is unstable.

The *stable manifold* of the fixed point $\mathbf{x}_{FP}$, written $SM(\mathbf{x}_{FP})$, is an invariant set $\{\mathbf{x}_0\ :\ \lim_{t\to\infty}\mathbf{X}(t,\mathbf{x}_0)=\mathbf{x}_{FP}\}$.

The *basin of attraction* of a set A is $\{\mathbf{x}_0\ :\ \lim_{t\to\infty}\mathbf{X}(t,\mathbf{x}_0)\in A\}$.

For the Navier-Stokes equations shifted about the laminar shear flow solution, as in (1), the linear operator L is stable, having eigenvalues all with negative real part, but is non-normal, i.e. L does not commute with its adjoint. These properties of L imply that any solution of the linearized problem $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=L\mathbf{x}$ has $\lim_{t\to\infty}|\mathbf{x}(t)|=0$, but some solutions can exhibit a large transient growth before decay. The non-linear function N is quadratic in $\mathbf{x}$ and conserves the energy $\mathcal{E}$.

The *Linear Optimal Perturbation* (LOP) $\mathbf{x}_{LOP}$ satisfies

$$\max_t\{|\mathbf{X}^{(lin)}(t,\mathbf{x}_{LOP})|\}\ \geq\ \max_t\{|\mathbf{X}^{(lin)}(t,\mathbf{x})|\}\qquad\forall\ |\mathbf{x}|=1 \tag{2}$$

where $\mathbf{X}^{(lin)}(t,\mathbf{x}_0)$ is the solution to the linearized problem $\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t}=L\mathbf{x}$, and having $\mathbf{X}^{(lin)}(0,\mathbf{x}_0)=\mathbf{x}_0$. In fact there are two LOPs, with $-\mathbf{x}_{LOP}$ also satisfying the above. In words, $\mathbf{x}_{LOP}$ maximizes the furthest distance from the origin obtained by the orbit of the linearized problem. When L is non-normal, transient growth of $|\mathbf{X}^{(lin)}(t,\mathbf{x}_0)|$ is possible even when all eigenvalues have negative real part (e.g. Farrell, 1988).

We define the *Non-Linear Optimal Perturbation* (NLOP) $\mathbf{x}_{NLOP}$, for a given d and T, to satisfy $|\mathbf{x}_{NLOP}|=d$ and

$$|\mathbf{X}(T,\mathbf{x}_{NLOP})|\ >\ |\mathbf{X}(T,\mathbf{x})|\qquad\forall\ |\mathbf{x}|=d. \tag{3}$$

That is, the NLOP maximizes, over all initial conditions with magnitude d, the distance from the origin of the orbit at time T.

The *Edge* E is the boundary of Ω, the basin of attraction of the origin. Recall that the boundary of a set Ω is the set defined as $\{\mathbf{x}\ :\ \forall\ \epsilon>0,\ B_\epsilon(\mathbf{x})\cap\Omega\neq\varnothing$ and $B_\epsilon(\mathbf{x})\cap\overline{\Omega}\neq\varnothing\}$, where $B_\epsilon(\mathbf{x}_0)=\{\mathbf{x}\ :\ |\mathbf{x}-\mathbf{x}_0|<\epsilon\}$, and $\overline{\Omega}$ is the compliment of Ω. That is, for a point $\mathbf{x}\in E$ and any $\epsilon>0$, the ball of radius ϵ centred at $\mathbf{x}$ contains a point in Ω and another point not in Ω. For all problems considered here, it will be a co-dimension one invariant manifold.

A *Strong Edge* is an edge, or subset of an edge, which separates orbits on one side which return to the origin from orbits on the other side which do not.

A *Weak Edge* is an edge, or subset of an edge, for which the orbits of initial conditions on either side both return to the origin but do so in a qualitatively different way: orbits started on one side return to the origin directly while orbits started on the other side take a more circuitous route and/or require more time to return to the origin.

The Edge is, in general, the union of the Strong Edge and the Weak Edge.

The *Minimal Seed* $\mathbf{x}_{MS}$ is the initial condition on the Edge that is closest to the laminar fixed point (the origin): $\mathbf{x}_{MS} \equiv \text{argmin}_{\mathbf{x}}\{|\mathbf{x}| \ : \ \mathbf{x} \in E\}$.

1.3 A 2D Example

An illustrative two-dimensional dynamical system was explored by Lebovitz (2012). The equations are:

$$\frac{\mathrm{d}x_1}{\mathrm{d}t} = -\delta x_1 + x_2 + x_1 x_2 - 3x_2^2, \tag{4}$$

$$\frac{\mathrm{d}x_2}{\mathrm{d}t} = -\delta x_2 - x_1^2 + 3x_1 x_2 \tag{5}$$

where $\delta = 1/R$. The origin is a fixed point for all R. A saddle-node bifurcation occurs at $R = 2$: hence for $R > 2$ an additional pair of fixed points exists, the lower branch point $\mathbf{x}_{LB}$ and the upper branch point $\mathbf{x}_{UB}$. For a full bifurcation analysis, see Lebovitz (2012).

Let us consider this example system with $R = 2.45$, as shown in Figure 1. The LOP shows transient growth: its orbit at a particular time is actually further from the origin than where it began at a radius d from the origin. Most other orbits started at radius d do not have this property. Note that because $L(a\mathbf{x}) = aL\mathbf{x}$ for any scalar a, trajectories of the linearized system have a scale invariance, and hence so too does the LOP.

Four NLOPs are shown. In each case it is seen that its orbit (magenta curve) finishes (at time T) further from the origin than other orbits (black curves) started the same distance from the origin.

Here, the stable manifold of the lower branch point $SM(\mathbf{x}_{LB})$ extends from $\mathbf{x}_{LB}$ in two directions: positive in x_1 and negative in x_1, roughly speaking. The first extends to $+\infty$ in x_1, while the second winds around $\mathbf{x}_{UB}$ and hence does not extend to $-\infty$ in x_1. Thus, $SM(\mathbf{x}_{LB})$ does not divide phase space into two separate regions. Orbits started on either side of $SM(\mathbf{x}_{LB})$ eventually return to the origin, but they do so by qualitatively and significantly different paths. Thus, $SM(\mathbf{x}_{LB})$ forms the weak edge. Orbits started below the weak edge return to the origin with at most one instance of transient growth, which can be thought of either as due to the linearized dynamics or due to the orbit, started near $SM(\mathbf{x}_{LB})$, must move to near to $\mathbf{x}_{LB}$. Orbits started above $SM(\mathbf{x}_{LB})$ must wind their way around $\mathbf{x}_{UB}$ before returning to the origin.

The strong edge is, clearly, the periodic orbit P surrounding $\mathbf{x}_{UB}$. Orbits started outside P eventually return to the origin, while those started inside P eventually reach $\mathbf{x}_{UB}$.

The edge, being the union of its strong and weak components, is therefore the union of $SM(\mathbf{x}_{LB})$ and P.

The Minimal Seed $\mathbf{x}_{MS}$ is that point on the edge with least distance from the origin. This distance $|\mathbf{x}_{MS}|$ happens to be approximately 0.2384. Furthermore, $\mathbf{x}_{MS}$ is the NLOP for $d = |\mathbf{x}_{MS}|$, since $\mathbf{X}(T, \mathbf{x}_{MS})$ will be near $\mathbf{x}_{LB}$ while any other $\mathbf{x}$ with $|\mathbf{x}| = |\mathbf{x}_{MS}|$ has $\mathbf{X}(T, \mathbf{x})$ near the origin, for sufficiently large T. In fact, the $d = 0.2384$ ball is slightly below the edge everywhere, and hence the orbit of its NLOP (third magenta curve moving radially outward from the origin) goes towards $\mathbf{x}_{LB}$ then decreases towards the origin. Increasing this NLOP by a multiplicative factor of 1.0005 yields a new initial condition that is slightly above the edge: its orbit (cyan curve) goes to $\mathbf{x}_{LB}$ then increases (eventually returning to

the origin given enough time, but here it is only shown up to $T = 16$). Thus the exact minimal seed lies somewhere near these points and has magnitude between 0.2384 and 1.0005×0.2384. These bounds can be tightened considerably, but for illustrative purposes we shall leave it at that.

For small d where the linearized dynamics well approximate the full dynamics, the NLOP tends to be near to the LOP (scaled by d). This is not in general true for larger d. Note that it is merely coincidental that here the LOP scaled by $|\mathbf{x}_{UB}|$ is in fact $\mathbf{x}_{UB}$. It is also not in general true that the LOP scaled by 0.2384 is the NLOP with $d = 0.2384$.

2 Optimization Algorithm

In this section we describe the optimization technique, which is the subject of a recent review paper (Kerswell *et al.*, 2014), that we employ to find the NLOP. The basic idea is described in the next section, followed by a more careful discussion of the convergence criteria. We then discuss how this algorithm fits into a bisection of energy levels to find the minimal seed. The reader may find it useful to keep Figure 1 in mind throughout this section.

2.1 Basic Algorithm

The technique described here is an algorithm to determine the NLOP, i.e. the initial condition $\mathbf{x}_0 \equiv \mathbf{x}(0)$ having $|\mathbf{x}_0| = d$ which maximizes $|\mathbf{x}(T)|$. The leftmost branch of the flowchart in Figure 2, beginning with "K" and ending with "A", "U", or "MS", shows the algorithm. The underlying idea is to maximize a Lagrangian

$$\mathcal{L} = \mathcal{L}(\mathbf{x}, \boldsymbol{\nu}, \lambda; d, T) = |\mathbf{x}(T)|^2 + \int_0^T \boldsymbol{\nu} \cdot \left(\frac{\mathrm{d}\mathbf{x}}{\mathrm{d}t} - L\mathbf{x} - N(\mathbf{x}) \right) \mathrm{d}t + \lambda \left(|\mathbf{x}(0)| - d \right) \tag{6}$$

where $\boldsymbol{\nu}$ and λ are the Lagrange multipliers which respectively ensure that $\mathbf{x}$ is a solution of the dynamical system (1), and that the initial condition lies on a sphere of radius d from the origin, i.e. has a specified energy.

Taking the variation of $\mathcal{L}$ with respect to $\mathbf{x}$ gives

$$\delta\mathcal{L} = \underbrace{[\mathbf{x}(T) + \boldsymbol{\nu}(T)]}_{A} \cdot \delta\mathbf{x}(T) - \int_0^T \underbrace{\left[\frac{\mathrm{d}\boldsymbol{\nu}}{\mathrm{d}t} + \boldsymbol{\nu} \cdot \frac{\partial \mathbf{F}}{\partial \mathbf{x}} \right]}_{B} \cdot \delta\mathbf{x} \, \mathrm{d}t + \underbrace{[\lambda\mathbf{x}(0) - \boldsymbol{\nu}(0)]}_{C} \cdot \delta\mathbf{x}(0) \tag{7}$$

When $\mathcal{L}$ has been maximized, the perturbation $\delta\mathcal{L} = 0$ for an arbitrary perturbation $\delta\mathbf{x}(t)$, $t \in [0, T]$, and thus the quantities labelled A, B, and C must all vanish. Thus, the optimization procedure is as follows:

Step 0: Guess an initial condition $\mathbf{x}^{(n)}(0)$ with $n = 1$.

Step 1: Integrate the dynamical system (1) forwards from $\mathbf{x}^{(n)}(0)$ to obtain $\mathbf{x}^{(n)}(T)$.

Step 2 (Apply $A = 0$): Set $\boldsymbol{\nu}^{(n)}(T) = -\mathbf{x}^{(n)}(T)$.

Step 3 (Apply $B = 0$): Integrate the dual dynamical system,

$$\frac{\mathrm{d}\boldsymbol{\nu}^{(n)}}{\mathrm{d}t} = -\boldsymbol{\nu}^{(n)} \cdot \frac{\partial \mathbf{F}}{\partial \mathbf{x}}, \tag{8}$$

backwards in time[4] from $\boldsymbol{\nu}^{(n)}(T)$ to obtain $\boldsymbol{\nu}^{(n)}(0)$.

Step 4 (Apply C to maximize $\mathcal{L}$): Having now ensured $A = 0$ and $B = 0$, $\mathcal{L}$ varies with $\mathbf{x}(t)$ only through $\mathbf{x}(0)$. Therefore, use

$$\frac{\delta \mathcal{L}}{\delta \mathbf{x}^{(n)}(0)} = \lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \tag{9}$$

to maximize $\mathcal{L}$ by a maximum ascent method: for some small step-size $\epsilon^{(n)}$, the next iteration's guess for the NLOP is

$$\mathbf{x}^{(n+1)}(0) = \mathbf{x}^{(n)}(0) + \epsilon^{(n)} \left[\lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \right]. \tag{10}$$

Note that, although $\mathcal{L}$ is being maximized for arbitrary values of the Lagrange multipliers $\boldsymbol{\nu}$ and λ, their specific values are important, and here we must simultaneously choose λ to satisfy

$$d = \left| \mathbf{x}^{(n+1)}(0) \right| = \left| \mathbf{x}^{(n)}(0) + \epsilon^{(n)} \left[\lambda \mathbf{x}^{(n)}(0) - \boldsymbol{\nu}^{(n)}(0) \right] \right| \tag{11}$$

Using the squared L_2 norm $|\mathbf{x}| = \|\mathbf{x}\|_2^2 = \mathbf{x} \cdot \mathbf{x}$, this condition on λ yields an easily solved quadratic equation (dropping (n) superscripts for clarity):

$$(\epsilon^2 \ \mathbf{x}_0 \cdot \mathbf{x}_0)\lambda^2 + (2\epsilon \ \mathbf{x}_0 \cdot \mathbf{x}_0 - 2\epsilon^2 \ \mathbf{x}_0 \cdot \boldsymbol{\nu}_0)\lambda + (\epsilon^2 \ \boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0 - 2\epsilon \ \mathbf{x}_0 \cdot \boldsymbol{\nu}_0) = 0. \tag{12}$$

The choice of the step-size ϵ is critically important for convergence and is discussed next. The choice of T is also important and may need to be adjusted as d changes, as discussed in the subsequent section.

2.2 Convergence

How one chooses the step size ϵ is of great practical importance. With ϵ too small, the method will take too long to feasibly complete, while too large an ϵ can lead to $\Delta \mathbf{x}_0 = \mathbf{x}(0)^{(n+1)} - \mathbf{x}(0)^{(n)}$ also being so large that $\Delta \mathcal{L} = \mathcal{L}^{(n+1)} - \mathcal{L}^{(n)} \approx (\delta \mathcal{L}/\delta \mathbf{x}_0) \cdot \Delta \mathbf{x}_0$ is a poor approximation and $\mathcal{L}$ actually decreases. We therefore employ some additional measures, applied in each iteration of the algorithm, to determine a good step-size ϵ, so as to prevent both the above problems.

First, we require $\epsilon \leq \epsilon_{\max,\lambda}$, where the latter is defined as the maximum ϵ for which λ can be real. Requiring $\lambda \in \mathbb{R}$ in turn requires the discriminant of (12) to be non-negative, giving another quadratic equation, this time for ϵ, the solution of which is

$$\epsilon_{\max,\lambda} = \left(\frac{(\mathbf{x}_0 \cdot \mathbf{x}_0)^2}{(\mathbf{x}_0 \cdot \mathbf{x}_0)(\boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0) - (\mathbf{x}_0 \cdot \boldsymbol{\nu}_0)^2} \right)^{1/2}, \tag{13}$$

having chosen the positive root so as to maximize, not minimize, $\mathcal{L}$. This criterion is applied at each iteration (n).

[4]This requires knowing $\mathbf{x}(t)$ at each time step (and sub-steps) of the ODE integration method. For high dimensional dynamical systems such as a direct numerical simulation of the Navier-Stokes equations, the full forward integration of $\mathbf{x}(t)$ cannot be stored in memory, and therefore a check-pointing procedure is required, as discussed by Kerswell *et al.* (2014), wherein $\mathbf{x}(t)$ is saved only at some times t_j and these $\mathbf{x}(t_j)$'s are used, when needed in solving for $\boldsymbol{\nu}(t)$, to initiate another forward integration of $\mathbf{x}$ from t_j to t_{j+1}.

