Numerical Linear and Nonlinear Investigation of Vertical Slot Convection

ALAN H. McALLISTER
Institute for Fusion Studies
The University of Texas at Austin
Austin, Texas 78712

April 1991

THE UNIVERSITY OF TEXAS

AUSTIN
NUMERICAL LINEAR AND NONLINEAR INVESTIGATION OF VERTICAL SLOT CONVECTION

APPROVED BY
SUPERVISORY COMMITTEE:

[Signatures]

D. S. Stein
Wendell Horton
Henry L. Carpenter
NUMERICAL LINEAR AND NONLINEAR
INVESTIGATION OF VERTICAL SLOT
CONVECTION

by

ALAN HOLLAND McALLISTER, B.S.

DISSERTATION
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

DOCTOR OF PHILOSOPHY

THE UNIVERSITY OF TEXAS AT AUSTIN
May, 1991
Copyright by
Alan Holland McAllister
1991
Dedicated to Chiwa, Madoka, Takumi, and Baba
Acknowledgments

I grateful to my advisors, Toshiki Tajima and Richard Steinolfson, for their support and guidance, and in particular for allowing me the see through a project whose end was not always clear. I would like to recognize Dr. J. Mizushima for his long work on the VSC problem, and in particular for laying out the framework of the nonlinear code on which our work is based. and his support by JIFT and the IFS DOE grant while he was here in Austin. Maurice Lebrun was a great help with programming matters and an inspiration to try for elegant and clear coding. The rest of my friends at the IFS have made working here a much richer and enjoyable experience than it might otherwise have been. I am indebted to Greg Nagao for many illuminating conversations on the more general theoretical aspects of the problem. To Jim Dibble and Buff Miner I own many thanks for helping to actually produce the thesis. The onerous task of proofreading the rough draft was professionally done by Mrs. M.H. McAllister who must get credit for all the mistakes that used to be in the text but now are not. I, of course, am to blame for those that have reappeared or been created anew since she finished. My apologies to the reader.

On the home front I am forever indebted to my family for their support, and tolerance through all the long nights of working at school rather than being at home and to my parents who made it financially possible to be in school and raise a family on a graduate student salary.

This work was partially supported by NSF grant ATM8811128, DOE grant 80ETS53088, and NASA grants W-846, 5-846, 5-701, and W-1324.
Finally I wish to acknowledge the San Diego Super Computing Center, where most of the computational work was carried out.

ALAN HOLLAND MCALLISTER

The University of Texas at Austin

May, 1991
Abstract

A stressed hydrodynamical system undergoes a series of changes in its behavior as the control parameter (proportional to the stress) is increased. We take the specific case of a slab of fluid bounded by rigid vertical walls horizontally, but of infinite extent in height and depth, that is subject to a vertical gravitational field and a horizontal temperature gradient imposed by maintaining different uniform temperatures on the side walls. This is called the vertical slot convection (VSC) problem. This system exhibits a one dimensional buoyancy-driven convective flow at any finite temperature difference. This flow, in turn, becomes unstable as the dimensionless control parameter, the Grashof number, is increased.

We derive 2-D vorticity/stream function equations in the Boussinesq approximation (eliminating pressure waves, but allowing buoyancy forces). We linearize these and perform a thorough linear stability analysis, obtaining accurate eigenvalues, neutral stability curves, and eigenfunctions for the VSC problem and several related and limiting cases, including the Benard problem. This demonstrate the role of each term in the equations.

We constructed a fully nonlinear pseudo-spectral semi-implicit fluid simulation code using a Fourier-Chebychev decomposition and working at a Prandtl number of 7.5 (water), we use the linear eigenfunctions at the linearly critical wave number, $\alpha = 1.383$, as perturbations to the 1-D basic shear flow. Above the critical Grashof number, $G_c$, the system becomes unstable to the onset of secondary vortices. The Landau expansion theory is applied
and checked in the near supercritical regime. At higher Grashof numbers a second bifurcation (Hopf) leading to oscillatory behaviors is found. Above $G_h$ the oscillations become anharmonic and evolve into nonlinear thermal relaxation oscillations which are analyzed in detail. These show similarities to the well-known 'sawtooth' oscillations found in tokamak discharge.

After a probable third bifurcation we find an apparently chaotic attractor that resembles the well-known Lorenz attractor. Preliminary results are presented to support this interpretation. Finally we examine extensions to other parameter regimes and consider future extensions and applications to plasma physics in solar and astrophysical problems.
Table of Contents

Acknowledgments

Abstract

Table of Contents

List of Figures

1. Introduction
   1.1 A Brief Review of Previous Work on the VSC Problem
   1.2 Outline of the Dissertation

2. The Model
   2.1 General Equations
   2.2 The Boussinesq Approximation
   2.3 The Normalizations
   2.4 The Vorticity/Stream Function Formulation
   2.5 The Boundary Conditions
   2.6 Comparison With The Benard Case
   2.7 The Numerical Methodology

3. The Hierarchy
   3.1 The Laplacian
   3.2 The Navier-Stokes Equation

ix
3.3 The $Pr = 0$ Limit ....................................................... 27
3.4 The Full Vertical Slot Problem ........................................ 30
3.5 The Benard Problem ...................................................... 33
3.6 Summary ................................................................. 34

4. The Linear Problem ...................................................... 36
   4.1 A General Discussion of Linear Stability Analysis ................ 36
   4.2 The linearization of the Jacobian .................................. 38
   4.3 The Spectral Analysis ............................................... 39
   4.4 The Linear Hierarchy ................................................. 40
       4.4.1 The Navier-Stokes Equation .................................. 41
       4.4.2 The $Pr = 0$ Limit .............................................. 42
       4.4.3 The Zero-Velocity Initial State ................................ 43
       4.4.4 The Benard Problem ............................................. 44
       4.4.5 The Benard Configuration with Shear Flow .................. 45
       4.4.6 The VSC Problem ................................................. 45
   4.5 The Jacobian Terms for the 1-D Background ....................... 45

5. The Linear Computational Results ..................................... 47
   5.1 The General VSC Eigenvalue Structure ............................ 47
   5.2 The Neutral Stability Curve ........................................ 50
   5.3 Critical Point Sweeps vs. the Prandtl Number .................... 55
   5.4 $\lambda(\alpha, G)$ and $\omega(\alpha, G)$ .............................. 57
       5.4.1 The High $Pr$ Complex Mode ................................. 61
   5.5 Other Cases and Discussion ........................................ 63
       5.5.1 The Navier-Stokes Case ....................................... 63
5.5.2 The \( Pr = 0 \) Limit ........................................... 64
5.5.3 The Benard Problem ........................................ 64
5.5.4 Benard with Shear Flow .................................... 67
5.5.5 The Zero-state VSC Case .................................... 70
5.6 Eigenmode zoology ............................................ 72
  5.6.1 Analytic Tools ........................................... 73
  5.6.2 The Navier-Stokes Case .................................... 78
  5.6.3 The Benard Problem ........................................ 79
  5.6.4 The Zero-state Case ........................................ 82
  5.6.5 The \( Pr=0 \) Limit ........................................ 84
  5.6.6 The VSC Problem .......................................... 84
  5.6.7 Sample VSC Modes. ...................................... 90
  5.6.8 Benard with Shear Flow ................................... 100

6. Theoretical Background to the Nonlinear Analysis .......... 101
  6.1 Notes on Bifurcation Theory ................................ 101
    6.1.1 Equilibria and Dynamics ................................ 105
  6.2 Expansion Theory ........................................... 107
    6.2.1 The Expansion for the VSC Problem ...................... 108
    6.2.2 The Energy Equation: Landau Theory .................... 111
    6.2.3 Solutions of the General Expansion ..................... 112
    6.2.4 A Note on Stable Solutions ............................. 118
    6.2.5 \( G - G_c \) Dependence ................................ 119
  6.3 A General Theoretical Context .............................. 120
    6.3.1 Notes on the Ginzburg-Landau Equation .................. 120
    6.3.2 Phase Invariance ..................................... 122
6.4 Diagnostic Tools ................................................. 124
   6.4.1 The Nusselt Number ........................................... 124
   6.4.2 The Gravitational Potential Energy ......................... 125
   6.4.3 The Energy Integrals ......................................... 125
   6.4.4 The Perturbation Energy Balance ............................ 129

7. The Initial Bifurcation and the Primary Branch .......... 130
   7.1 An Outline of the Solutions for \( G < 2500 \) at \( Pr = 7.5 \) .......... 131
      7.1.1 A Note on the Measured Quantities ......................... 133
   7.2 The Primary Branch ............................................ 133
      7.2.1 The Primary Critical Point .................................. 133
      7.2.2 Other Perturbative Tests ................................... 135
      7.2.3 The Flows Near \( G_e \) ...................................... 137
      7.2.4 The Upper Primary Branch Flows ............................. 141
      7.2.5 Basic Landau Theory ...................................... 142
      7.2.6 Extending the Landau Theory ................................. 147
      7.2.7 The Expansion Coefficients ................................. 149
      7.2.8 The Perturbation Dynamics ................................. 152
      7.2.9 The Upper Primary Branch Dynamics ......................... 158
   7.3 Fourier Mode Analysis of the Primary Branch ............ 160
      7.3.1 The Main Fourier Modes .................................... 161
      7.3.2 Expansion of the Coupled Equations ....................... 163
      7.3.3 The Mode Dynamics .......................................... 168
8. The Hopf Bifurcation and Thermal Relaxation Oscillations 173

8.1 The Approach to $G_h$ ................................. 173
8.2 The Second Bifurcation at $G_h$ .......................... 179
8.3 The Regular Relaxation Oscillations Above $G_h$ .............. 184
  8.3.1 Detailed analysis of the oscillations. .................. 185
  8.3.2 A Theoretical Discussion. .......................... 193
8.4 Comparison with Other Thermal-relaxation modes. .......... 196

9. Three Mode Oscillations 199

9.1 A Third Bifurcation Near $G = 2000$? ..................... 199
9.2 Detailed Analysis at $G = 2250$ .......................... 202
9.3 Phase-space Portraits ................................... 207
  9.3.1 Stream Function-Temperature Phase Plots .............. 207
  9.3.2 Stream Function Mode-Mode Phase Plots .............. 210
9.4 Oscillatory Regimes when $\alpha = 1.6$ ...................... 212
  9.4.1 The $G = 2000, \alpha = 1.6$ Case .................... 213
  9.4.2 Lorenz's Attractor .................................. 214
  9.4.3 The Attractor at $G = 1750, \alpha = 1.6$ .............. 216

10. Future Directions 220

10.1 The Spikes ........................................ 220
10.2 Variations in $\alpha$ .................................. 227
10.3 Breaking the Horizontal Symmetry ........................ 227
10.4 Other Possibilities .................................. 228
11. Experimental Summary and Comparison

11.1 A Summary of Experimental Results ........................................ 229
11.2 Suggestions for Experiment and Simulation .............................. 234
11.3 A Note on Experiment and Simulation .................................... 236

A. Theoretical Appendices ...................................................... 238

A.1 The 1-D Solution .................................................................. 238
A.2 The High G Solution ......................................................... 242
A.3 The Boussinesq Approximation .......................................... 243
  A.3.1 The Validity Criterion ................................................. 243
  A.3.2 The Physical Significance of These Conditions ............... 249
A.4 Scale Invariance ................................................................ 250
  A.4.1 Energy Scaling .......................................................... 252
  A.4.2 Global Invariance ..................................................... 252
A.5 Solutions to $\hat{A}$ Expansions ........................................... 253
  A.5.1 Linear ........................................................................ 254
  A.5.2 Quadratic .................................................................... 254
  A.5.3 Cubic ......................................................................... 255
  A.5.4 Higher Orders ............................................................ 257

B. Numerical Appendices ........................................................ 258

B.1 Approximation Theory ....................................................... 258
  B.1.1 The Minimax Criterion ................................................. 259
  B.1.2 The Least Squares Technique ....................................... 260
B.2 Fourier Series ...................................................................... 262
  B.2.1 The Fourier Transforms .............................................. 264
B.3 Chebychev Polynomials ........................................ 266
  B.3.1 As Solutions that Minimize Error ....................... 267
  B.3.2 Chebychev Series Representation ....................... 269
  B.3.3 The Discrete Representation .......................... 270
  B.3.4 The Chebychev Transforms .............................. 272
  B.3.5 Chebychev Properties .................................. 274
B.4 Convolutions .............................................. 277
B.5 Aliasing .................................................... 280

C. Computational Appendices .................................... 284
  C.1 The Linear Eigenvalue Problem ............................ 284
    C.1.1 The Two-Variable Formulation ....................... 286
  C.2 The Operators ........................................... 287
    C.2.1 The Algebraic Operators ............................ 288
    C.2.2 The Derivative Operators ........................... 291
  C.3 The Nonlinear Numerical Methods ......................... 295
    C.3.1 The Differencing Techniques ......................... 295
    C.3.2 The Coefficient Equations ........................... 297
  C.4 Accuracy Tests .......................................... 302
    C.4.1 The FFTs ........................................... 302
    C.4.2 The Algebraic Operators ............................ 306
    C.4.3 The Derivative Operators ............................ 306
    C.4.4 The Nonlinear Terms and De-aliasing ............... 314

BIBLIOGRAPHY .................................................. 318

Vita ........................................................................... xv
List of Figures

2.1 The Vertical Slot Convection Geometry ............... 11

5.1 VSC eigenvalues at $Pr = 7.5$, $\alpha = 1.383$. .......... 48
5.2 VSC eigenvalues at $Pr = 50.0$, $\alpha = 0.9$. .......... 50
5.3 The VSC neutral stability curve to $G = 10,000$, $\alpha = 5$. .... 51
5.4 VSC critical value mode-number dependence, $\alpha_c$ vs. $G_c$. ...... 53
5.5 VSC neutral stability curves with 8 to 32 modes. .......... 54
5.6 VSC stationary mode critical values vs. Prandtl number. .... 56
5.7 VSC $\lambda_{max}(G, \alpha)$ contours to $G = 10,000$ and $\alpha = 10$ at Prandtl numbers 0.001, 0.25, 0.5, 7.5, 15, and 100. The data ranges from $-.1031e+3$ to $+.1955e-1$ and the spacing between contours is .004. ................. 58
5.8 VSC $\omega(G, \alpha)$ contours to $G = 10,000$ and $\alpha = 10$ at Prandtl numbers 0.001, 0.2, 0.25, 0.5, 2.0, and 7.5. The data range from 0 to .5902 in the first four frames and to .6336 and .6367 in the fifth and sixth respectively and the contour spacing is .02. .... 60
5.9 VSC complex mode critical values vs. Prandtl number. .... 62
5.10 The $\lambda(G, \alpha)$ contours at $Pr = 50$ to $G = 1000$, $\alpha = 3$. The data range from -.1474e+1 to .1009e-1 in increments of .0008. .... 63
5.11 Benard eigenvalues at $Pr = 7.5$, $\alpha = 1.5585$. .......... 65
5.12 Benard critical values vs. Pr. .......................... 66

5.13 Benard $\lambda(G, \alpha)$ contours at $Pr = 10$ to $G = 100, \alpha = 10$. The
data ranges from -1.024e-2 to +.3604e-1 in increments of .004. 66

5.14 The Benard-with-shear-flow critical values. ................. 67

5.15 Benard with shear flow $\lambda(G, \alpha)$ contours. For details see the text. 68

5.16 Benard with shear flow $\omega(G, \alpha)$ contours. For details see the text. 69

5.17 The Zero-state critical values. .............................. 71

5.18 Zero-state $\lambda(G, \alpha)$ Contour at $Pr = 0.75$ to $G = 10,000, \alpha = 10$.
The data range from -1.03e+3 to .545e-2 with interval .004. 71

5.19 Sample modes of different horizontal wavelength. ............ 74

5.20 Examples of Phase, 0 and $\pi/2$. .......................... 76

5.21 A complex pair of fields. ................................. 77

5.22 A typical Navier-Stokes eigenmode. .......................... 78

5.23 The Navier-Stokes phase variation to $G = 10,000, \alpha = 10$. Data
from -3.142 to +3.142. ...................................... 79

5.24 A typical near critical Benard eigenmode. ..................... 80

5.25 The Benard low $Pr$ phase variation. ......................... 81

5.26 A sample Zero-state mode. .................................. 82

5.27 The high $Pr$ Zero-state phase structure. ..................... 83

5.28 The overall phases of the VSC modes. ......................... 85

5.29 The Internal Phase Structure of the VSC Modes. ............... 87

5.30 The $G_c$ and $\alpha_c$ curves with mode property regimes. ....... 89

xvii
5.31 A low \(Pr\), real VSC mode. ........................................ 91
5.32 A low \(Pr\), complex VSC mode. ................................. 91
5.33 The arrowhead mode. ............................................. 92
5.34 A long thin mode. .................................................. 93
5.35 A proto half wavelength mode. ................................... 93
5.36 A half wavelength mode. .......................................... 94
5.37 The thermal half wavelength mode. ............................. 95
5.38 A zero relative phase mode. ...................................... 95
5.39 The VSC linear critical mode at \(Pr=7.5\). .................... 97
5.40 A typical high \(\alpha\) high \(Pr\) complex mode. .............. 98
5.41 A typical low \(\alpha\) high \(Pr\) complex mode. ................. 98
5.42 The complex critical modes. .................................... 99

6.1 a: Supercritical b: Subcritical c: Transcritical 104
6.2 Quasi-static and dynamic evolution .............................. 106
6.3 a: Linear growth b: Quadratic growth .......................... 113
6.4 a: Cubic growth b: More cubic growth .......................... 116
6.5 a: Quartic growth b: Fifth order growth .......................... 117
6.6 Log-log plot of \(\lambda\) vs. \(G - G_c\). ............................. 119
6.7 The Chebychev mode amplitudes of the basic state at \(G = 2000, Pr = 7.5\), and \(\alpha = 1.383\) .......................... 123
6.8 The Chebychev mode amplitudes of the full flow at \(t = 8128.5\) .......................... 124
7.1 VSC bifurcation diagram including the conductive state. .... 131
7.2 VSC bifurcation diagram for $\Delta$ kinetic energy. .......... 132
7.3 The bifurcation from the 1-D State. ......................... 134
7.4 Perturbation decay at $G = 475$, with $\epsilon = 0.001$. ...... 136
7.5 The primary branch perturbations for $G = 495, 510, 600$ at $Pr = 7.5$ and $\alpha = 1.383$. ......................... 138
7.6 The primary branch flow for $G = 600$. ..................... 139
7.7 The Flow for $G = 1000$ ................................. 141
7.8 The $G_c$ bifurcation of the thermal variables. ............... 143
7.9 The Log-log plots vs. $G - G_c$. .......................... 144
7.10 The Log-log plots vs. $\lambda$. ............................. 146
7.11 $\lambda/A'_{asm}$ for the change in kinetic energy. .......... 149
7.12 $\lambda/A'_{asm}$ for the change in Nusselt number. .......... 150
7.13 $\lambda/A'_{asm}$ for the stream function. .................. 151
7.14 $\lambda/A'_{asm}$ for the temperature. ...................... 152
7.15 The perturbation time histories at $G = 492, 525, 600, 1000$. 153
7.16 The exponential approach to $A_{asm}, \gamma_i$ to $G = 600$. . 155
7.17 The initial evolution of fields and perturbations for $G = 1000$. 159
7.18 $A_{asm}$ for the $k = 0, \alpha, 2\alpha$ Modes, to $G = 1250$. .... 161
7.19 Log-log plots of $A_{asm}$ for the $k = 0, \alpha, 2\alpha$ modes to $G = 1250$. .. 162
7.20 The initial growth of the $k = 0, \alpha, 2\alpha$ modes at $G = 495$. ... 169
7.21 The Growth of the $k = 0, \alpha, 2\alpha$ Modes at $G = 495$  
7.22 The growth of the $k = 0, \alpha, 2\alpha$ modes at $G = 600$.  
7.23 The mode amplitudes at $G = 1000$.  
8.1 The log of the rate of decay to the asymptotic steady state.  
8.2 The perturbation dynamics at $G = 1360$.  
8.3 The mode dynamics at $G = 1360$.  
8.4 A detailed study of the oscillations at $G = 1360$.  
8.5 A detail of the kinetic energy and Nusselt number at $G = 1360$.  
8.6 The fields near the minima and maxima of the oscillations in the perturbation fundamental modes at $G = 1360$.  
8.7 The real frequency of oscillations around the primary branch.  
8.8 The oscillation amplitude: stream function.  
8.9 The kinetic energy bifurcation diagram above $G_h$.  
8.10 The normalized period.  
8.11 The mode histories at $G = 1500$.  
8.12 The detailed mode study of an oscillation at $G = 1500$.  
8.13 The kinetic energy and Nusselt number at $G = 1500$.  
8.14 The fields near the maxima and minima, at $G = 1500$.  
8.15 Detailed mode study of an oscillation at $G = 1750$.  
8.16 The kinetic energy and Nusselt number at $G = 1750$.  
8.17 The stream function and temperature perturbations at $G = 1750$.  

xx
8.18 Sample Field Plots at $G = 1750$ ........................................... 193
8.19 The Nusselt number and the various regimes ............................... 195

9.1 The minimum total kinetic energy. ............................................. 201
9.2 The height of the initial Nusselt number peak. ............................... 202
9.3 The kinetic energy, Nusselt number, and perturbation amplitudes at $G = 2250, \alpha = 1.383$ ............................................. 203
9.4 The diagnostics for a single oscillation at $G = 2250$ .......................... 204
9.5 The modes for a single oscillation at $G = 2250$ ............................... 205
9.6 Selected frames during an oscillation at $G = 2250$ ........................... 206
9.7 Stream function-temperature phase portraits for $G = 1000, 1360, 1400, 1750, 2000,$ and $2250$ ............................................. 208
9.8 Stream function $k = 0$ vs. $K = \alpha$ phase-space portraits at $G = 1400, 1750, 2000,$ and $2250$. ............................................. 210
9.9 The $G = 2250$ phase portrait, including the initial transient. .......... 212
9.10 The $G = 2000, \alpha = 1.6$, period doubled, phase portraits. .......... 213
9.11 The $G = 2000, \alpha = 1.6$, diagnostics. ........................................ 215
9.12 The $G = 2000, \alpha = 1.6$, mode decomposition. ............................. 215
9.13 The $G = 1750, \alpha = 1.6$, phase portraits. ..................................... 216
9.14 The potential energy at $G = 1750, \alpha = 1.6$. ................................ 217
9.15 The attractor's peak to peak plot. ........................................... 218
9.16 The the other peak to peak plots. ........................................... 219

xxi
10.1 The kinetic energy, Nusselt number, and perturbation amplitudes at $G = 5000$, $\alpha = 1.383$. ........................................ 221

10.2 Sample field plots. .................................................. 223

10.3 Large scale mixing across the slot: the full temperature field. . . 224

10.4 Waves traveling along the walls: the full temperature field. . . 225

10.5 Waves traveling along the walls: the full stream function field. . 226

B.1 Illustration of Aliasing .............................................. 281

C.1 Error vs. number of modes for a $\sin(x) + \cos(x)$ initial field. . . 307

C.2 Error vs. cutoff at $N_X = 32$ and 64 for a $\sin(x) + \cos(x)$ initial field. .................................................. 309

C.3 Error vs. wavenumber (1-8) for a $\sin(x) + \cos(x)$ initial field. . 310

.71.0 .31.0 .20.0
Chapter 1

Introduction

The study of hydrodynamical flows for the last one hundred years has generally been formulated in terms of their stability. Starting with some configuration the fluid is put under stress, for instance, by applying pressure (Poiseuille flow) or thermal (the Benard problem) gradients or subjecting it to the differential acceleration caused by rotation. (Couette flow). As the stress is increased the fluid will exhibit a wide range of flows depending on the exact system parameters and the magnitude of the stress. These are investigated in terms of their form and their stability which indicates the circumstances under which they may be expected to appear in a physical setting. The foundations of this field were laid out more than a century ago by Helmholtz, Kelvin, Rayleigh and Reynolds.[1]

These problems are of interest for their practical applications in many types of problems in hydrodynamic and plasma physics, engineering, meteorology, oceanography, and geo- and astrophysics. When the driving stress on a fluid body in a gravitational field is a temperature gradient the result is buoyancy driven thermal convection. The classic arena for studying thermal convection is the Benard problem in which a horizontal fluid layer is heated from below[2]. In our work the slab is rotated so that we have a vertical fluid layer, bounded horizontally by rigid vertical planes, and subjected to a vertical gravitational field and a horizontal temperature gradient. This change in
geometry, which adds a basic 1-D shear-flow\(^1\) to the buoyancy of the Benard case, increases the complexity of the problem. It is this problem that we have focused on, studying it both linearly and non-linearly.

Physically there are two sorts of questions that have drawn attention in the past. The earliest experimental workers where interested in the heat transport across the slot\(^2\) while the theorists focused on the breakdown of the initial shear-flow as the temperature difference is increased.\(^3\) The question of heat transport, or conversely of thermal insulation, is at the heart of many of the applications of this problem. These include insulation in double windows and walls (cited by nearly everyone), of internal walls in the case of fires\(^6\), and between the absorber and cover plates on solar collectors\(^7\). Cooling effects are of prime importance in the cooling of nuclear fuel assemblies and shipping casks\(^6\), of high speed turbine blades by thermosyphons\(^8\), and of integrated circuits.

On the theoretical side these problems, when viewed in a general context are perhaps even more significant. Hydrodynamics, owing to a lack of strong restoring forces, generally scales in size and energy (see Section 6.3.2). As a result typical hydrodynamic modes, with the exception of phonons,\(^4\) lack characteristic frequencies or have nearly zero frequency. This is in contrast to high frequency plasma or particle physics where there are strong restoring forces and a corresponding hierarchy of structure and energies.\(^5\) The restoring forces make possible high frequency regular oscillations with characteristic frequencies

---

\(^1\)See Appendix A.1.
\(^2\)Eckert and Carlson\(^3\) cite Nusselt\(^4\) as the earliest experimental paper on this problem. They give a list of other experimental results up to the 1950's.
\(^3\)See Squire\(^5\) for a review of work to about 1930.
\(^4\)Even phonon modes are absent in the incompressible case.
\(^5\)See Tajima \(^9\) for plasma physics and Perkins for particle physics \(^10\).
(and energies) that dominate the physics at higher energies. In plasma physics, for example, the plasma frequency arises from the electrostatic restoring force, the photon frequency from Maxwell's stress force, and the gyrofrequency from the magnetic force. These oscillations are described in terms of elementary excitations[11] or 'quasi-particles', which for the plasma forces are plasmons, photons, and magnons. The interactions of physics with different characteristic frequencies and energies yields hierarchical behavior with each type influenced by and influencing the behaviors at the next higher and lower frequencies.

At the lower end of this hierarchy of frequencies, however, the restoring forces are less important and the medium more plastic so that displacements can be very large, sometimes arbitrarily so. This behavior approaches that found in hydrodynamics and the modes are closely related. It is probable that low energy particle physics also corresponds in some way with hydrodynamics (see Section 6.3). Furthermore it is plausible that these long wavelength modes exhibit greater nonlinearity than the more highly localized 'elementary excitations' of high energy physics. In plasma physics it is believed that transport properties are dominated by low frequency physics for precisely this reason.

The last decade or two has seen the rapid rise of computational nonlinear physics, allowing the study of bifurcations, nonlinear mode interaction, chaos and the onset of turbulence. This has been particularly true in the relatively uncomplicated realm of hydrodynamics where these subjects can be studied in detail by increasingly powerful computational means. The insights gained into the inherently nonlinear around critical points and phase transitions can then be applied in the low energy limits of other fields that are mathematically more complex. For instance, many aspects of 2-D convective flow
problems have direct parallels in plasma physics. In other areas the parallels are less direct but of potentially great interest nonetheless.

Our interest in the VSC problem is therefore two fold. Most immediately it is to understand more thoroughly the physics of this specific hydrodynamic problem, but more generally this is a relatively simple system in which to study general features of nonlinear physics that may be of use in many other fields as well.

1.1 A Brief Review of Previous Work on the VSC Problem

Most of the work into the 50's seems to have been experimental, but after the mid sixties theoretical interest picked up and quite a bit of work was done with various expansions of the basic equations. While there have continued to be theoretical and a few experimental papers in the last two decades there has been a noticeable increase of computational work in the 70's and 80's as computational power has increased and as the more accessible linear questions have been generally sorted out.

In examining the literature of the problem there are several parameters that usefully classify the various investigations. We shall first sketch these and the general regimes of each, followed by a quick outline of references. Much of this body of material has been more specifically referenced throughout the discussions of our current work and so there is no need for a detailed review here.

---

*The 2-D Navier-Stokes equation for the stream function is mathematically the same as the 2-D ExB magnetized plasma equation for the electric potential.*[12]
The ratio of the height over the width of a cavity is called the aspect ratio, $h$. In treating the vertical slot problem there are several distinct regimes that produce different physics and tend to be favored by different methods of investigation. Much of the theoretical work has been done in infinite aspect ratio slots, allowing exact solutions of the basic shear-flow.\footnote{Chait and Korpela\cite{13} give a nice discussion of the advantages and disadvantages of taking $h$ infinite in a simulation, including the need to fix the wavenumber.} The modern theoretical literature seems to date back to a paper by Batchelor \cite{14} and was then taken up in the mid 60's, \cite{15}\cite{16} \cite{17}, and the early 70's, \cite{18}\cite{19}\cite{20} \cite{21}, making use of various limits to expand the equations and obtain the linear stability characteristics of the basic 1-D background shear-flow. By the late 70's, with more powerful computers the linear results were more exact\cite{22} and began to go beyond the linear\cite{23}. Nagata and Buss\cite{24} who concentrated on 3-D perturbations of the 2-D secondary flow were part of a move towards a study characteristics of the secondary flows in greater detail \cite{25}\cite{26}\cite{13}, while Gotoh, Mizushima, and Fujimura have begun investigations of the dynamic role played by mode-mode interactions with a small number of harmonics added to the fundamental frequency \cite{27}\cite{28}\cite{29}.

With infinite $h$ and medium to high $G$, there is a second analytic solution to the equations based on assuming a constant vertical temperature gradient in the center of the slot. This case has been studied to determine the dependence of the secondary flows on the magnitude of the gradient\cite{30} \cite{31}\cite{32}\cite{33}.

On the other hand this value of $h$ is impossible in the lab and most experimental work has been done for $h = 10 - 60$. It has been found that the end affects of the slots do not propegate very far (Elder\cite{34} estimated not
more than 10% at either end was affected in slots with \( h = 10 - 60 \), so that the results of infinite \( h \) theoretical and numerical studies are believed to apply to much of the height of a tall thin slot.\(^8\) This approximation certainly seems not to hold for \( h < 10 - 12 \) where the basic flow ceases to show the secondary vortices found at larger \( h \). Finally, simulations tend to be for low \( h \) that match the experimental results, and not infrequently taking \( h = 1 \) which is in some senses a different problem [8][37][38][39][40].

The experimental work, and to some extent the theoretical and numerical work depends on the type of fluid in use, as reflected in the value of the Prandtl number, \( Pr \). Eckert and Carlson[3] used air (\( Pr \sim .7 \)), as did [41] in a small \( h \) slot, and [36] and [7] (in slots with variable slopes). Elder used water (\( Pr \sim 7 \)) in his high \( G \) experiments[42], and oil (\( Pr \sim 1000 \)) in his low \( G \) work [34]. Seki et.al.[43] also made use of both water and heavy oils (at \( Pr =480 \) and 12,500) and Vest and Arpac[44] used oils at \( Pr =25 \) and 1000. The theoreticians also on occasion make use of the limit \( Pr = 0 \) to decouple the heat equation from the equation of motion[45] [28].

Numerical work on the vertical slot seems to date from the late 60's. MacGregor and Emery[41] give a good review of the early work in their paper on low aspect ratio \( h = 1,10 \) experimental and simulation results. Low \( h \) results were also reported by [46](\( h = 5 \)), [47](\( h = 10 \)), and [6](\( h = 1 - 15 \)). Medium \( h \) numerical and simulation results include some of those in [6] and those in [48] and [49](\( h = 15 \)). Raithby and Wong [35] give results for \( h \) up to 80, but these high \( h \) results are computationally expensive. Little 3-D work has yet been done for the same reason, although 3-D studies have been done by

\(^{8}\)The question of when the end effects are negligible has also been addressed in [35] and [36].
Hart[50] (in a sloping slot) and Nagata and Busse[24]. Mallinson and de Vahl Davis have done 3-D computations of fluid element movement in small aspect boxes, e.g.[8].

In relevant extensions to the basic vertical slot problem, the non-Boussinesq problem has been looked at by Chenoweth and Paolucci [51] [52], while Choi and Korpela[53], and Lauriat and Desrayand [54] have broken the horizontal mid-plane symmetry with interesting results. This was done by using an annulus, adding differential curvature at the boundaries, and by adding radiation at the boundaries which can be asymmetric.

Finally, although there do not seem to be any general review articles on this subject, we can mention a few papers that do include review and discussion of previous work. Early reviews which lean to experimental work are found in Eckert and Carlson(1961)[3], Elder(1965)[34], Vest and Arpaci(1969)[44], and MacGregor and Emery(1969)[41]. Hollands and Konicek(1973)[36] look at previous work in sloped slots, Pepper and Harris(1977)[6] cover earlier computational work and Raithby and Wong(1981)[35] treat finite h simulations. Good general reviews can be found in Bergholtz(1978)[32], Patterson and Imberger(1980) [38] who organize their's by h value, and Lee and Korpela(1983) [49]. Lastly Chenoweth and Paolucci(1985)[51] list earlier non-Boussinesq studies.

1.2 Outline of the Dissertation

In our work we have treated the infinite aspect ratio vertical slot problem, in The Boussinesq approximation, using a periodic Fourier decomposition.

---

*We shall comment further on these results in Section 10.3.*
vertically and a Chebychev decomposition horizontally. The specific formulation of the problem is outlined in Chapter 2, with reference to the appendices for the details. In Chapter 3 we discuss the physics involved by building the full equations up piece by piece. We then formulate the linear problem for the vertical slot problem with 1-D shear-flow and several related and limiting cases in Chapter 3. The results of a thorough investigation of the linear problem are presented in Chapter 4, including comparison with previous results. Special attention is paid to the linear eigenmodes which are classified and presented for all reasonable values of $G$, $Pr$, and $\alpha$ (the perturbing wave number).

Armed with some insight, and with the linear eigenmodes and growth rates we turn to study the dynamic growth of two dimensional perturbation of the basic one dimensional shear flow. In Chapter 6 we present some theoretical background, including elementary bifurcation theory and application of Landau expansion theory to the vertical slot problem. We point out the formal similarities to the Ginzburg-Landau Equation and note some of its many uses in treating phase transitions and critical phenomena in other branches of physics.

Choosing to focus on Prandtl number $Pr = 7.5$, we use a nonlinear pseudo-spectral computer code to investigate the secondary flows reached by the growing linear eigenmodes (using the critical wave number $\alpha = 1.383$) and explore the limits of the Landau expansions. We look at both the secondary stable states and the dynamical evolution by which they are reached. These results are presented in Chapter 7. Moving beyond the range of expansion theory, $G > 1000$, we investigate further bifurcations and the appearance of nonlinear thermal relaxation oscillatory states that appear to be temporally chaotic for some values of $G$ and $\alpha$. These oscillations are studied in detail.
using mode decompositions and preliminary comparison is made to similar phenomena in plasma physics and astrophysics.

Finally in Chapter 10 we discuss some preliminary results at high $G$, at lower $Pr$, and at lower $\alpha$ where the fundamental becomes unstable to perturbation by its first harmonic. In conclusion we indicate various promising directions for continued investigation, improvement of the codes, and possible extensions of the physical problem.

The appendices include many mathematical and computational details that are not really appropriate to the main discussion, but are integral to the actual implementation and interpretation of the analytical and computational analysis.
Chapter 2

The Model

We examine a fluid slab unbounded in two directions but with rigid thermally conducting boundaries in the third; a fluid slab between two walls. We fix different uniform temperatures on each wall, which provides the driving stress on the system. This configuration is placed in a uniform gravitational field which induces a uniform pressure gradient.

In the classic special case, the Benard problem, the walls are horizontal, the fluid is heated from below, and the gravitational field is perpendicular to the walls and pressure gradient is anti-parallel to the driving temperature gradient. We have focused, however, on the special case in which the gravitational field is parallel to vertical walls and the pressure gradient is perpendicular to the temperature gradient. This Vertical Slot Convection (VSC) problem is illustrated in Figure 2.1.

In both cases the cold wall, at temperature $T_1$, is at $x = -L$ and the hot wall, at temperature $T_2$, is at $x = +L$. In the VSC case the gravitational force is in the negative $z$ direction, $g = (0, 0, -1)$. These two cases have nearly identical equations but there is a difference in the symmetry that produces quite different physics.
2.1 General Equations

We start from the general hydrodynamic equations\(^1\) expressing respectively the conservation of density, momentum, and energy. These are specifically;

The continuity equation:
\[ \frac{D}{Dt} \rho = -\rho \nabla \cdot \mathbf{u} \quad (2.1) \]

where \( \rho \) is the density, \( \mathbf{u} \) is the fluid velocity, and \( \frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla \) is the convective derivative.

The equations of motion:
\[ \frac{D}{Dt} (\rho \mathbf{u}) = -\nabla P + \mu \nabla^2 \mathbf{u} + \rho \mathbf{g} \quad (2.2) \]

where \( P \) is the isotropic pressure and \( \mu \) is the coefficient of kinematic viscosity.

\(^1\)See [2] or [55].
The heat conduction equation:

\[
\frac{D}{Dt}(\rho C_p T) = \nabla \cdot (k \nabla T) - P \nabla \cdot u + \frac{\mu}{2} (\partial_i u_j + \partial_j u_i)^2
\]  \hspace{1cm} (2.3)

where $T$ is the temperature, $k$ is the coefficient of heat conduction and $C_p$ is the specific heat at constant pressure. We will also need an equation of state.

### 2.2 The Boussinesq Approximation

Starting from the above general equations we apply the Boussinesq approximation\(^2\):

*All variations of the thermodynamic parameters ($\mu$, $\alpha$, $k$, $C_p$) are ignored. The pressure driven variations in the density are ignored everywhere while the thermally driven variations are retained only in the buoyancy force term in the equation of motion (2.2).*

First we take $\rho = \rho_o + \rho_h + \rho'$ and $P = P_o + P_h + P'$, where $o$ indicates a spatially constant base value, $h$ indicates the hydrostatic part, and $'$ the temporally varying part.

In the incompressible case with $\rho = \rho_o$ the equation of continuity (2.1) reduces to:

\[
\nabla \cdot u = 0,
\]  \hspace{1cm} (2.4)

Introducing the gravitational potential $\Phi$ such that $g = -\nabla \Phi$, $\rho_o g = -\nabla \rho \Phi$ and we can absorb it into the pressure gradient term, where it cancels $P_h$. We

---

\(^2\)See Appendix A.3 for details of this process and the criterion governing the validity of the approximation.
also assume that $\rho_h g$, the pressure driven density variation, is a negligible term. The equation of motion (2.2) becomes the Navier-Stokes Equation

$$\rho_0 \frac{D}{Dt} u = -\nabla P' + \mu \nabla^2 u + \rho' g$$

(2.5)

Using Equation (2.4) to eliminate the compression term in Equation (2.3), and ignoring the viscous dissipation term, the equation of heat conduction (2.3) reduces to

$$\rho_o C_p \frac{D}{Dt} T'' = k \nabla^2 T'$$

(2.6)

where $T' \equiv T - T_o = T - (T_1 + T_2)/2$ is the deviation from the mean base temperature $T_o$. The appropriate equation of state is

$$\rho = \rho_o (1 - \alpha T'')$$

(2.7)

where $\alpha$ is the volume expansion coefficient and we continue to ignore all pressure driven density variations.

2.3 The Normalizations

We have already performed an overall energy normalization\(^3\) by letting

$$T \rightarrow T' = T_{tot} - T_o$$

(2.8)

We also define $\Delta T \equiv (T_2 - T_1)/2$, such that the temperature gradient is $\Delta T/L$.

In contrast to the Benard configuration, the VSC configuration does not have a ground state conductive solution. For any $\Delta T > 0$ its ground state is a single convective cell with fluid rising along the hot wall and falling along

---

\(^3\)For more detail and a discussion of the ramifications of the normalizations outlined here see Appendix A.4.
the cold wall. For a slot with an infinite aspect ratio there is a 1-D steady-state solution\(^4\) describing this flow. The vertical velocity of this flow, resulting from the balancing of the viscous shear force and the buoyancy force, serves as an appropriate reference velocity:

\[ W_o = \alpha g L^2 \Delta T / \nu \quad \text{where } \nu = \mu / \rho_o. \]

We renormalize our variables to make them dimensionless using \( W_o, \Delta T, L, \) and \( \rho_o \):

\[
\begin{align*}
\tilde{x} &= \frac{x}{L} \\
\tilde{T} &= T' \frac{1}{\Delta T} \\
\tilde{u} &= u \frac{1}{W_o} \\
\tilde{t} &= t \frac{W_o}{L} \\
\tilde{P} &= P \left( \frac{1}{\rho_o W_o^2} \right). 
\end{align*}
\]

Using these normalizations our equations are rewritten in terms of two dimensionless parameters:

The Grashof number: \( G \sim \alpha g L^3 \Delta T / \nu^2 = W_o L / \nu \)

The Prandtl number: \( Pr = \nu / \kappa \) \quad \text{where } \kappa = k / C_p \rho_o

The primary definition of the Grashof number, \( G \), shows explicitly the dependence on the driving temperature gradient and is related to the Rayleigh number by \( Ra = G Pr \). The definition of \( G \) in terms of \( W_o \) shows it to be a Reynold’s number, and is more meaningful at small Prandtl numbers when the thermal equation decouples from the equation of motion.\(^5\)

The Prandtl number, \( Pr \), is the ratio of the relative responsiveness

---

\(^4\)For a detail derivation of this solution see Appendix A.1.

\(^5\)See Section 3.3.
of the fluid to kinetic and thermal stresses. As we will see below it is a key parameter that couples the thermal physics into the kinetic flows.

Using these definitions and henceforth dispensing with the tildes, we have the normalized equations: \(^6\)

\[
\nabla \cdot \mathbf{u} = 0
\]

\[
\frac{D}{Dt} \mathbf{u} = -\nabla P + \frac{1}{G} \nabla^2 \mathbf{u} + \frac{1}{G} T \hat{z}
\]

\[
\frac{D}{Dt} T = \frac{1}{PrG} \nabla^2 T
\]

2.4 The Vorticity/Stream Function Formulation

It is well known that for shearflow between parallel walls the initial instabilities are two dimensional.\(^7\) Confining ourselves to two dimensions, and assuming an incompressible fluid, we can recast the equation of motion in terms of a stream function and a vorticity. To do so we confine our investigation to the \(x-z\) plane and take the \(y\) component of the curl of Equation (2.10):

\[
\frac{\partial}{\partial t} \omega + \frac{\partial (\omega, \psi)}{\partial (x, z)} = -\frac{1}{G} \frac{\partial T}{\partial x} + \frac{1}{G} \nabla^2 \omega
\]

where the stream function \(\psi\) is defined by

\[
\nu_x = \frac{\partial \psi}{\partial z} \quad \nu_z = -\frac{\partial \psi}{\partial x},
\]

the vorticity \(\omega\) by

\[
\omega = \nabla \times \mathbf{u}
\]

\(^6\)This is now explicitly for the VSC configuration. For the Benard problem replace \(\hat{z}\) with \(-\hat{z}\). See also Sections 2.6 and 3.5

\(^7\)This result was proved by Squire [5].
and the general form of the Jacobian is

$$\frac{\partial(\phi, \psi)}{\partial(x, z)} = \frac{\partial\phi}{\partial x} \frac{\partial\psi}{\partial z} - \frac{\partial\phi}{\partial z} \frac{\partial\psi}{\partial x}.$$  \hspace{1cm} (2.15)

The Jacobian derives from using definition (2.13) to substitute for \( u \) in the advective part of the total derivative \( \frac{D}{Dt} \) of Equation (2.11).

Similarly we can rewrite the equation of heat conduction as:

$$\frac{\partial}{\partial t} T + \frac{\partial(T, \psi)}{\partial(x, z)} - \frac{1}{PrG} \nabla^2 T = 0 \hspace{1cm} (2.16)$$

We note that the continuity equation (2.4) is now identically satisfied but has in practice\(^8\) been replaced by the relationship between \( \omega \) and \( \psi \) which can be derived by substituting definition (2.13) into Equation (2.14) to give

$$\omega = \left( \frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial z^2} \right) \psi.$$  \hspace{1cm} (2.17)

Thus our model will consist of Equations (2.12), (2.16), and (2.17) dependent on the definitions (2.13), (2.14), and (2.15).

2.5 The Boundary Conditions

Physically we are modeling a fluid between two vertical walls. These are taken to be fixed rigid boundaries where the parallel velocity satisfies a no-slip condition. Each wall is held at a uniform fixed temperature, with one wall hotter than the other, and they are thermally conductive. This has the effect of placing the fluid between to infinite heat baths at somewhat different temperatures.

\(^8\)This holds as long as we use both \( \omega \) and \( \psi \) as variables. We could also use Relation (2.17) to eliminate \( \omega \) and write Equation (2.12) solely in terms of \( \psi \). The reasons for not doing so are largely numerical. See Appendices C.1 and C.3.
In the normalized units these conditions are maintained when

\[ T = \pm 1 \quad \text{at} \ x = \pm 1 \quad \text{for all} \ z, \]

\[ \psi = \text{a constant} \quad \text{at} \ x = \pm 1 \quad \text{for all} \ z, \]

\[ \partial_x \psi = 0 \quad \text{at} \ x = \pm 1 \quad \text{for all} \ z. \]

The constant value for \( \psi \) is assumed to be zero, and Definition (2.13) guarantees that the conditions on the velocity hold. We note that this provides the 6 conditions sufficient to specify the solutions to Equations (2.12), (2.16), and (2.17) which are all of second order.⁹

Vertically we treat the infinite aspect ratio case which applies in the limit of very tall, thin, slots, and to the central regions of shorter slots,¹⁰ whenever end effects can be ignored. The top and bottom boundaries are taken to be periodic. This condition is automatically maintained by the vertical Fourier decomposition discussed in Appendix C.3.

2.6 Comparison With The Benard Case

Although the equations are very similar, the physics of the Benard problem differs significantly from that of the VSC case. When the \( \frac{\partial T}{\partial z} \) term in Equation (2.12) is replaced by \( \frac{\partial T}{\partial \zeta} \) then we have the Benard Equations. Otherwise the sets of equations are the same. This substitution reflects the change in the direction of \( g \) from \( -\zeta \) to \( \zeta \), affecting the symmetry of the problem and of its solutions.

⁹Newell and Schmidt [48] lay these out clearly.
¹⁰See discussion in Chapter 1.
2.7 The Numerical Methodology

In solving the equations for an incompressible fluid one may treat the solenoidal condition, Equation (2.4), either implicitly or explicitly. It is a common approach\(^\text{11}\) to take the divergence of the Navier-Stokes Equation (2.5) and, making use of Equation (2.4) produce a Poisson Equation for the pressure:

\[
\nabla^2 P = \nu \nabla \cdot \nabla^2 u \tag{2.18}
\]

This equation contains the solenoidal condition implicitly.

Starting with an initial velocity field the pressure is solved for and then used along with the same velocity in the Equation (2.5) for the new velocity. In the thermal equation one would likewise pick an initial temperature field and then use this and the velocity at each time step to calculate the temperature at the next timestep.

This is a naturally explicit technique, the variables at each timestep being calculated using the values at the previous timestep. It has, however, both physical and numerical drawbacks. The boundary conditions needed are those on the pressure and the initial temperature and velocity fields. Not only are these somewhat unphysical but the pressure boundary conditions are overspecified, forcing us to choose either Neumann or Dirichlet conditions which in general give different solutions. Furthermore, if we try to take advantage of the numerical accuracy of spectral methods we run into convergence problems and fail to obtain useful results. This is related to the failure to satisfy both sets of boundary conditions at the same time.\(^\text{12}\) Lastly explicit solution of the

\(^{11}\)See discussion in Moin and Kim [56].
\(^{12}\)This discussion follows [56]. See also Chenoweth & Paolucci [52] and Kublbeck et.al. [39] on complexities of using the Poisson Equation for the Boussinesq VSC problem.
diffusion leads to a prohibitively small time step.\textsuperscript{13}

In order to avoid these problems we follow a method outlined for us by J. Mizushima,\textsuperscript{14} based on that given by Moin & Kim\textsuperscript{[56]}, in which we solve the continuity equation explicitly (in two dimensions it is replaced by Equation (2.17)) and avoid the Poisson equation for the pressure. We then use Fourier-Chebychev spectral expansions\textsuperscript{15} to solve implicitly for the field variables at the next time step by writing a series of linear equations for the expansion coefficients. The nonlinear terms, which can not be handled this way, are evaluated using an explicit method (but we can still do the necessary derivatives spectrally).\textsuperscript{16}

Using this method we can specify boundary values for the velocity and the temperature, which is physically reasonable, with no need to specify pressure conditions on the walls. Numerically we avoid the special handling for boundary point calculations often needed in finite difference treatments (e.g. [48]) and are able to take full advantage of the speed and accuracy of using a spectral representation.

Lastly we should note that although in two dimensions the pressure is generally removed from the equations for any numerical treatment, Moin and Kim's method has the advantages of being more naturally suited to spectral representation and is fairly easily extendible to three dimensions, in which case the pressure is cannot be eliminated, but in this treatment it enters as just one more variable, being defined by the equations of motion without the need for

\textsuperscript{13}e Quere and de Roquefort [40] give a good review of these issues.
\textsuperscript{14}Personal communication.
\textsuperscript{15}A technique originated by S.A. Orszag in the early 1970's. For a thorough reference see [57].
\textsuperscript{16}For a complete description of the numerical techniques used see Appendix C.3.
a separate Poissonian.
Chapter 3

The Hierarchy

One way to understand the structure of the VSC equations is to build them up term by term, creating a hierarchy of equations that gradually include more and more physics. This is a useful exercise in that it leads to a clearer understanding of the physics of the full equations and at the same time displays most of the interesting special cases. Further this hierarchy of equations serves as a good framework for organizing the results of our linear investigations.

The following equations can be studied in two different ways. One is to look at the types of behavior they can describe; i.e., what physics is included, and what states they produce under time evolution. This means the approach to various equilibria and steady-state flows from either non-equilibrium flows or unstable equilibria that have been perturbed.¹

The second is to linearize the equations and then use them to perform stability analyses. This allows us to examine the stability of initial background base flows, which may or may not be solutions of the full equations, such as using a shear flow background in the Benard equations even though such a flow is not supported by the equations themselves except as a possible boundary condition. In other words we can expand around non-equilibrium as well as equilibrium states and must therefore pay careful attention to what we are

¹See Section 6.1.1 for a more detailed discussion of the ways in which we can use the full code.
doing in order to avoid confusion.

In this section we follow the former approach and take up the latter in Section 4 when we turn specifically to the linear problem and its formulation. It is useful to keep in mind that the only terms in Equations (2.12) and (2.16) that will change under linearization are the Jacobians.

3.1 The Laplacian

We start building the equation of motion (2.12) with the simple Laplacian term:

$$\frac{\partial \omega}{\partial t} = \frac{1}{G} \nabla^2 \omega$$  \hspace{1cm} (3.1)

This can be solved by assuming an exponential form for $\omega$:

$$\omega \propto e^{(kz+lx)+\lambda t}$$  \hspace{1cm} (3.2)

which upon substitution in Equation (3.1) yields

$$\lambda = -\frac{1}{G} (k^2 + l^2)$$  \hspace{1cm} (3.3)

The eigenvalue $\lambda$ is always less than zero for real $k$ and $l$ so that the steady-state solution is always stable and all perturbations will decay in time. The steady-state solution depends on the boundary conditions for $\omega$. In our formulation these boundary conditions are determined by Relation (2.14) from the boundary conditions for $\psi$. For the boundary conditions given in Section 2.5 the steady state solution is $\omega = 0$ throughout the fluid.