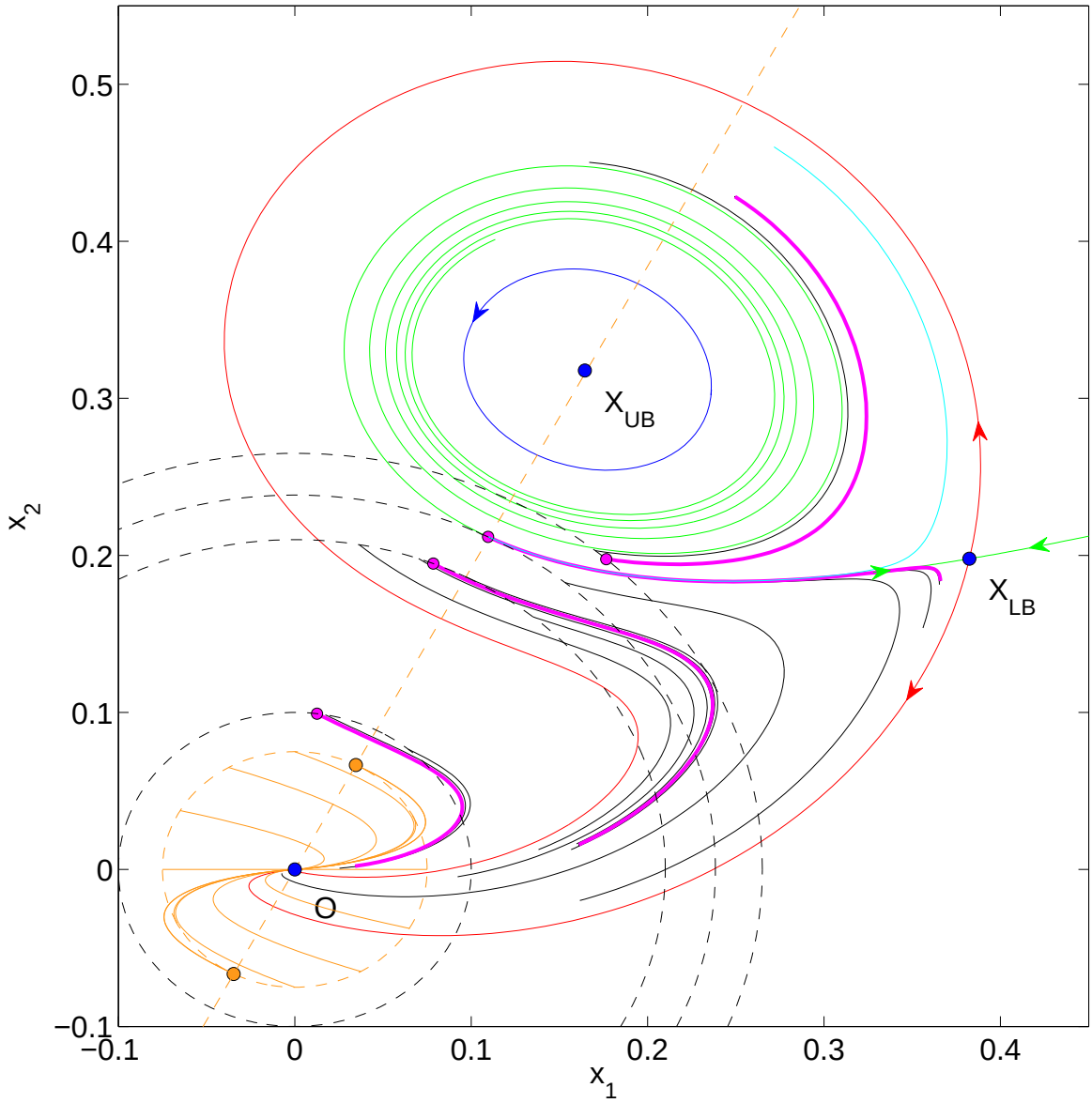


Figure 1: The system (4) with $R = 2.45$. Blue circles indicate the three fixed points: the origin O, the lower branch point $\mathbf{x}_{LB}$, and the upper branch point $\mathbf{x}_{UB}$. The unstable manifold of $\mathbf{x}_{LB}$ (red curves) consists of two orbits, beginning in opposite directions but both converging to O. The stable manifold of $\mathbf{x}_{LB}$ (green curves) winds around an unstable periodic orbit P (blue) which encloses the basin of attraction of the stable fixed point $\mathbf{x}_{UB}$. The LOP and its negative are indicated by the the orange circles, together with their orbits in the linearized system (orange curves). Other orange lines show orbits of the linearized system started at 30° intervals around the origin and at a radius of 0.075 (dashed orange circle). Any point on the dashed orange line is an LOP. The NLOP (magenta circles) is shown for each of $d = 0.1, 0.21, 0.2384, 0.265$ (black dashed circles), with $T = 8, 8, 16, 8$ respectively. The NLOP is calculated by the iterative procedure described in Section 2: successive orbits (black curves) begin with the orbit started at an angle from the $+x_1$ axis of 50° (for this illustration) and converge to the NLOPs orbit (magenta curves). The orbit started from the NLOP with $d = 0.2384$ scaled by a factor 1.0005 is shown (cyan curve).

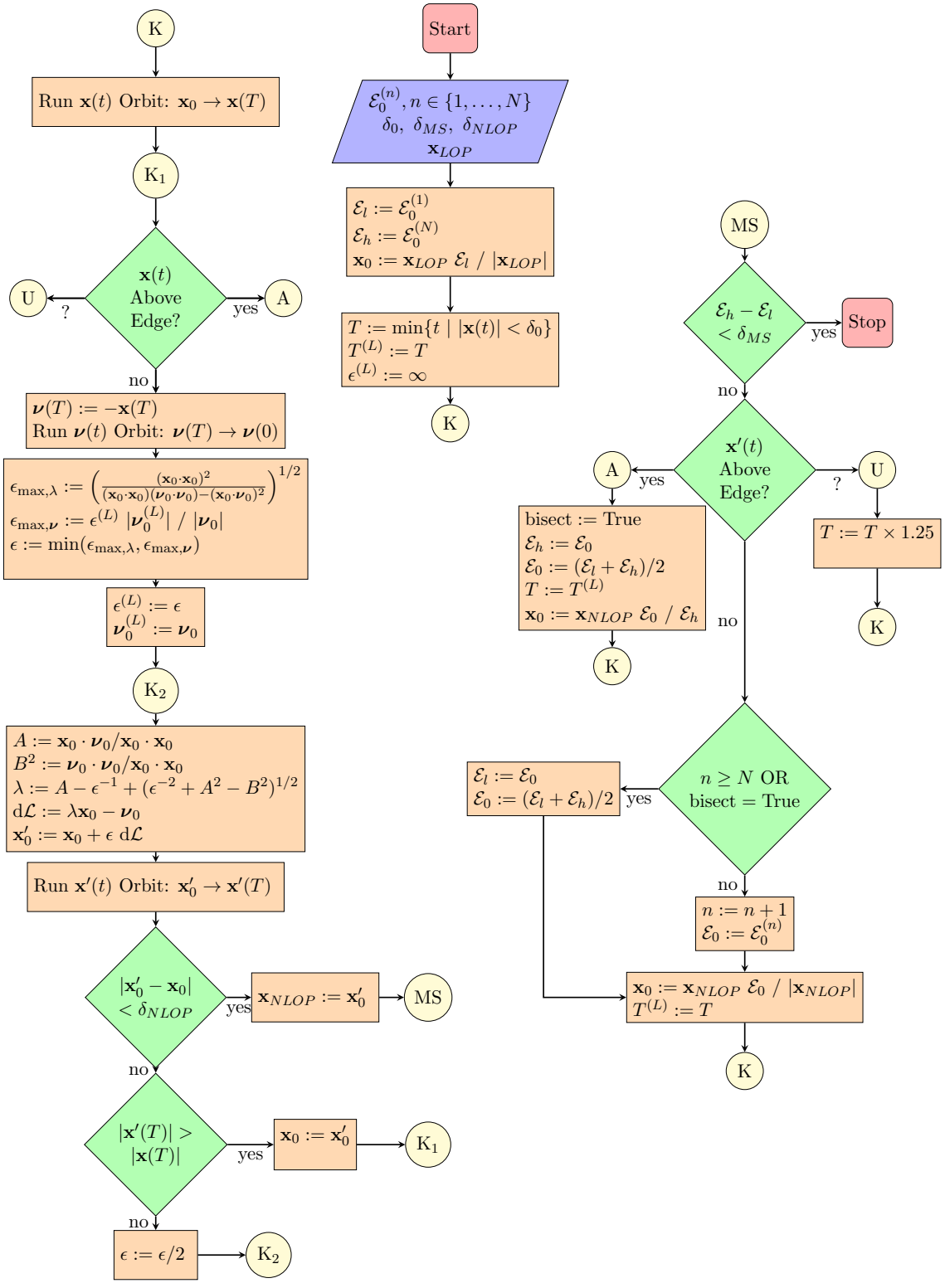


Figure 2: Flowchart showing the optimization procedure to calculate the NLOP (left branch, K to A, U, or MS), the energy incrementation and bisection procedure (right branch, MS to K or Stop), and the initialization for the whole algorithm (centre branch, Start to K).

The above criterion is useful when beginning the iterations, telling us how large a step size we can feasibly take. But it will not bring $\epsilon \to 0$, so we add another criterion which limits the contribution $\boldsymbol{\nu}$ can make to $\Delta\mathbf{x_0}$ in (10):

$$\epsilon^{(n)}_{\max,\boldsymbol{\nu}} = \epsilon^{(n-1)} \frac{|\boldsymbol{\nu}_0^{(n-1)}|}{|\boldsymbol{\nu}_0^{(n)}|}. \tag{14}$$

and require $\epsilon^{(n)} \leq \epsilon^{(n)}_{\max,\boldsymbol{\nu}}$.

Another criteria that was considered was to require that the jump in $\mathbf{x}_0$ between successive iterations does not increase. That is, $\left|\mathbf{x}_0^{(n+1)} - \mathbf{x}_0^{(n)}\right|^2 \leq \left|\mathbf{x}_0^{(n)} - \mathbf{x}_0^{(n-1)}\right|^2 \equiv D_0^{(n)}$. Using (10) this yields another upper bound for $\epsilon^{(n)}$, which is

$$\epsilon^{(n)}_{\max,\mathbf{x}_0} = \left(\frac{D_0(\mathbf{x}_0 \cdot \mathbf{x}_0) - D_0^2/4}{(\mathbf{x}_0 \cdot \mathbf{x}_0)(\boldsymbol{\nu}_0 \cdot \boldsymbol{\nu}_0) - (\mathbf{x}_0 \cdot \boldsymbol{\nu}_0)^2} \right)^{1/2}, \tag{15}$$

all evaluated at iterate (n). However, this was found to be too restrictive. Consider when the initial guess $\mathbf{x}_0$ begins in a fairly "flat" region of phase space (i.e. where $|\mathbf{x}(T)|$ changes little with changes in $\mathbf{x}(0)$) and is also not near the local maxima of $|\mathbf{x}(T)|$. Including this restriction forces all subsequent steps after the initial step in $\mathbf{x}(0)$, which is small owing to flatness, to be at least as small; hence convergence to the NLOP requires an enormous number of iterations. Without this condition, small initial steps are taken, followed by larger steps when $\mathbf{x}(0)$ has left the flat region.

The two criteria used above attempt to limit ϵ but do not themselves ensure that $\mathcal{L}$, and hence $|\mathbf{x}(T)|$, increases with each iteration. Since this is essential, one final condition is added: the dynamical system is integrated forward to obtain $\mathbf{x}^{(n+1)}(T)$ and $\left|\mathbf{x}^{(n+1)}(T)\right| > \left|\mathbf{x}^{(n)}(T)\right|$ is explicitly checked. This requirement becomes more stringent as the iterations proceed. If this check fails, $\epsilon^{(n)}$ is halved, $\mathbf{x}_0^{(n+1)}$ re-calculated from (10), and this criteria checked again. This halving process is repeated until the condition passes, or some maximum number of halvings is reached, currently set to 20.[5]

Not only does this ϵ-halving technique give us confidence that we are finding a local maximum of $\mathcal{L}$, numerical experiments suggest that it may even hasten the whole optimization algorithm by preventing overshooting of the NLOP and thereby reduce the number of iterations (n) required.

The final statement regarding convergence regards when the NLOP is deemed to be found and the full optimization algorithm finishes. During the forward integrations of the ϵ-halving procedure (which also serve as performing Step 1 in Section 2.1 above; see also Figure 2) we determine whether $\left|\mathbf{x}^{(n+1)}(0) - \mathbf{x}^{(n)}(0)\right| < \mathcal{E}_0^{\text{tol}}$, for some small number $\mathcal{E}_0^{\text{tol}}$, for which we tend to use a default value of 10^{-8}. When this is true we declare the NLOP found and finish the algorithm.

[5]For additional numerical safety, subsequent applications of the ODE solver for forward integration of $\mathbf{x}(t)$ at this iteration level (n) will do the following: sub-sample the output $\mathbf{x}(t)$ by a further factor of two, to help ensure the backwards integration of ν is accurate, and raise by tenfold the accuracy of the ODE solver (which reduces the size of the time steps).

2.3 Energy Incrementation and the Two Sides of the Edge

Having a procedure to find the NLOP on a given energy ball $\{\mathbf{x} \ : \ |\mathbf{x}| = \mathcal{E}_0\}$ (and for a given integration time T), we now desire an algorithm that applies the above procedure on successively larger energy spheres, in order to find the minimal seed. The basic idea here is to increase the size of the energy spheres until the edge is passed, and then to refine the energy by bisection down to a desired precision.

2.3.1 Which Side of the Edge?

This procedure will require a function that determines whether a particular orbit is above or below the edge, or whether this is unclear from the information available. The information available is the full orbit $\mathbf{x}(t)$, $t \in [0, T]$, as well as, in principle, all previously calculated orbits, although in practice only the orbits of the previously calculated NLOPs are needed and generally just some metric (such as the gain) of them.

This function will vary between problems. Clearly, if the problem has a strong edge (the basin of attraction of $\mathbf{x}_{UB}$) but not a weak edge, a criteria could be the following: if $|\mathbf{x}(T)| < \delta_O$ then $\mathbf{x}_0$ is below the edge; if $|\mathbf{x}(T) - \mathbf{x}_{UB}| < \delta_{UB}$ then $\mathbf{x}_0$ is above the edge; otherwise, it is unclear whether $\mathbf{x}_0$ is above or below the edge.

However, if the problem has a weak edge which is sought then a different function must be used. One approach is use the ratio of the gain $G \equiv \max_t\{|\mathbf{x}(t)|\} \ / \ |\mathbf{x}(0)|$ between the current orbit and the orbit of the last NLOP determined to be below the edge. If this ratio exceeds a certain value[6] then $\mathbf{x}_0$ is above the edge; if $|\mathbf{x}(T)| < \delta_O$ then $\mathbf{x}_0$ is below the edge; otherwise, it is unclear whether $\mathbf{x}_0$ is above or below the edge.

This approach has been successful in simple problems such as the 2D system studied by Lebovitz (2012) (Section 1.3), and also in the 2D example problem used by Kerswell *et al.* (2014) which only has a strong edge.

However, the geometry of the problem may prevent this gain-based approach from working. This is easily seen if we had $|\mathbf{x}_{LB}| > |\mathbf{x}_{UB}|$. Then two orbits started nearby but on either side of the edge would, as per the usual understanding, go near $\mathbf{x}_{LB}$; then one goes directly back to the origin O while the other goes around $\mathbf{x}_{UB}$ before returning to O. But the gain of both orbits would be approximately $|\mathbf{x}_{LB}|/|\mathbf{x}(0)|$, with no obvious difference between them.

For the W97 4D model, we do have the desirable $|\mathbf{x}_{LB}| < |\mathbf{x}_{UB}|$ but the separation is not so large and a formulation in terms of the gain, as above, was not robustly successful. The following, more reliable approach, was found instead. Find all local maxima of the function $|\mathbf{x}(t)|$, then for numerical reasons discard all those having $|\mathbf{x}(t)| < \delta_0$. If there are one or fewer local maxima—the orbit either directly relaminarizes or does so after a transient growth as per a non-normal linear system—the orbit is deemed below the edge. If there are two or more local maxima—the orbit first undergoes transient growth, then undergoes secondary growth when repulsed on the far side of the saddle that is $\mathbf{x}_{LB}$—and their magnitudes exceed some specified thresholds (for each maxima), the orbit is deemed

[6] This value need not be much great than 1. For the 2D problem of Section 1.3, we choose a value of 1.2. If the incremental increase of the energy spheres is small enough then, for two spheres below the edge, the ratio of the gains of their NLOPs will be nearly 1, while for two spheres straddling the edge this ratio grows without bound as the distance between the energy spheres shrinks.

above the edge. The thresholds chosen were 0 for the first peak (any size peak will do), and $|\mathbf{x}_{LB}| + \max(\delta_{LB}, 0.2 \cdot |\mathbf{x}_{UB} - \mathbf{x}_{LB}|)$, where δ_{LB} is defined below. This function, together with another improvement detailed below, is illustrated as a flowchart in Figure 3.

When, in all examples above, it is deemed unclear whether the orbit is above or below the edge, the correct thing to do is to choose a larger T. Clearly this works as desired in the first example above when all of phase space (except for a measure zero set, the stable manifold of the lower branch point) is in the basin of attraction of either O or $\mathbf{x}_{UB}$. In the second example, the added time may allow further gain to be realized (supporting an above-edge result), or it may simply bring the orbit closer to the origin (supporting a below-edge result). Thus, increasing T is still the correct response, but one must have carefully chosen the threshold for the ratio of the gains, based on previous experience with the particular problem.