Physically the Laplacian term describes viscous dissipation through which an arbitrary flow relaxes, in the absence of additional driving terms, to
a uniformly motionless fluid.\(^2\)

By Fourier expanding \(\omega\) in \(z\) to derive an evolution equation for each mode,

\[
\frac{\partial \omega_k}{\partial t} = \frac{k^2}{G} \nabla^2 \omega_k
\]

(3.4)

we see that this term damps high frequency modes more heavily than low frequency modes.\(^3\) The Laplacian's general character to damp a field to an equilibrium with a minimum of curvature (a field has curvature when derivatives of second order or greater are non-zero) is reflected in this greater damping of higher frequencies, which may be considered to have greater curvature.\(^4\) This characteristic plays a key part in the oscillations we discuss in Chapter 8.

Note also that, in the limit \(Pr \to 0\) and \(G \to 0\), the VSC problem reduces to two uncoupled Laplacians, one for the vorticity and one for the temperature. This represents the case of an undriven slot filled with a fluid of Prandtl number approaching zero, which exhibits pure damping of all perturbations.

### 3.2 The Navier-Stokes Equation

If we add the nonlinear terms to Equation (3.1) we get a purely kinetic (no outside forces) vorticity-stream function version of the Navier-Stokes

---

\(^2\)This will not be true if one enforces boundary conditions such that \(\omega\) varies along the boundary.

\(^3\)This attenuation of higher frequencies is an important property that we shall come back to when discussing the topological results of using an insufficient number of modes at higher Grashof number.

\(^4\)Graphically this is straightforward. A zero curvature field graphs, in cross-section, as a straight line.
Equation (2.5);
\[
\frac{\partial \omega}{\partial t} = -\frac{\partial (\omega, \psi)}{\partial (x, z)} + \frac{1}{G} \nabla^2 \omega
\]  

(3.5)

These new terms derive from the advective part of the total derivative \( \frac{D}{Dt} \) and take the form of a Jacobian in \( \psi \) and \( \omega \) which couples existing modes together to produce new ones, thus making possible the evolution of new forms of motion in the fluid. As we will discuss in greater detail in Section 3.4, the Grashof number is a measure of the related importance of these two terms and in the absence of the heat conduction equation is a measure of the nonlinearity of the system.\(^5\)

Just as there are two different, though related, types of modes, physical modes and Fourier or Chebychev spectral modes, there will be two different, though related, types of coupling. The first is the nonlinear coupling of various the Fourier-Chebychev frequencies that make up the mode space representation of each of these physical modes. We take the \( x \) and \( z \) dependence of two fields \( \psi \) and \( \phi \) as in Equation (3.2) and write each Fourier-Chebychev mode as \( \phi_{lm} \delta(k - l)\delta(k - m) \) etc., where \( \phi_{lm} \) is an amplitude and the \( \delta \)s fix the wave numbers. Making use of the results of Appendix B.5 we have,

\[
\partial_x \phi_{lm} \partial_z \psi_{lm}^* - \partial_x \phi_{lm} \partial_z \psi_{lm}^* = (lm' - l'm) \phi_{lm} \psi_{lm}^* \\
[\delta(k - l - l') + \delta(k - l + l')] [\delta(k - m - m') + \delta(k - m + m')] 
\]  

(3.6)

Here we can see explicitly the frequency mode coupling induced by the nonlinear term which as modes grow or the order parameter increases beyond critical will modify the linear eigenmodes creating new nonlinear modes.

\(^{5}\)This holds in the kinetic equation and in part in the heat conduction equation, however, in the latter the role of the Prandtl number as arbitrator between the convection and conduction terms is also significant. See discussions in Sections 3.3 and 3.4.
When we consider frequency mode coupling, we see that there will be an asymmetrical propagation of energy in the direction opposite to the gradient of the mode amplitudes with respect to the frequency $\frac{\partial \phi}{\partial k}$. In most instances this leads to a propagation of energy up the frequency spectrum. This is especially true for initial states that are described mainly by a few low $k$ modes, while the viscous damping term drains energy from the high $k$ modes faster than the small $k$ modes. Since the Jacobian actually couples the derivatives of the fields rather than the fields themselves, the zero frequency mode, which represents the constant part of the real field, will not couple to itself to generate other higher frequency modes. This means there is a lower bound on the mode frequency and that coupling starts with only the $k = 1$ mode. There is, however, no upper bound. In cases where the amplitude versus wave number spectrum is weighted towards high frequencies, such as can develop when there is either a numerically induced upper cutoff, the nonlinear term can cause propagation of energy towards lower frequency modes.

While the general tendency in coupling is for energy to move up the spectrum, an important exception is the ability of modes to couple to themselves and thus feed energy directly back to the fundamental mode. We will come back to this subject again in discussing the nonlinear results in Chapter 7.

The second type of mode coupling is that of different physical modes,

---

6Detailed energy mode analysis by Chait and Korpela [13] indicates that this is true for even the main secondary modes, at least in the near critical regime.

7This is also true for the thermal diffusivity term in the thermal equation.

8We are interested here in spatial scale and thus consider only the magnitude of $k$.

9This may be the mechanism by which random fluctuations of small scale and high frequency can suddenly generate large scale flows as a system passes through a critical point.
such as linear eigenmodes or full nonlinear modes, which are separately solutions of the equations but which can interact with each other through the Jacobian. This will allow mode resonances, oscillatory coupling of modes and other sorts of nonlinear behavior. We will return to this subject when discussing the nonlinear simulation results.

Physically the nonlinear terms, coupling the field gradients, add the ability to model shear-driven motions. In particular, a bounded shear flow is well known to give rise to secondary vortices.\textsuperscript{11} As discussed by Brown[59] when there is an inflection in the velocity profile the instability can be viewed as a generalized Kelvin-Helmholtz instability. Application of Squire’s theorem [5] shows that one must expect two-dimensional rather than three-dimensional secondary flows to form first. These two-dimensional vortices are referred to as ‘cat’s eye’ cells. Chait and Korpela [13] in a detailed energy balance analysis, show that kinetic energy for these cells is generated at the center of the slot by the shearing of the background flow and, to a lesser extent, by the buoyancy term (not included here). It is then carried to the walls where the dissipation is greatest. This process decreases the total kinetic energy of the flow, reduces the central shear, and lessens the stress at the walls.\textsuperscript{12}

The coupling to higher frequencies that the nonlinear terms allow models the shear driven formation of smaller and small vortices leading possibly to turbulence. Recalling our discussion of the Laplacian above, we see that there is a competition between the dissipative and nonlinear terms, one coupling energy into the higher modes and the other taking it out. This is

\textsuperscript{11}For a general discussion of instability criterion in shear flows including the inflection instability see [58].
\textsuperscript{12}See Nagata and Busse [24].
the mechanism behind thermalization and dissipation of kinetic flows. It also enables the nonlinear saturation of the initial exponential growth of linearly unstable perturbations.

Finally, as we will see below, Equation (3.5) applies in the $Pr \to 0$ limit with $1/G$ small.

### 3.3 The $Pr = 0$ Limit.

Next we add in the thermal physics. The coupling takes place in the equation of motion (2.12) by adding the gradient of the temperature perpendicular to $g$ to the Navier-Stokes equation (3.5). This is the buoyancy force term, and it breaks the symmetry of the equations in the $z$ direction\(^{13}\) just as the thermal gradient, $\Delta T$, breaks it in the $x$ direction. It is the forcing or driving term and is the last term needed to complete the equation of motion:

$$\frac{\partial \omega}{\partial t} = -\frac{\partial (\omega_1, \psi)}{\partial (x, z)} - \frac{1}{G} \frac{\partial T}{\partial x} + \frac{1}{G} \nabla^2 \omega. \quad (2.13)$$

If we now consider the equation of heat conduction

$$\frac{\partial T}{\partial t} = -\frac{\partial (T, \psi)}{\partial (x, z)} + \frac{1}{G Pr} \nabla^2 T, \quad (2.17)$$

we can see that the size of $Pr$ determines the relative importance of the two terms on the right. As $Pr \to 0$ the heat conduction term dominates, while for $Pr \to \infty$ it is negligible. For $Pr$ small $G$ still measures the nonlinearity of the system. Here we take $Pr = 0$ thus reducing Equation (2.16) to a Laplacian

$$\frac{\partial T}{\partial t} = \frac{1}{G Pr} \nabla^2 T \quad (3.7)$$

---

\(^{13}\)Recall that before we took the curl of the Navier-Stokes equation this term was proportional to the vector $g$ and anti-parallel to the $z$ axis.
representing pure thermal diffusion (see Gill[60]). This form is produced whenever \( GPr \) becomes very small. It is, in particular, true for \( Pr \to 0 \), a regime found physically in liquid crystals \([49]\).\(^{14}\)

In contrast to the boundary conditions on the Laplacian for \( \omega \) in Section 3.1, \( T \) is held at non-zero values at \( x = \pm L \). We then have the steady-state solution

\[
T = x,
\]

(3.8)
towards which thermal diffusion always tries to drive the temperature field. The curvature of this solution is zero, as there are no non-zero derivatives above first order. This steady-state solution describes the basic temperature gradient that drives the various convective motions that occur in our model, and forms the background thermal profile for all \( Pr \) and \( G \) in the VSC problem (Benard as well). It is referred to as the conductive temperature profile as it is due solely to thermal conduction.\(^{15}\)

With \( Pr = 0 \), we always have the conductive solution for the temperature. No other variations in the temperature can develop as the thermal diffusivity is too large, and rapidly reduces any fluctuations to zero. Physically, this is a fluid that is always, instantly, in thermal equilibrium. For \( T = x \) the coupling term in Equation (2.12) is simply the constant \( -\frac{1}{G} \), and for large \( G \) this will be a small constant. This small term is, however, important in that it allows the kinetic fields to feel the presence of the temperature gradient.

\(^{14}\)This is of course an asymptotic limit and we can never explicitly solve this equation with \( Pr = 0 \), as then the diffusivity is infinite. We can, however, empirically replace it with the large \( t \) solution for this equation with \( GPr \) very large, thereby assuming that as \( Pr \to 0 \) \( T(t) \to T(\infty) \) for all \( t \). This is true for low \( G \) as well.

\(^{15}\)A good illustration of this solution was produced by Eckert & Carlson [3] who performed interferometer visualizations on air \( (Pr = 0.71) \) at a medium aspect ratio \( (h = 47) \) with \( G = 16 \), far below critical. This simple solution for \( T \) can be seen nearly filling the slot.
Equations (2.12) and (3.7) are, therefore, consistent in the sense that they can support the complete 1-D solution of the VSC equations.\textsuperscript{16} With a horizontal temperature gradient $T = x$ the fluid cannot be in hydrostatic equilibrium. The fluid near the cold wall sinks and that at the hot wall rises, driving a single convection cell that fills the whole slot. For a slot with an infinite aspect ratio (height/width $= \infty$) this is a 1-D flow\textsuperscript{17}

$$W = \frac{1}{6} x (1 - x^2), \quad (3.9)$$

where $W$ has been normalized by $W_o$.

It is a flow with horizontal shear and a velocity inflection in the center of the slot and this can generate secondary flows, as we mentioned in Section 3.2 and will see in our results. With $Pr = 0$, however, these flows are restricted by the necessity of maintaining the conductive temperature profile.

This 1-D solution is of particular importance as it serves as the 'ground state' or 'basic solution'\textsuperscript{18} for the VSC problem in the sense that for any $G > 0$ this exists as a steady-state solution. It is one of only two analytic solutions to the full set of equations, although these are not, of course, the only steady-state solutions.\textsuperscript{19}

We see that the $\frac{\partial T}{\partial x}$ term is the main driving term in the equations, being the term through which the gravitationally driven buoyancy force is incorporated in the system. It arises from the interaction of the main thermal

\textsuperscript{16}Note that we could drop the Jacobian from Equation (2.12) and still support this solution.
\textsuperscript{17}This solution was first given by Batchelor [14]. For the full derivation see Appendix A.1.
\textsuperscript{18}See definitions in Section 6.1.
\textsuperscript{19}The other widely used 'exact' solution is found by assuming a constant vertical temperature gradient as part of the temperature solution. While this reflects experimental evidence for such a gradient at large $G$, it does impose artificial constraints on the problem. For this solution see Appendix A.2.
gradient with the pressure gradient and drives the basic single cell convective flow that in turn drives the higher order flows through shearing, and through the interaction of the shearing with the thermal gradient itself. In particular at low and mid Pr \((Pr < 12.5)\) this drives the ‘cat’s eye’ secondary circulation cells. It is important to consider that without a gravitational force this term would vanish no matter what thermal gradients were applied. Thus the \(\frac{\partial T}{\partial z}\) term serves two functions; it couples the temperature into the equation of motion, but it also introduces the externally enforced pressure gradient. There is, however, as yet no coupling at all of the kinetic variables back into the heat conduction equation.

Finally the thermal differential included in the boundary conditions for \(T\) may be thought of as a hidden driving term in the heat conduction equation.\(^{20}\)

### 3.4 The Full Vertical Slot Problem

Taking a general Pr, we can now look at the complete set of equations with the thermal effects fully coupled in:

\[
\frac{\partial \omega}{\partial t} = -\frac{\partial (\omega, \psi)}{\partial (x, z)} - \frac{1}{G} \frac{\partial T}{\partial x} + \frac{1}{G} \nabla^2 \omega,
\]

(2.12)

\[
\frac{\partial T}{\partial t} = -\frac{\partial (T, \psi)}{\partial (x, z)} + \frac{1}{GPr} \nabla^2 T.
\]

(2.16)

Including the thermal Jacobian term as well, we can expect that, just as the kinetic Jacobian term makes possible new shear-driven modes, this term makes

\(^{20}\)Note that there is another possible ‘driving’ term for the heat conduction equation, the ignored viscous dissipation term.
possible new modes driven by this product of velocity and temperature gradients. It also induces more complex temperature profiles which can drive new modes through the buoyancy term.

For $G \to 0$, we can see that the nonlinear terms become unimportant and leave a system dominated by the thermal gradient and the dissipation and diffusion terms. This system is generally stable with decaying oscillatory modes most prominent, except possibly at high $Pr$ where the nonlinear terms in Equation (2.16) can become important enough to drive unstable modes. Except for very viscous fluids a small thermal stress will not produce new modes. On the other hand for high $G$ the equations tend to decouple,\textsuperscript{21} unless the thermal gradients are also very large, with the Jacobian term being dominant in Equation (2.12). One may thus expect less impact from variations in $Pr$ at high $G$.

As we mentioned in Section 3.3 the value of $Pr$ determines the relative importance of the two terms in Equation (2.16). There are essentially three regimes: a low $Pr$ regime in which heat conduction dominates (as discussed in Section 3.3), a high $Pr$ regime where the convection term dominates and a transition range in the middle where all the terms are significant and where the $Pr$ dependence of the system is most complex.\textsuperscript{22} Physically it was pointed out by Gill [60] that for $Pr \neq 1$ that the kinetic and thermal boundary layers should be of different widths, the thermal layer being wider at small $Pr$ and the kinetic layer wider at large $Pr$. This is supported by experimental observation (Eckert and Carlson [3]) and computational results (Ruth [22], Chait and Korpela [13] and our own results, see Section 5.6.6). This double layer structure is also

\textsuperscript{21}As exploited by Gotoh & Satoh [15] to generate a power expansion in $\frac{1}{k_0}$.

\textsuperscript{22}See also Section 5.3.
discussed by Patterson and Imberger [38].

In the mid and high Pr ranges the thermal Jacobian and the buoyancy term cause changes in the complex modes. The most interesting of these new modes are traveling modes at low $G$ and low $\alpha$ that arise from the upstream buoyancy of the background shear flow, with kinetic energy generation peaking at the background velocity profile peaks.\(^{23}\) In these modes dissipation peaks at the center of the slot when the modes are in phase, but vanishes when they are out of phase. At all times there are smaller peaks nearer the walls. These traveling modes are not present in the Benard configuration.\(^{24}\) even though thermal Jacobian terms are present there as well. They are directly dependent on the combination of the shear flow and the vectorial relationship of the applied pressure and thermal gradients which yield the specific form of the buoyancy term in the equation of motion. Choi and Korpela [53] speculate that since these modes are driven by the boundary layers they are unlikely to combine to form standing waves. In support of this Elder[42] measured traveling wave onset at a Rayleigh number commensurate with that measured for a single heated wall. Especially in the high $G$ regime it is often assumed that the flow is restricted to narrow boundary layers along the walls.

Finally, as $Pr \rightarrow 0$ and Equation (2.16) reduces to a Laplacian for $T$, it decouples from the equation of motion (2.12). Thus it can be useful to think about $Pr$ as a coupling parameter or, for $Pr$ near 0, even as an expansion parameter. As this parameter increases, the thermal physics is coupled into the kinetic equation and breaks its symmetry through the buoyancy term. In

\(^{23}\)See energy analysis in Chait & Korpela [13].

\(^{24}\)This is true for the linear solutions. Nonlinearly, however, any regimes introducing complex eigenvalues can have similar traveling modes appear in the Benard configuration as well.
this view the thermal effects are taken as perturbations to the purely kinetic solutions of the equation of motion. As $Pr$ increases, there are two types of changes. One is due to the increasingly strong linking of the two equations and the second to the shifting balance of the terms within the thermal equation. Thus there is a qualitative change in the thermal field that is being linked into the kinetic equation. This second effect is the result of the increasing effect of the kinetic terms on the thermal solution through the increased important of the thermal Jacobian term. This point of view is of particular interest in making contact with field theoretical techniques.

3.5 The Benard Problem

In trying to understand the results for the vertical cases it will be useful to make some comparisons with the Benard case in which the temperature gradient is parallel to the gravitational pressure gradient, changing the buoyancy force term in the forced Navier-Stokes equation to give,

$$\frac{\partial \omega}{\partial t} = -\frac{\partial (\omega, \psi)}{\partial (x, z)} - \frac{1}{G} \frac{\partial T}{\partial z} + \frac{1}{G} \nabla^2 \omega,$$

(3.10)

but leaving the heat conduction equation the same,

$$\frac{\partial T}{\partial t} = -\frac{\partial (T, \psi)}{\partial (x, z)} + \frac{1}{GP_r} \nabla^2 T.$$

(2.17)

Since the pressure gradient must now be overcome before there is convection there is a low $\Delta T$ (low $G$) regime with a stable motionless solution. Only above a certain critical thermal forcing does convection occur. The ‘ground state’ is thus one with zero velocities, and unless it is imposed, the shear flow of the 1-D VSC solution is absent. This means that many of the flow patterns that one
sees in the VSC problem are absent from the Benard case as the symmetries of the problem are different.

In contrast, the VSC configuration has the thermal gradient perpendicular to the pressure gradient. In this case we found (Section 3.2) that there is convection for any $\Delta T > 0$ and that the symmetry of the problem is broken in two dimensions rather than only one. Correspondingly, the symmetry of the initial, low $\Delta T$ solution, our 1-D flow, is less as well. One might picture that in the VSC case there is an initial bifurcation at $\Delta T = 0$, $G = 0$ and we are then interested in a secondary bifurcation at finite $G$, whereas in the Benard case the initial bifurcation takes place at a finite value of the forcing (or order) parameter. Because of these differences one needs to be careful in making comparisons between these two cases. As we shall discuss in Section 4, there are special cases that can be set up in the linearized equations to make more direct comparisons.

Lastly, it should be noted here that the reduced cases of Sections 3.1 and 3.2 are simultaneously reduced cases for both the VSC and the Benard problems.

3.6 Summary

One can give a brief outline of the physics involved in the problem.

In the infinite aspect ratio VSC problem the interaction of the gravitational force and the temperature gradient combining in the buoyancy force term give rise to a single long convection cell, which is the 1-D solution. This flow has a velocity inflection in the center of the slot, and the instability there gives rise to shear-driven to secondary convective cells. That this is so can be seen by looking at the Benard configuration with flow. When the shear is strong
enough in the center of the slot, small convection cells appear superimposed on the background shear flow. We will see that at higher Grashof numbers there are even tertiary counter-rotating vortices formed from the vertical shear between the secondary cells. Lastly at large \( Pr \) there are regimes with dominant complex modes that are the result of the symmetry breaking performed by the gravitational force that are absent in the Benard case even when a shear flow is imposed. We will discuss these further when we examine the linear eigenmodes and the high \( G \) experimental results.
Chapter 4

The Linear Problem

Our linear analysis of the VSC problem and related special cases is based on linearizing the convective terms, the Jacobians, and uses a spectral analysis to obtain a complex eigenvalue problem that will yield the main eigenvalues and their associated eigenfunctions.

4.1 A General Discussion of Linear Stability Analysis

The general idea of a linear stability analysis is to examine the stability of a specific steady-state solution. To do this, one linearizes the full nonlinear equations by adding to this initial state a small general perturbation and after inserting this into the equations, dropping all terms with more than one power of the perturbation.

In our analysis we assume that the perturbations have an exponential dependence on time of the form $\exp(\lambda + i\omega t)$, where $\lambda$ and $\omega^1$ are respectively the real and the imaginary parts of the solution to the eigenvalue problem derived from the linear equations. Physically $\lambda$ is the linear growth rate of the associated eigenmode and $\omega$ is the real frequency (the imaginary frequency is given by $-\lambda$). We can also solve for the associated eigenfunction which will give the specific spatial dependence of the perturbation.

---

1In most of this text we will make use of $\omega$ as a symbol for the vorticity. On occasion, as here, it will also be used for the mode frequency. It will then be specifically indicated as such.
The sign of $\lambda$ determines the stability of the initial state to a perturbation with the spatial dependency of the associated eigenfunction. If it is positive, the perturbation grows exponentially, and the initial state is unstable. Likewise, if $\lambda$ is negative the perturbation decays and the initial state is stable. Of course, there can be many types of perturbations, and it only takes one with $\lambda > 0$ to cause the initial state to be unstable. The value of $\omega$ determines the character of this growth or decay of the perturbation. For any non-zero value ($\omega \neq 0$) the solution is oscillatory with $\omega$ being the frequency of the oscillation.

When the set of unlinearized equations contains Jacobian terms, these terms are reduced in different ways, when linearized with respect to different initial conditions. For some initial states they may vanish partially or even completely and therefore become irrelevant to the question of stability, even though they would in fact affect the evolution of any flow that arose. This is an important point, as the evolution of the perturbations under the full nonlinear equations remain the same as that under the linear equations only as long as the nonlinear terms are small. Once the nonlinear terms become significant, they alter both the spatial form of the new flow and, more importantly, the growth rate.

Because the stability question and the longer-term evolution of perturbations are separate issues, we can impose 'background' flows as part of the initial states by including them in the linearized Jacobian terms (see Section 4.2 for details) without worrying about how they would be maintained under time evolution of the full nonlinear equations. This is true as long as the background flows are solutions of the full equations and can therefore be subtracted off.
4.2 The linearization of the Jacobian

As we mentioned above we start the linearization process by separating all the variables, here $\psi$, $\omega$, and $T$, into a known solution to the full equations, the 'background', and a small perturbation term. Thus we have

$$
\psi \Rightarrow \psi_b + \hat{\psi}
$$

$$
\omega \Rightarrow \omega_b + \hat{\omega}
$$

$$
T \Rightarrow T_b + \hat{T}
$$

When these are inserted in the full nonlinear VSC equations (2.12) and (2.16) we have a set of terms for the background, or basic, solution, another set for the perturbations, and, where there are nonlinearities, sets of cross terms:

$$
\frac{\partial \omega_b}{\partial t} + \frac{\partial \hat{\omega}}{\partial t} = -\frac{\partial(\omega_b, \psi_b)}{\partial(x, z)} - \frac{\partial(\omega_b, \hat{\psi})}{\partial(x, z)} - \frac{\partial(\hat{\omega}, \psi_b)}{\partial(x, z)} - \frac{\partial(\hat{\omega}, \hat{\psi})}{\partial(x, z)}
\quad - \frac{1}{G} \frac{\partial T_b}{\partial \tilde{x}} - \frac{1}{G} \frac{\partial \hat{T}}{\partial \tilde{x}} + \frac{1}{G} \nabla^2 \omega_b + \frac{1}{G} \nabla^2 \hat{\omega},
$$

(4.1)

$$
\frac{\partial T_b}{\partial t} + \frac{\partial \hat{T}}{\partial t} = -\frac{\partial(T_b, \psi_b)}{\partial(x, z)} - \frac{\partial(T_b, \hat{\psi})}{\partial(x, z)} - \frac{\partial(\hat{T}, \psi_b)}{\partial(x, z)} - \frac{\partial(\hat{T}, \hat{\psi})}{\partial(x, z)}
\quad + \frac{1}{GPr} \nabla^2 T_b + \frac{1}{GPr} \nabla^2 \hat{T}.
$$

(4.2)

The terms involving the background solution can be subtracted out leaving the purely perturbative terms and the cross terms. In our case the Jacobians provide all the cross-terms and all the terms with more than one perturbation factor. The latter are now removed, thereby 'linearizing' the equations as follows:

$$
\frac{\partial \hat{\omega}}{\partial t} = -\frac{\partial(\omega_b, \hat{\psi})}{\partial(x, z)} - \frac{\partial(\hat{\omega}, \psi_b)}{\partial(x, z)} - \frac{1}{G} \frac{\partial \hat{T}}{\partial \tilde{x}} + \frac{1}{G} \nabla^2 \hat{\omega},
$$

(4.3)

$$
\frac{\partial \hat{T}}{\partial t} = -\frac{\partial(T_b, \hat{\psi})}{\partial(x, z)} - \frac{\partial(\hat{T}, \psi_b)}{\partial(x, z)} + \frac{1}{GPr} \nabla^2 \hat{T}.
$$

(4.4)
We can see here how the cross-terms in the Jacobians cause the linearized equations to depend on the particular state with respect to which they have been linearized. For the cases we are interested in, the steady-state base flows are independent of $z$, being either a motionless state or the 1-D VSC solution. When the base flow is independent of $z$ the $z$ partial derivatives of $\psi_b$, $\omega_b$, and $T_b$ are zero and Equations (4.3) and (4.4) simplify to:

$$\frac{\partial \dot{\omega}}{\partial t} = - \frac{\partial \omega_b}{\partial x} \frac{\partial \dot{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \dot{\omega}}{\partial z} - \frac{1}{G} \frac{\partial \dot{T}}{\partial x} + \frac{1}{G} \nabla^2 \dot{\omega},$$  \hspace{1cm} (4.5)

$$\frac{\partial \dot{T}}{\partial t} = - \frac{\partial T_b}{\partial x} \frac{\partial \dot{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \dot{T}}{\partial z} + \frac{1}{GP_r} \nabla^2 \dot{T}. \hspace{1cm} (4.6)$$

### 4.3 The Spectral Analysis

Having linearized the equations we use a spectral decomposition to set up an eigenvalue problem. We assume the following form for the perturbations:

$$\dot{\psi}(x, z, t) = \dot{\psi}(x) \exp(\lambda + i \omega t) \exp i \alpha z$$

where $\lambda$, $\omega$, and $\alpha$ are real,\(^2\) and similarly for $\dot{\omega}(x, z, t)$ and $\dot{T}(x, z, t)$. Inserting these into Equations (4.5) and (4.6) we get:

$$\lambda \dot{\psi}(x) = \frac{1}{G} (\partial_x^2 - \alpha^2) \dot{\omega}(x) - i \alpha \partial_x (\omega_b) \dot{\psi}(x) + i \alpha \partial_x (\psi_b) \dot{\omega}(x) - \frac{1}{G} \partial_x T$$  \hspace{1cm} (4.7)

$$\lambda \dot{T}(x) = \frac{1}{P_{Pr} G} (\partial_x^2 - \alpha^2) \dot{T}(x) - i \alpha \partial_x (T_b) \dot{\psi}(x) + i \alpha \partial_x (\psi_b) \dot{T}(x)$$  \hspace{1cm} (4.8)

with constraints:

$$(\partial_x^2 - \alpha^2) \dot{\psi}(x) - \omega \dot{\psi}(x) = 0 \hspace{1cm} (4.9)$$

and

$$\dot{T}(\pm 1) = \dot{\psi}(\pm 1) = \partial_x \dot{\psi}(\pm 1) = 0 \hspace{1cm} (4.10)$$

\(^2\)Note dual usage of $\omega$ here.
We now replace each \( x \) dependent perturbation with a Chebychev series in \( x \):

\[
\hat{\psi}(x) = \sum_{n=0}^{N} T_n(x) \psi_n
\]

and similarly for \( \hat{\omega}(x) \) and \( \hat{T}(x) \). Finally, we replace the fixed functions of \( x \) representing the base solution and the remaining partial derivatives with matrix operators. We thereby formulate a complex generalized eigenvalue problem that can be solved numerically for \( \lambda \), and the functions \( \hat{\psi}(x), \hat{\omega}(x), \) and \( \hat{T}(x) \).\(^3\)

### 4.4 The Linear Hierarchy

In this section we take a look at the linearized version of the hierarchy of equations that was introduced in Section 3. We examine the special cases with particular attention given to the specific initial conditions that we have used in our study. The results for the models outlined below are presented in Section 5.

Using a Chebychev-Fourier analysis, it is not possible to analytically solve for the exact eigenvalues.\(^4\) We can, however, make some observations that will help in understanding the general features of the solutions.

There are two characteristics of the eigenvalues and the eigenmodes that we particularly want to examine here. These are the reality or complexity of the eigenvalues and the phase relationships of the eigenmodes. The key to sorting this out lies in the derivatives. Our two (three, including the definition of the vorticity in Equation (2.14)) equations are of the form

\[
\lambda \hat{\phi} = A \hat{\phi} + B \hat{\Xi},
\]

\((4.11)\)

\(^3\)See AppendixC.1 for a more detailed explanation of this method.
\(^4\)They can be solved for using a Fourier-Cosine analysis.
where A and B are differential operators, $\hat{\phi}$ is one of our variables, and $\hat{\mathcal{E}}$ represents a vector of the other variables in the system. The two operators determine, respectively, the reality or complexity of the eigenvalue $\lambda$ and the phase relations of $\hat{\phi}$ with the variables represented by $\hat{\mathcal{E}}$. If A is real, the eigenvalue is real. This happens whenever the operator includes only even numbers of derivatives.

Recalling that each of our variables is assumed to have the form given in Equation (3.2), we can see that each derivative will pull down a factor of $i$. The number of derivatives in B acting on one of the variables in $\hat{\mathcal{E}}$ serves to produce a relative phase equal to that number times $\frac{\pi}{2}$. Thus for instance in all cases Equation (2.14) implies that the vorticity and the stream function fields are exactly opposite in phase.

4.4.1 The Navier-Stokes Equation

We consider first the Navier-Stokes equation linearized about a null initial state.

$$\frac{\partial \hat{\omega}}{\partial t} = \frac{1}{G} \nabla^2 \hat{\omega} \tag{4.12}$$

Since this is of exactly the same form as Equation (3.1) in Section 3.1 above, the analysis and solutions given there are valid here as well. These solutions are all stable (decaying) real modes.\(^5\)

The Navier-Stokes equation with a shear-flow initial state, which would include the Jacobian terms, has essentially the same solutions as the

\(^5\)Consider that the Chebychev polynomials can be written as cosines, which in turn can be written as combinations of complex exponentials.

\(^6\)Here and hereafter references to 'real modes' mean modes with real $\lambda$ and $\omega = 0$. 
$Pr = 0$ case given next. It is not, however, consistent (as it is missing the thermal gradient term necessary for the 1-D solution) whereas the $Pr = 0$ case is consistent in this sense.

### 4.4.2 The $Pr = 0$ Limit

This case is both the $Pr = 0$ limit\(^7\) of the full VSC problem, and the equivalent to the Navier-Stokes case with the VSC 1-D solution as the initial condition. The addition of the temperature Laplacian does not affect the kinetic solutions. The relevant equations are

\[
\frac{\partial \hat{\omega}}{\partial t} = -\frac{\partial \omega_b}{\partial x} \frac{\partial \hat{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \hat{\omega}}{\partial z} \left[ -\frac{1}{G} \frac{\partial \hat{T}}{\partial x} \right] + \frac{1}{G} \nabla^2 \hat{\omega},
\]

(4.13)

\[
\frac{\partial \hat{T}}{\partial t} = \frac{1}{G Pr} \nabla^2 \hat{T}.
\]

(4.14)

Since Equation (4.14) is a Laplacian for the temperature perturbation, it prevents any growth and holds $\hat{T} = 0$. This, in turn, means that for any significant $G$ the thermal gradient term $[-\frac{1}{G}]$ in the equation of motion (4.13) is essentially zero and the kinetic solutions are indeed the same as those for the Navier-Stokes equation (when it is also expanded around the 1-D kinetic solution). When linearized Equation (4.13) without the gradient term becomes the Orr-Sommerfeld equation ([44],[18],[45]).\(^8\) In the full equations (both VSC and Benard) there are often both thermal and kinetic gradients driving the

---

\(^7\)In the sense that it is taken here this limit results from the increase of the thermal diffusivity $\kappa \rightarrow \infty$ while the viscosity $\nu$ remains finite.

\(^8\)The Orr-Sommerfeld equation can also be derived in the limit $G \rightarrow \infty$, [18][19], in which case the buoyancy term is dropped while the viscous term is kept, although it is also proceeded by $\frac{1}{G}$. We also note that with a different basic state it describes linearized Poiseuille flow [61].
fluid motions. This equation is useful for checking which motions are purely kinetic as opposed to those driven by thermal gradients.

4.4.3 The Zero-Velocity Initial State

This case is the VSC equivalent of the standard Benard problem. The initial state consists of a thermal gradient, $T = x$, but has null kinetic fields. The reduced equations are

$$\frac{\partial \omega}{\partial t} = -\frac{1}{G} \frac{\partial \hat{T}}{\partial x} + \frac{1}{G} \nabla^2 \hat{\omega}, \quad \text{(4.15)}$$

$$\frac{\partial \hat{T}}{\partial t} = -\frac{\partial \psi}{\partial z} + \frac{1}{G Pr} \nabla^2 \hat{T}. \quad \text{(4.16)}$$

It is only consistent for $G = 0$, since it is only for $G = 0$ that the kinetically null initial state is a solution of the unperturbed form of these equations. This is due to the non-zero value of the term $\frac{1}{G} \frac{\partial T_k}{\partial x}$. This initial state, although unstable for all $G > 0$, is equivalent to the Benard initial equilibrium, and serves as the base state from which the 1-D equilibrium bifurcates (as suggested in Section 3.5). Thus besides the direct mathematical comparison to the Benard problem, with which it differs solely in that the $z$ derivative of $T$ in that case is replaced here by an $x$ derivative of $T$, this case describes the establishment of the 1-D shear flow after a sudden switching on of the temperature gradient. In this regard it is suited to yield information about the transients involved in the establishment of such a flow.\footnote{The issue of transients and their implication in applications such as crystal growth were investigated experimentally by Ivey(82).}

The phase structure in this case holds for the rest of our case studies as well. The thermal gradient in Equation (4.15) leads to a phase difference of
\( \frac{\pi}{2} \) between the temperature and the vorticity, while the gradient of the stream function appearing in Equation (4.16) leads to a similar difference between the stream function and the temperature.\(^{10}\) Thus, in general, the phase of the temperature leads that of the vorticity by \( \frac{\pi}{2} \) and lags that of the stream function by the same amount. The modes in this case are real as there are no contributions from the Jacobian terms to the operator \( A \) defined above.

### 4.4.4 The Benard Problem

The classic Benard problem, where the initial state is a motionless fluid with a conductive temperature profile, \( T = x \), is represented by

\[
\frac{\partial \omega}{\partial t} = -\frac{1}{G} \frac{\partial \hat{T}}{\partial z} + \frac{1}{G} \nabla^2 \hat{\omega},
\]  

(4.17)

\[
\frac{\partial \hat{T}}{\partial t} = -\frac{\partial \hat{\psi}}{\partial z} + \frac{1}{GPr} \nabla^2 \hat{T}.
\]  

(4.18)

Although these equations are nearly the same as Equations (4.13) and (4.14) the partial of \( T \) with respect to \( z \) (rather than \( x \)) is zero for \( T_0 = z \), and the null kinetic solution as a consistent initial condition. This case has real modes and the same relative phasing as the zero-velocity initial state case. Since the gradient term in Equation (4.17) differs numerically from that in Equation (4.15) there may be differences in the overall phase at large \( Pr \).

\(^{10}\)There seem to be some linear eigenvector solutions where this is not true, but they are quite rare and may well be unphysical.
4.4.5 The Benard Configuration with Shear Flow

The Benard Equations expanded around the 1-D solution, so that we have added a shear flow to the thermal gradient in our initial conditions, give

\[
\frac{\partial \dot{\omega}}{\partial t} = -\frac{\partial \omega_b}{\partial x} \frac{\partial \dot{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \dot{\omega}}{\partial z} - \frac{1}{G} \frac{\partial \dot{T}}{\partial z} + \frac{1}{G} \nabla^2 \dot{\omega}, \tag{4.19}
\]

\[
\frac{\partial \dot{T}}{\partial t} = -\frac{\partial T_b}{\partial x} \frac{\partial \dot{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \dot{T}}{\partial z} + \frac{1}{G Pr} \nabla^2 \dot{T}. \tag{4.20}
\]

Although this is another inconsistent case, in that the background state is not a solution to the full Benard equations unless the boundary conditions are changed (see Section 3.5), it is the best horizontal comparison to the full VSC case, indicating which modes are purely shear driven and which depend as well on the perpendicularity of the gravitational force. Some eigenvalues are real and others complex, with the same relative phasing as outlined in Section 4.4.3 above.

4.4.6 The VSC Problem

For completeness we can refer to Equations (4.5) and (4.6) and say that there will be both real and complex modes, but again the relative phases may be expected to be the same as given above.

4.5 The Jacobian Terms for the 1-D Background

It is useful for the cases where the background is given by the asymptotic 1-D solution to re-express the derivatives of \( \omega_b \) and \( \psi_b \) in the Jacobian terms as derivatives of the 1-D velocity \( W_b \) instead. As given in Appendix A.1 we have

\[
\psi_b = \frac{1}{24} (x^2 - 1)^2
\]
\[ \omega_b = \frac{1}{6} (3x^2 - 1) \]
\[ W_b = \frac{1}{6} x (1 - x^2) \]

while from Definition (2.14) we get
\[ \omega_b = \frac{\partial^2 \psi_b}{\partial x^2} \]

Combining these we get
\[ \frac{\partial \psi_b}{\partial x} = -W_b \]
\[ \frac{\partial \omega_b}{\partial x} = -\frac{\partial^2 W_b}{\partial x^2} \]

With these relationships we use can write the Jacobians in terms of the quantity \( W_b \) and the perturbations.
Chapter 5

The Linear Computational Results

In this section we will present a summary of the results of our linear study of the related VSC and Benard convection problems. These results are built on an algorithm and code that solve the eigenvalue problem posed by linearizing the governing equations, as outlined in Section 4.2 and Appendix C.1. We have been able to construct various graphic outputs by repeated use of the ability to solve a given problem for different values of the controlling non-dimensional parameters.

We first present graphic and quantitative results for the Vertical Slot Problem, which is the main focus of our interest, introducing the types of results we have obtained. We then discuss these in the context of supplementary results for the various special cases, which serve to clarify the significance of the VSC findings. Together, these results will serve to illustrate points made in our previous discussions of these equations and to provide specific new qualitative and quantitative measures of the effects of the different terms and parameters in the full set of equations. Comparison with previous results also serves as a good test of our numerical methodology.

5.1 The General VSC Eigenvalue Structure

We start with the raw data, the eigenvalues obtained for the VSC problem. Doing a sweep versus the Grashof number for a given Prandtl number
and wave number, chosen here as $Pr = 7.5$ and $\alpha = 1.383$, we plot the real parts of the eigenvalues in Figure 5.1. This plot does not show the entire verticle slot problem.

![Figure 5.1: VSC eigenvalues at $Pr = 7.5$, $\alpha = 1.383$.](image)

eigenvalue structure but, rather, only the top portions which are of the most interest. The horizontal line near the top of the plot marks $\lambda = 0$ indicating that most of these modes are decaying modes that would appear physically as short-lived transients. We must keep in mind as well that due to our spectral approach we will produce a large number of solutions to a problem with only six degrees of freedom, indicating that other than the top few these modes are purely mathematical products of our methodology.

Figure 5.1 illustrates several general features. First at low $G$ most of the modes are real modes, meaning that the eigenvalue is real ($\omega = 0$); however

---

1 Although in our linear study we will explore all the parameter space defined by the Grashof number $G$, the Prandtl number $Pr$, and the perturbation vertical wave number $\alpha$ (see the perturbation definition in Section 4.3), we will often use these specific values of for $Pr$ and $\alpha$, respectively the Prandtl number for water and the linearly critical wave number at this Prandtl number.

2 In a test we increased the number of Chebychev decomposed modes from 32 to 64 and found that at larger $G$ only the real mode and the top six complex modes remain the same while at lower $G$ not more that the top dozen real or complex modes are unaltered.
as $G$ increases these combine two at a time to form pairs of complex modes with eigenvalues of the form $\lambda \pm i\omega$. Physically this represents the change from stationary modes, which may grow or decay but do not move in the slot, to pairs of traveling modes that travel up and down the hot and cold walls of the slot respectively.

Practically we are most interested in two modes. The first is the sole stationary mode to persist as $G$ increases, becoming unstable at $G = 491.8$ for $Pr = 7.5$ it is the main growing for all Prandtl numbers less than $Pr \sim 12.5$. The second is the top complex mode which at $Pr = 7.5$ and for $G > 150$ has a small, almost constant, negative real part to its eigenvalue. This mode is the primary transient for $Pr < 12.5$. Also of note is the point where these two top modes cross, a sort of exchange of instability. There are some indications that the addition of noise to the full nonlinear simulations can produce a resonance linking these modes above this crossing point at $G \sim 405$ resulting in sub-critical instability.[63]

At larger Prandtl numbers the real part of the complex mode’s eigenvalue, $\lambda_c$, becomes positive for some values of $G$ and $\alpha$. Above $Pr \sim 12.5$ $\lambda_c$, becomes positive at a value of $G$ that is less than the critical value for the real mode so that there the critical modes are traveling modes. In Figure 5.2 we show the eigenmode structure at $Pr = 50$. Here we can see again the pairing of real modes as complex pairs. The top complex mode becomes unstable for roughly the range $70 < G < 440$ while the stationary mode can be seen rising across all the complex modes, becoming unstable for $G > 600$. Note that as pointed out by Birikh et.al.[21] this real mode arises from complex modes that

---

3This general feature of the linear spectrum was first pointed out by Rudakov [16].
in turn form from pairing real modes. Birikh's group also pointed out that the small real part of the eigenvalue of the unstable complex mode limits it to large aspect slots. If the slot is too short the mode cannot grow before it reaches the top of the slot. Comparing the time to travel up the slot given by the slot height and the wave speed, with the exponential growth time one can only expect to see these waves if the former is significantly larger than the latter. The specific value for this minimum aspect ratio depends on the parameters $G$, $Pr$, and $\alpha$.

5.2 The Neutral Stability Curve

One of the main uses of a linear study is to identify which modes are stable and which are not. In particular one may define a curve $G_c(Pr, \alpha)$ at which $\lambda_{\text{max}}$, the real part of the largest eigenvalue, is zero. Combining Equations (4.7) and (4.8) to eliminate $\dot{T}$ and setting $\lambda = 0$ one gets

$$\left[(\partial_z^2 - \alpha^2)(\partial_z^2 + \alpha^2 + i\alpha G \partial_z(\psi_b) - i\alpha G \partial_z(\omega_b)\right].$$

$^4$One can alternatively eliminate $\dot{\psi}$, obtaining an identical equation in terms of $\dot{T}$. 
\[
\left( \frac{d^2}{d\xi^2} - \alpha^2 + i\alpha GPr \partial_\xi(\psi_b) \right) + i\alpha GPr \partial_\xi(T_b) \right) \hat{\psi} = 0. \tag{5.1}
\]

Using the appropriate boundary conditions this equation may be solved numerically to yield the desired curve \(G_c(Pr, \alpha)\). This is called the neutral stability curve and separates the unstable modes \((\lambda_{\text{max}} > 0)\) and the stable modes \((\lambda_{\text{max}} < 0)\). Figure 5.3 shows the neutral stability curve for \(Pr = 7.5\).

Figure 5.3: The VSC neutral stability curve to \(G = 10,000\), \(\alpha = 5\).

Rudakov[16] seems to be the first to have published such curves, sketches for \(Pr \leq 0.2\) and \(G \leq 500\). Vest and Arpaci[44] present a one for general \(Pr\) that reaches larger \(G\), and Mizushima and Gotoh[23] show one for \(Pr = 7.5\) that ours matches well. The shape of this curve is essentially unchanged for all \(Pr < 12.5\), moving slightly up and down with \(Pr\). As the neutral curve for the stationary mode it holds for high \(Pr\) as well. Mizushima[26] shows neutral curves for high \(Pr\) that drop significantly. However they are also spread out as

\[\text{In our work however we found this curve using a search routine based on the eigenvalue solver.}\]
happens when there is insufficient resolution at large values of $Pr$. Our study of mode truncation at lower $Pr$, see Figure 5.4, indicate that $G_c$ is sharply reduced when there are too few modes.\(^6\)

Thus for low and mid $Pr$ the vertical slot is linearly unstable to stationary modes with a wave number\(^7\) less than 2.0 at $G = 1000$ and not more than 2.5 at high $G$. This upper limit on $\alpha$ holds for all $Pr > 1000$ where the traveling mode is unstable up to $\alpha \sim 2.7$.\(^8\) In dimensionless units the slot width is 2, and this cap on $\alpha$ limits the wavelength of any unstable mode to be greater than $\sim 2.5$ while the most unstable modes have an aspect ratio of $\sim 1.26$. There is also an upper limit on the wavelength as all modes are stable for $\alpha \to 0$, which is stable for all $G$. Calculation of the mean-flow as a function of $x$ for various wavenumbers at a fixed Grashof number\(^24\) shows that it is maximally reduced when the wave number is in the range 1.3-1.9 so that the wavenumber is selected that will minimize the mean flow. The Prandtl number independence is a sign of the kinetic nature of these shear driven modes.

The stationary mode neutral stability curve is nearly but not quite symmetric around the line through the critical wave number at $\alpha = 1.38348$, as it rises more steeply on the large $\alpha$ side. The critical Grashof number is $G = 491.8187$. These figures are derived by a routine that searches specifically for the critical point, the smallest value of $G$ for which $\lambda = 0$ holds. These values agree well with previous published results which are mostly older rough estimates on the general order of $G \sim 490$ and $\alpha \sim 1.4$. Vest & Arpac\(^{44}\) gave

\(^6\)If one solves the linear problem linearizing about the high $G$ solution (see Appendix A.2) rather than the 1-D background the neutral curves do seem to drop significantly in $G$ at high $Pr$. See Daniels \cite{48} and also \cite{44}.

\(^7\)Having defined the wave number as $\alpha = 2\pi/\lambda_{pert}$, where $\lambda_{pert}$ is the perturbing wavelength, its normalisation factor must be the scale length $L$. See Section 2.3.

\(^8\)See also Figure 5.10.
more accurate values for a general Pr as $G = 492.5$ and $\alpha = 1.325$ respectively. Ruth [22] gave the most accurate values published to date, of $G_c = 491.776$ and $\alpha_c = 1.3835$. Note that Mizushima & Gotoh[23] gave $G = 492.3$ at $\alpha = 1.414$. This not the critical point but it is a point on the neutral stability curve, when we assume this value for $\alpha$ and then solve for $G$ our result is in good agreement.\(^9\) In more recent studies Dr. Mizushima working with Dr. Fujimura\(^10\) has obtained results of a high order of accuracy, with which our results agree to five decimal places in $\alpha_c$ and differing slightly in the third place for $G_c$.\(^11\)

Figure 5.4: VSC critical value mode-number dependence, $\alpha_c$ vs. $G_c$.

In the interests of testing the numerical stability of these results we did a series of neutral stability curve calculations varying the number of Chebychev modes, $N$. In figure 5.4 we show the variation of $G_c$ and $\alpha_c$ with the number of modes ($N$ goes from 8 to 64, $Pr = 7.5$) and also with the variation of the Prandtl number (from 0.00001 to 15). We found that when we reduced the $N$, 

\(^9\) In a private communication Mizushima has confirmed that this is what they did.

\(^10\) Private communication 1990.

\(^11\) For critical values at other Prandtl numbers see Section 5.3.
\( \alpha_c \) increased while \( G_c \) decreased, giving the long horizontal sweep in the figure. The point marked by the arrow indicates the results for \( N = 24 - 64 \), all on top of each other.\textsuperscript{12} The Prandtl number sweep (shown for comparison) starts at the upper left, rises in \( \alpha \) while dipping temporarily in \( G \), and finally moves back to lower \( \alpha \), ending up just a little below the \( Pr = 7.5 \) accumulation point.

Figure 5.5: VSC neutral stability curves with 8 to 32 modes.

The increase in \( \alpha_c \) with decreasing \( N \) reflects a general tendency for all curves and contours of \( \lambda(G, \alpha) \) to spread towards larger \( \alpha \) when the number of modes is insufficient to resolve the equations. This is especially true at high \( G \) and high \( Pr \), suggesting that it reflects the dwindling influence of the diffusion terms when the number of modes is too small. Note that the variation with \( N \) is much greater than that with \( Pr \) and that increasing \( Pr \) has generally the

\textsuperscript{12}The values for 32 and 64 differ by not more than .001.
same effect as reducing $N$.\textsuperscript{13}

In Figure 5.5 we have the VSC neutral stability curves for $N$ from 8 to 32. Most of the change takes place at large $\alpha$, only the $N=8$ curve deviates significantly at small $\alpha$. The neutral stability curves for small mode numbers besides dropping lower in $G$ show new modes appearing on the high $\alpha$ side of the main mode. These modes appear at high $G$ first and descend towards small $G$ with decreasing $N$. This is similar to the descent of the high $Pr$ low $\alpha$ complex mode, which also appears first at high $G$ and slides down to low $G$ with increasing $Pr$. We see that for $N<12$ there are unstables modes at all $\alpha$ up to at least 5. For 8 modes there is a low $G$, low $\alpha$ `foot' extending to smaller $G$, but also that for some parameters there is a new found stability.

In general, for $Pr$ up to about 10.0, 32 modes seems to be sufficient unless precise quantitative results are needed. For larger $Pr$ however the number of modes increases rapidly. By $Pr = 1000$ even 64 modes is too few to prevent spreading of the neutral curves.

5.3 Critical Point Sweeps vs. the Prandtl Number

From the results of varying the Prandtl number we can see that although the qualitative form of the stationary mode neutral stability curve changes very little, there is quantitative variation, especially in $G$. These changes are presented in Figure 5.6 where we see plotted the critical Grashof and wave numbers for the real mode from $Pr = 0.001$ to $Pr = 100$. We observe that there are asymptotic regimes at low and high $Pr$ where, as discussed

\textsuperscript{13}Rudakov\cite{17} shows a case at $Pr = 10$ where reducing the number of modes will cause eigenvalues to become positive artificially.
Figure 5.6: VSC stationary mode critical values vs. Prandtl number. above (Section 3.4), the dissipative and the nonlinear terms dominate, while in the central region there is competition between them. We note again that at higher $Pr$ there is also an unstable complex mode which we shall come back to later.