When, by bisection of energy levels, the orbits under question are very near the minimal seed, they will track the stable manifold of $\mathbf{x}_{LB}$ until near $\mathbf{x}_{LB}$ itself; this point being a fixed point, velocities near $\mathbf{x}_{LB}$ are small so the orbit lingers for much time near $\mathbf{x}_{LB}$. If increasing T sufficiently to escape $\mathbf{x}_{LB}$ is infeasible, a complimentary condition may be added that applies when the orbit enters a small ball $\delta_{LB} \ll 1$ of $\mathbf{x}_{LB}$ wherein the linearized dynamics dominate. First, linearize F about $\mathbf{x}_{LB}$: $F_{LB} = \partial F/\partial \mathbf{x}\Big|_{\mathbf{x}_{LB}}$. Second, find all stable eigenvectors of F_{LB} and (as by Gram-Schmidt) orthonormalize them to get $\mathbf{v}_1, \ldots, \mathbf{v}_{n-1}$. Third, calculate the vector rejection of the unstable eigenvector $\mathbf{v}_n$ of F_{LB}:

$$\mathbf{v}_O = \mathbf{v}_n - \sum_{i=1}^{n-1} (\mathbf{v}_n \cdot \mathbf{v}_i)\, \mathbf{v}_i. \tag{16}$$

This $\mathbf{v}_O$ is orthogonal to the hyperplane formed by the stable eigenvectors. Fourth, calculate the orbits starting from $\mathbf{x}_{LB} + \delta_{LB}\mathbf{v}_O$ and $\mathbf{x}_{LB} - \delta_{LB}\mathbf{v}_O$ and decide which is below the edge and which is above. The orbit below the edge may be said to be the one that requires less time to reach a δ_0-ball around the origin[7], or the one that has a lesser gain. If $\mathbf{x}_{LB} - \delta_{LB}\mathbf{v}_O$ is the one below the edge, then re-assign $\mathbf{v}_O := -\mathbf{v}_O$. Hence $\mathbf{v}_O$ points, following the dynamical system's trajectories, from $\mathbf{x}_{LB}$ towards the origin. Now, for any orbit $\mathbf{x}(t)$, find the time t_c which minimizes $D_{LB} = |\mathbf{x}(t_c) - \mathbf{x}_{LB}|$. If $D_{LB} > \delta_{LB}$ then it is unclear from this analysis whether the orbit is above or below the edge. If $D_{LB} < \delta_{LB}$ then the orbit $\mathbf{x}(t)$ is deemed below the edge if

$$S = (\mathbf{x}(t_c) - \mathbf{x}_{LB}) \cdot \mathbf{v}_O \tag{17}$$

is positive, and above the edge if S is negative.

2.3.2 Energy Incrementation

Now the task is to increase the size of the energy spheres upon which the NLOP is calculated, until a sphere intersects the edge. The fundamental outline of the algorithm is, in words, as follows.

[7]There are systems, such as the 2D system of Lebovitz (2012), where the orbit from $\mathbf{x}_{LB}$ that winds around $\mathbf{x}_{UB}$, thereby obtaining a much larger gain, actually relaminarizes (reaches the δ_0-ball centred at the origin for $\delta_0 \ll 1$) in less time.

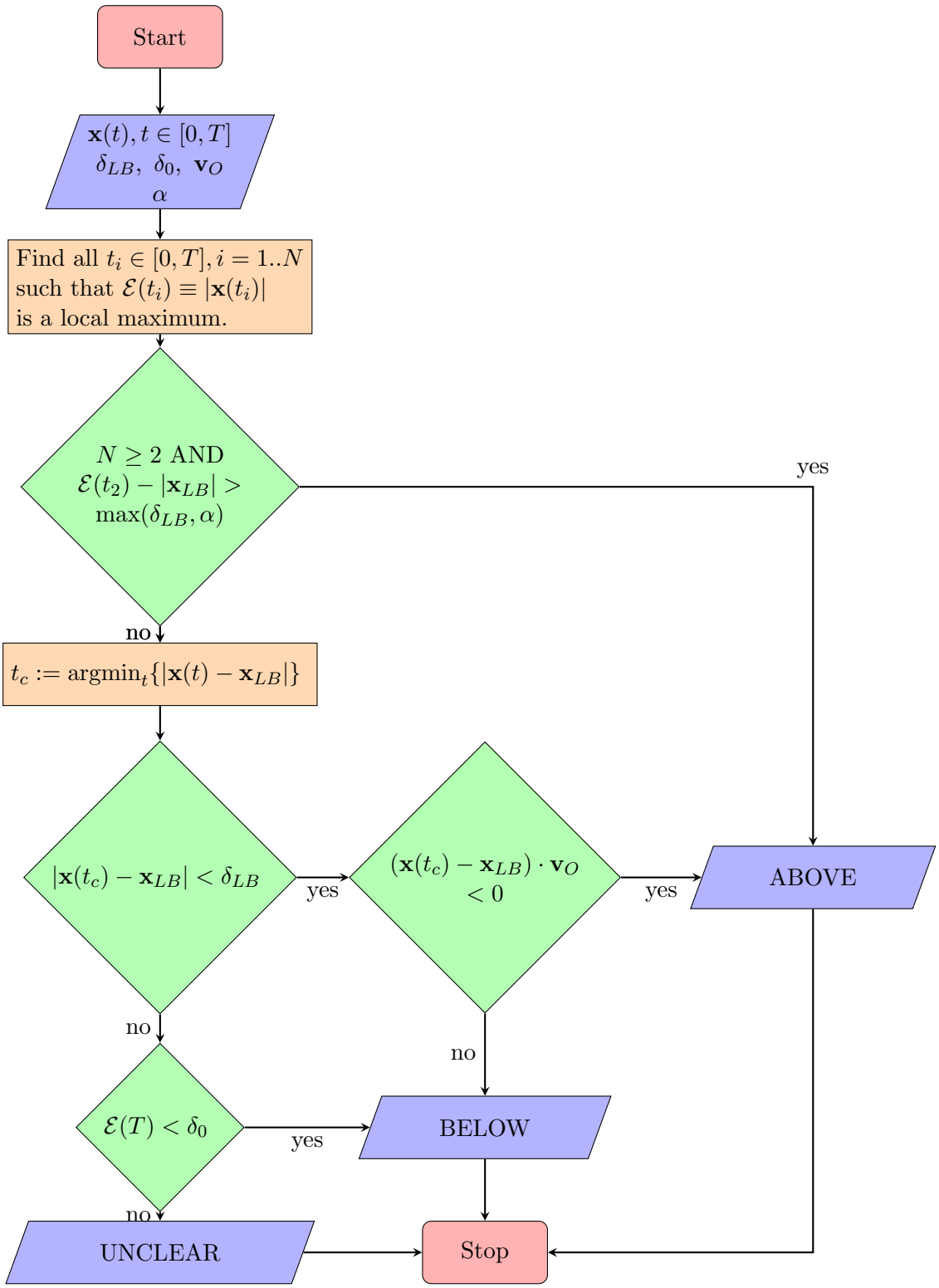


Figure 3: Flowchart to determine whether an orbit $\mathbf{x}(t)$ is above or below the edge, or whether this is unclear. This formulation was used to work on the W97 problem, with $\delta_{LB} = 10^{-2}$, and $\alpha = (|\mathbf{x}_{UB}| - |\mathbf{x}_{LB}|)/5$. δ_0 was initially chosen as 10^{-2} but can be made larger (as discussed in Section 2.3.2) once the minimal seeds for some values of R are available and an approximate scaling law for the minimal seed can be made.

Step 0: Choose a series of energy spheres $\mathcal{E}_l \equiv \mathcal{E}_1, \ldots \mathcal{E}_N \equiv \mathcal{E}_h$. Choose δ_0: an orbit entering the δ_0-ball from outside is deemed going to the origin and below the edge. Set $\mathbf{x}_0 := \mathcal{E}_l\ \mathbf{x}_{LOP}$, where $\mathbf{x}_{LOP}$ is the Linear Optimal Perturbation with unit magnitude. Run the orbit starting at $\mathbf{x}_0$ until the smallest time T such that $|\mathbf{x}(T)| < \delta_0$.

Step 1: Calculate, by the optimization algorithm above, $\mathbf{x}_{NLOP}$.

Step 2: Check whether the orbit from $\mathbf{x}_{NLOP}$ is above or below the edge, or if this is unclear.

If unclear: increase T.

If above: set $\mathcal{E}_h := \mathcal{E}_0$; reduce $\mathcal{E}_0$ by bisection between $\mathcal{E}_h$ and $\mathcal{E}_l$; scale $\mathbf{x}_0$ to $\mathbf{x}_0\ \mathcal{E}_0\ /\ |\mathbf{x}_0|$.

If below: set $\mathcal{E}_l := \mathcal{E}_0$; increase $\mathcal{E}_0$ by bisection between $\mathcal{E}_l$ and $\mathcal{E}_h$ if previously an orbit was found above the edge, or if the $\mathcal{E}_i$ set is exhausted; otherwise set $\mathcal{E}_0$ to the next element in $\mathcal{E}_i$; set $\mathbf{x}_0$ to $\mathbf{x}_{NLOP}\ \mathcal{E}_0\ /\ |\mathbf{x}_{NLOP}|$.

Step 3: Check convergence. If $\mathcal{E}_h - \mathcal{E}_l < \delta_{MS}$ for some specified δ_{MS}, exit the algorithm; otherwise, go to Step 1.

There are several additional points which can improve the basic algorithm, which we now discuss. The flowchart in Figure 2 includes many of these improvements.

In all problems considered so far, the condition for an orbit to be below the edge rests on whether $|\mathbf{x}(T)| < \delta_0$. When beginning this energy incrementing algorithm, this δ_0 must be chosen very small so that there is complete confidence that any orbit which crosses from outside to inside of the δ_0-ball centred at the origin will continue directly to the origin, i.e. with no later growth of $|\mathbf{x}(t)|$. However, this δ_0 can be increased to $\mathcal{E}_l$, the size of the energy sphere upon which the NLOP was last found to be below the edge: if the NLOP on the $\mathcal{E}_l$-ball goes to the origin and an orbit started on the $\mathcal{E}_0$-ball passes from outside to inside of the $\mathcal{E}_l$-ball, then this orbit must also go to the origin.

When an orbit is above the edge, we don't need to find the NLOP for that energy sphere. We simply decrease the energy and try again. Thus, at each iteration in the optimization algorithm, the orbit is passed to the function that determines whether it's above or below the edge. If it's above the edge, then the energy sphere is immediately shrunk ("A" in Figure 2).

Relatedly, if it is unclear whether the orbit is above or below the edge, but the orbit is in fact below the edge, then since further iterations of the optimization procedure will only increase $|\mathbf{x}(T)|$, the orbits of all further iterations will also be deemed unclear. Hence, T should be immediately increased and the optimization procedure restarted. Similarly, if it is unclear whether the orbit is above or below the edge, but the orbit is in fact above the edge, then increasing T will likely accelerate the whole algorithm and at the very least will not lead to an incorrect answer. Hence, shortcut "U" is taken in Figure 2.

When the NLOP is found to be below the edge, it is scaled radially outward to the next energy sphere. But note, when an $\mathbf{x}_0$ is detected above the edge (shortcut "A"), the NLOP is not found and so it is the last NLOP that is scaled to the next energy sphere. $\mathbf{x}_0$ should similarly not be updated when it is unclear whether the orbit of the optimization procedure is above or below the edge (but this case can only occur on the first iteration of the optimization procedure, before $\mathbf{x}_0$ could be updated.)

In the W97 model, the fourth and final component of an orbit that relaminarizes decays to 0 much faster than the other components. This component can quickly become less than

the absolute error tolerance of the ODE solver used to integrate the dynamical system—a situation that can cause numerical issues. Thus, it was found extremely helpful to nudge the initial condition $\mathbf{x}_0$ given to the optimization algorithm. In this case we chose to increase the fourth component to a minimum of 10^{-4}, while scaling down the other components of $\mathbf{x}_0$ such that the total energy $|\mathbf{x}(0)|$ is unchanged.

3 In Search of the Minimal Seed

In this section we give the results obtained by applying the optimization algorithm of Section 2 to the W97 four-dimensional model. We begin by describing this model, then discuss how the optimization technique operates for this model, then show our results and briefly compare them to the previous results of Cossu (2005).

3.1 The Waleffe (1997) 4D Model

While significant progress has been made recently, a complete understanding of shear turbulence arising from the full Navier-Stokes equations, even in simple geometries such as plane Couette flow or plane Poiseuille flow, remains out of reach. When viewed as a dynamical system, such a system would have roughly 10^5 degrees of freedom, obfuscating visualization as well as the implementation of many numerical techniques. A lower dimensional system that still captures the important features of a potentially turbulent shear flow is therefore desired.

One such low dimensional model for plane Couette flow was developed in the 1990's by Waleffe and collaborators, culminating in a 1997 paper (Waleffe, 1997). One feature of this work was to recognize that organized, coherent structures may exist within a turbulent shear flow. These structures, uncovered in their analysis, are streamwise roll vortices, streaks (spanwise perturbations of the streamwise flow), and a streamwise wave dominated by spanwise velocity perturbations. The streamwise rolls advect the mean shear, pulling positive streamwise flow down and lifting negative streamwise flow up, thereby creating streaks. Even for roll vortices that are weak (relative to the mean shear), the streaks represent significant spanwise variation of the streamwise flow and can be linearly unstable, leading to the growth of the streamwise wave. Through careful analysis, Waleffe showed that these waves non-linearly force a streamwise roll vortex very near to the original rolls, thereby amplifying the original perturbation. Thus, this is an energetically "self-sustaining process", thereby avoiding the decay to the laminar state.

In addition to the physical understanding of these structures and their interaction, this work led to a four-dimensional dynamical system. It has been much studied and seems to capture many important features of the full Navier-Stokes equations for shear turbulence.

Writing $\mathbf{x} = (m, u, v, w)^T$, the Waleffe's model in the general form (1) has

$$L = \begin{pmatrix} -k_m^2 R^{-1} & 0 & 0 & 0 \\ 0 & -k_u^2 R^{-1} & \sigma_u & 0 \\ 0 & 0 & -k_v^2 R^{-1} & 0 \\ 0 & 0 & 0 & -k_w^2 R^{-1} \end{pmatrix} \tag{18}$$

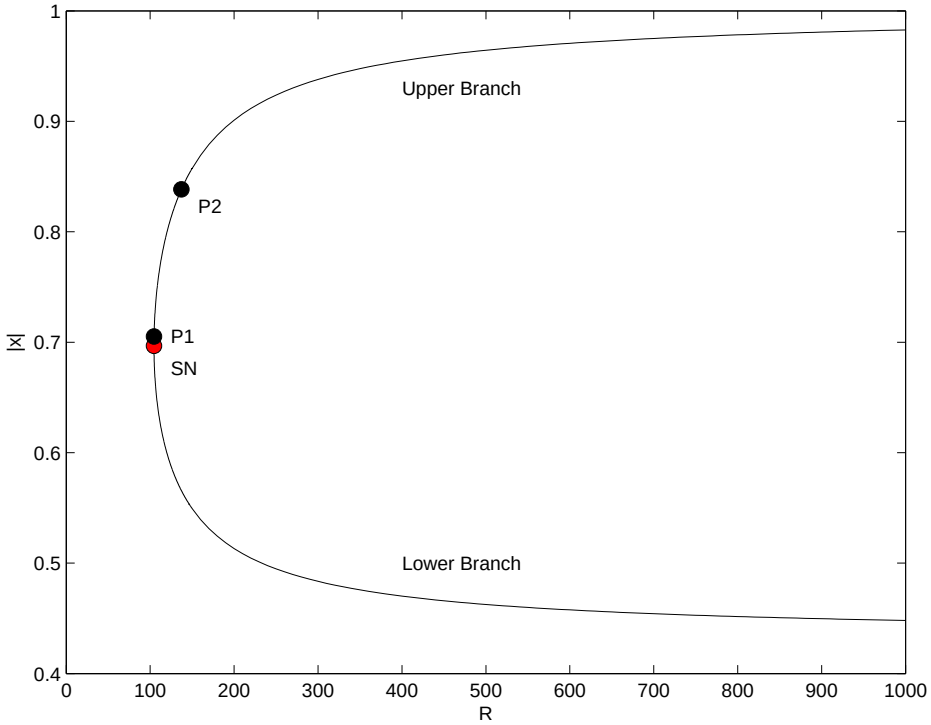


Figure 4: Magnitude of fixed points of the W97 model over a range of Reynolds numbers R. See text for details.

and

$$N(\mathbf{x}) = \begin{pmatrix} \sigma_m w^2 - \sigma_u uv \\ -\sigma_w w^2 + \sigma_u mv \\ \sigma_v w^2 \\ (\sigma_w u - \sigma_m m - \sigma_v v)w \end{pmatrix}. \tag{19}$$

The non-linear terms conserve the energy, recycling it between the different modes. The linearized problem admits a globally attracting fixed point at the origin—the laminar state. The same is true for the full non-linear problem when R is small.

For values of $R > R_{SN} = 104.5467$ where a saddle-node bifurcation occurs, two additional fixed points are born: the lower branch fixed point, $\mathbf{x}_{LB}$, and the upper branch fixed point $\mathbf{x}_{UB}$. Their magnitudes are shown in Figure 4. A sub-critical Hopf-type bifurcation occurs at $R_{P2} = 137.2569$: an unstable periodic orbit is born and for $R > R_{P2}$, $\mathbf{x}_{UB}$ becomes a stable attractor—it has two complex pairs of eigenvalues, both pairs having negative real parts. For $104.6434 = R_{P1} < R < R_{P2}$, $\mathbf{x}_{UB}$ one pair of complex eigenvalues has positive real part, so $\mathbf{x}_{UB}$ is unstable. For $R_{SN} < R < R_{P1}$, these eigenvalues become real but remain positive and hence $\mathbf{x}_{UB}$ is unstable here as well.