These plots are in good agreement with similar ones in Rudakov [17], Ruth [22], and by Mizushima(1990).\textsuperscript{14} Our results at specific values of $Pr$ are also in good agreement with previous results; at $Pr = 0$, where $\alpha_c = 1.34414$ and $G_c = 495.6284$ [44][24], at $Pr = .7$ (air) where $\alpha_c = 1.4049$ and $G_c = 502.5889$ [44][49], and at $Pr = 6.7$ (also water) $\alpha_c = 1.3836$ and $G_c = 491.7479$ [50]. We will return to discuss the correlation of the Prandtl number variation of the critical values with the physics of the modes when we discuss the VSC eigenmodes (see Section 5.6.6).

\textsuperscript{14}Ruth's is the most precise in publication, and also summerizes some earlier results. Mizushima's results where provided in a personal communication.
5.4 $\lambda(\alpha, G)$ and $\omega(\alpha, G)$

In order to get a better picture of the modes that form solutions to the VSC problem for various values of $Pr$, one can plot the maximum real eigenvalue $\lambda_{\text{max}}$ as a function of $G$, and $\alpha$. In Figure 5.7 we show a sequence of contour plots of $\lambda_{\text{max}}(G, \alpha)$ at Prandtl numbers 0.001, 0.25, 0.5, 7.5, 15, and 100. The range in $G$ is 0 to 10,000 and in $\alpha$ 0 to 10.0. The solid lines indicate positive values and therefore unstable regions, while the dashed lines are used for negative values or stable regions ($\lambda = 0$ is the first dashed line). The data ranges from $-.1031e + 3$ to $+.1955e - 1$ and the spacing between contours is .004.\textsuperscript{15} The main feature in all these cases is the unstable real mode at low $\alpha$. The eigenvalues for this mode remain relatively unchanged over a wide range of Prandtl numbers, especially at large $G$. Since this plot shows only the value of the maximum eigenvalue, it represents a surface made up of features corresponding to more than one mode. Aside from the real mode the others are all complex, and for $Pr < 12.5$ they are all stable. Above $Pr \sim 12.5$ there is an unstable complex mode at low $G$, low $\alpha$ which can be seen in the bottom two frames of Figure 5.7. We treat this separately in Section 5.4.1.

At $Pr = 0.001$ there is a change at approximately $\alpha = 2.4$ to a series of complex modes that fill up the high $\alpha$ region. The complexity of the contours here arises from a patchwork of different modes that are largest for different regions of the parameter space. As the Prandtl number increases this patchwork is replaced by a single complex mode, which fans out from the vertical boundary with the real mode. For $Pr > .7$ this mode is the largest for all $G$ at high $\alpha$.

\textsuperscript{15}The lower limit rises for $Pr \geq 7.5$, and there is considerable numerical noise at low $G$ and large $\alpha$ at $Pr = 100$. 

Figure 5.7: VSC $\lambda_{\text{max}}(G, \alpha)$ contours to $G = 10,000$ and $\alpha = 10$ at Prandtl numbers 0.001, 0.25, 0.5, 7.5, 15, and 100. The data ranges from $-1.031e+3$ to $+1.955e-1$ and the spacing between contours is .004.
While the plot of $\lambda(G, \alpha)$ shows clearly the most important physical property of the stability of the system and gives the rates of growth or decay, it is not so clear what the range of each mode is and of what type it is. In order to explore the nature of the different eigenmodes further and clarify the presentation of their character, we used contour plots of two more functions.

The first is the imaginary part of the eigenvalue, $\omega(G, \alpha)$. This is the mode frequency, which is essentially zero for the stationary modes, and is related to the wave speed and phase velocity of the traveling modes, $v_p = \omega/\alpha$. In Figure 5.8 we show contour plots of $\omega(G, \alpha)$ at Prandtl numbers 0.001, 0.2, 0.25, 0.5, 2.0, and 7.5. The data range from 0 to .5902 in the first four frames and to .6336 and .6367 in the fifth and sixth respectively and the contour spacing is .02. Looking at the results for various values of $Pr$ in Figure 5.8, we see that the clear division of real and complex modes at low $\alpha$ starts breaking down about $Pr = 0.5$ when a complex mode first appears at low $G$, and $\alpha \sim 1.2$. As $Pr$ increases, this presence grows and for $Pr \gtrsim 2.0$ there is a complex mode at low $G$ for all $\alpha$ except those near zero. Also noticeable in this figure are the discontinuities in $\omega$ that break the $(G, \alpha)$ plane into regions. Comparison with Figure 5.7 shows that these discontinuities coincide with the abrupt slope changes seen in the contours of $\lambda$ and examination of the eigenmodes shows that they are caused by the change from one mode to another with a different value of $\omega$. Lastly we should note that as $Pr$ increases into the middle range where the effect of the nonlinear terms begins to be felt, there is a sweeping out towards higher $\alpha$ of one of these discontinuities that is related to a process by which a single mode comes to dominate in the whole complex part of the $(G, \alpha)$ plane.\(^\text{16}\)

\(^{16}\)We shall come back to this in describing the eigenmodes and their categorization.
Figure 5.8: VSC $\omega(G, \alpha)$ contours to $G = 10,000$ and $\alpha = 10$ at Prandtl numbers 0.001, 0.2, 0.25, 0.5, 2.0, and 7.5. The data range from 0 to 0.5902 in the first four frames and to 0.6336 and 0.6367 in the fifth and sixth respectively and the contour spacing is .02.
Calculations of the critical wave speed have shown that it is very close to \( W_0 \), the maximum vertical velocity in the background shear flow, ranging from \(.93W_0 \) at \( Pr = 20 \) to \( 1.02W_0 \) at \( Pr = 100 \)\(^{[21]} \) and equally exactly \( W_0 \) at \( Pr \sim 38 \)\(^{[45]} \). Bergholtz \(^{[32]} \) suggests that the wave speed is buoyancy enhanced at large \( Pr \) as it is generally less than \( W_0 \) for inviscid parallel shear flows.

### 5.4.1 The High \( Pr \) Complex Mode

We now turn to discuss the complex mode that shows up subcritically at about \( Pr = 0.5 \). As \( Pr \) increases, this traveling mode\(^{[17]} \) becomes unstable, starting at high \( G \) and then at progressively lower \( G \) until, at \( Pr = 12.4542 \) when \( G_c = 492.056 \), the critical number for this complex mode (at \( \alpha_c = 0.3424 \)) is equal to that for the stationary real mode (at \( \alpha_c = 1.38317 \)).\(^{[18]} \) As \( Pr \) increases further, the complex mode becomes unstable for increasingly small \( G \) and increasingly large \( \alpha \). In Figure 5.9 we show roughly\(^{[19]} \) how the critical values for this mode change with \( Pr \) up to 1000. The Grashof number starts large and drops with increasing Prandtl number, reaching \( \sim 20 \) at \( Pr = 1000 \).\(^{[20]} \) The initial wave number for the traveling mode is at much lower \( \alpha \) than that for the stationary mode as it first appears along the \( \alpha = 0 \) axis. As \( Pr \) increases further, however, the mode curls around under the stationary mode (see Figures 5.7 and 5.10) and \( \alpha_c \) for the traveling mode grows until it is in the same range as

---

\(^{[17]}\) Traveling wave modes where first described by Elder\(^{[42]} \) and first treated theoretically by Gill\(^{[60]} \).

\(^{[18]}\) These values are from a recent communication from J. Mizushima.

\(^{[19]}\) Due to the double valued nature of the neutral curve for this mode (see Birikh et.al.\(^{[21]} \) for neutral curves for \( Pr = 12.5 \) to 40) it is difficult numerically to search for the neutral curve and the critical points.

\(^{[20]}\) Elder\(^{[42]} \) and Vest & Arpaci\(^{[44]} \) give values of \( \sim 20 \) and \( 19.5 \) respectively. Our estimate, from a \( \lambda(G,\alpha) \) plot done with \( N = 100 \), is 15.7.
Figure 5.9: VSC complex mode critical values vs. Prandtl number.

As can be more clearly seen when the critical values for both modes are plotted together[45][21] the changes in the ranges in Figure 5.9 are much greater than those for the real mode in Figure 5.6. In Figure 5.10 we show in detail the $\lambda(G, \alpha)$ contours at $Pr = 50$. The data range from -.1474e+1 to .1009e-1 in increments of .0008. For reference the vertical line indicates the wavenumber at which the data in Figure 5.2 were calculated. We can see clearly here how the complex mode, rather than replacing the real mode, fills the space around it at low $\alpha$ and low $G$. Note also that the growth rates for the complex mode are generally much less than those for the real mode. Finally we can see the upper limit in $G$ at which the mode becomes stable again. Increasing $G$ even further the stationary mode in turn becomes unstable. At some wave numbers there is a direct exchange, at

---

21 Vest & Arpaci[44] give $\alpha = .925$ at $Pr = 1000$ but our data indicate a value of $\alpha = 1.2$ there.
Figure 5.10: The $\lambda(G, \alpha)$ contours at $Pr = 50$ to $G = 1000$, $\alpha = 3$. The data range from $-1.474e+1$ to $1.009e-1$ in increments of $0.0008$.

others a gap, only at small $\alpha$ does this mode persist at high $G$ and there only with a very small growth rate.

5.5 Other Cases and Discussion

In order to clarify these results, we can run a similar analysis on the other cases in the hierarchy described in sections 3 and 4.4 above. By comparing results with those for the full problem we can often see quite clearly the physical consequences of the different terms in the full equations.

5.5.1 The Navier-Stokes Case

The results of the Navier-Stokes equations with no background shear flow may be compared to the $Pr = 0$ asymptotic limit for the Benard case rather than the VSC. As one would expect, this is an entirely stable system,
all modes being real with negative $\lambda$.

### 5.5.2 The $Pr = 0$ Limit

This case is essentially a limiting case of the full VSC problem and shows which modes are driven by the kinetic terms, the result of the interaction of the Jacobian and the damping terms in the kinetic equation. It is important to remember that we are assuming a background flow of the form of the 1-D solution to the full equations. While this is a consistent solution to the asymptotic full equations which still support a $T = x$ solution and influence the kinetic equation through the $-\frac{1}{c} \partial_x T$ term, even if it is only a constant. In the case of the Navier-Stokes equation we get these same results but only by starting with the 1-D kinetic solution with its sheared flow.

This demonstrates that the main stationary mode is primarily driven by the central shearing of the 1-D kinetic solution without any direct input from the thermal gradient. This is the situation that is ‘perturbed’, as we increase the Prandtl number and increasingly couple in the thermal effects.

### 5.5.3 The Benard Problem

As a check on our work as well as for comparison to this closely related problem we also examined the same quantities for the Benard problem. Although the nonlinear equations for this problem are nearly identical the change in coupling between them, which changes the symmetries of the problem, has significant effects. The simplification of the linear physics can be related to two factors. The first is that the ‘ground state’ is now a motionless fluid, so that there is no initial background flow and no initial shearing.

Figure 5.11 shows the eigenvalues for the Benard problem. These are
Figure 5.11: Benard eigenvalues at $Pr = 7.5$, $\alpha = 1.5585$. all real, stationary modes as we expected.\textsuperscript{22} The first positive mode peaks at $G \sim 75$ and thereafter decreases, while others become positive as well. The critical point at $Pr = 7.5$ was found to be $G_c = 14.2313$ and $\alpha_c = 1.55816$ which is in good agreement with Chandrashekar[2].

In Figure 5.12 we see that the critical wave number is constant in $Pr$, while $G$ varies inversely.\textsuperscript{23} This implies that the critical Rayleigh number $Ra_c = G_c \times Pr$ is also constant in $Pr$, as it is known to be. In fact the whole solution for $\lambda(G, \alpha)$ scales with $Pr$, so that the contour plot for $Pr = 10$ shown in Figure 5.13\textsuperscript{24} can be scaled in $G$ to give that for any value of $Pr$.\textsuperscript{25}

Although the scales are very different from Figure 5.7 comparison shows that this system is unstable for much larger values of $\alpha$ as well as for much smaller values of $G$. We see that the result of losing the Jacobian terms leads to real modes and a very simple variation with $Pr$.

\textsuperscript{22}See discussion of linear hierarchy Section 4.4.
\textsuperscript{23}The spikes at low $Pr$ in the $\alpha_c$ plot are due to numerical inaccuracies and are very small compared to $\alpha_c$ itself.
\textsuperscript{24}The contour interval is .004.
\textsuperscript{25}As we shall see below this does not hold for the phases of the eigenmodes for all $Pr$. 
Figure 5.12: Benard critical values vs. Pr.

Figure 5.13: Benard $\lambda(G, \alpha)$ contours at $Pr = 10$ to $G = 100$, $\alpha = 10$. The data ranges from $-.1024e+2$ to $+.3604e-1$ in increments of .004.
5.5.4 Benard with Shear Flow

A second difference between the Benard and the VSC problems lies in their respective symmetries. In order to do a more realistic comparison and display the result of changing the configuration from a vertical to a horizontal slot, we assume a background shear flow in the Benard configuration. Physically this can be applied to many problems in fluid pipe flows or atmospheric physics[64].

The results in this case are considerably more complex. For low $Pr$ they are very similar to those for the VSC problem, again indicating that these modes are kinetic, depending on the shearing in the center of the slot and not on the thermal equation or how it couples to the kinetic equation. We can see

Figure 5.14: The Benard-with-shear-flow critical values.

from the plots of the critical values against the Prandtl number, Figure 5.14, that in a qualitative way the values for $Pr < 0.1$ are similar to those for the VSC problem (see Figure 5.6). At higher $Pr$, however, these curves are quite
different. \( G_c \) falls farther and never rises again approaching an asymptotic limit of \( \sim 305 \) at large \( Pr \). Meanwhile the critical wave number plot, while still

Figure 5.15: Benard with shear flow \( \lambda(G, \alpha) \) contours. For details see the text.

having two bumps, rises much higher at the initial peak (near \( Pr = 0.1 \)) and only slightly at the second (around \( Pr = 1.0 \)). This is the inverse of the VSC case. Also in the present case the large \( Pr \) value \( \alpha_c = 1.327 \) is slightly lower, rather than quite a bit higher, than the low \( Pr \) limit as was true for the VSC
case.

In Figures 5.15 and 5.16 we show contour plots of $\lambda(G, \alpha)$ and $\omega(G, \alpha)$ for Prandtl numbers of 0.1, 7.5, 15.0, and 50.0. In both figures the Grashof number ranges up to 10,000 in the first two and 1,000 in the last two. In the top two $\alpha$ ranges up to 10, and in the bottom two, up to 5. In Figure 5.15 the data ranges to .1991e-1 in the top two (compare to .1958e-1 in the VSC plots) and
to .1360-1 in the bottom two and the spacing is .004 and .001 respectively. In Figure 5.16 the data ranges to .5901, .6350, .3092 and .3134 successively. The spacing is .002 in all four frames. Note that the \( Pr = 0.1 \) and 7.5 maximum values again compare closely to those in the VSC case.

We see clearly that the low \( Pr \) behavior is that same as that for the VSC configuration (compare to Figures 5.7 and 5.8). But although with increasing \( Pr \) the complex modes reduce to a single dominant one, and although this one mode does curl around under the stationary mode at higher \( Pr \), it does not become unstable! Therefore, while one might expect to see stationary modes develop in this system, there is nothing to break the horizontal symmetry along the slot and thus lead to traveling modes.

We see from these comparisons that the VSC stationary mode arises from the center-slot shear and is a kinetic mode, driven only indirectly by the thermal gradient. The gradient that gives rise to the background shear flow and is therefore the indirect source of energy for the stationary mode. The development of a supercritical complex, or traveling mode, at larger \( Pr \) is seen to depend on the specific geometry of the buoyancy force, rather than the interaction of the thermal gradient with the shear flow.

5.5.5 The Zero-state VSC Case

The results of this case parallel those for the Benard problem, as was expected; although \( \alpha_c = 1.266 \) is slightly smaller, it is constant, while \( G_c \) is again inversely proportional to \( Pr \), although roughly 4 times larger at a given value of \( Pr \).

As in the Benard case, the plot for \( \lambda(G, \alpha) \) shown in Figure 5.18 can be scaled in \( G \) to give the correct plot for any value of \( Pr \). The qualitative
Figure 5.17: The Zero-state critical values.

Figure 5.18: Zero-state $\lambda(G, \alpha)$ Contour at $Pr = 0.75$ to $G = 10,000$, $\alpha = 10$. The data range from $-1.03e+3$ to $0.545e-2$ with interval $0.004$. 
form of this plot is similar to that of Figure 5.13, although the magnitude of \( \lambda \) is generally smaller. This implies that in the growth of background shear-flow in the VSC case occurs with a longer time-scale than does the growth of cells in the Bernard configuration. Comparing the graphs of \( \lambda(G, \alpha) \) for the sheared cases with those that have a motionless initial state, we can see that the shearing stabilizes the high \( \alpha \) modes, limiting the wave number to \( \alpha \leq 2.5 \). This yields a maximum cell aspect ratio of \( \sim 1.26 \) as we saw in Section 5.2. Where there is no shear flow the wavenumber can be much larger, especially at high \( G \), and thus we can have much smaller cell aspect ratios. If there is a background shearing it will stretch out any secondary flow causing it to have a longer wavelength but a smaller wave number.

### 5.6 Eigenmode zoology

We now come to the eigenmodes that go with the eigenvalues. As was the case with the eigenvalues, we work, at any given set of values for the parameters \( \{G, Pr, \alpha\} \), with the mode that has the largest value of \( \lambda \) (the real part of the eigenvalue). It is for this reason that we talk of different modes at different values of the parameters \( G \) and \( \alpha \), even though mathematically they may all be present over the whole \( G-\alpha \) plain.\textsuperscript{26}

There are, especially at low \( Pr \), quite a variety of different complex modes and even apparently of real modes (when overall mode phase is taken into consideration). In trying to sort these out we have made use of several properties, based on one last diagnostic, a measure of the phase of the mode.

\textsuperscript{26}We will usually refer to a complex pair of eigenmodes as a single mode, except when it is relevant to compare them to each other. Thus at any given set of parameters there is only one 'mode' under consideration.
It should be understood that although mode properties can be used in combination to categorize the different modes, these 'modes' while retaining a general characterization do vary with changing $G$ and $\alpha$. And while there are instances of abrupt changes in these properties from one mode to the next, there are also cases of a slow transformation with no distinguishable discontinuity.

We shall first define and outline the properties used for the classification and then give an outline of the type and range of modes with emphasis on general characteristics and parameter dependencies. This outline will roughly follow the hierarchy of equations discussed in Section 4.4 but arranged here so as to move from the simplest modes to the more complex.

Lastly we note that all mode plots are contour plots of the respective field with a width equal to that of the slot and a height equal to a single vertical wavelength. Thus although they are plotted as squares they represent aspect ratios of $\lambda/2 = \pi/\alpha$.

### 5.6.1 Analytic Tools

- **Horizontal Wavelength.**

  The horizontal wavelength of the mode refers to the width of the positive and negative vortices in the eigenmode. A single vortex the width of the slot (e.g. Figure 5.19 on the left) has a wavelength of twice the slot width. This is true for all the real modes and many of the complex modes, but there are also modes with wavelengths equal to the slot width, one half the slot width, or even less. As used here this is not a strict uniform wavelength, as adjacent vortices in the same mode may have different widths. In the complex modes one may find same-sign vortices next to each other and still talk of the 'wavelength' even though the cross section
is more like two half waves than a single full wave.

In Figure 5.19 we show two examples. The first is the full slot vortex referred to above and the second is a complex mode that has a half slot wavelength. These are both stream function eigenmodes, as this is the

![Stream Function Diagrams](image)

Figure 5.19: Sample modes of different horizontal wavelength.

field we use for determining the 'wavelength'. The vorticity fields always have a wavelength that is less than that of the stream function, while the temperature modes are always of the same wavelength.\(^27\)

In the wavelength halving process the two halves of the mode may be pictured as coming out of phase (vertically) with each other as the parameters vary. When they are sufficiently out of phase (on the order of \(\frac{\pi}{2}\)) then the mode appears to have doubled its horizontal frequency. It is not presently clear to what extent this internal phasing remains the same as the modes evolve and to what extent these parts all move independently.

Recalling that the complex modes come in pairs, we usually show plots of only one half of the pair. Thus, as in the right-hand side of Figure 5.19, the mode is considered as only filling half the slot width, even though it

\(^{27}\)There is one exceptional case at \(Pr = 1.0\) of a small group of modes whose temperature field also has half-slot wavelength.
tails all the way across.

**Frequency Doubled Modes:** These are those complex modes where each member pair has a wavelength equal to half the slot, as in the right-hand plot in Figure 5.19.

**Stretching:** When we talk of 'stretching' or 'pulling back' of modes this refers to changes in the horizontal distribution of the mode in the slot in which the fields pull back towards the walls, stretching out in the middle of the slot. There is a general, though not absolute, tendency for modes to pull back with higher $G$ and higher $\alpha$, reflecting the evolution towards boundary layer physics in these regimes. At large $Pr$ the thermal fields pull in but the kinetic fields remain the same, or expand slightly.

- **Phase Properties.**

The second property of the modes is what we have called their phase. There are actually several properties that all relate to the phase of the eigenmodes, which is derived from the ratio of the real and imaginary parts of the fourier-physical representation of the modes. This means that having solved for the mode in spectral space we transform it out of Chebychev space horizontally, but before doing the vertical Fourier transform we find the largest element and take its phase. The 'phase' of a mode (including all three fields) is defined from the temperature field, as this is the only one that always has a single clear peak horizontally. This is because all its complex modes have a horizontal wavelength equal to the slot, only half a wavelength appears in each of the pair of modes.

Note that in comparing the phases of the different variable fields of a single complex mode we always take the phase of the $\lambda + i\omega$ half of the
pair, rather than the $\lambda - i\omega$ half. This is because the phase flipping\textsuperscript{28} of one of the fields but not the other will give the negative $\omega$ fields a different relative phase. The phase of the $\lambda - i\omega$ half can always be deduced from the relative phase of the positive $\omega$ fields and the phase flipping of the mode.

**Overall phase:** This is the phase of the temperature field which is used as a base for comparing modes to each other, as opposed to comparing the fields within each mode (see below). We adopted this analysis after

![Figure 5.20: Examples of Phase, 0 and $\pi/2$.](image)

noticing that with small variation there is a strong tendency for the modes to have half-integral $\pi$ vertical phases. While the overall phase in an infinitely long slot may not have a physical significance, it does indicate something mathematically about the solutions, and the limited number of cases make this a useful cataloging tool. In Figure 5.20 we show two fields with phases of zero and $\pi/2$ respectively. Multiplying these fields by $-1$ would give fields with phases of $\pi$ and $-\pi/2$.

**Relative Phases:** The relative phases are the phase differences between

\textsuperscript{28}For phase flipping see below.
the different fields of the same mode. The relative phases of these fields are more likely be of physical relevance than the overall phase. As discussed in Section 4.4.3 there is for most modes a fixed pattern of phase differences, and in fact the phases of the kinetic variables always differ by $\pi$ as we expected. But while there are some modes where the phase difference between the $\psi$ and $T$ fields is $\frac{\pi}{2}$, there are a few where there is no difference. These are call in-phase modes.

Flipped Phase: A third type of phase comparison involves the complex pairs of the complex modes. These pairs can be compared by taking one and flipping it about the vertical and horizontal centerlines of a cell one wavelength high and the width of the slot. This procedure is equivalent to rotating the mode by 180 degrees about a point in the center of its unit cell. When we compare the rotated mode with its partner it is found to be always either in phase or out of phase by half a wave length. In all cases either the kinetic pairs are in phase and the thermal fields not, or vice versa. In Figure 5.21 we show a pair of $\psi$ fields that are, in the sense

![Figure 5.21: A complex pair of fields.](image)

of flipped phase, out of phase with each other. Note that the downward traveling mode on the left side of the slot is always the one associated with a positive $\omega$ while the upward traveling mode on the right-hand side
has a negative $\omega$.

5.6.2 The Navier-Stokes Case

The mode structure for the Navier-Stokes equations with no background flow, $W_b = 0$, is simple. As explained in Section 4.4.1, there are only real stable modes, and we have solutions for only the kinetic variables. The $\psi$ fields all have a horizontal wavelength of twice the slot width, while the vorticity has a wavelength that is half this. In Figure 5.22 we show a typical example.

![Figure 5.22: A typical Navier-Stokes eigenmode.](image)

This is essentially the mode for this Navier-Stokes case. These fields can be well approximated with the proper combination of sines and cosines.\(^{29}\) There is a stretching out of the fields as $\alpha$ increases, but they are unaffected in this way by variation in $G$. This stretching is most noticeable in the vorticity where the quarter wavelength features near the walls narrow as $\alpha$ increases and the

\(^{29}\)We must remember that these functions are distorted by the non-uniformity of the grid.
Figure 5.23: The Navier-Stokes phase variation to $G = 10,000$, $\alpha = 10$. Data from -3.142 to +3.142.

central features widen to fill nearly the whole slot. The Navier-Stokes eigenmodes have predominantly the same phase as the mode of Figure 5.22, where the $\psi$ field has a phase of zero. For $\alpha > 6.67$, however, the phases change by $\pi$ as shown in Figure 5.23.\footnote{The lozenges in this plot are due to numerical noise.} Note that for these large $\alpha$ modes the magnitude of $\omega$ changes proportionately to that of $\psi$ rather than inversely. This is related to a minus sign due to the change in phase.

5.6.3 The Benard Problem

The mode structure in the Benard case\footnote{Note that the plots of the Benard configuration cases are still oriented so that the slot appears vertical, not horizontal. The gravitational field in these cases, however, is also horizontal. To view these in their proper physical orientation one must turn them so that the left side is up.} is also simple as there is still no background shear flow. It is nearly the same as the Navier-Stokes case.
described above but with the addition of an equally simple temperature field. Again all the modes are real, although they are no longer uniformly stable. The $\psi$ and $T$ fields all have a horizontal wavelength of twice the slot width

![Eigenmodes Diagram](image)

Figure 5.24: A typical near critical Benard eigenmode.

while the vorticity has a wavelength that is half this. In Figure 5.24 we show a typical example.\(^\text{32}\) Comparison with the Navier-Stokes case above shows the main difference being a shift in the phase of the kinetic fields. This is essentially the mode for the Benard case. These fields can be well approximated with the proper combination of sines and cosines as is done in the well-known Lorenz formulation [65]. The horizontal stretching out of this mode has the same variation as the Navier-Stokes case.