This 4D model has been a popular choice for study, but so far only one attempt has been made to find the minimal seed systematically over a range of R values. That study, by Cossu (2005), found that $|\mathbf{x}_{MS}|$ scaled closely with R^{-1}. However, the orbits started

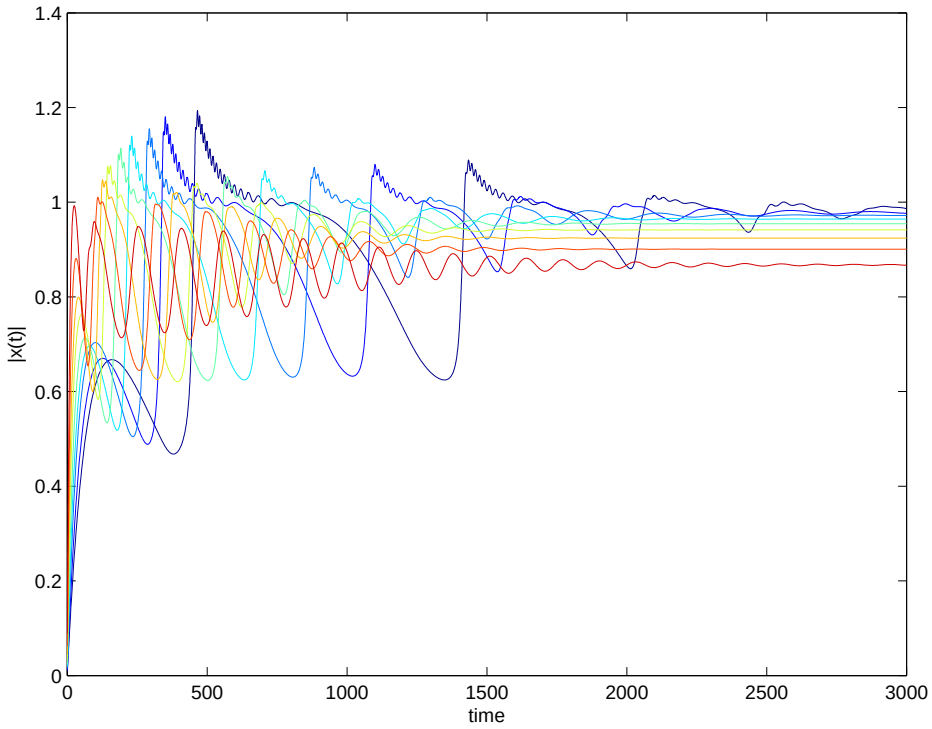


Figure 5: Magnitude of orbits beginning from the minimal seeds found by Cossu (2005). Colours are for different R values. All minimal seeds have orbits that are attracted to the upper branch point. Hence, they are within the basin of attraction of the upper branch point, not the origin, and are on the far side of the strong edge.

from these minimal seeds are seen (Figure 5) to go to the upper branch point[8] rather than to return to the origin. Thus, Cossu's results are for the minimal distance from the origin to the strong edge. If there exists a weak edge, it is entirely possible and perhaps likely that it contains a point closer to the origin than any point on the strong edge, in which case the minimal seed will be smaller than stated by these previous results. We will give results from applying the search method of Section 2 to the W97 model, which first requires a discussion of local maxima.

3.2 Optimizing to the Global Maximum

The optimization procedure of Section 2.1 finds a local maximum of $\mathcal{L}$, i.e. it maximizes $|\mathbf{x}(T)|$. A common difficulty with such optimization algorithms lies in whether the result is the global maximum or merely a lesser, local maximum.

We therefore calculate $|\mathbf{x}(T)|$ for two degrees of freedom of initial conditions $\mathbf{x}(0)$, constrained to have a specified w and a specified $|(m, u, v)| = \mathcal{E}_0$ (Figure 6). When these constraints place all $\mathbf{x}(0)$ well below the edge (top panel), $|\mathbf{x}(T)|$ appears a smooth function of initial conditions with only two local maxima, and for very small $\mathcal{E}_0$ these are nearly equal, as expected from the linearized problem which has a -1 symmetry (i.e. $\mathbf{X}(t, \mathbf{x}_0) = -\mathbf{X}(t, -\mathbf{x}_0)$;

[8] Cossu (2005) only studied values of R for which the upper branch point is a stable attractor.

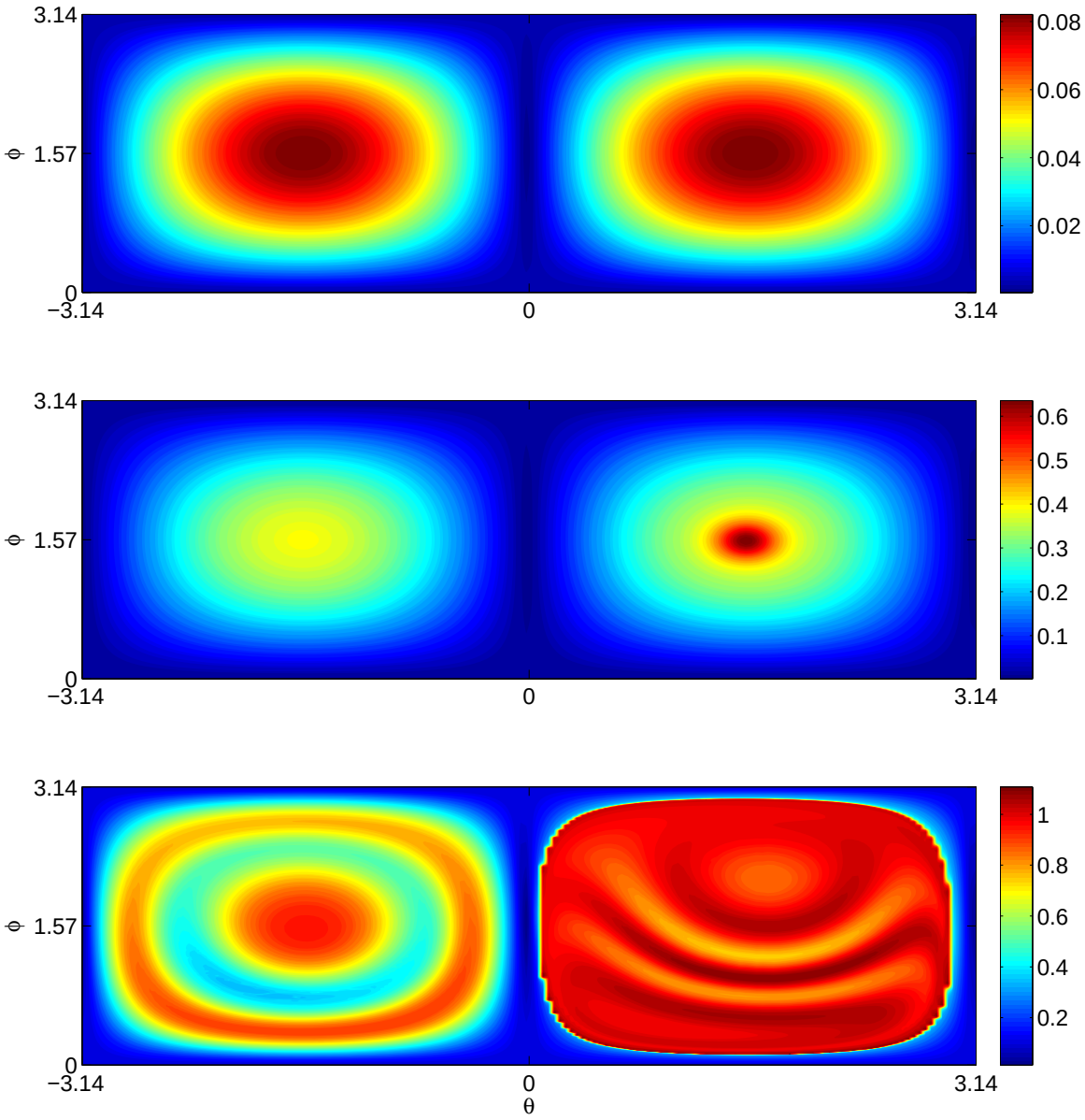


Figure 6: The final energy $|\mathbf{x}(T)|$ (colours), as a function of $\phi \in [0, \pi]$ and $\theta \in [-\pi, \pi]$ which define an initial condition $\mathbf{x}(0) = (\mathcal{E}_0 \cos\phi, \mathcal{E}_0 \cos\theta \sin\phi, \mathcal{E}_0 \sin\theta \sin\phi, W_0)$. The W97 system with $R = 200$ and $T = 80$ then gives $\mathbf{x}(T)$. The values $\mathcal{E}_0$, W_0 are respectively (top) 10^{-2}, $2.1894 \cdot 10^{-3}$, (middle) $4.3 \cdot 10^{-2}$, $9.4142 \cdot 10^{-3}$, and (bottom) 0.3, $6.5681 \cdot 10^{-2}$. The chosen $\mathcal{E}_0$ place the top, middle, and bottom panels well below, just below, and well above the edge, respectively. To complement this W_0 was chosen as $w_{MS} \cdot (\mathcal{E}_0/|\mathbf{x}_{MS}|)$, but the qualitative results are not sensitive to small changes of W_0. Well below the edge (top) there are two nearly equal local maxima of $|\mathbf{x}(T)|$, one the negative of the other, as expected from the non-normal linearized problem. Just below the edge (middle), there is one clearly global maxima and one lesser, local maxima. Well above the edge (bottom) there are many local maxima.

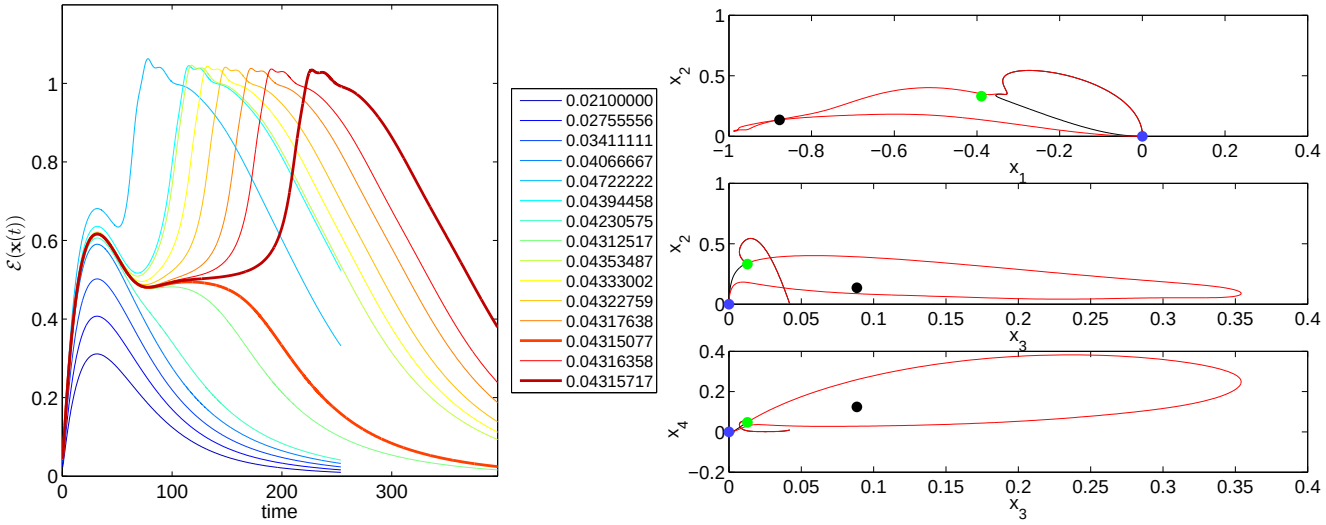


Figure 7: Left: The energy, $\mathcal{E}(\mathbf{x}(t)) \equiv |\mathbf{x}(t)|$, of successive orbits $\mathbf{x}(t)$ in the iterative procedure to find the minimal seed, for the W97 model with $R = 200$. Each curve corresponds to the orbit started from an initial condition with energy indicated by colour. For orbits below the edge, the initial condition is the NLOP. Thick lines show the orbits of those initial conditions straddling the edge. Right: The two thick-lined orbits from the left panel, shown in phase space. Both orbits begin near the origin (blue circle) and track each other closely until they near the lower branch point (green circle), at which point the black orbit (below the edge) returns directly to the origin while the red orbit (above the edge) circumnavigates the upper branch point (black circle) before returning to the origin. The top, middle, and bottom panels show x_1 vs. x_2, x_3 vs. x_2, and x_3 vs. x_4, respectively.

also note that the W97 model is symmetric about the $w = 0$ hyperplane). For larger $\mathcal{E}_0$ up to and below the minimal seed, one local maxima dominates and becomes the global maxima (middle panel). For $\mathcal{E}_0$ above the minimal seed (bottom panel), $SM(\mathbf{x}_{LB})$ folds (back and forth, or perhaps around $\mathbf{x}_{UB}$) across this domain of $\mathbf{x}(0)$. In the limit of $T \to \infty$, all orbits started off of $SM(\mathbf{x}_{LB})$ will have $|\mathbf{x}(T)| \approx 0$ or $|\mathbf{x}_{UB}|$, while those started on $SM(\mathbf{x}_{LB})$ will have $|\mathbf{x}(T)| \approx |\mathbf{x}_{LB}|$, suggesting the folds of $SM(\mathbf{x}_{LB})$ represent different local maxima of $|\mathbf{x}(T)|$ with finite but large T. Thus, for such high $\mathcal{E}_0$, many local maxima of $|\mathbf{x}(T)|$ emerge.

These calculations provide evidence that the optimization procedure of Section 2, applied to successively larger energy spheres and initiated in the correct hemisphere (the right hemisphere in Figure 6) will find the global maximum when below the minimal seed. Above the minimal seed, there is no guarantee that the optimization procedure will find the global maximum, which is why the algorithm was chosen so that finding the NLOP for energy spheres above the minimal seed is not necessary.

3.3 Results

We have found the minimal seed $\mathbf{x}_{MS}$ for the W97 model for values of R ranging from 105 to 2000. Let us first illustrate the search: the left panel of Figure 7 shows the progression of orbits in the search for the minimal seed when $R = 200$. The orbit of an NLOP below the edge shows transient growth with one well-defined maximum in energy. The growth is transient as would be had in the linearized system, with the orbit returning to the origin, hence producing a single maximum.

An orbit started above the edge shows similar initial transient growth followed by at least one further growth episode. When the initial condition is near the edge, the initial transient growth sends the orbit into a region of phase space where it is attracted towards the lower branch point. This much of the orbit, so far, has therefore been on or near the non-linear stable manifold of the lower branch point. The lower branch point being a saddle point, it has one unstable eigenvector $\mathbf{v}_u$. An orbit that began just above the edge will, when it gets near to $\mathbf{x}_{LB}$, be pushed off the stable manifold of $\mathbf{x}_{LB}$ by its unstable dynamics $\mathbf{v}_u$ and *away* from the origin, generating a second local maxima in the orbit's magnitude. An orbit started just below the edge, however, would be pushed *towards* the origin by the unstable dynamics of $-\mathbf{v}_u$. This weak edge, which is the stable manifold of the lower branch point, separates initial conditions whose orbits follow qualitatively different routes back to the origin (right panel of Figure 7). These orbits that straddle the edge also linger for a long time near $\mathbf{x}_{LB}$ itself ($t \sim 80$—125 in Figure 6): since $\mathbf{x}_{LB}$ is a fixed point, the velocity through phase space in nearby regions is small.

Note that there are often additional small maxima of energy at large time (not visible), owing to numerical inaccuracies when the orbit becomes extremely close to the origin. This necessitates the condition involving α in Figure 3.

Before showing the main result of this work—how the minimal seed scales with R—we ask how the minimal seed scales relative to $|\mathbf{x}_{LB}|$. In this model we know that $|\mathbf{x}_{LB}|$ asymptotes to a non-zero value for large R, and that the minimal seed $|\mathbf{x}_{MS}|$ asymptotes to zero for large R. Hence the ratio $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$ should go to ∞ as $R \to \infty$. But how does it behave for finite R, in particular for $R \to R_{SN}$. This is shown in Figure 8 over the range of R tested. The relationship is remarkably linear over almost all R, shallowing only for R very close to R_{SN} where the lower branch point is born and its magnitude changes rapidly with R.

Now in Figure 9 we show the scaling of the minimal seed against R. First, note that for each R the size of the minimal seeds found here is less than that found by Cossu (2005). This is indeed because we have found the minimal distance to the weak edge, which lies closer to the origin than does the strong edge. Second, we are now able to add to the discussion regarding how the minimal seed should scale with R. Cossu found reasonably good agreement with an asymptotic R^{-1} scaling, and we do too. For comparison, we found that the asymptotic (i.e. requiring the $R = 2000$ data point to be matched perfectly) scaling with a minimal least squares fit is $R^{-1.0361}$—reasonably close to R^{-1}. However, the exceptionally linear relationship shown in Figure 8 suggests we can do even better if we include not just R but also the (easily attainable) information of $|\mathbf{x}_{LB}|$: we find the scaling $|\mathbf{x}_{LB}|R^{-1}$ fits especially well. Indeed, $|\mathbf{x}_{LB}|$ is asymptotically flat at large R (recall Figure 4), so this scaling is equivalent to the R^{-1} scaling in the large R limit. However, the increase of

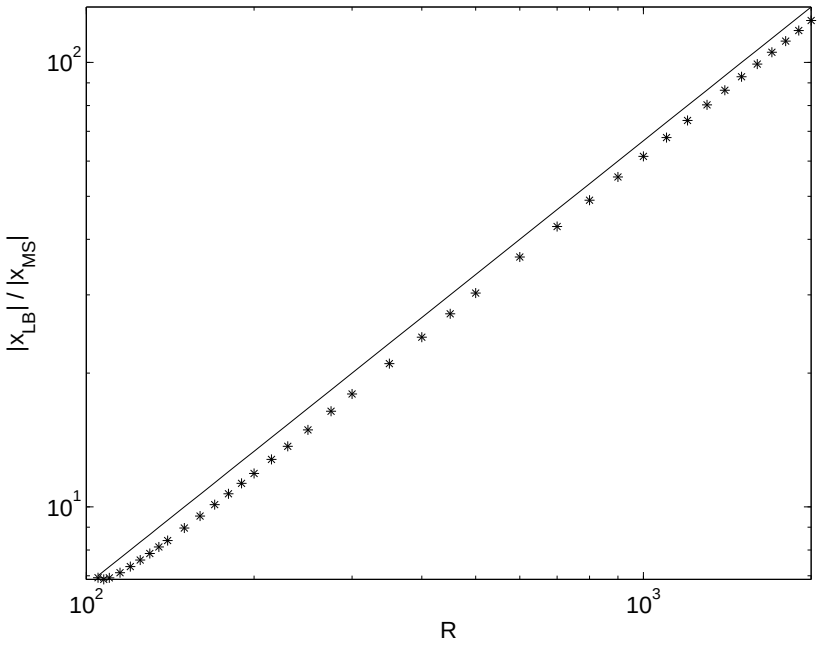


Figure 8: The ratio between the magnitude of the lower branch point $|\mathbf{x}_{LB}|$ and the minimal seed $|\mathbf{x}_{MS}|$ shown in black stars over a range of R, for Waleffe's 4D model. The upper solid line with a slope of 1 has been added to guide the eye, indicating a linear relationship between R and $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$.

$|\mathbf{x}_{LB}|$ going to smaller R allows this scaling to capture a great deal of how the minimal seed grows (as R decreases) faster than the R^{-1} scaling in the finite, non-asymptotic R regime.

Finally, we show in Figure 10 how the first, second, and fourth components of the minimal seed compare against its third component v, over a range of R. Recall that in the W97 model, v represents the amplitude of streamwise rolls. That v is always the largest component confirms that, in this model, perturbing the laminar flow with streamwise rolls is, broadly speaking, the most effective way to transition to turbulence. However, the other components being non-zero, the minimal seed—the most effective perturbation—is somewhat different from a pure streamwise roll perturbation, having finite but smaller perturbations from the laminar state in all modes (components of $\mathbf{x}$).

3.4 Cossu's Necessary Condition

Cossu (2005) gives a necessary condition for an initial condition $\mathbf{x}_0$ to be that point with minimal magnitude $\mathcal{E}(\mathbf{x}_0)$ which satisfies

$$\lim_{t\to\infty} \mathcal{E}(\mathbf{X}(t, \mathbf{x}_0)) \neq 0, \tag{20}$$

i.e. for $\mathbf{x}_0$ to be what Cossu had called the minimal seed. Following his argument, consider a candidate $\mathbf{x}_0$ for the minimal seed, and perturb it by following its orbit forward for a

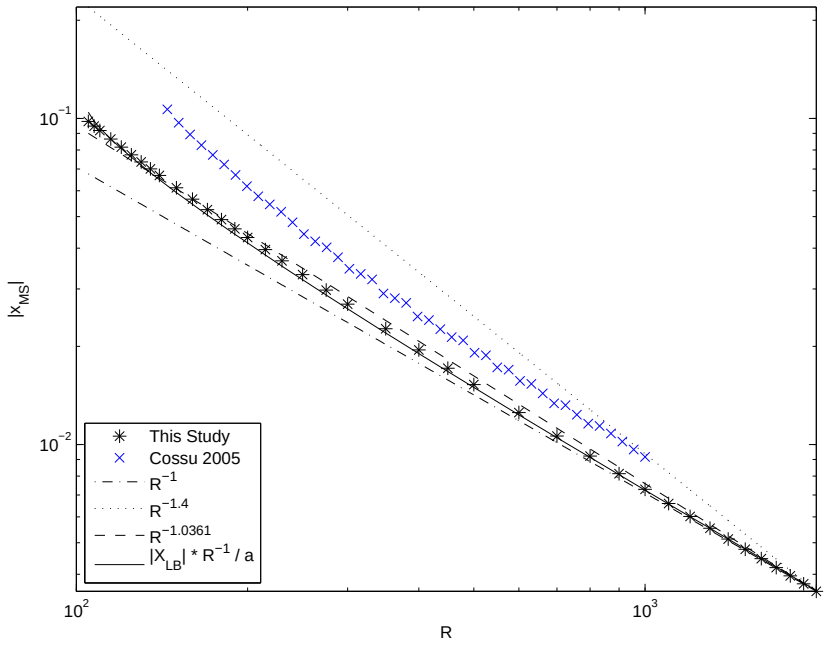


Figure 9: The Minimal Seed as a function of R for Waleffe's 4D model. Previous calculations by Cossu (2005) are shown in blue crosses, while our data is shown in black stars. The dash-dot and dotted lines show the R^{-1} and $R^{-1.4}$ scalings, respectively, while the dashed line shows the R power law with a least squares fit; since these are asymptotic ($R \to \infty$) scalings they are all fit to match the largest R data point. The solid line is the curve $|\mathbf{x}_{LB}|R^{-1}/a$ where a is the least-squares fit to $|\mathbf{x}_{LB}|/|\mathbf{x}_{MS}|$ vs. R from Figure 8.

short time Δt. The perturbation energy, in a Taylor series expansion, is

$$\mathcal{E}_0(\Delta t) \equiv \mathcal{E}(\mathbf{X}(\Delta t, \mathbf{x}_0)) = \mathcal{E}_0(0) + \Delta t \left(\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\bigg|_{\Delta t=0} \right) + \frac{\Delta t^2}{2} \left(\frac{\mathrm{d}^2\mathcal{E}_0}{\mathrm{d}t^2}\bigg|_{\Delta t=0} \right) + \mathcal{O}(\Delta t^3). \quad (21)$$

If $\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\Big|_{\Delta t=0} < 0$, then $\mathcal{E}(\mathbf{X}(\Delta t, \mathbf{x}_0)) < \mathcal{E}(\mathbf{x}_0)$. So $\mathbf{X}(\Delta t, \mathbf{x}_0)$, since it shares an orbit with $\mathbf{x}_0$, must also satisfy (20), and hence $\mathbf{x}_0$ is not the minimal seed. Similarly if $\frac{\mathrm{d}\mathcal{E}_0}{\mathrm{d}t}\Big|_{\Delta t=0} > 0$, then $\mathbf{X}(-\Delta t, \mathbf{x}_0)$ shows that $\mathbf{x}_0$ is not the minimal seed. Therefore, it is necessary that the minimal seed $\mathbf{x}_{MS}$ must satisfy

$$\frac{\mathrm{d}\mathcal{E}(\mathbf{X}(t, \mathbf{x}_{MS}))}{\mathrm{d}t}\bigg|_{t=0} = 0 \quad \text{and} \quad \frac{\mathrm{d}^2\mathcal{E}(\mathbf{X}(t, \mathbf{x}_{MS}))}{\mathrm{d}t^2}\bigg|_{t=0} \geq 0 \quad (22)$$

This argument also applies when finding the minimal seed when the edge contains a weak component, i.e. when the minimal seed is on the stable manifold of, in this case, the lower branch point: $\mathbf{x}_{MS} \in SM(\mathbf{x}_{LB})$. But note, however, that this argument does not apply so easily when searching for the NLOP subject to $\mathcal{E}(\mathbf{x}_0)$ equal to some specified d.

These ideas afford us a way to check our results. First, for the W97 model and any

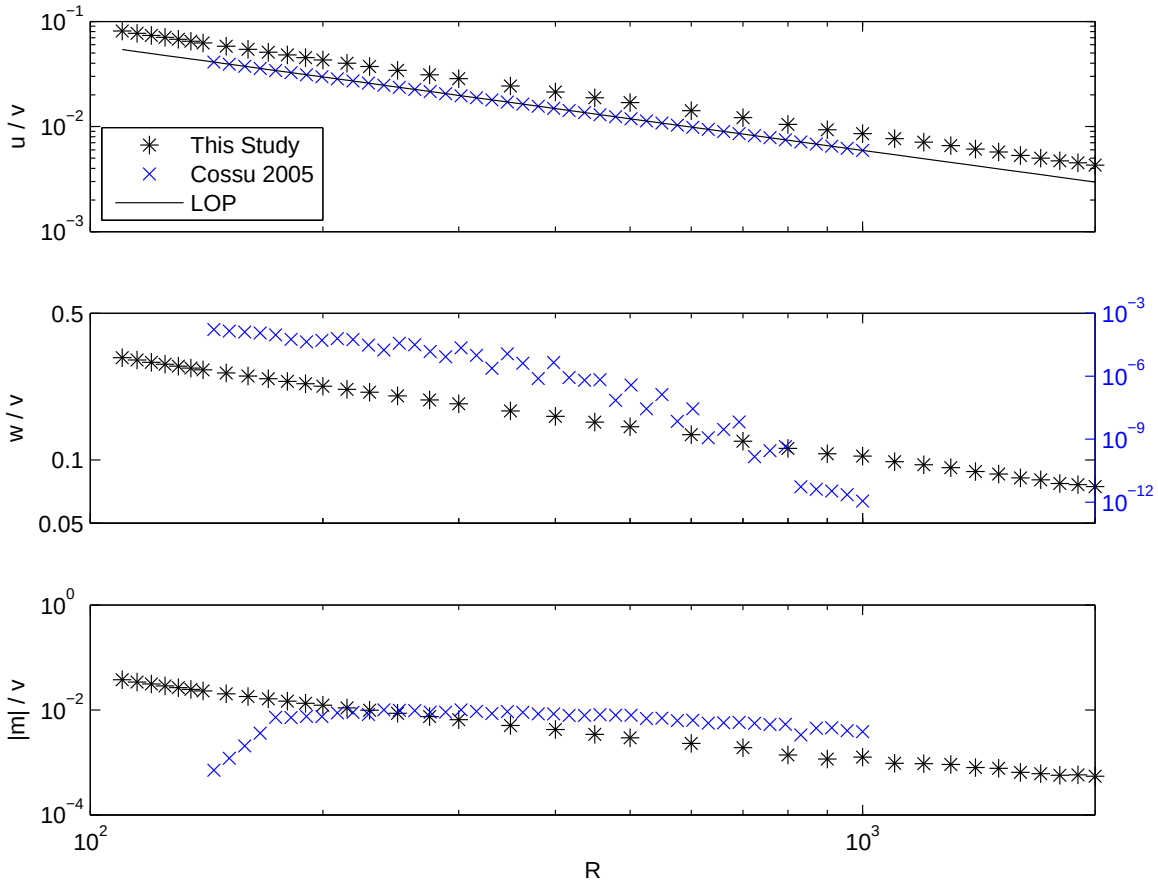


Figure 10: Ratio of the first (m), second (u), and fourth (w) components of the minimal seed to the third component v. We found $m > 0$ for all minimal seeds, but Cossu (2005) had $m < 0$ for most R, hence $|m|$ is shown here.

other in which the non-linear terms are energy preserving,

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{2}\mathcal{E}(\mathbf{x})^2\right) = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\mathbf{x}\cdot\mathbf{x} = \mathbf{x}\cdot(L\mathbf{x}) + \mathbf{x}\cdot(N(\mathbf{x})) = \mathbf{x}\cdot(L\mathbf{x}). \tag{23}$$

This measure of the error is immediately calculable and shown in Figure 11, scaled by $\mathbf{x}_{MS}\cdot\mathbf{x}_{MS}$, over the range of R tested. That there is no particular pattern suggests the error is due to numerical noise.

However, what is an acceptable magnitude for the above error? If we shift $\mathbf{x}_{MS}$ along its orbit by a small time Δt, then $\mathcal{E}^2/2$ will change by approximately $\Delta t\ \mathbf{x}_{MS}\cdot(L\mathbf{x}_{MS})$. But the size of Δt is presently unknown: too large a Δt and higher order terms in (21) will become significant. Thus, for a given R and estimate of $\mathbf{x}_{MS}$, we follow the orbit from $\mathbf{x}_{MS}$ backwards, or forwards, in time until $\mathrm{d}\mathcal{E}/\mathrm{d}t = 0$ is had, and therefore a more minimal $\mathbf{x}_{MS}$ is found. Carrying this out on the $\mathbf{x}_{MS}$ for $R = 120$ manages to reduce $|\mathbf{x}_{MS}|$ by only $5\cdot10^{-12}$, or $2\cdot10^{-11}$ for $R = 800$, or $2\cdot10^{-13}$ for $R = 2000$, suggesting our estimate of $\mathbf{x}_{MS}$ is quite accurate.

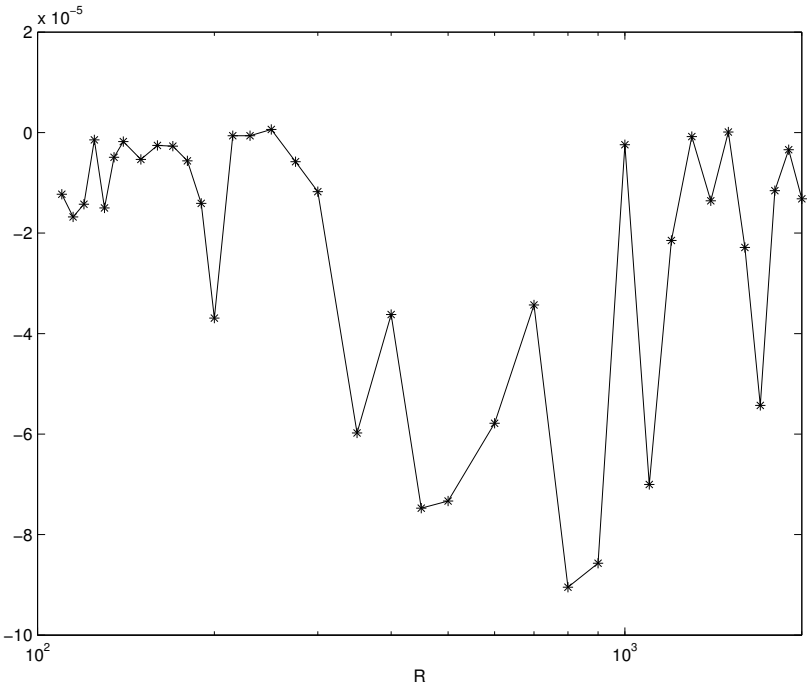


Figure 11: $(\mathbf{x}_{MS} \cdot (L\mathbf{x}_{MS}))$ / $(\mathbf{x}_{MS} \cdot \mathbf{x}_{MS})$ as a function of R. Deviations from zero indicate error in the estimation of $\mathbf{x}_{MS}$.

4 Conclusion

We have undertaken a search, across a range of Reynolds numbers R, for the minimal seed in the low-dimensional model of W97. One previous attempt (Cossu, 2005) at this was made. Cossu kindly shared his data with us in private communication, from which we discovered that his results for the minimal seed are in fact the least distance from the origin (the laminar flow state) to the boundary of the basin of attraction of the upper branch fixed point (a flow state that exists within turbulence yet itself is stable). The latter marks the ("strong") edge of the basin of attraction of the laminar fixed point, separating orbits that eventually return to the laminar flow state from those that never return. However, we have shown that in the W97 model there exists a boundary that is internal to the basin of attraction of the laminar fixed point: this so-called "weak" edge separates orbits that both return to the laminar flow state given sufficient time but do so by qualitatively different routes: orbits below this "weak" edge return fairly directly to the origin, while orbits above undergo a circuitous route that must wind around the upper branch fixed point before returning to the origin. The latter orbit, it would be said, did indeed transition to turbulence, despite that it ultimately relaminarizes. Thus the true minimal seed lies on this "weak" edge, and is in fact smaller than any point on the "strong" edge. We have found a more minimal seed than was previously known.

We have applied the relatively new non-linear optimization technique, reviewed by Kerswell *et al.* (2014). In this problem, the demarcation between states that are above and those that are below the weak edge is more subtle than between states that are above or below

the strong edge in which the demarcation is simply whether the orbit eventually relaminarizes. This marks new territory for this optimization technique and some modifications were required.

Considering the magnitudes of the minimal seeds we found over a range of R, our results support the hypothesis for a R^{-1} scaling law of the minimal seed in the W97 model. However, we discovered an even better scaling can be obtained if an additional piece of information is used: the magnitude of the lower branch point. Now, the minimal seed is the point on the stable manifold of the lower branch point closest to the origin. At large R the lower branch point is relatively insensitive to increases in R, and the decreasing minimal seed with increasing R is due to this stable manifold shifting closer to the origin. At finite R however, the lower branch point moves towards the origin with increasing R, translating its stable manifold closer to the origin along with it. This shifting of the lower branch point, it seems, causes the minimal seed to decrease somewhat faster than R^{-1}. In fact, we found that the ratio of the lower branch point to the minimal seed was exceptionally linear with R, leading to a new scaling for the minimal seed that goes as the product of the magnitude of the lower branch point with R^{-1}. The agreement of this scaling law with our results is quite encouraging, but whether such a scaling law is unique to the particular model studied here is not presently known.

5 Acknowledgements

I am deeply grateful for the opportunity to participate in the Geophysical Fluid Dynamics Program. Norman Lebovitz deserves my deepest thanks, for developing a summer research project that was both intriguing and accessible to me, and for skillfully guiding me through it. I also thank Colm-cille Caulfield for many useful discussions related to this research project. I would like to thank Glenn Flierl, Raf Ferrari, and Antonello Provenzale for organizing and directing the 2014 school, Geoff Vallis and Kerry Emanuel for their lectures and mentorship, and all the seminar speakers for sharing their knowledge particularly with the fellows. Lastly, thanks to all the fellows for your camaraderie and joviality.

References

J. S. Baggett and L. N. Trefethen, *Low-dimensional models of subcritical transition to turbulence*, Physics of Fluids (1994-present), 9 (1997), pp. 1043–1053.

S. J. Chapman, *Subcritical transition in channel flows*, Journal of Fluid Mechanics, 451 (2002), pp. 35–97.

C. Cossu, *An optimality condition on the minimum energy threshold in subcritical instabilities*, Comptes Rendus Mécanique, 333 (2005), pp. 331 – 336.