The Benard eigenmodes have a slightly more complicated phase structure. As with the Navier-Stokes case one mode dominates. This mode is shown

\(^{32}\text{This particular example is the critical mode for } Pr = 7.5.\)
Figure 5.25: The Benard low $Pr$ phase variation.

in Figure 5.24 and has an overall phase of zero. For low $Pr$ (starting at about $Pr = 1.0$), however, there are some low $G$ and high $\alpha$ modes with different overall phases. A detailed examination at $Pr = 1.0$ down to $G = 1.0$ reveals non-zero phase modes but only for very small $G$. Thus we expect that this is roughly the upper limit of these modes. For higher $Pr$ they are absent, while for lower $Pr$ they cover more of the $(G, \alpha)$ plane and below about $Pr = 0.5$ maintain the pattern seen in Figure 5.25, which shows the phases of the Benard low $Pr$ temperature eigenmodes for $Pr = 0.1$.\textsuperscript{33}

The low $G$, low and medium $\alpha$ mode has an overall phase of $\frac{\pi}{2}$, while the high $\alpha$ mode has an overall phase of $-\frac{\pi}{2}$. This pattern is identical at $Pr = 0.01$ and does not scale with $Pr$, unlike the values of $\lambda(G, \alpha)$. The transitions from one mode to the next are sharp discontinuities.

\textsuperscript{33}In this figure the modes with non-zero phase exist at values of $G(\alpha)$ about ten times as large as those reached at $Pr = 1.0$. 
5.6.4 The Zero-state Case

The main difference between the eigenmodes in this case and those of the Benard case is that these have a non-zero horizontal phase gradient, and appear tilted.\textsuperscript{34} Although there is no background vertical flow here the buoyancy term coupling the vertical gravity with the horizontal thermal gradient causes the fluid near the warm wall (on the right) to rise as it circulates across the slot, while the fluid near the cool wall (on the left) falls as it circulates. This leads to a component with odd horizontal symmetry and the tilting of

![Eigenmodes](image1)

![Stream Function](image2)

![Temperature](image3)

![Vorticity](image4)

Figure 5.26: A sample Zero-state mode.

the modes, as seen in Figure 5.26. Comparing with the Benard mode shown in Figure 5.24, we see that aside from its tilt the Zero-state mode is qualitatively the same. The fact that $G$ is so large for the Zero-state mode shown here leads to the dramatic tilt. For $G \rightarrow 0$ the tilt also goes to zero.

\textsuperscript{34}We refer to this as the 'pitch' of a mode. It can be determined by the ratio of the odd over the even horizontal components of the mode.
The low $Pr$ phase structure seen in the Benard case is also found in the Zero-State problem. It is the exactly the same as that in Figure 5.25 for all $Pr$ up to $\sim 0.5$ above which it starts retreating to smaller $G$, until it disappears between $Pr = 0.75$ and $Pr = 1.0$, slightly lower than in the Benard case. On the other hand at larger $Pr$ (and probably for very large $G$ and $\alpha$

![Figure 5.27: The high $Pr$ Zero-state phase structure.](image)

even at smaller $Pr$), there are new modes with a phase of $-\frac{\pi}{2}$. Unlike the low $Pr$ non-zero phase modes, the high $Pr$ phase structure does scale with $Pr$, like the eigenvalues.

In Figure 5.27 we show the Zero-state phase diagram at $Pr = 7.5$ with a band of $-\frac{\pi}{2}$ phase modes. Scaling $G$ inversely with $Pr$ will give the structure for other values of $Pr$. If the pitch of the modes is plotted the phase boundaries of Figure 5.27 are found to lie along equal pitch lines. These curves drop rapidly at low $\alpha$, then bottom out and turn upwards again, rising with increasing speed at larger $\alpha$. A line can be drawn from the origin through the turning points of these equal pitch lines which is roughly parabolic ($G \propto \alpha^2$). The width of this
parabola varies as $Pr$. The pitch varies smoothly as $G$ and $\alpha$ are increased or decreased with no discontinuity (other than due to measurement technique) at the phase discontinuities.

5.6.5 The Pr=0 Limit

This case is both the limiting case of the full VSC problem as $Pr \to 0$, and at the same time equivalent to the Navier-Stokes problem with an initial state given by the VSC 1-D solution. Since the modes in this case are the same as those for low Pr in the full VSC case we shall not discuss them separately here.

5.6.6 The VSC Problem

With the inclusion of the background shear flow so that the rest of the Jacobian terms are non-zero, we obtain complex solutions as well as real ones. This gives us a new type of mode, traveling modes, made up of complex pairs of modes that ride on the shear-flow background, traveling up the warm side and down the cool side of the slot.

In the following discussion please refer back to Figures 5.7 and 5.8 in Section 5.4 above for the relevant values of $\lambda$ and $\omega$. The plots of $\lambda$ show the stable and unstable regions as discussed above. Likewise the plots of $\omega$ indicate the regions of real and complex modes which are similar to but not identical with the stable/unstable regions.

Mode Diagnostics.

In this section we add Figures 5.28 and 5.29 presenting plots of the overall phases and the structure of the relative and flipped phases described
Figure 5.28: The overall phases of the VSC modes.
in Section 5.6.1. The first of these figures shows the regions of the four main phases; zero in yellow, \( \frac{\pi}{2} \) in blue, \( \pi \) in red, and \( -\frac{\pi}{2} \) in green. Finally there are regions at low \( Pr \) where the phases are not half integrals of \( \pi \) but are somewhere in between. These are indicated by the striped regions that are especially frequent around \( Pr = 0.03 \).

At low \( Pr \) there are a profusion of different phases the number of which decrease as \( Pr \) increases until by \( Pr \sim 0.76 \) the only remaining modes have zero phase.\(^{35} \) This zero-phase mode can be seen near the real/complex boundary at high \( G \) in the \( Pr = 0.001 \) plots, and with increasing \( Pr \), it spreads out first at low \( \alpha \) and then to high \( \alpha \) until these zero-phase modes are the largest everywhere. This expansion correlates with the movement of the \( \omega \) discontinuities as seen in Figure 5.8. The low \( G \) and high \( \alpha \) structure for \( Pr = .1 \) to \( \sim .7^{36} \) is similar to that seen in the Benard and Zero-state cases (see Figure 5.25).

We next turn to Figure 5.29 which serves to describe internal phase properties of the complex modes. The complex modes can in all cases be divided into those which when flipped are \( \pi \) out of phase in \( T \) (indicated in red) and those for which the kinetic variables are out of phase (indicated in yellow). These later modes make up the majority, and at \( Pr \sim 0.5 \) and above all the modes have this property, while the \( T \) flipped phase modes are most numerous at small \( Pr \). The next property is wavelength halving (indicated by the green outline). All the real modes and those complex modes that appear

\(^{35} \)At low \( Pr \) the exact mode structure is sensitive to the number of modes used in the eigenmode problem. The general structure is constant, especially at low \( \alpha \), but the details vary.

\(^{36} \)The reference is to the \( \frac{\pi}{2} \) phase (blue) mode at low \( G \), the \( \frac{\pi}{2} \) phase (green) mode at lower \( G \) and high \( \alpha \) and the main zero phase (yellow) mode in Figure 5.28.
Figure 5.29: The Internal Phase Structure of the VSC Modes.
for $\alpha \approx 2.4$ at all $Pr$ have a horizontal wavelength that is equal to the width of the slot. This is also true for low $Pr$ up to about 0.1 for all modes except a small vertical strip that corresponds to a complex mode at $G > 5000$ and $\alpha \approx 2.5$. This mode, which lies on the real/complex boundary, has wavelength that is equal to half the width of the slot. Above $Pr \sim 0.1$ it spreads out to higher $\alpha$ and lower $G$, until at larger $Pr$ all modes for $\alpha > 2.0$ have this property. Note that the low $\alpha$ unstable complex modes that appear at high $Pr$ do not have this property.

Finally there are modes where the temperature field and the stream function field are in-phase (shown in blue stripes) rather than being out of phase by $\frac{\pi}{2}$ as we had expected and as is the case most of the time. These modes appear in a band to the right (higher $\alpha$) of the wavelength halved modes and are swept out in front of them as they unfold with increasing $Pr$.

This completes the study of the properties which combine to give a fairly clear description of the modes. There is a general tendency for modes and properties that appear at high $G$ along the real/complex boundary at low $Pr$ to sweep out and down as $Pr$ increases, first the zero phase, then the in-phase band and finally the wavelength halving property.

In Figure 5.30 we plot the critical numbers again (for scales see Figure 5.6) and indicate with bars the ranges of the properties and mode types discussed above. The low-$\alpha$ expansion of the zero-phase mode starts near $Pr = 0.01$, as soon as the effects of the nonlinear terms in the thermal equation begin to be felt and pull the critical values away from their low $Pr$ values. The jumble of high $G$, mid-$\alpha$ complex modes also begins to change at this point and the pitch of the critical real mode begins to decrease. At $Pr \sim 0.03$ the in-phase modes first appear and the zero-phase mode begins its sweep to the
Figure 5.30: The $G_c$ and $\alpha_c$ curves with mode property regimes. At $Pr = .07$ the low-$\alpha$ modes are all zero-phase and the pitch of the critical mode temperature field goes through zero and starts to grow with the opposite sign. Ruth [22] decomposed eigenmode solutions into horizontally odd and even components the ratio of which gives the mode pitch. He found the odd component of the temperature field passing through zero just below $Pr = .07$ in agreement with our observation of the mode pitch. At $Pr = 0.1$, the $G_c$ minimum, Ruth's inversion of the odd thermal component is complete (the reversed pitch is now equal in magnitude to its low $Pr$ value), and the expansion of the half-slot wavelength complex modes begins. As $G_c$ increases again Ruth finds a further increase in the odd thermal field until at $Pr = 0.5$, where $G_c$ peaks it has the same relative magnitude as that of the stream function field. We find that overall the pitch of the critical mode stream function is essentially independent of $Pr$, while that of the temperature changes continuously. For $Pr > 0.5$ it is larger in magnitude than the kinetic pitch, but it is inverted, sloping up to the hot wall. The $G_c$ peak also coincides with the completion of the zero-phase mode expansion, and the disappearance of the
in-phase and T-out-of phase complex modes. While many of the low and mid
Pr modes are gone the low α complex modes start to appear subcritically at
Pr ~ 0.5, although they don’t become unstable until Pr ~ 12.5. Lastly the
zero-phase mode expansion is complete at Pr ~ .75 the peak in $\alpha_e$ and the half
wavelength expansion is finished by $Pr = 1.0$.\textsuperscript{37}

It is not surprising to see that we may describe three regimes; a low
Pr regime, up to $Pr \sim 0.01$, where the flows are primarily kinetic, a high Pr
regime, above $Pr \sim 0.75$ in which the thermal effects are fully coupled in and
there is a single stationary and two complex modes, all with zero overall phase.
In between there is a transition regime where the shifting balance of terms
leads to a series of changes in the modes with changing Pr. This regime can be
classified again into two sub-regimes associated, respectively, with the decline
and the rise of $G_e$ below and above $Pr = 0.1$.

5.6.7 Sample VSC Modes.

We will now give some illustrative examples of the specific modes and
their range of form, starting with the small Pr kinetic modes and noting how
they change as Pr increases.

$Pr=0.001$:

In Figure 5.31 we show a typical real mode which is similar to those of
the zero-state case, although the pitches here tend to be much less. While some
of the real modes have a different phase, they are (at low Pr) all of this general

\textsuperscript{37}Ruth found a rapid reversal in sign of all mode components at the $Pr \sim 2.0$ minimum in
$\alpha_e$. We however find no noticeable change in the modes at this point. In particular a change
in overall phase by $\pi$, which is equivalent to his sign change, is not found there, although
there is one near $Pr = .07$ along with the zero in the pitch of the temperature mode. As we
noted above this signals the passage of the front of the expanding zero-phase mode.
Figure 5.31: A low $Pr$, real VSC mode.

Figure 5.32: A low $Pr$, complex VSC mode.
form with the exception of a vertical strip between the neutral stability curve ($\alpha \sim 2.0$) and the real/complex boundary ($\alpha \sim 2.5$) where the real modes have a reverse pitch. At low $G$ there is a continuity in form between these modes and the $\lambda - i\omega$ modes at slightly higher $\alpha$. This reverse pitch is also found to some extent in most complex modes, especially those near the real/complex boundary.

![Eigenmodes and Stream Function](image)

**Figure 5.33:** The arrowhead mode.

Figures 5.32 and 5.33 show the range of variation of the typical full-slot wavelength complex mode that is common at lower Prandtl number. With increasing $\alpha$ the reverse pitch decreases, while increasing $\alpha$ and/or increasing $G$ as well causes the form to become more compact, centering on the line where the background vertical velocity is greatest. We can also see that this line moves closer to the walls at larger $G$.

In Figure 5.34 we see a variation found near the real/complex boundary which has been drawn out vertically. Note the strong reverse pitch of the
EIGENMODES
The Vertical Slot Problem

Figure 5.34: A long thin mode.

Figure 5.35: A proto half wavelength mode.
kinetic fields, which fades at higher $\alpha$ and is only found at high $G$. In Figure 5.35 we have another mode near the real/complex boundary, but at high $G$, that is drawn out in the other direction and its pitch is 'normal' rather than 'reversed'. Comparison with the half-slot wavelength mode in Figure 5.36 shows that this is a transition mode with ill-defined 'wavelength'. In fact, with

![Figure 5.36: A half wavelength mode.](image)

a proper path in parameter space (and excluding the shift from one phase to another) these mode forms can be seen to change continuously from one to another.

*Mid-Range $Pr$ Special Cases:*

The proceeding survey of modes at $Pr = 0.001$ gives a good summary of most of the fields encountered at larger $Pr$ as well. There are however several special cases found for middle range $Pr$ that should be mentioned. In general a look at the 'characteristics' of a mode as outlined above and reference to the examples given will yield a good idea of what it looks like. The one example
Figure 5.37: The thermal half wavelength mode.

Figure 5.38: A zero relative phase mode.
of a completely different mode is shown in Figure 5.37. It is an example of a small group of modes at $Pr = 0.1$ for which the temperature as well as the $\psi$ field has half-slot wavelength. We mentioned above that for a middle range of $Pr$, from perhaps 0.025 to 0.75, there are a band of modes with a zero relative phase between the $T$ and $\psi$ modes. One of these, from $Pr = 0.25$ is illustrated in Figure 5.38.

$Pr=7.5$ Modes

In studying the VSC problem the Prandtl number for water is often chosen both for its intrinsic interest and as a good basis for comparison to previous work. It is high enough to be out of the mid range where all the complexities lie and thus for our purposes here serves to illustrate the typical high $Pr$ modes.

Here there are essentially three types of modes, one real mode that is mostly unstable and two complex modes, one of which becomes unstable for slightly higher values of $Pr$. In Figure 5.39 we show the linearly critical mode (note the value of $\lambda$) which agrees very well with previous results.\footnote{For comparison see [17] and [20].} This form is typical of the real modes for higher $Pr$. The pitch is strongly reversed in the temperature field, which is also confined to the center of the slot. In the stream function field we can see that there are now two maxima on either side a slight central minima. Ruth’s study—citeRu:79 indicates that these changes in profile start from $Pr \sim 0.5$. The central concentration of the temperature perturbation reflects the decreasing width of the full flow thermal boundary layers (see Section 3.4). Comparison with Figure 5.31, coupled with the discussion of mode properties at the end of Section 5.6.6, allows one to
Figure 5.39: The VSC linear critical mode at $Pr=7.5$. Extrapolate the form of the real modes for the whole middle range of $Pr$. Note that the differences in the vorticity field have more to do with the difference in $G$ than that in $Pr$.

The two complex modes may be called 'low $\alpha$' and 'high $\alpha$' with the dividing line at about $\alpha = 2.0 - 2.5$. It is important to note that even for parameters where these modes are stable, they will appear as transients. The complex modes appear in pairs of right and left, upward and downward, traveling modes.

The high $\alpha$ mode is a typical half-wavelength mode with kinetic flipped-phase. The temperature field is an arrowhead type. In comparison to the $Pr = 0.001$ example, Figure 5.36, the strength of the two sets of vortices in the stream function field (one along the wall and the other bounding the center-slot line) has shifted so that the wall vortices are more prominent than before. This mode is illustrated in Figure 5.40. Comparison with the
Figure 5.40: A typical high $\alpha$ high $Pr$ complex mode.

Figure 5.41: A typical low $\alpha$ high $Pr$ complex mode.
$Pr = 0.001$ mode shown in Figure 5.32 again shows how the temperature field has pulled in to the wall while the kinetic boundary layer remains about the same (note that the wave number is slightly higher in Figure 5.40, contributing to the reduced width of both fields).

Finally the low $\alpha$ mode, seen in Figure 5.41, is the one that becomes unstable first at $Pr > 12.5$. The exact eigenmode shown here, however, being at lower $Pr$ is stable, linearly and nonlinearly. It forms the first transient seen in the nonlinear simulations.

![Eigenmodes](image)

Figure 5.42: The complex critical modes.

For values of $Pr > 7.5$ the modes are essentially the same, although they hold more tightly to the lines of maximum $W_b$. As our last example in Figure 5.42 we show the complex critical mode pair at $Pr = 12.45$ when this mode has the same critical Grashof number as the real mode. This serves as an illustration of a complex pair of modes as they would appear together in a

---

39 Compare traveling modes in [17] and [50].
simulation or experiment. One would expect at these values of $G$ and $Pr$ to actually see the stationary mode as well and that there might be resonances between them.

5.6.8 Benard with Shear Flow

This case presents no new modes. It has the same modes as the VSC problem with the exception of the high Pr, high $\alpha$ modes we had above.
Chapter 6

Theoretical Background to the Nonlinear Analysis

In the previous chapters we have examined the stability of the 1-D VSC solution to perturbation using a linear stability analysis. We have found the modes that have the largest real eigenvalues and thus are expected to be the visible growth modes or decaying transients in the full nonlinear system. We will now turn our attention to the nonlinear VSC problem seeking its steady state solutions and the dynamical behavior of the system, the way in which unstable solutions evolve to secondary states.

Before we proceed to discuss our simulation results we need to review some theoretical methods that will provide a framework for discussion and analysis.

6.1 Notes on Bifurcation Theory

The goal in all our work is to be able to model physical systems, to be able to predict and explain behavior that can be realized in nature or the laboratory. In order to do this one wishes first to know what the steady state solutions to the problem are, solutions which will persist in time, that are supported by the equations after the initial conditions have been forgotten and transients have died away. Bifurcation theory attempts to classify and characterize these solutions, along with their stability or instability to various perturbations and, in the case of instability, the nature of the transitions from
one state to another.\footnote{See D. Joseph [66].}

In its most abstract form this theory arises in the study of the solutions of systems of differential equations. Such a system can be written schematically as $\dot{x} = f(x)$, where $f$ is a matrix differential operator. In our case $x$ is a vector made up of $\psi$, $\omega$, and $T$. The operator $f$ is often a function of several parameters (in our case $G$, $Pr$, and in the numerical problem, implicitly, $\alpha$), which as they are varied will change the form of the operator.

A common starting point in studying such a set of equations, and a very natural one when we are interested in their physical content, is a look at the solutions for $f(x) = 0$ which are call \textit{fixed points} and represent equilibria or stationary (steady-state) solutions.\cite{67} Such a stationary solution that exists for all values of the parameters of the problem is called a \textit{basic} solution.\cite{68} This solution, however, is not usually stable for all parameter values. In particular we chose one parameter which physically is related to the forcing of the system, in our case the Grashof number which is proportional to the temperature gradient. This is called a control parameter. At small values of this parameter the basic flow is stable and, at least in hydrodynamic problems, it is usually a unique solution with the maximal symmetry allowed by the geometry of the problem and its boundary conditions.\cite{66} As the control parameter is increased, this solution loses stability, and other solutions come to exist. Usually these new solutions split off from the old ones, a process that Poincaré called \textit{bifurcation}.\footnote{This is not always the case. There can be \textit{isolated} solutions that do not branch off from other solutions and cannot be found by tracing bifurcations. One should also note that in many physical cases bifurcations decouple, so that rather than splitting into two or three solutions, one is always preferred, and the others can only be reached by finite perturbations \cite{69}.} There is now a large body of theory that allows the mapping of the stationary
solutions versus the control parameter to produce what is called a bifurcation diagram. Analyzing the stability of each branch, the points where new solutions split off, called bifurcation points,\(^3\) can be found and characterized and these new branches can be mapped in turn.

In order to explore the structure of the solutions of a set of equations,\(^4\) we write them in the form

\[
\frac{\partial x}{\partial t} = F(\mu, x) = A_{ij}x_j + B_{ijk}x_jx_k
\]  

(6.1)

showing the dependence on the control parameter \(\mu\). When the basic solution is given by \(x = 0\) the linear problem is contained in the matrix \(A_{ij}\), the eigenvalues of which are the linear growth rates.\(^5\) Graphing these on the complex plane one finds that for small values of the control parameter they are all to the left of the complex axis, having negative real parts. As the control parameter increases they move to the right, until one or more of them crosses the complex axis and the real parts become positive. This value of the control parameter is called the critical value and marks a bifurcation of the basic solution. The character of the eigenvalue that first crosses the complex axis determines the type of bifurcation. If it is purely real, \(\omega(\mu_c) = 0\), then the basic solution loses its stability to a new stationary solution in a process called an exchange of stability. This is the type of bifurcation seen at \(G_c\) in the VSC problem. It is also possible that a pair of

\(^3\)These points correspond to the physical critical points, and we will use both terms depending on the context.

\(^4\)Iooss and Joseph [70] provide a good mathematical introduction to stability and bifurcation theory.

\(^5\)When this is not true, or when investigating secondary and higher bifurcations, then Equation (6.1) can be rewritten in terms for new variables

\[
\dot{x} = \bar{x} - x_b
\]

where \(x_b\) describes the primary, or higher, branch solution which we are stability-analyzing.
complex conjugate eigenvalues, having $\omega(\mu_e) \neq 0$, can cross the complex axis simultaneously in what is called a Hopf bifurcation. In this case the bifurcated solutions are not stationary but oscillatory. Secondary bifurcations are often of this type as is that for the VSC problem.\footnote{See Chapter 8.}

There is a second type of classification of bifurcations based on defining an 'order' parameter, $\epsilon$, such as the magnitude of one of the variables $x$.\footnote{The system kinetic energy, the Nusselt number, and the perturbation amplitudes are examples of order parameters we will discuss below.} Solving the steady-state, time-asymptotic, problem for the order parameter

$$F(\mu, \epsilon) = 0$$

it is possible to define the function $\epsilon(\mu)$ which when plotted gives the bifurcation diagram. The form that this function takes at a critical value of $\mu$ determines the type of bifurcation. In Figure 6.1 we show the most common types of double-point bifurcations.\footnote{A double point of $F(\mu, \epsilon) = 0$ is a singular point ($F_\mu = F_\epsilon = 0$) through which two and only two solutions pass. See [70].}

![Figure 6.1: a: Supercritical b: Subcritical c: Transcritical types of double-point bifurcations.](image)

The supercritical or 'pitch fork' bifurcation reflects symmetries in a problem such that odd-order derivatives of the inverted function $\mu(\epsilon)$ all vanish[70]. The subcritical bifurcation exhibits finite
amplitude instabilities in regions that are stable to infinitesimal perturbations and also hysteresis loops formed by following one solution as \( \mu \) increases below the critical point and the other as it decreases. Hysteresis may also arise from the asymmetrical transcritical bifurcation.

Lastly we should note that these exact bifurcations can be broken, for instance by noise in a system, so that the solutions become isolated and while approaching each other at the critical value of \( \mu \) do not actually join or cross there. The study of isolated solutions as perturbations of bifurcation ones is known as imperfection theory.

6.1.1 Equilibria and Dynamics

In the use of our code and the study of fluid problems in general. There are two ways to approach a given secondary state. For experimentally one starts at a low value of the control parameter where there is a known basic solution and then increases it slowly watching how the flows change. This approach is equivalent to tracing out the bifurcation diagram starting at the base state and following it until it branches, then following the branch until it branches, etc. Of course the diagram can get complicated, especially at high control parameter values, and in the laboratory bifurcations will usually be imperfect (only one branch accessible through smooth transitions), and there may well be hysteresis and other complications. But the usual technique is to work with slow changes, smooth transitions, and remain near equilibrium states. This can be called quasi-static evolution.

In contrast with our code we generally set the Grashof number and

---

9 For a similar discussion see [61].
starting at the basic state, set it off through a host of non-equilibrium states to find a point on the nearest stable branch. This is something like cutting cross-country rather than sticking to the roads. Graphically it means tracing vertically up and down the bifurcation diagram rather than moving more or less horizontally along the branches. This is illustrated in Figure 6.2. Each

![Diagram of Stream Function]

Figure 6.2: Quasi-static and dynamic evolution

of these techniques has its own uses and teaches different lessons. They each have their own sets of issues, problems, and advantages, and we will point these out when appropriate. Most important, the first deals with equilibria and the second with non-equilibrium dynamics.

We introduce the distinction in methodology here because just as we can use our code in these two different ways we may also make use of our expansion analysis in two related ways. The first is to tell us something about how the perturbations grow and evolve by examining the form of the solutions $A(t)$. The second is to examine the $G$ dependence of the secondary states, in particular $A_{asm}$, and to study how the increase of $G$ affects the number of terms that are significant and the values of the coefficients.
6.2 Expansion Theory

In Chapters 4 and 5 we set up linearized equations and assumed that the perturbation amplitudes satisfied an equation of the form

\[ \frac{dA}{dt} = \lambda A, \]  

(6.2)

which had the solution \( A(t) \propto e^{\lambda t}. \)

We solved for the eigenvalues, \( \lambda, \) and for the form of the corresponding eigenmodes. In the process we discovered that above a certain critical Grashof number \( G_c, \) the basic solution is unstable to perturbation by modes with positive real \( \lambda. \) Initially these unstable modes will grow exponentially as indicated by Equation (6.2) but we know that in a nonlinear world they do not grow forever; they eventually saturate and settle into some new steady state flow.