B. F. Farrell, *Optimal excitation of perturbations in viscous shear flow*, Physics of Fluids (1958-1988), 31 (1988), pp. 2093–2102.

B. Hof, a. Juel, and T. Mullin, *Scaling of the turbulence transition threshold in a pipe*, Physical Review Letters, 91 (2003), p. 244502.

R. R. Kerswell, C. C. T. Pringle, and A. P. Willis, *An optimization approach for analysing nonlinear stability with transition to turbulence in fluids as an exemplar*, Reports on Progress in Physics, 77 (2014), p. 085901.

G. Kreiss, A. Lundbladh, and D. S. Henningson, *Bounds for threshold amplitudes in subcritical shear flows*, Journal of Fluid Mechanics, 270 (1994), pp. 175–198.

N. R. Lebovitz, *Boundary collapse in models of shear-flow transition*, Communications in Nonlinear Science and Numerical Simulation, 17 (2012), pp. 2095 – 2100. Special Issue: Mathematical Structure of Fluids and Plasmas Dedicated to the 60th birthday of Phil Morrison.

S. A. Orszag, *Accurate solution of the orr-sommerfeld stability equation*, Journal of Fluid Mechanics, 50 (1971), pp. 689–703.

L. N. Trefethen, A. E. Trefethen, S. C. Reddy, and T. A. Driscoll, *Hydrodynamic stability without eigenvalues*, Science, 261 (1993), pp. 578–584.

F. Waleffe, *Hydrodynamic stability and turbulence - beyond transients to a self-sustaining process*, Studies In Applied Mathematics, 95 (1995), pp. 319–343.

——, *Transition in shear flows. nonlinear normality versus non-normal linearity*, Physics of Fluids (1994-present), 7 (1995), pp. 3060–3066.

——, *On a self-sustaining process in shear flows*, Physics of Fluids (1994-present), 9 (1997), pp. 883–900.

Vortex Filament Dynamics

Jim Thomas

August 16, 2015

1 Introduction

Formation of large coherent vortices is a recurring theme in two-dimensional turbulence investigations [1]. DNS simulations and lab experiments have demonstrated this time and again. Once these large vortices form, like-signed vortices merge together in an inelastic process. During this merger process filamentation occurs leading to the formation of thin strips of vorticity. While some filaments remain attached to the large coherent vortices others are expelled during the merger and remain detached from the large vortices. Under normal conditions one would naturally expect these filaments to be unstable to perturbations, the growth of instabilities leading to roll up of these filaments. But surprisingly, it has been demonstrated that in the presence of other coherent vortices, the rolling up of these filaments is inhibited [2]. The strong external shearing and straining fields prevents the rolling up of these filamnets. As a result, the filaments form a characteristic feature of fully developed 2D turbulence. This work aims to capture the dynamics of these filaments using a simple new model. An integro-differential equation is derived for this purpose, which is then solved for a variety of cases.

In Sec. 2 derivation of the relevant equations and some of its properties are discussed. Sec. 3 involves studying the effect of disturbances on filaments and response of filaments to externally imposed velocity fields. Possibility of equilibrium states for filaments in the presence of external fields is discussed in Sec. 4 . Some filament-point vortex and filament-filament interactions are investigated in Sec. 5 .

2 Mathematical Model

A vortex filament (VF hereafter) is a very thin strip of vorticity in the $x - y$ plane. The small thickness assumption will be used to characterize the filament by its center line curve. Mass (area) conservation and material transport of vorticity are used to describe variation in local thickness and vorticity.

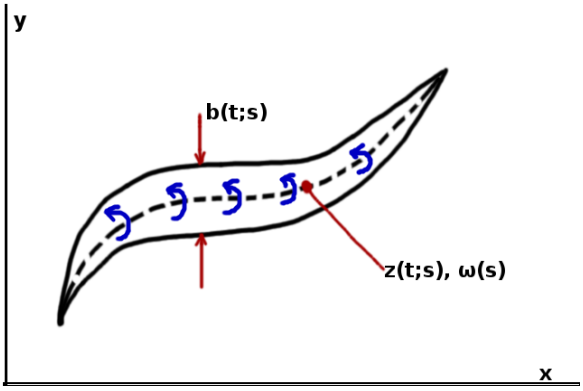


Figure 1: A vortex filament: z(s,t) is a curve in the complex plane parametrised by s, varying with time, and b(s,t) is the local filament thickness.

2.1 Derivation of the VF equation

The VF equation is derived from first principles here.

Vorticity at any point is

$$\boldsymbol{\omega} = \nabla \times \boldsymbol{v}$$

$$\nabla \cdot \boldsymbol{v} = 0 \quad \Rightarrow \boldsymbol{v} = \hat{\boldsymbol{z}} \times \nabla\psi \ .$$

Combining above two equations gives

$$\Delta\psi = \omega$$

If vorticity field is compact, this can be uniquely solved to get

$$\psi(\boldsymbol{x}) = \frac{1}{2\pi} \int\limits_{\boldsymbol{x'}\in\mathbb{R}^2} \omega(\boldsymbol{x'}) \ln|\boldsymbol{x} - \boldsymbol{x'}| d\boldsymbol{x'}$$

which leads to

$$\boldsymbol{v}(\boldsymbol{x}) = \frac{1}{2\pi} \int\limits_{\boldsymbol{x'}\in\mathbb{R}^2} \frac{\hat{\boldsymbol{z}} \times (\boldsymbol{x} - \boldsymbol{x'})}{|\boldsymbol{x} - \boldsymbol{x'}|^2} \omega(\boldsymbol{x'}) d\boldsymbol{x'}$$

This is the 2D version of the infamous Biot- Savart law.

Using small thickness approximation, we write $d\boldsymbol{x'} = b(s')dl(s'), \quad \omega = \gamma(s')$

This is analogous to thin airfoil approximation in aerodynamics [3].

Thus, self-induced velocity of the filament at any point is

$$\boldsymbol{v}(\boldsymbol{x}, t) = \frac{1}{2\pi} \int\limits_{L} \gamma(s') \frac{\hat{\boldsymbol{z}} \times (\boldsymbol{x}(s,t) - \boldsymbol{x}(s',t'))}{|\boldsymbol{x}(s,t) - \boldsymbol{x}(s',t)|^2} b(s',t) dl(s')$$

The sum of this and the external velocity is the total velocity at any point on the filament, i.e.

$$\frac{\partial \boldsymbol{x}(s,t)}{\partial t} = \boldsymbol{U}(\boldsymbol{x}(s),t) + \frac{1}{2\pi}\fint_L b(s',t)\gamma(s')\frac{\hat{z}\times(\boldsymbol{x}(s,t)-\boldsymbol{x}(s',t'))}{|\boldsymbol{x}(s,t)-\boldsymbol{x}(s',t)|^2}dl(s')$$

Or in complex variable notation,

$$\frac{\partial \boldsymbol{z}^*(s,t)}{\partial t} = \boldsymbol{U}^*(\boldsymbol{z}) + \frac{1}{2\pi\iota}\fint_L \frac{b(s',t)\,|\boldsymbol{z}_s(s',t)|\,\gamma(s')}{\boldsymbol{z}(s,t)-\boldsymbol{z}(s',t)}ds'$$

where $\boldsymbol{z} = x + \iota y$

To close the system, we need an evolution equation for thickness. This can be derived by demanding local area conservation of the VF.

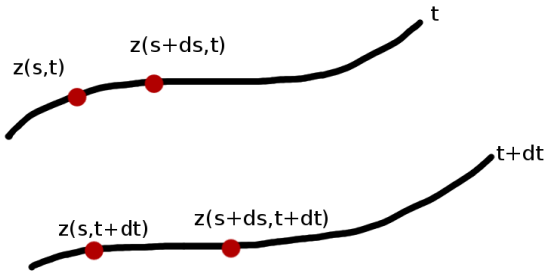


Figure 2: Centerline curve of a VF at two different times

On demanding that local area remains same at t and $t+\delta t$

$$b(s+\delta s/2,t)\,|\boldsymbol{z}(s+\delta s,t)-\boldsymbol{z}(s,t)| = b(s+\delta s/2,t+\delta t)\,|\boldsymbol{z}(s+\delta s,t+\delta t)-\boldsymbol{z}(s,t+\delta t)|$$

$$b(s,t)\,|\boldsymbol{z}_s(s,t)\delta s| = (b(s,t)+\delta t b_t(s,t))\,|\boldsymbol{z}_s(s,t)\delta s + \boldsymbol{z}_{st}(s,t)\delta s\delta t| + O(\delta s^2)$$

$$\frac{1}{b}\frac{\partial b}{\partial t} + \frac{1}{2|\boldsymbol{z}_s|^2}\frac{\partial |\boldsymbol{z}_s|^2}{\partial t} = 0$$

After some simplifications, this yields

$$b(s,t)\,|\boldsymbol{z}_s(s,t)| = b(s,0)\,|\boldsymbol{z}_s(s,0)| = b^0(s)|\boldsymbol{z}_s^0|$$

This is a remarkable result since it implies that thickness in this model is slaved to the curve. As a consequence, one need not solve for the thickness simultaneously with the curve points. Once the curve is known, thickness can be calculated and updated.

The final equation is then

$$\frac{\partial \boldsymbol{x}(s,t)}{\partial t} = \boldsymbol{U}(\boldsymbol{x}) + \frac{1}{2\pi}\int\limits_{L} b^0(s')|\boldsymbol{x}_s^0|\gamma(s')\frac{\hat{\boldsymbol{z}}\times(\boldsymbol{x}(s,t)-\boldsymbol{x}(s',t))}{|\boldsymbol{x}(s,t)-\boldsymbol{x}(s',t)|^2}ds'$$

which is in complex notation

$$\frac{\partial \boldsymbol{z}^*(s,t)}{\partial t} = \boldsymbol{U}^*(\boldsymbol{z}) + \frac{1}{2\pi\iota}\int\limits_{L} \frac{b^0(s')|\boldsymbol{z}_s^0|\gamma(s')}{\boldsymbol{z}(s,t)-\boldsymbol{z}(s',t)}ds' \tag{1}$$

This formulation can be extended to include multiple filaments as given below:

$$\begin{aligned}\frac{\partial \boldsymbol{x}_i(s_i,t)}{\partial t} &= \boldsymbol{U}(\boldsymbol{x}_i)\\ &+ \frac{1}{2\pi}\int\limits_{L_i} b_i^0(s_i')|\boldsymbol{x}_{i\,s_i}^{\,0}|\gamma_i(s_i')\frac{\hat{z}\times(\boldsymbol{x}_i(s_i,t)-\boldsymbol{x}_i(s_i',t'))}{|\boldsymbol{x}_i(s_i,t)-\boldsymbol{x}_i(s_i',t)|^2}ds_i'\\ &+ \sum_j \frac{1}{2\pi}\int\limits_{L_j} b_j^0(s_j')|\boldsymbol{x}_{j\,s_j}^{\,0}|\gamma_j(s_j')\frac{\hat{z}\times(\boldsymbol{x}_i(s_i,t)-\boldsymbol{x}_j(s_j',t))}{|\boldsymbol{x}_i(s_i,t)-\boldsymbol{x}_j(s_j',t)|^2}ds_j'\end{aligned}$$

The last summation over all other filaments is the mutual interaction term.

Analogous to Contour Dynamics (CD) we shall call this model Filament Dynamics (FD hereafter)

As a special singular limit, let $b \to 0, \gamma \to \infty$ holding $b\gamma$ fixed and $L \to \infty$. Then (1) reduces to:

$$\frac{\partial \boldsymbol{z}^*(s,t)}{\partial t} = \frac{1}{2\pi\iota}\int\limits_{-\infty}^{\infty} \frac{d\Gamma(s')}{\boldsymbol{z}(s,t)-\boldsymbol{z}(s',t)}$$

This is the Birkhoff-Rott equation used to model the evolution of a vortex sheet [4,5].

2.2 Conservation laws for self interacting filament (U=0)

Equation (1) has a Hamiltonian formulation.

Hamiltonian,

$$H[\boldsymbol{x}] = -\frac{1}{2\pi}\int\limits_{s}\int\limits_{s'} b^0(s)\left|\boldsymbol{x}_s^0(s)\right|\gamma(s)b^0(s')\gamma(s')\left|\boldsymbol{x}_s^0(s')\right| ln\left|\boldsymbol{x}(s,t)-\boldsymbol{x}(s',t)\right| ds'ds$$

Thus, (1) can be written as Hamilton's equations

$$\frac{\partial x}{\partial t} = \frac{1}{\zeta}\frac{\delta H}{\delta y}$$

$$\frac{\partial y}{\partial t} = -\frac{1}{\zeta}\frac{\delta H}{\delta x}$$

where $\zeta = \gamma(s) b^0(s) \left| \boldsymbol{x}_s^0(s) \right|$

Thus, the Hamiltonian (Energy) is an invariant of the system, i.e. $H_t = 0$

The Hamiltonian can be modified to include the effect of external velocity field using the stream function.

An invariant center of vorticity $\overline{\boldsymbol{x}}$ can be defined for a VF such that

$$\oint_L b^0(s') |\boldsymbol{x}_{s'}^0| \gamma(s') \frac{\hat{z} \times (\overline{\boldsymbol{x}} - \boldsymbol{x}(s', t))}{|\overline{\boldsymbol{x}} - \boldsymbol{x}(s', t)|^2} ds' = 0$$

The center of vorticity $\overline{\boldsymbol{x}}$, is the point where the velocity induced by the whole filament is zero.

Equation (1) conserves angular momentum.

Angular Momentum,

$$L = \oint_s b^0(s) \left| \boldsymbol{x}_s^0(s) \right| \gamma(s) \left| \boldsymbol{x}(s,t) \right|^2 ds$$

2.3 Numerical solutions

Examples are calculated by a numerical scheme in which the VF is discretized into rectangular strips with N points on it. The system of $2N$ equations are then solved using RK-4. The integral is replaced with a desingularised summation [6]

$$\sum_j \frac{1}{|\boldsymbol{x}_i - \boldsymbol{x}_j|^2 + \delta^2}$$

with midpoints between two points used to evaluate the sum. Thickness is at the midpoint between two points. After each time step (n), the thickness is updated using

$$b_j^{n+1} l_j^{n+1} = b_j^n l_j^n$$

3 Effect of Disturbances on VFs

It is well known that thin strips of vorticity are unstable to disturbances and tend to roll up. Here we test this using our model. We study two cases: a parabolic VF and a circular VF.

A parabolic VF has thickness variation

$$b(s) = b_{max}(1 - s^2), \quad s \in [-1, 1]$$

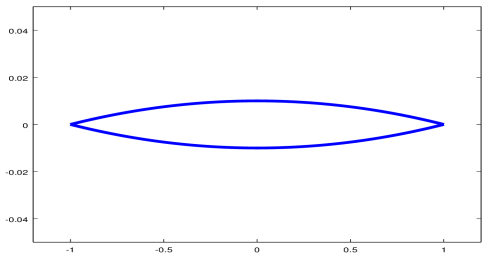


Figure 3: A parabolic VF

This is now disturbed by a finite amplitude sinusoidal perturbation and allowed to evolve. In the configuration that we study, maximum thickness, $b_{max} = 10^{-5}$

The initial and long time states are shown below:

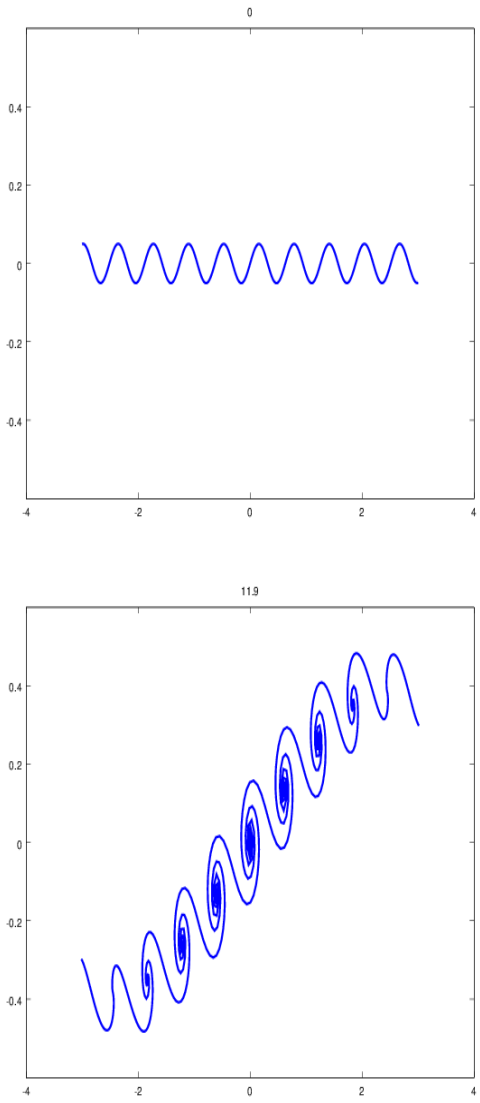


Figure 4: Initial and long time state of a perturbed parabolic VF

The second case we study is a slightly perturbed circular VF. The circular VF is of constant thickness, $b = .02$

The initial and long time states are shown below:

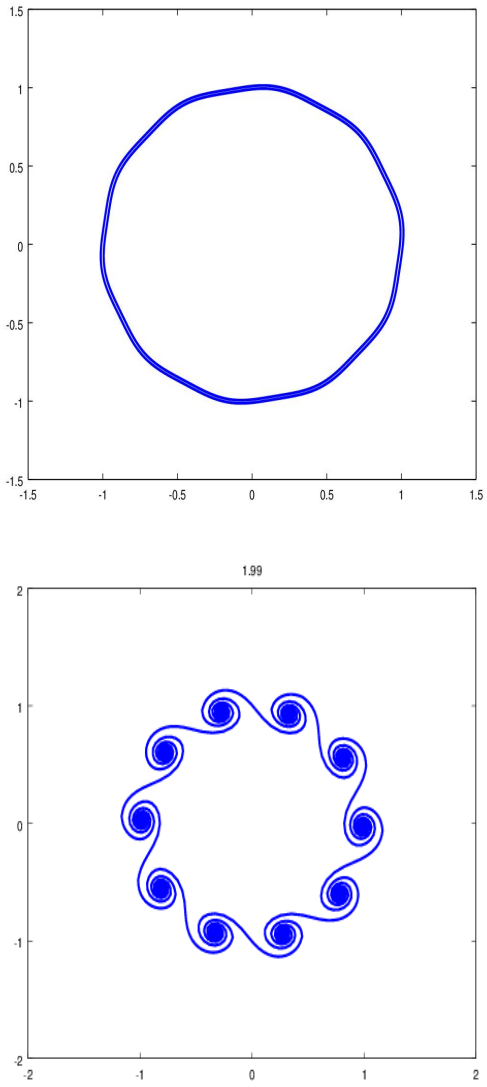
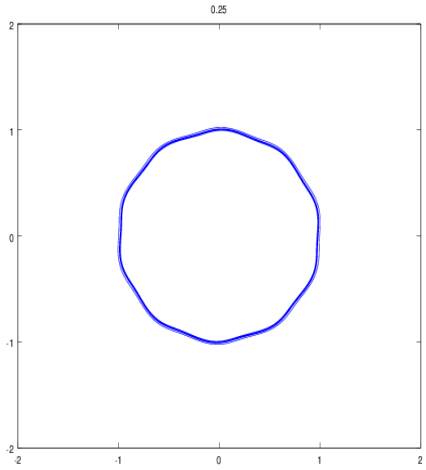
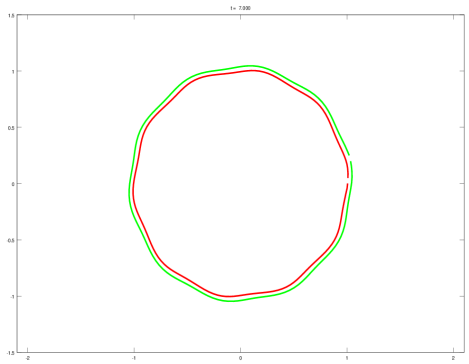
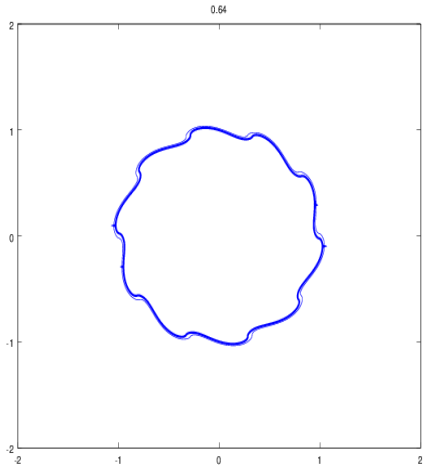


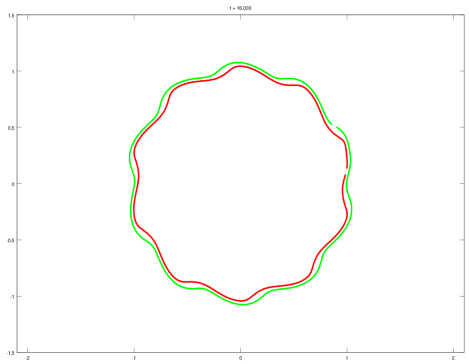
Figure 5: Initial and long time state of a perturbed circular VF

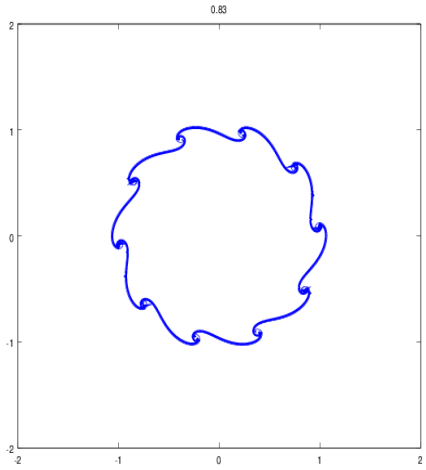
Thus VFs, as expected do roll up on being disturbed. For the case of a slightly disturbed circular VF, we compare this model with a Contour Dynamics calculation. Various stages of evolution are as shown:

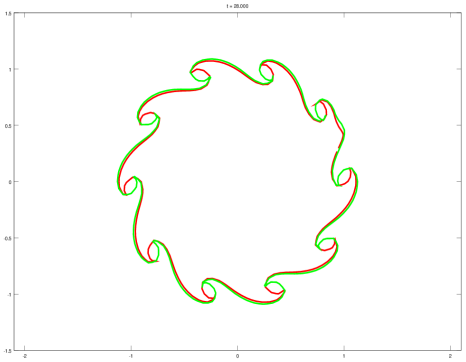

0.25

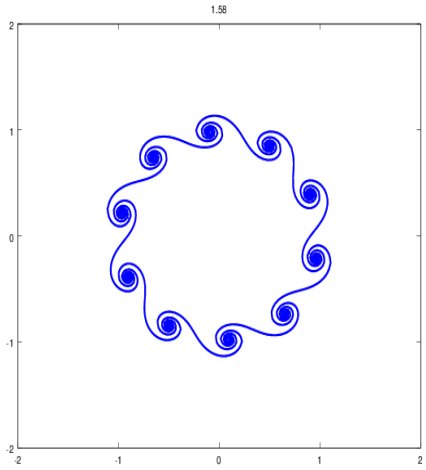


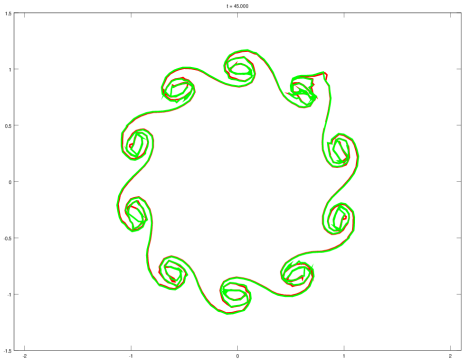

0.64




0.83




1.58



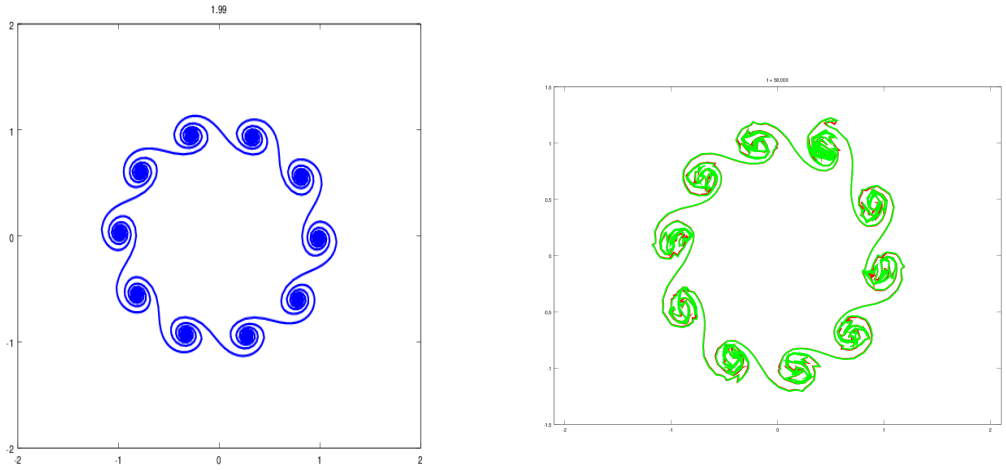


Figure 6: Comparison between FD and CD

Note that disturbances propagate and grow in both models. Thickening happens in certain regions leading to the formation of beads of vorticity. However, the FD model is unable to maintain regions of constant vorticity as CD and it rolls up into a spiral in the final stages as seen above.

3.1 Shape of VFs and their influence on dynamics

The dynamics of VFs depend a lot on their shapes. It is seen that an elliptical VF $\left((\frac{x}{a})^2 + (\frac{y}{b})^2 = 1\right)$ rotates without change of shape.

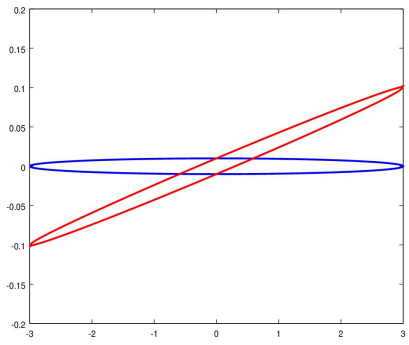


Figure 7: Elliptic VF-red is initial configuration and blue is state after a small time

Hyper-ellipses $((\frac{x}{a})^n + (\frac{y}{b})^n = 1, \quad n > 2)$ rotate with their tips leading the main body of rotation.

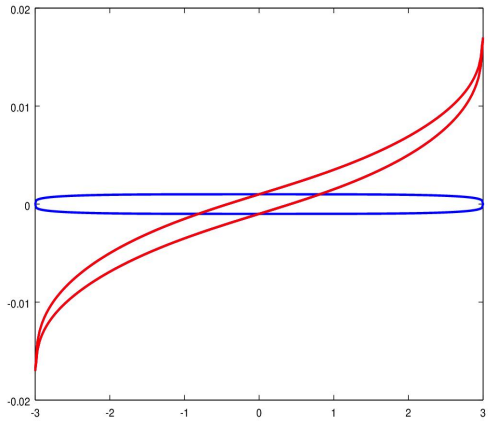


Figure 8: Hyper-elliptic VF-red is initial configuration and blue is state after a small time

A parabolic filament ($b = b_{max}(1 - s^2)$) rotates with its tip lagging behind the main body of rotation.

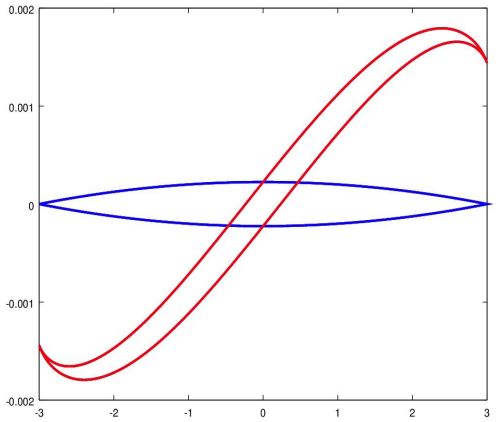


Figure 9: Parabolic VF-red is initial configuration and blue is state after a small time

Further investigations indicate that sharper filament tips tend to lag while blunter filament tips have a tendency to lead the main body of the filament.

3.2 Effect of external velocity field on VFs

Here we impose an external velocity field and study the response of the VF. The previous case of perturbed parabolic VF is used but now a velocity field equivalent to the presence of a point vortex at (0,3) is imposed. i.e., $v_r = 0, \quad v_\theta = \frac{\Gamma}{2\pi|\boldsymbol{x}-3\widehat{\boldsymbol{y}}|}$

Two cases are considered. Case of weak external field ($\Gamma/(2\pi b_{max} L\gamma) \ll 1$) and strong external field ($\Gamma/(2\pi b_{max} L\gamma) \gg 1$) as shown below:

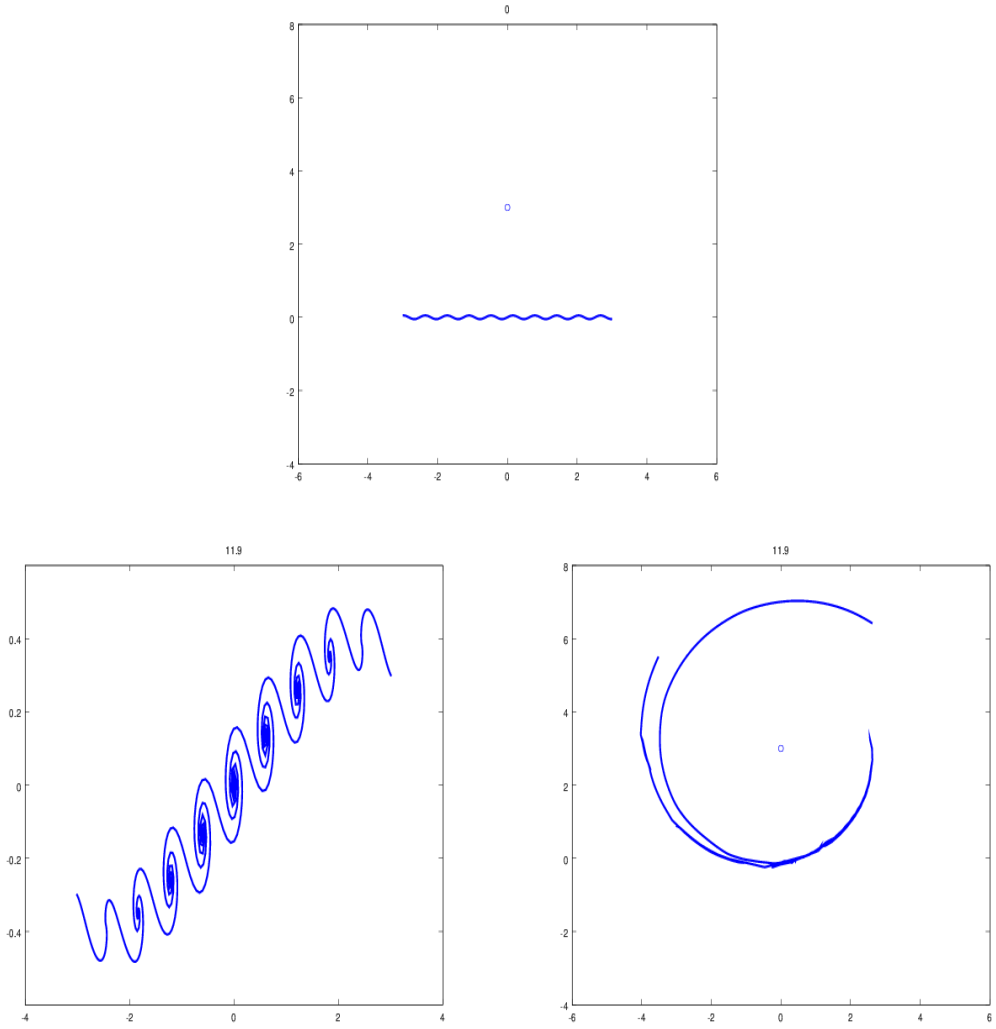


Figure 10: Counterclockwise from top: Initial state, long time state in the absence of external field, long time state in the presence of external field

In the case of weak external velocity, rolling up of the VF occurs as expected. But in the presence of strong external velocity field, the filament is extremely stretched out and flattened, preventing all rolling up activities.

Further numerical experiments indicate that a strong external field can stretch the VF so as to prevent rolling up from occurring. On the other hand, a weak external flow is unable to impact the VF. This leads to the natural question of equilibrium. Are there critical velocity fields that can hold the whole VF in equilibrium?

4 Equilibrium states of VFs

The relevant equations for equilibrium states are

$$\boldsymbol{U}^*(\boldsymbol{z}) + \frac{1}{2\pi\iota}\int_L \frac{b(s')\gamma(s')}{\boldsymbol{z}(s) - \boldsymbol{z}(s')}\left|\boldsymbol{z}_s(s')\right| ds' = 0$$

Closed filaments

For the class of closed filaments, a circular filament of constant thickness and vorticity is in self equilibrium , i.e. U=0. This is easily verified by using $z = e^{\iota\theta}$ to evaluate the integral.

Open filaments

Consider straight filaments placed on the x-axis with constant vorticity along the curve. Thus $z = s$ is the equation of the VF. We prescribe various thickness distributions and find the corresponding flow fields based on the equation:

$$\boldsymbol{U}(s) = \frac{\iota\gamma}{2\pi}\int_L \frac{b(s')}{s - s'} ds' \tag{2}$$

1. $b(s) = b_{max}\sqrt{1 - (s/a)^2}$
   .

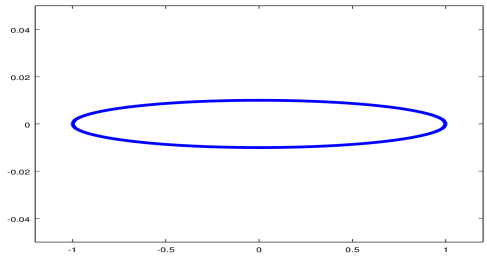


Figure 11:

This is an elliptic thickness variation as shown. For this $b(s)$, the critical velocity field is

$$u = ky, \quad v = -kx \Rightarrow \omega = -2k \quad (k = \gamma b_{max}/2a)$$

This is a flow with constant vorticity that can hold an elliptic VF in equilibrium. Half the vorticity of this external field is exactly the rotation rate of an elliptic VF.

Consider an ellipse $\frac{x^2}{a^2} + \frac{y^2}{c^2} = 1$

In the filament approximation, we obtain the rotation rate as

$$\Omega_f = \gamma\lambda$$

where $\lambda = \frac{c}{a}$ is the aspect ratio

If we had an elliptic patch of constant vorticity [7],

$$\Omega = \gamma \frac{\lambda}{(1+\lambda)^2}$$

If $\lambda \ll 1$, binomial expansion gives $\Omega \sim \Omega_f + O(\lambda^2)$.