Near \( G_c \) the amplitudes of the secondary fields, even at saturation, are small in comparison to the basic solution. In fact, even at values of \( G \) considerably above \( G_c \) this is still true.\(^{10}\) This suggests an approach used by various authors [61][19] [55] in which we represent our perturbation fields as the product of a time-dependent amplitude and a spatial form factor:

\[ \hat{\psi}_1(x, t) = A(t)\hat{\psi}_1(x) \]  

(6.3)

where \( \hat{\psi}_1(x) \) is taken to be the linear eigenfunction at a given \( G \) and \( \alpha \) and \( A(t) \) is the amplitude of the perturbation at time \( t \) after the perturbation takes place.\(^{11}\)

---

\(^{10}\)The amplitude \( A(t) \) is limited by its time-asymptotic value, \( A_{assm} \), which is of order \( \sqrt{\lambda} \) and proportional to \( \sqrt{G-G_c} \) (see Section 6.2.3). At \( G_c \) these are all zero.

\(^{11}\)This holds directly if we use the eigenfunctions as perturbations, while a general perturbation will rapidly take on the same form as the eigenfunction as the transients die out.
For small enough $A$ we interpret Equation (6.2) as giving the first term in a series expansion for the time change of $A$ in powers of $A$, the higher order terms modeling the nonlinear evolution while $A$ remains small enough. Just above $G_c$ we can neglect most of the higher order terms and arrive at useful expressions for the amplitude of the time-asymptotic secondary states $A_{asm}$, for the coefficients of the expansion terms, and for $A(t)$, that give insight into how the perturbations evolve as they grow into a secondary state. As we increase $G$ and move farther from $G_c$ we add more terms that will bring with them new solutions and make possible increasingly complex temporal behavior. When $A(t)$ becomes large enough however this expansion will no longer converge. Stuart[61] suggests that convergence will occur whenever $\lambda(\alpha G)^{\frac{1}{2}}$ is sufficiently less than unity.\footnote{Stuart suggests this in the context of Poiseuille flow which obeys the a similar Navier-Stokes equation.}

### 6.2.1 The Expansion for the VSC Problem

To derive the form of Equation (6.12) that is appropriate for our VSC problem,\footnote{Note that many of the results presented here are generally true for hydrodynamic problems, and in some cases for all formally similar dynamic problems.} reconsider, for a moment, the VSC perturbation equations:

\[
\frac{\partial \hat{\omega}}{\partial t} = -\frac{\partial (\hat{\omega}, \hat{\psi})}{\partial (x, z)} - \frac{\partial \omega_b}{\partial x} \frac{\partial \hat{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \hat{\omega}}{\partial z} - \frac{1}{G} \frac{\partial \hat{T}}{\partial x} + \frac{1}{G} \nabla^2 \hat{\omega} \tag{6.4}
\]

\[
\frac{\partial \hat{T}}{\partial t} = -\frac{\partial (\hat{T}, \hat{\psi})}{\partial (x, z)} - \frac{\partial T_b}{\partial x} \frac{\partial \hat{\psi}}{\partial z} + \frac{\partial \psi_b}{\partial x} \frac{\partial \hat{T}}{\partial z} + \frac{1}{GPr} \nabla^2 \hat{T} \tag{6.5}
\]

For a perturbation at a single wavenumber such that $\hat{\psi}_1 = A e^{ikz} \psi_1(x)$ and $\hat{T}_1 = B e^{ikz} T_1(x)$ we can write symbolic equations for $\hat{A}$ and $\hat{B}$, the funda-
mental mode amplitudes, as

\[
\frac{dA}{dt} = \lambda A - \alpha_1 J(\hat{\psi}_1, \hat{T}_1) \tag{6.6}
\]

and

\[
\frac{dB}{dt} = \lambda B - \alpha_2 J(\hat{\psi}_1, \hat{T}_1). \tag{6.7}
\]

In these equations the terms linear in the perturbations combine to yield the $\lambda A$ and $\lambda B$ terms, as we know from linear theory. The remaining nonlinear Jacobian term is second order in the perturbations. For the moment we shall assume that the magnitude of $\hat{T}_1$ grows at the same rate as the kinetic perturbations and focus on the kinetic equations.\textsuperscript{14}

Although at first glance the Jacobian terms would appear to be proportional to $A^2$, giving us equations of the form

\[
\frac{dA}{dt} = \lambda A + \alpha A^2 \quad \alpha < 0, \tag{6.8}
\]

a more detailed examination shows that the second-order terms give contributions to a change in the background ($k = 0$) and to the first harmonic of the fundamental, perturbation ($k = 2k_z$), but do not contribute to the fundamental itself. In turn however higher order contributions to the fundamental come from running these second-order terms back through the Jacobian, along with the fundamental again. This gives terms that are third-order in $A$ and contribute to the fundamental ($k = k_z$) and the second harmonic ($k = 3k_z$).\textsuperscript{15} The Jacobian generates higher and higher orders in $A$, producing terms that contribute to a change in the base flow ($k = 0$), to the fundamental ($k = k_z$),

\textsuperscript{14}The results will apply equally to the change of $\hat{\psi}_1$ and $\hat{T}_1$, as long as the equations are considered separately.

\textsuperscript{15}For a nice description of this process see [29].
and to all the higher harmonics of the fundamental \((k = nk_x)\), limited only by the time scale of the coupling and the time elapsed from when the perturbation was applied. Formally these results are obtained by Fourier expanding the perturbations (which initially consist of a single mode), inserting them into Equations (6.4) and (6.5), and collecting terms which contribute to the same harmonic. This produces an infinit series of equations, one for each mode. Here we concentrate on the equation for the fundamental.

Thus the \(J(\dot{\psi}, \dot{\psi})\) term contributes, not a single term of order \(A^2\), but by a series of terms starting with the third-order and including all higher odd-orders. There are, however, no even-order contributions to the fundamental mode. The even-order terms arising from the Jacobian will contribute to the equations for the even harmonics while the odd-order terms contribute to the odd harmonics.

The proper expansion for the fundamental mode is a series including only odd powers of \(A\)

\[
\frac{dA}{dt} = \lambda A + \frac{\alpha}{2} A^3 + \frac{\beta}{2} A^5 - \frac{\gamma}{2} A^7 \ldots. \tag{6.9}
\]

In general a similar equation holds for the even harmonics, while the odd harmonics are described by even-order series. In general for the \(n^{th}\) harmonic only terms of order \(\geq n\) in \(A\) would be present, and, correspondingly, to any given order \(n\) in \(A\) only harmonics with \(k \leq (n - 1)k_x\) will be present.

A similar derivation gives the same expansion for \(\frac{dF}{dt}\) although the constants may be expected to differ.
6.2.2 The Energy Equation: Landau Theory

In order to derive an energy equation we multiply through Equation (6.9) by $A^*$ to get

$$\frac{dA^2}{dt} = 2\lambda A^2 + \alpha A^4 + \beta A^6 + \gamma A^8 \cdots$$  \hspace{1cm} (6.10)

which is a series in even powers of $A$. This new expansion represents the time change of the energy of the fundamental mode, and as long as the contributions to the perturbation come mainly from the fundamental frequency, it is valid for the total perturbation energy as well. The $k = 0$ component to the perturbation produced by the nonlinear term will be fairly strong, however. It is the one component not subject to dissipation and, since its $z$ derivative is zero, couples weakly back into the higher harmonics as well. And, unlike the other modes which couple to themselves and feed energy into the $k = 0$ component, the $k = 0$ mode cannot lose energy by coupling to itself. As noted above, the $k = 0$ mode can be represented as a series in even powers of $A$:

$$\frac{d\psi_0}{dt} = a_1 A^2 + a_2 A^4 + a_3 A^6 \cdots$$

its kinetic energy contribution is given by

$$\frac{d\psi_0^2}{dt} = b_1 A^4 + b_2 A^6 + b_3 A^8 \cdots$$

which can be added to Equation (6.10) to give a better approximation to the kinetic energy of the perturbation without changing the dependence on $A$. This causes a shift in the coefficients only from the $A^4$ term upwards but does not affect $\gamma$ as the $A^2$ term comes solely from the fundamental. Similarly we can square the series for the other harmonics and always yielding contributions to the kinetic energy that are even in $A$. These are of increasingly high order in $A$ and thus of increasingly small importance.
This expansion for the perturbation energy in even powered terms in the amplitude was first suggested by Landau [71] in examining the stability of general hydrodynamic flows. Taking the product of $A^*$ with a general power series in $A$ and time averaging, he reduced the result to a series in even powers of $A$ only: \(^{16}\)

$$\frac{dA^2}{dt} = 2\lambda A^2 - \alpha A^4 + \beta A^6$$  \hspace{1cm} (6.11)

In his analysis Landau assumed a complex eigenvalue $\lambda_c = \lambda_r - i\omega_r$, \(^{17}\) and that $\omega \gg \lambda$. In our case, however, $\omega = 0$, and so his procedure does not directly apply. We can achieve the same result, however, by considering the vertical invariance of our problem and the assumed periodicity in $z$. The spatial factor $\psi(x)$ will have a $z$ dependence $\propto e^{ikz}$. \(^{18}\) In doing the volume integral to calculate the total kinetic energy, the integral over this factor will play the same role that the time dependent $e^{i\omega t}$ did in Landau's analysis. Only terms which are even in $e^{ikz}$, will contribute. These terms are also even in $A(t)$.

\hspace{1cm} \hspace{1cm} \hspace{1cm} 6.2.3 Solutions of the General Expansion

In the previous section we have derived various series expansions that are all special cases of the general form:

$$\frac{dA}{dt} = \lambda A + \alpha A^2 + \beta A^3 + \gamma A^4 \cdots$$  \hspace{1cm} (6.12)

Note as well that if we make a change of variables in Equation (6.10) and take $A' \equiv A^2$ as the basic variable (e.g. the kinetic energy rather than the velocity), it takes the form of Equation (6.12) with all its terms. When convenient, one

---

\(^{16}\)In [71] he only discussed the linear and first nonlinear terms. In a more recent version[55] the $A^6$ term is also included. The coefficient $\alpha$ is called the Landau constant.

\(^{17}\)Note that following Landau the sign on $\omega$ is opposite to our general use elsewhere.

\(^{18}\)This is the result of the Fourier expansion described in Section 6.2.1.
can solve the lower-order equation in $A'$ and then, taking the square root, restate the results in terms of $A$.$^{19}$

In this section we outline the analytical solutions to successive truncations of Equation (6.12) and some of their expected characteristics. These results will be useful in treating the low order expansions that describe the nonlinear results in the supercritical regime just above $G_c$. $^{20}$ Note that we shall use $A$ as the expansion variable throughout, but that here it is not to be taken as anything more than a mathematical, generic, variable with no specific physical meaning.

It is useful to plot $f(A) \equiv \frac{dA}{dt}$ as a function of $A$, as we have Figure 6.3a, for the linear case of Equation (6.2). In this type of graph stationary solutions are found whenever $f(A) = 0$. $^{21}$ They are stable if the slope is negative and unstable if the slope is positive.

\[ \begin{align*}
&\text{a: Linear growth} \\
&\begin{array}{c}
\frac{dA}{dt} \\
\lambda > 0 \\
\lambda < 0
\end{array}
\end{align*} \]

\[ \begin{align*}
&\text{b: Quadratic growth} \\
&\begin{array}{c}
\frac{dA}{dt} \\
\lambda > 0 \\
\lambda < 0
\end{array}
\end{align*} \]

$^{19}$For instance when $A^2$ is expanded to second order ($A^4$), we expect its value on the primary branch to go as $\frac{A}{a}$, while the equivalent third order equation for $A$ will yield $A_{\text{rms}} = \sqrt{\frac{A}{a}}$. See Section 6.2.3.

$^{20}$See Chapter 7.

$^{21}$We will in the following discussion limit our comments to the solutions for $A > 0$, although the plots will show all solutions for both positive and negative $A$. 
For completeness start again with the linear truncation of Equation (6.12),

$$\frac{dA}{dt} = \lambda A, \quad (6.2)$$

which had the solution $A(t) \propto e^{\lambda t}$.

We can see that for the first-order case there is a single stationary point at the origin which is stable or unstable depending on the sign of $\lambda$.

Next we add on a second order term

$$\frac{dA}{dt} = \lambda A + \alpha A^2 \quad \alpha < 0. \quad (6.12)$$

The solutions for the general quadratic case are plotted in Figure 6.3b in which we can see that the inclusion of the higher-order term allows a new stationary point. This second stationary point represents a primary branch solution. It is stable for $\lambda > 0$ and $\alpha < 0$. For $\lambda < 0$ and $\alpha > 0$ there is also be a second stationary point, but it is unstable. Moreover, in this case the origin is stable, and so only a large perturbation would allow the secondary point to be reached.

The solution to this equation$^{23}$ is given by:

$$A(t) = \frac{-\frac{\lambda}{\alpha}}{1 - \left(1 + \frac{\lambda}{\alpha A_0}\right) e^{-\lambda t}} \quad (6.13)$$

The new stationary point can be found by letting $t \to \infty$ when the initial growth$^{24}$ has reached saturation.

In this limit we find $A_{asm} = \frac{\lambda}{|\alpha|}$, for $\alpha < 0$.

$^{22}$The solution for the $\alpha < 0$ case is marked by an asterisk (*).

$^{23}$See Appendix A.5 for details.

$^{24}$As noted in Appendix A.5, Equation (6.13) reduces to the exponential growth found with Equation (A.60) in the small $t$ limit.
It is also clear from Equation (6.13) that the approach to this secondary state is an exponential decay as long as \( \lambda \) is purely real. When using a second-order energy equation in \( A' \) this models the action of the diffusive terms reducing the transients left after the transition to the secondary state.

Replacing the quadratic term with a cubic term to give

\[
\frac{dA}{dt} = \lambda A + \beta A^3
\]  

(6.14)

merely changes the slopes of the curves in Figure 6.3b,\textsuperscript{25} steepening them and pulling the primary branch closer to the origin. Solving Equation (6.14) we have\textsuperscript{26}

\[
A^2(t) = \frac{-\lambda}{1 - \left(1 + \frac{\lambda \bar{A}}{\beta A^2}\right) e^{-2\lambda t}},
\]  

(6.15)

so that now \( A_{arm} = \sqrt{\frac{\lambda}{\beta}} \) for \( \beta < 0 \). We see also that the time scale has been shortened by a factor of two.

To add the possibility of a third positive stationary point we must include both the quadratic and cubic terms:

\[
\frac{dA'}{dt} = \lambda A' + \alpha A'^2 + \beta A'^3
\]  

(6.16)

Now the relative signs of the coefficients determine how many stationary points there are. When they are all of the same sign we have either explosive growth or damping but no stationary points to the right side of the origin. A positive second stationary point is formed when \( \lambda \) and \( \beta \) are of opposite signs; \( \alpha \) can be of either sign, pushing the point to larger \( A \) when it has the same sign as \( \lambda \) and pulling it to smaller \( A \) when it agrees with that of \( \beta \). These possibilities are illustrated in Figure 6.4a. In order to see the third stationary point we have

\textsuperscript{25}This holds for the positive \( A \) halves; the negative \( A \) sides would change in a more complicated fashion.

\textsuperscript{26}See Appendix A.5.
to have either $\lambda > 0$, $\alpha < 0$, and $\beta > 0$, or the inverse. As shown in Figure 6.4b, although this condition ensures a second extremum, we must also have $\alpha^2 > 4\beta \lambda$ before we have three stationary points.

Looking at the $t \to \infty$ solution for $A(t)$, derived in Appendix A.5,

$$A_{asm} = -\frac{\alpha}{2\beta} \pm \frac{1}{2} \sqrt{\frac{\alpha^2}{\beta^2} - 4\frac{\lambda}{\beta}},$$

(6.17)

we can see that this condition is equivalent to having a real solution.

At $\alpha^2 = 4\beta \lambda$ there will be a single stationary point at $A = -\frac{\alpha}{2\beta}$ which is stable to perturbations reducing $A$ but unstable to perturbations increasing $A$. When the inequality holds there will be two solutions, as given in Equation (6.17), but only the smaller will be stable.

The full temporal solution for $A(t)$ is

$$A(t) = -\frac{\alpha}{2\beta} \{\} \pm \frac{1}{2} \sqrt{\frac{\alpha^2}{\beta^2} \{\}^2 - 4\frac{\lambda}{\beta} \{\}},$$

(6.18)

where

$$\{\} \equiv \left\{ \frac{1}{\left[1 - \text{cnst } e^{-2\gamma t}\right]} \right\} \quad \{\} = \left[ \frac{A + \frac{\alpha}{\beta} - b}{A + \frac{\alpha}{\beta} + b} \right]^{-\frac{1}{2} \gamma},$$

and

$$b = \sqrt{\frac{\alpha^2}{\beta^2} - 4\frac{\lambda}{\beta}}.$$
In general in order to increase the number of solutions as we add more terms to the series of Equation (6.12) we must take the coefficients as alternately greater than and less than zero. This will ensure the increase of the number of turning points in \( f(A) \) and also the number of solutions of \( f(A) = 0 \). These solutions are given by the roots of a polynomial of increasing order, always one less than the order of \( f(A) \),\(^{27}\) formed in the fashion indicated in Appendix A.5. As these new solutions are added they will be alternately stable and unstable to growing perturbations. Thus in order to add the next stable solution we need to add two terms to our series. For \( \lambda < 0 \) the odd terms will give new stable roots, while for \( \lambda > 0 \) it will be the even terms.

For fourth-order series and fifth-order series shown in Figure 6.5 we show the extremes (for \( A \geq 0, \ A = 0 \) being unstable) of a lone root a \( A = 0 \) and of four and five roots respectively. In both cases there is the possibility of another stable root for \( A > 0 \) and thus the possibility of interactions between the secondary solutions.

\(^{27}\)Having divided out the \( A = 0 \) solution.
6.2.4 A Note on Stable Solutions

To summarize some of the results obtained above, and relate them specifically to the perturbation and energy expansions for the VSC problem we outline the stable solutions for a sequence of series that include an increasingly large number of terms.

- For a series in $A$ to third order: there are two stable solutions for $A$, both greater than zero.

- For the same series but with odd terms only: there are two stable solutions, symmetric at some $\pm A$, thus only one greater than zero.

- For the corresponding series in $A'$ to second order: there is one stable solution, corresponding to the one positive solution in $A$.

Thus we can see that to lowest order there are two symmetric solutions for the perturbation and one for the energy. Extending this:

- For a series in $A$ to fifth order (odd terms only): there are still only two stable solutions as before, the new solutions being unstable.

- For the corresponding series in $A'$ to third order: the one new solution is similarly unstable.

It should be noted that for the odd series in $A$ the solutions are always symmetric about $A = 0$. Extending again:

- For a series in $A$ to seventh order (odd terms only): there are now two new stable solutions, again paired.
• For the corresponding series in $A'$ to fourth order: there are now two stable solutions.

Thus to get new stable solutions we need to add four orders in $A$ to get two new terms in the expansion. This corresponds to one new energy solution and two new energy terms, but only two orders in $A'$. (Recall $A' = A^2$.) There are always the appropriate number of solutions for the energy, one half the number for the perturbations.

6.2.5 $G - G_c$ Dependence

Returning to the VSC expansions, we can follow Landau's analysis one step further and assume that $\lambda$, and in fact the other coefficients, can be expanded in powers of $(G - G_c)$. As can be seen in Figure 6.6, near $G_c$ we need take only the first-order term to get an excellent approximation for $G - G_c$ up to 60. If we use this in the relation $A_{as\alpha} \propto \sqrt{\lambda}$, we have Landau’s well-known

![Figure 6.6: Log-log plot of $\lambda$ vs. $G - G_c$.](image)

result that the magnitude of the secondary field varies as the squareroot of the
control parameter, $A_{asm} \propto \sqrt{G - G_c}$.

It should be noted that although most treatments follow Landau in proposing an expansion in powers of $G - G_c$, our results here show that $\lambda$ is given rather by successive roots of $G - G_c$. The relationship can be given by a power series in $\lambda$:

$$(G - G_c) = \alpha \lambda + \beta \lambda^2 + \gamma \lambda^3 + \cdots$$

Furthermore we can solve for the other coefficients of Equation (6.10) one at a time starting with the Landau constant: $\alpha = \lambda / A_{asm}$. Near $G_c$ this will be independent of $G - G_c$. Since we can measure $A_{asm}$ from our simulation results and take $\lambda$ from the linear analysis, we can find $\alpha$ and test its $G$ dependence as well. In a similar fashion we can calculate the other series coefficients one after another. Coupled with the known dependence of $\lambda$ on the control parameter we will be able to show how both the variables and the expansion coefficients vary with $G - G_c$.

6.3 A General Theoretical Context

We have now finished our theoretical outline and will come back to all these matters as they come up in our presentation of the simulation results below. But before we move on we outline some of the theoretical similarities between our problem and some other areas of physics.

6.3.1 Notes on the Ginzburg-Landau Equation

The Ginzburg-Landau Equation originated in work by V.L. Ginzburg and L.D. Landau[72] to describe the phase transition at the onset of superconductivity. As described in [73] the requirement that the order parameter
(in this case a quantity $\psi$ proportional to the condensate wave function) be invariant under a phase transformation $\psi \rightarrow \psi e^{i\alpha}$ leads to an expansion of the free energy in terms of even powers of $\psi$

$$\frac{dF}{dV} = \frac{\hbar^2}{4m} |\nabla \psi|^2 + a|\psi|^2 + b|\psi|^4$$

(6.19)

where $F$ is the change in free energy per volume from the normal state. This expansion is similar to that of Equation (6.10), but the diffusion term has been left explicit. The coefficient $a$ is proportional to $T_c - T$, which is the control parameter here, just as $\lambda$ is proportional to $G - G_c$, while $b$ is independent of $T$ just as we shall find $\alpha$ to be independent of $G$. We note finally that the equilibrium value below $T_c$ is given by

$$|\psi|^2 = \frac{a}{b}$$

in perfect analogy to the asymptotic value of $|A|^2$ found above. It is plausible, not only that the hydrodynamic saturation amplitude is analogous to that in critical phenomena, but that the parallel extends to a many other properties near $G = G_c$, such as the time scaling near the critical point. We will refer to this again in the discussion of our results.

Similar equations have since been used in many areas that involve phase transitions and broken symmetries. Landau and Lifshitz [74](p. 473) derive a similar equation for the fluctuations of a single order parameter describing the change in symmetry at a general second order phase transition. Aranson et al.[75] discuss a generalized Ginzburg-Landau equation in the context of a wave instability in a binary fluid mixture

$$\frac{\partial u}{\partial t} = -u + \beta |u|^2 u - |u|^4 u - (k_0^2 + \nabla^2)^2 u + \mu \phi \ldots$$

which corresponds to Equation (6.9). The last term here functions as the $\frac{dF}{ds}$ term in our equations.
Ma[76], using a discussion of the theory of spins to illustrate the
general properties of critical phenomena uses a Ginzburg-Landau equation to
give the Hamiltonian of an array of spin blocks with spin $\sigma$ in an external
magnetic field $h$:
\[
\frac{H[\sigma]}{T} \equiv \int dv \left[ a_0 + a_2 \sigma^2 + a_4 \sigma^4 + c(\nabla \sigma)^2 - h \cdot \sigma \right]
\] (6.20)

From this brief survey it is clear that our problem possesses certain
generic aspects that are common to many problems of phase transition and
symmetry braking at critical points. Since we can with our nonlinear code run
precise numerical experiments and 'measure' the resulting data in detail, we
have a useful medium for testing theoretical results in any problem that can
be mapped onto ours, and vice versa.

6.3.2 Phase Invariance

The derivation of the Ginzburg-Landau Equation reflected the phase
invariance of the order parameter. In our problem the vertical translational
invariance that allows Fourier decomposition can be viewed as an invariance
under a change in $A$ such that $\psi_k \rightarrow \psi'_k = \psi_k e^{i\delta_z}$. The series given in
Equations (6.9) and (6.10) are of the correct form to remain invariant. For the
linear equations any phase transformation can simply be divided out leaving the
equations invariant. However when the non-linear terms are added (or become
important as the perturbations grow) then invariance can only be maintained
when the phase is the same in each term of the expansions. The phase must
therefore satisfy the condition that $e^{i\delta_z} = \pm 1$, restricting $\delta$ to 0 and $\pi$. We
shall see the effects of this with the appearance of a mode that is $\pi$ out of phase
with the perturbing mode at high $G$. 
There is an exciting parallel in formalism between the breaking of the phase invariance here and that of electromagnetic gauge invariance in a superconductor.\textsuperscript{28} It is far beyond the scope of this present work to explore this possibility, but would be of great interest if such a correspondence could be shown rigorously.\textsuperscript{29}

Another way to look at this result is to say that it maintains the parity of the fundamental, and the proper relative parity of its harmonics. While in the vertical direction the periodic boundary conditions do not impose a specific parity on the solution, this is not true horizontally. Here we have $\psi = 0$ on both walls, for even parity, while $T = \pm 1$ on the two walls, for odd parity. This will tend to suppress the odd modes in $\hat{\psi}(x)$ and the even modes in $\hat{T}(x)$. In fact the eigenmode is almost purely the $(0,2,4)$ modes for $\hat{\psi}$ and the $(1)$ mode for $\hat{T}$. This is illustrated in Figure 6.7. This parity selection is maintained

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{stream_function}
\includegraphics[width=\textwidth]{temperature}
\includegraphics[width=\textwidth]{chebyshev_energy_spectrum}
\caption{The Chebychev mode amplitudes of the basic state at $G = 2000$, $Pr = 7.5$, and $\alpha = 1.383$}
\end{figure}

\textsuperscript{28}See for instance Weinberg[77] where this breakdown is discussed from the point of view of field theory and particle physics.

\textsuperscript{29}Some work on the general correspondence between hydrodynamics and field theory concepts has already been done. See for instance [78].
Figure 6.8: The Chebychev mode amplitudes of the full flow at $t = 8128.5$ even in the fully non-linear regime as we can see in Figure 6.8 in which we can see that the perturbation is of significant size. These profiles were done after 35,000 timesteps at a dimensionless time of $t = 8128.5$.