Thus, in the filament approximation, the ellipse tends to rotate faster, the difference being $O(\lambda)$.

2. $b(s) = b_{max}(1-s^2)^{3/2}$

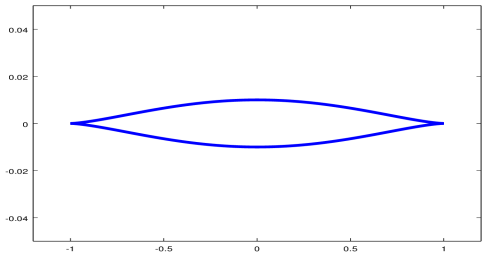


Figure 12:

The critical velocity field is

$$u = k_o\left(\frac{3}{2}y + y^3 - 3x^2y\right), \quad v = -k_o\left(\frac{3}{2}x - x^3 + 3y^2x\right)$$

3. $b(s) = b_{max}(1-s^2)$

This is a parabolic thickness variation.

$$b(s) = b_{max}(1-s^2)$$

The critical velocity field in complex notation is

$$\boldsymbol{U} = -k_p\left((1-z^2)\ln\frac{1+z}{1-z} + 2z\right)$$

The k_i's are constants that depend on the central region thickness of the filament and similar other parameters.

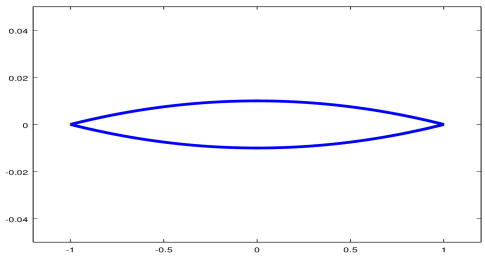


Figure 13:

Stability of equilibria

Here the stability of equilibria of VF is tested. We shall consider small local perturbations on an elliptic VF and study its evolution.

The disturbance equations are:

$$\frac{\partial \widehat{\boldsymbol{z}}^*(s,t)}{\partial t} = \widehat{\boldsymbol{U}}^*(\boldsymbol{z}_e, \widehat{\boldsymbol{z}}) - \frac{1}{2\pi\iota}\int_L \frac{b_e(s')\,|\boldsymbol{z}_e(s')|\,\gamma^0(s')}{(\boldsymbol{z}_e(s)-\boldsymbol{z}_e(s'))^2}(\widehat{\boldsymbol{z}}(s,t)-\widehat{\boldsymbol{z}}(s',t))ds'$$

Consider small perturbation on a straight elliptical filament with constant vorticity distribution in an equilibrium flow field. Then $\boldsymbol{z}_e = s,\ \boldsymbol{U} = \iota\beta\boldsymbol{z},\ b_e(s) = \sqrt{1-s^2}$ and after dropping $\widehat{\ }$ ' s leads to

$$(\boldsymbol{z}_t(s,t) + \iota\beta\boldsymbol{z})^* = -\frac{\gamma^0}{2\pi\iota}\int_{-a}^{a} b_0 \frac{\sqrt{1-(s'/a)^2}}{(s-s')^2}\left(\boldsymbol{z}(s,t)-\boldsymbol{z}(s',t)\right)ds'$$

Setting $s = a\xi$

$$(\boldsymbol{z}_t + \iota\beta\boldsymbol{z})^* = -\frac{\beta}{\pi\iota}\int_{-1}^{1} \frac{\sqrt{1-\xi'^2}}{(\xi-\xi')^2}\left(\boldsymbol{z}(\xi,t)-\boldsymbol{z}(\xi',t)\right)d\xi'$$

where $\beta = \gamma^0 b^0/(2a)$ which can be written as

$$(\boldsymbol{z}_t + \iota\beta\boldsymbol{z})^* = -\frac{\beta}{\pi\iota}F[z]$$

Looking for solutions in terms of the normal modes,

$$z = \sum A_n e^{\sigma_n t}\phi_n$$

where ϕ_n satisfies $F[\phi_n] = \lambda_n\phi_n$ and for the functional $F[z]$, one can show that the eigenvalues are $\lambda_n = n\pi$

Thus,

$$\sigma_n = \pm\beta\sqrt{n^2-1}$$

which implies that all modes except $n = 0$ and 1 are unstable.

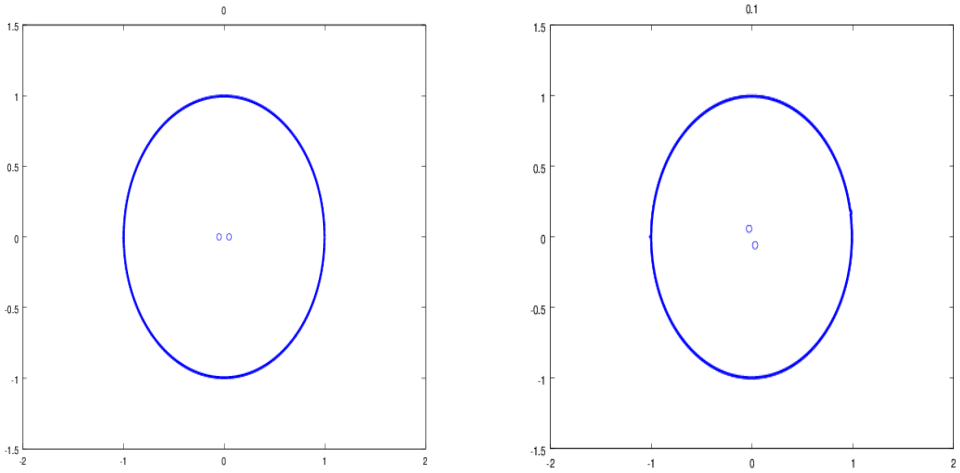
Hence, in spite of having critical velocity fields being capable of holding the whole VF in equilibrium, these states are in general unstable.

5 Interaction experiments

Here we study two types of interactions with point vortices and VF-VF interaction.

5.1 Interaction with point vortices

The original VF equation can be modified to include the effects of point vortices. Here we place two point vortices symmetrically inside a circular VF. Initially the point vortices just rotate about the center of the circular VF. Due to numerical instabilities the center of point vortices slightly shifts from the center of the circular VF. Once this happens, the point vortices destabilizes the VF by increasing thickness (hence vorticity) in certain regions at the expense of other regions and the region then rolls up as shown in the figures above.



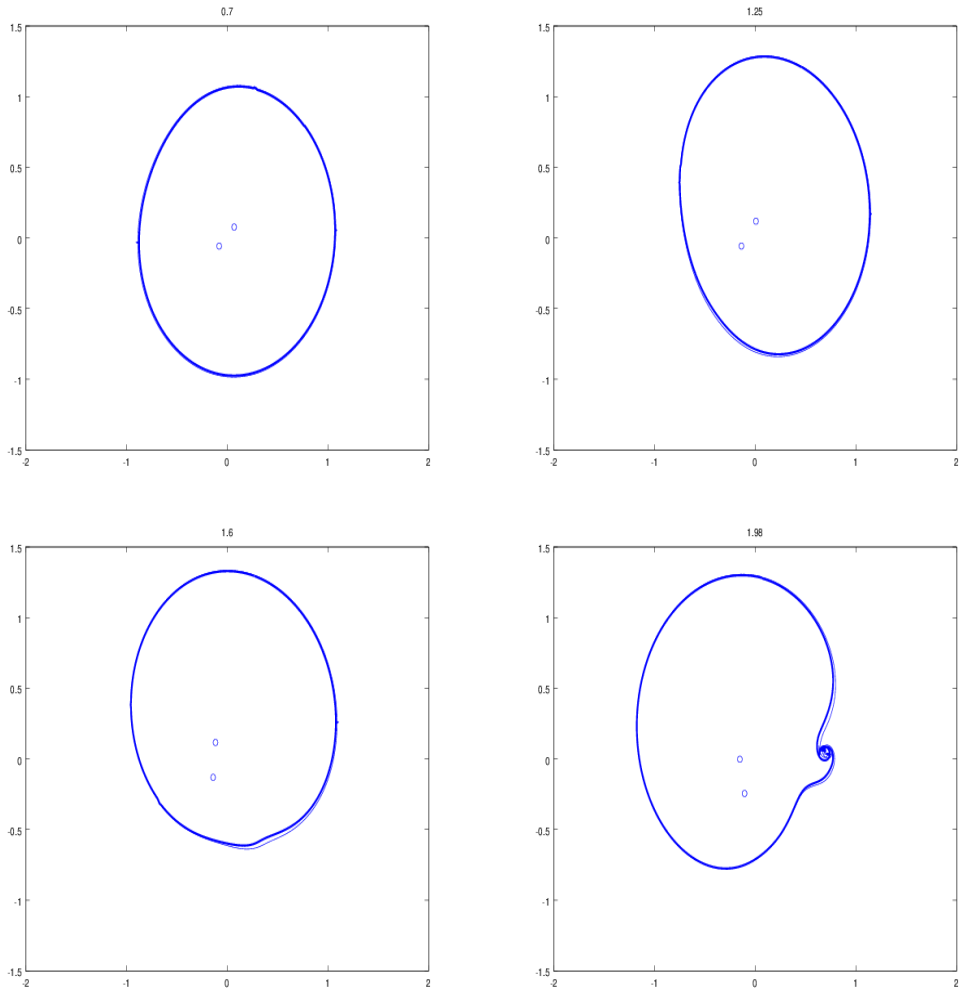
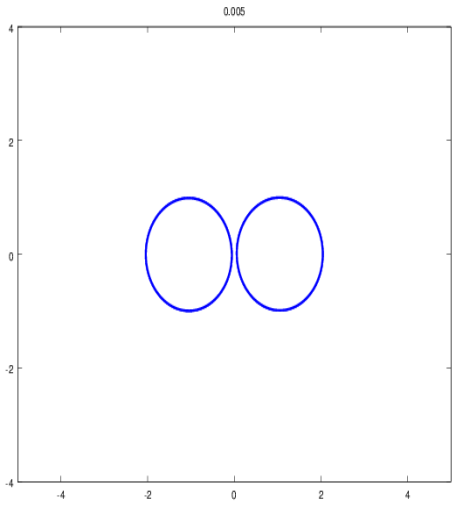
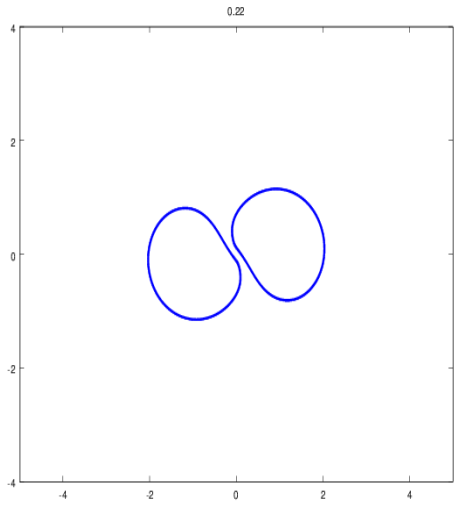


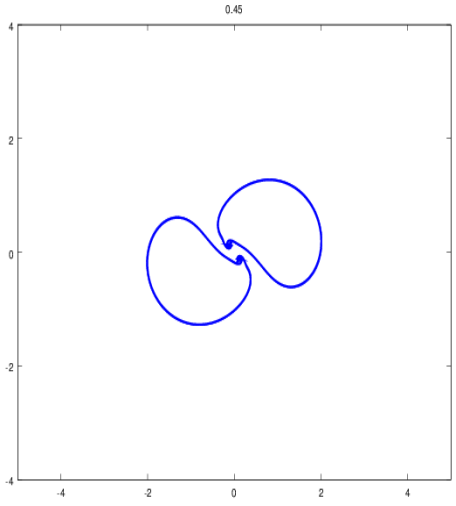
Figure 14: Interaction between circular VF and 2 point vortices

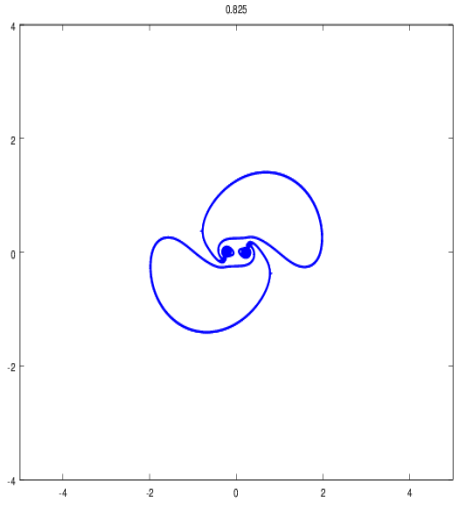
5.2 Interaction between VFs

The interaction between two circular VFs is investigated here. Two circular VFs of same strength and thickness are kept at a very small separation. The resulting interaction is shown above. Each VF rotates around the other, shears each other and sweeps thickness to a certain region which then starts to roll up. The spirals of roll up then merge together to form a complex pattern at the center. Rolling up is also initiated at other regions on the VFs.


0.005
4
2
0
-2
-4
-4
-2
0
2
4


0.22
4
2
0
-2
-4
-4
-2
0
2
4


0.45
4
2
0
-2
-4
-4
-2
0
2
4


0.825
4
2
0
-2
-4
-4
-2
0
2
4

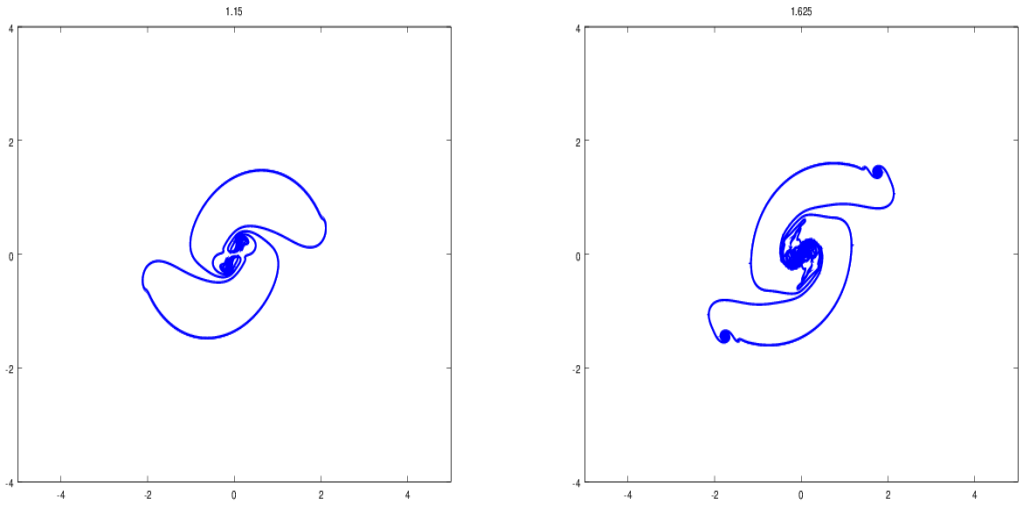


Figure 15: Interaction between 2 circular VFs

6 Conclusions

This work was undertaken with the goal of constructing a new reduction of the 2D Euler equations to model thin strips of vorticity or VFs. The new model (Filament Dynamics or FD) solves for a curve in the plane but has thickness slaved to it. Qualitative features of thin filaments such as rolling up and response to straining fields seems to be captured by the model. Comparisons with Contour Dynamics indicate similar behavior. Quantitative details and comparisons of instabilities in this model vs instabilities in the CD model remain to be examined.

CD often has difficulties capturing dynamics of very thin filaments, more so if the tips are sharp. CD tries to smooth the tips by adding in more points, taking up the role of viscosity. The number of points required also becomes larger with reduced thickness. These aspects force one to use techniques such as contour surgery[8]. As an application of the new model, one could imagine building a hybrid CD-FD code, where if the thickness of a certain strip reduces below a certain limit, the code would switch from CD to FD.

7 Acknowledgments

I would like to thank Philip Morrison for initial encouragement to pursue the idea when I first suggested this to him. His support throughout the work is gratefully acknowledged. Stephen Meacham was always available for discussions throughout the program andprovided much fruitful advice. Glenn R. Flierl is thanked for the CD code used in this work and for many helpful suggestions.

8 References

1. McWilliams, J.C., The emergence of isolated, coherent vortices in turbulent flow. *J. Fluid Mech.*, 146 (1984), 21-43.

2. Kevlahan, N.K. and Farge, M., Vorticity filaments in two-dimensional turbulence: Creation, stability and effect. *J. Fluid Mech.*, 346 (1997) 4976

3. Katz, Joseph and Plotkin, Allen, *Low-Speed Aerodynamics*, Cambridge Aerospace Series (Book 13), 2001.

4. G. Birkhoff, Helmholtz and Taylor instability , *Proc. Symp. Appl. Math.*, XII , Amer. Math. Soc. (1962) pp. 5576

5. N. Rott, Diffraction of a weak shock with vortex generation. *J. Fluid Mech.*, 1 (1956) pp. 111

6. R. Krasny, Desingularization of periodic vortex sheet roll-up. *J. Comput. Phys.*, 65 (1986) pp. 292313

7. G. Kirchhoff, *Vorlesungen uber mathematische Physik, Mechanik*, Leipzig: Teubner, 1876.

8. D. Dritschel, Contour surgery: A topological reconnection scheme for extended integrations using contour dynamics. *J. Comp. Phys.*, 77 (1988) pp. 240-266.