6.4 Diagnostic Tools

This section will define various quantities that we will use in analyzing the results of the non-linear VSC code.

6.4.1 The Nusselt Number

The Nusselt number defined as

$$N = \frac{1}{\lambda} \int_0^\lambda \left( \frac{\partial T}{\partial x} \right)_{x=-1} \, dz$$

(6.21)

which gives the ratio of the total heat transfer over the heat conduction. Thus in the basic conductive state the Nusselt number is unity, and any variation from this is due to convection. It is a wildly used order parameter as it measures the convective heat transport which is of great interest in many applications (as discussed in Section 1).

There is another consequence of the form of Equation (6.9) that was
pointed out by Batchelor[14]. Returning to the full expansion (Equation (6.12)) and write it for \( \hat{T} \). The integral in \( z \) will eliminate all the odd terms in \( B \) leaving only the even terms and yield a dependence similar to that for the kinetic energy. Note that this also means only the \( k = 0 \) and odd harmonic parts of the perturbation contribute to the change in the Nusselt number, the fundamental does not.

6.4.2 The Gravitational Potential Energy

It is also of use to calculate the gravitational potential energy of the flows.

\[
P E = \alpha g \int_V dzzT(x,t) \tag{6.22}
\]

6.4.3 The Energy Integrals

It is useful in understanding the physical behavior of the fluid flows generated in convection problems to derive the equations for the energy balance of the system under study.\(^{31}\) The result of deriving these equations will be terms that measure the generation, redistribution, and dissipation of energy. This analysis can provide useful diagnostic tools yielding valuable physical insight into the physics involved in the processes we observe.

We start by writing the VSC Equations (2.12) and (2.16) in the form:

\[
\frac{\partial}{\partial t} \nabla^2 \psi = -\frac{\partial (\nabla^2 \psi, \psi)}{\partial (x, z)} - \frac{1}{G} \frac{\partial T}{\partial x} + \frac{1}{G} \nabla^4 \psi \tag{6.23}
\]

\[
\frac{\partial}{\partial t} T = -\frac{\partial (T, \psi)}{\partial (x, z)} + \frac{1}{PrG} \nabla^2 T. \tag{6.24}
\]

\(^{30}\)Batchelor's expansion was in terms of powers of \( G \), which is proportional to our expansion in \( A \) while \( A \propto G - G_c \).

\(^{31}\)For a parallel discussion of the Benard case see [79].
Now taking the volume integrals \( \int \psi \partial(\nabla^2 \psi, \psi) dV \) and \( \int T \partial\nabla^2 \psi dV \) we have equations for the change of kinetic and available potential energy [79] (or the temperature variance [13]), as follows:

\[
\frac{\partial}{\partial t} \int \frac{1}{2} \psi \nabla^2 \psi dV = - \int \psi \frac{\partial(\nabla^2 \psi, \psi)}{\partial(x, z)} dV - \frac{1}{G} \int \psi \frac{\partial T}{\partial x} dV + \frac{1}{G} \int \psi \nabla^4 \psi dV \tag{6.25}
\]

where we have used the fact that, with two partial integrations and our given boundary conditions,

\[
\int \frac{\partial \psi}{\partial t} \nabla^2 \psi = \int \psi \nabla^2 \frac{\partial \psi}{\partial t},
\]

and

\[
\frac{\partial}{\partial t} \int \frac{1}{2} T^2 dV = - \int T \frac{\partial(T, \psi)}{\partial(x, z)} dV + \frac{1}{Pr} \int T \nabla^2 T dV. \tag{6.26}
\]

In Equation (6.25) we integrate the quadratic \( \psi \) terms by parts, integrating twice for the kinetic diffusion term:

\[
\frac{\partial}{\partial t} \int \frac{1}{2} \psi \nabla^2 \psi dV \rightarrow - \frac{\partial}{\partial t} \int \frac{1}{2} |\nabla \psi|^2 dV
\]

\[
\frac{1}{G} \int \psi \nabla^4 \psi dV \rightarrow \frac{1}{G} \int |\nabla^2 \psi|^2 dV
\]

We have used the fact that \( \psi = \nabla \psi = 0 \) on the \( x \) boundaries, while they are periodic, and thus equal, on the \( z \) boundaries, to eliminate the boundary terms:

\[
\psi \nabla \psi|_{\text{bnd}}, \quad \psi \nabla^2 \psi|_{\text{bnd}}, \quad \text{and} \quad \psi \nabla^3 \psi|_{\text{bnd}}.
\]

Next we rewrite the term derived from the kinetic Jacobian:

\[
- \int \psi \frac{\partial(\nabla^2 \psi, \psi)}{\partial(x, z)} dV = - \int \psi \frac{\partial \nabla^2 \psi}{\partial x} \frac{\partial \psi}{\partial z} dV + \int \psi \frac{\partial \nabla^2 \psi}{\partial z} \frac{\partial \psi}{\partial x} dV =
\]

\[
- \int \frac{\partial}{\partial z} \left( \frac{1}{2} \psi^2 \frac{\partial \nabla^2 \psi}{\partial x} \right) dV - \int \frac{1}{2} \psi^2 \frac{\partial^2 \nabla^2 \psi}{\partial x \partial z} dV
\]

\[
+ \int \frac{\partial}{\partial x} \left( \frac{1}{2} \psi^2 \frac{\partial \nabla^2 \psi}{\partial z} \right) dV + \int \frac{1}{2} \psi^2 \frac{\partial^2 \nabla^2 \psi}{\partial z \partial x} dV.
\]
The second and fourth terms cancel, while the first and third are now exact differentials in $z$ and $x$ respectively. Integrating these gives

$$- \int \psi \frac{\partial (\nabla^2 \psi, \psi)}{\partial (x, z)} dV = - \left[ \int \frac{1}{2} \psi^2 \frac{\partial \nabla^2 \psi}{\partial x} dx \right]_{z \text{bd}} + \left[ \int \frac{1}{2} \psi^2 \frac{\partial \nabla^2 \psi}{\partial z} dz \right]_{z \text{bd}}$$

which equals zero by the boundary conditions.

Putting this all together we have (with a change of signs)

$$\frac{\partial}{\partial t} \int \frac{1}{2} |\nabla \psi|^2 dV = \frac{1}{G} \int \psi \frac{\partial T}{\partial x} dV - \frac{1}{G} \int |\nabla^2 \psi|^2 dV \quad (6.27)$$

This equation for the time change of the kinetic energy shows clearly that for a steady state the dissipation of kinetic energy that is always present must be balanced by the generation of kinetic energy by the buoyancy forces. While the dissipative term is always negative, it is important to note that the work done by the buoyancy forces is only positive when $\psi$ and $\frac{\partial T}{\partial x}$ are positively correlated.

Turning to Equation (6.26) the term derived from the thermal Jacobian is treated in the same manner as its kinetic counterpart, with the same results.

$$- \int T \frac{\partial (T, \psi)}{\partial (x, z)} dV = - \int T \frac{\partial T}{\partial x} \frac{\partial \psi}{\partial x} dV + \int T \frac{\partial T}{\partial z} \frac{\partial \psi}{\partial z} dV =$$

$$- \int \frac{\partial}{\partial z} \left( \frac{1}{2} \psi \frac{\partial T}{\partial x} \right) dV - \int \frac{1}{2} \psi \frac{\partial^2 T}{\partial x \partial z} dV + \int \frac{\partial}{\partial x} \left( \frac{1}{2} \psi \frac{\partial T}{\partial z} \right) dV + \int \frac{1}{2} \psi \frac{\partial^2 T}{\partial z \partial x} dV$$

Again the second and fourth terms cancel while the first and third are integrated to give

$$- \int T \frac{\partial (T, \psi)}{\partial (x, z)} dV = - \left[ \int \frac{1}{2} \psi \frac{\partial T^2}{\partial x} dx \right]_{z \text{bd}} + \left[ \int \frac{1}{2} \psi \frac{\partial T^2}{\partial z} dz \right]_{z \text{bd}},$$

which is also zero by the boundary conditions.

Lastly we integrate the thermal diffusion terms by parts to get

$$\frac{1}{PrG} \int T \nabla^2 T dV \to - \frac{1}{PrG} \int |\nabla T|^2 dV + \frac{1}{PrG} \int \frac{1}{2} \frac{\partial T^2}{\partial x} \bigg|_{x = 1} dz$$
This last term is the one boundary term we cannot neglect. The $z$ periodicity causes the upper and lower boundary terms to cancel, but on the $x$ boundaries, although $\partial_z T$ is zero, $\partial_x T$, and for our definitions $T$ itself, are both non-zero. Since $T = \pm 1$ on these boundaries, this term can be written as

$$\frac{1}{Pr} \int \frac{\partial T}{\partial x} \bigg|_{z=\pm 1} dz + \frac{1}{Pr} \int \frac{\partial T}{\partial x} \bigg|_{z=1} dz,$$

twice the average heat flow into and out of the slot. When these are equal, as is usually the case, this will be simply twice the total heat flow through the slot.

In a treatment of the problem in which the conductive temperature gradient is taken as part of the background and the equations are written for the perturbations around $T = x$ (as is often done for the Benard problem), the boundary term will not appear, as then $T'$ is zero on the walls.

Substituting these results into $Pr$ times Equation (6.26) gives us an equation for the change of the available potential energy in time:

$$\frac{\partial}{\partial t} \int \frac{Pr}{2} T^2 dV = -\frac{1}{G} \int |\nabla T|^2 dV + \frac{1}{G} \int \frac{\partial T^2}{\partial x} \bigg|_{z=\pm 1} dz \quad (6.28)$$

Here we have a balance between the thermal diffusivity and the heat flow through the slot. This later could only be negative if the heat flow was reversed so that it was going from the cold side to the hot side of the slot.

Combining these two balance equations, we have for the whole system:

$$\frac{\partial}{\partial t} \int \frac{1}{2} |\nabla \psi|^2 + \frac{Pr}{2} T^2 dV =$$

$$\frac{1}{G} \int \psi \frac{\partial T}{\partial x} dV + \frac{1}{G} \int T \frac{\partial T}{\partial x} dV \bigg|_{z=\pm 1} - \frac{1}{G} \int |\nabla^2 \psi|^2 + |\nabla T|^2 dV \quad (6.29)$$

Thus the combined kinetic and potential energy will be constant when the increase in energy due to the buoyancy force, which draws on the gravitational field, and to the flux of heat through the system is balanced by dissipation.
and thermal diffusivity. It is of note that the nonlinear terms do not affect the energy of the system as a whole, but they do effect a reapporitionment of the energy between variables and modes. In particular as pointed out in [13], the Reynolds stress in the kinetic equation and the interaction of the mean temperature gradient with the convective heat transfer can shift energy in and out of the background initial flow.

6.4.4 The Perturbation Energy Balance

In order to derive a similar set of equations for the secondary flow alone we start from Equations (4.5) and (4.6) and proceed in a manner paralleling the derivation in the last section.\(^{32}\)

\[
\frac{\partial}{\partial t} \int \frac{1}{2} |\nabla \hat{\psi}|^2 dV = \int \hat{\psi} \frac{\partial \nabla^2 \psi_b}{\partial x} \frac{\partial \hat{\psi}}{\partial z} dV + \int \hat{\psi} \frac{\partial \nabla^2 \psi_b}{\partial x} \frac{\partial \psi_b}{\partial x} dV + \frac{1}{G} \int \hat{\psi} \frac{\partial \hat{T}}{\partial x} dV - \frac{1}{G} \int |\nabla^2 \hat{\psi}|^2 dV
\]  
(6.30)

The new terms\(^{33}\) here represent the energy transfer from the mean flow due to the Reynolds stress.

\[
\frac{\partial}{\partial t} \int \frac{Pr}{2} \hat{T}^2 dV = -Pr \int \hat{T} \frac{\partial \psi_b}{\partial x} \frac{\partial \hat{T}}{\partial x} dV + Pr \int \hat{T} \frac{\partial \hat{T}}{\partial z} \frac{\partial \psi_b}{\partial z} dV - \frac{1}{G} \int |\nabla \hat{T}|^2 dV
\]  
(6.31)

\(^{32}\)We have not forgotten the terms nonlinear in the perturbation, but they do not contribute here, just as the nonlinear terms in the full fields did not contribute above, and so we might as well start from the linear equations.

\(^{33}\)In Chait and Korpela [13] only the first of these terms is given. Bergholtz [32], however, has both.
Chapter 7

The Initial Bifurcation and the Primary Branch

In order to study the nonlinear VSC problem we have constructed a psuedo-spectral semi-implicit computer code to solve for the time evolution of the VSC equations derived in Chapter 2. The linear terms are handled are spectrally analyzed using a Fourier decomposition vertically and a Chebychev decomposition horizontally and solved for implicitly, while the nonlinear terms are treated explicitly in real space. The general architecture of this method was written by J. Mizushima following a method outlined by Moin and Kim.[56]. Discussion of Fourier and Chebychev series, and their relevant properties is included in Appendix B, and a presentation of the numerical methodology and a derivation of the spectral coefficient equations that are actually implemented on the computer can be found in Appendix C.3. Appendix C.4.3 contains some tests of the accuracy of the various operators and transforms used in the code and some analysis of alternative techniques.

The results presented in this chapter and the next two are all obtained at a Prandtl number of 7.5 (for water) for reasons mentioned in Section 5.6.6. Unless otherwise noted we perturb at the critical wavenumber \( \alpha = \alpha_c = 1.383 \). We start with an overview and then discuss the primary branch results in terms of the Landau theory. In Chapter 8 we describe the secondary Hopf bifurcation, \( G_h \sim 1360 \), and the resulting thermal relaxation oscillations. Then in Chapter 9 we take a look at a Lorenz-like attractor and apparent temporal chaos above what is probably a third bifurcation at \( G \sim 2000 \).
7.1 An Outline of the Solutions for $G < 2500$ at $Pr = 7.5$

In the VSC configuration the motionless conductive state can exist only at $G = 0$. There is from this point an immediate branching to the 1-D solution which we have taken as our basic solution, and which does indeed exist at all values of $G$ and $Pr$. Therefore branching from the 1-D solution is referred to as primary bifurcation and the solutions arising therefrom as primary branches. Further branching takes place at secondary bifurcation points and gives rise to secondary branches, etc. In Figure 7.1 we show a bifurcation diagram including the conductive state, showing that the 1-D solution is a non-zero base state.\footnote{In the dimensionless units we normally use, the kinetic energy of this basic state is independent of $G$. In this diagram we have plotted a dimensional velocity.} We can see here that there is 'room' for perturbations that reduce the kinetic energy and that there is a physical limit to the scale of these perturbations set by the kinetic energy in the background flow. In general we will ignore the 'conductive' state and present bifurcation diagrams and plots versus $G$ with the dimensionless 1-D base-state values taken as zero.

![Graph](image-url)  

Figure 7.1: VSC bifurcation diagram including the conductive state.
Figure 7.2: VSC bifurcation diagram for $\Delta$ kinetic energy.

In Figure 7.2 we show the a plot of the kinetic energy of the steady states reached at Grashof numbers up to 2500. To be more precise, the plot represents steady states up to the secondary bifurcation. This is a Hopf bifurcation above which there are oscillatory solutions. The two branches shown above indicate the extremes between which these oscillations take place. We will return later to the question of whether this represents a single oscillatory mode or an oscillation between two separate solutions.

This range can be broken up into several regimes which we will treat separately below. These are, in order, the primary branch from the primary critical point, $G_c = 491.818$, up to about $G = 1200$, the regime around the secondary bifurcation, roughly $G = 1250 - 1500$, and the oscillatory three mode regime above $G = 1500$ which becomes temporally unsteady above $G \sim 2000$.

---

2In many works the choice of scale for $L$ and $\Delta T$ are smaller by a factor of two. Thus the Grashof numbers quoted there are 16 times larger than ours. In that case our upper limit is $G = 40000$ which, at $Pr = 7.5$, corresponds to a Rayleigh number of $Ra = 300000$. 
7.1.1 A Note on the Measured Quantities

In the following analysis we shall present results for various physical and quasi-physical quantities, such as the kinetic energy, various specific energy terms, the Nusselt number, the perturbation amplitudes $\hat{\psi}$ and $\hat{T}$, and Fourier modes of these perturbations. In most cases these quantities are taken as volume integrals over $x$ and $z$. There are, however, three exceptions. The first is the Nusselt number, which, as shown in Section 6.4.1, is based on an integration over $z$ only. The other two are the perturbation 'amplitudes' which are calculated as $\psi_{\text{max}}^* - \psi_{\text{min}}^*$, and $T_{\text{max}}^* - T_{\text{min}}^*$. Thus these two quantities are not integrated at all, and this leads to their sometimes irregular behavior. Lastly we emphasize that the perturbation modes $\psi_0, \psi_1, \psi_2$, etc. are integrated over $x$ in Fourier space.

7.2 The Primary Branch

The primary branch starts with an exchange of stabilities at $G_c$ and lasts up to a Hopf bifurcation at about $G = 1350$. In examining it in detail we will find that, although in many ways it can be considered a single regime, there are some analyses that indicate subregimes. The behavior changes, when it changes, at $G \sim 550$ and $G \sim 1000$, dividing the range into perhaps three subregimes.

7.2.1 The Primary Critical Point

It is natural to look first at the behavior in the neighborhood of the linear critical point, which for $Pr = 7.5$ was found to be $G_c = 491.818$. Here there is a supercritical exchange of stabilities at a double-point bifurcation, a transition from one stationary ($\omega = 0$) state to another. The basic solution be-
comes unstable while a new stable solution branches off. We have investigated the range from $G = 450$ to $G = 550$ in some detail, especially for $G$ within $\pm 20$ of $G_c$, using the linear eigenmodes to give the spatial dependence of the perturbations, with a factor of $\epsilon = 0.0001$ to give their initial scale. These eigenmodes are stationary modes of the form given in Figure 5.39 of Section 5.6.6.\footnote{In all cases where we perturb the 1-D solution, we will make use of similar perturbations, unless otherwise noted.}

In Figure 7.3 we can see that this is a very clean bifurcation and indeed for $G < G_c$ the residual perturbations are damped by a factor of the order $10^3$ from their already small initial values. More significantly, the final values are independent of $G$ and can be assumed to be due to computational inaccuracy.
7.2.2 Other Perturbative Tests

In order to test the stability of the bifurcation to large noisy perturbations, we used a thermal sine-sine perturbation\(^4\) at \(G = 475\) and ran \(\epsilon\) up to 10%. In all cases the perturbations died rapidly, returning to the same final state that we had found using the stationary eigen functions. The transient behavior was also roughly the same. First one sees the complex traveling mode appear, but it decays rapidly \((\lambda_c = .0047)\), leaving the stationary mode which decays more slowly \((\lambda_c = .00078)\). The traveling mode shows up at first because, being a superposition of two separate modes, one on each wall, the combination has a larger amplitude than the stationary mode. This feature is seen also at supercritical values of \(G\) where a general perturbation leads first to a decaying pair of traveling modes and then to the stationary mode, which for \(G > G_c\) grows rather than decays.

In Figure 7.4 we show the time histories of the stream function and temperature perturbations for \(G = 475\) and \(\epsilon = 0.001\). Note the sharp bend in these curves that corresponds to the change in eigenvalue with the change in dominant transient. The initial oscillatory character of the stream function is due to the traveling mode, and dies out as the stationary mode becomes dominant.

As noted above, using a general perturbation for \(G > G_c\) causes the growth to be slower at first while the transients are dying and energy is transferred into the stationary mode. The final results, however, are the same. Here again the amplitude of the perturbation can be as much as 10% without disturbing the final results.

\(^4\)The temperature perturbation has spatial dependence \(\sin(\pi x)\sin(\pi z)\) \((x\ and\ z\ ranging\ from\ -1\ to\ +1)\), and the kinetic variables are not perturbed.
Figure 7.4: Perturbation decay at $G = 475$, with $\epsilon = 0.001$.

It is known from imperfection theory that perturbing the equations themselves can lead to changes in the form of the bifurcation[70]. Previous results [63] exhibited sub-critical coupling between the complex modes and the stationary mode, shifting the critical point from $G_c = 491.8$ to the mode crossing point at $G \sim 407$. It has since been realized that there was a subtle error in the code so that these results were derived with numerically perturbed equations. Several runs were therefore made with the exact equations using the complex traveling mode eigenfunctions as perturbations, but no evidence of such mode coupling was found. Runs were made at $G = 450$ and $G = 550$. In the former case, the usual final state was reached with no coupling at all to the stationary mode. In the latter case, the perturbation completely decays before the stationary mode has enough energy to appear and grow to its expected secondary configuration. This same growth of the stationary mode would occur in the absence of any perturbation due to machine noise.

In general it seems that in this regime near $G_c$ there is very little transfer of energy between the linear eigenmodes as long as their amplitudes are small. It remains to be seen whether the addition of a high $\kappa$ noise term to
the equations, which could be expected to produce an imperfect bifurcation, might not also serve to allow the linear modes to couple more easily and lead different behavior in the near subcritical region.

7.2.3 The Flows Near $G_c$

Before we go into a detailed analysis of the primary bifurcation, let us take a look at the form of the secondary flows that result on the primary branch. The perturbative parts of these flows are shown in Figure 7.5, where we can see that with increasing $G$ the primary branch solution evolves away from the linear eigenmode (see Figure 5.39) as the nonlinear terms become more important. At $G = 495$ the difference is slight, but at higher $G$ the magnitude of the secondary state increases the effect of the nonlinear terms, and we move towards a distinctly nonlinear mode. Gotoh and Ikeda [20] plotted results at $G = 600$ but they assumed that the spatial form of the mode was constant, so although their total flow is not that different from Figure 7.6 the perturbations given quite a bit different from ours. Chait and Korpela [13] plot the secondary state for a similar $G$, but at $Pr = .71$, where the form of the eigenfunction is a little different.\(^5\)

The most notable features of this mode are the large $k = 0$ contribution to the stream function and the increasing asymmetry of the lobes of the temperature perturbations. The latter change coincides with the development of a strong horizontal gradient\(^6\) (in the perturbation) in the center of the slot as the positive parts move left and the negative parts right. This results

---

\(^5\)Their results are in good agreement with eigenfunction plots we have done at $Pr = .7$.

\(^6\)A very important numerically, as it results in the need for a large number of Chebychev modes to properly resolve this gradient.
Figure 7.5: The primary branch perturbations for $G = 495, 510, 600$ at $Pr = 7.5$ and $\alpha = 1.383$. 
in a flattening of the center-slot horizontal temperature profile (for the total flow). It is a general tendency of shear flow instabilities that the perturbations flatten all gradients (see [80] cited in [81]). In the VSC configuration the fixed boundaries limit this effect to the center-slot, a precursor to the boundary layer physics at high $G$ and an integral factor in the physics of the thermal relaxation oscillations discussed in Chapter 8.

The changes to the total flow can be seen in Figure 7.6 for $G = 600$. We see here the emergence of the 'cat's eye' vortices in the stream function and

![Stream Function and Temperature Diagram]

data range from $-1.13E-15$ to $0.4129E-01$

data range from $-1.0E+01$ to $1.0E+01$

Figure 7.6: The primary branch flow for $G = 600$.

the wave in the temperature contours in the center slot. Both of these features evolve and become more prominent at larger $G$. These cells were first seen in experiments by Elder [34] (photos and plots) in oil ($Pr = 1000$). Vest and Arpaci [44] have pictures of cells in air ($Pr = .71$) and oil ($Pr \sim 900$) and Seki et al. [43] present pictures for $Pr = 480$ and 12500. Several authors including

---

This is a well-known result. See for instance [41] or [48] for experimental results.

Profile flattening was shown numerically by [53] (in a vertical annulus), and both [24] and [25] give the reduced velocity profiles with increasing $G$. We note that these profiles are similar (in reverse order to those seen as the background flow builds up from a motionless isothermal initial state[18]).
[49] and [24] have shown calculations that give similar results. Comparison with results at low aspect ratio show that these cells are similar to the primary flows formed in boxes with $1 < h < 5$ which never form secondary cells.

It is not surprising that although we can numerically 'see' the secondary flow at small $G - G_c$ the full flow doesn't change noticeably until at least $G = 550$. This implies that experimentally it would be hard to achieve accurate location of the critical point by purely visual observation, even for a perfect bifurcation.

The 'cat's eye' vortices are a particular instance of the general phenomenon of vortices produced in shear layers. As discussed by Pierrehumbert and Widnell[82] these vortices can be modeled independently using two dimensional Euler equations (no viscosity damping and no temperature) and compare reasonably well with steady states found by Stuart[80] that have a stream function given by

$$\psi = \frac{1}{2} \ln [\cosh 2x - \rho \cos 2x],$$

when $\rho = 0.25$.

In Section 6.1.1 we introduced the alternate notions of moving from one equilibrium to the next along a branch of the bifurcation diagram and of moving dynamically between branches. The flows given here, when taken in sequence, illustrate the first approach, although they were arrived at through dynamic simulation. Examination of the intermediate flows in the simulation shows that for these relatively low values of $G$ the dynamic path from the 1-D basic state to the primary branch traverses essentially the same intermediate states and is thus illustrated by this same series. As we shall see later this is not true at higher $G$, where the dynamic path includes states that are different from the equilibrium states along the primary branch.
7.2.4 The Upper Primary Branch Flows

Before continuing our discussion of the Landau theory and its extensions at larger $G$, we describe here the flows on the rest of the primary branch. Figure 7.7 shows both the full fields and the perturbations at $G = 1000$. The asymptotic forms of the stream function and the temperature fields remain qualitatively the same above $G = 600$ (compare with Figure 7.5), although the 'cat's eye' cells become increasingly prominent as the strength of the overall flow decreases. One can also see the increased clockwise rotation of the long axis of the perturbation vortex. Meanwhile the temperature continues its evolution to

Figure 7.7: The Flow for $G = 1000$
a more vertically uniform state that is increasingly polarized horizontally. The full temperature field has, by $G = 1000$, developed the hook feature associated with this mode.\footnote{For good comparison plots see \cite{20} and \cite{49}.} This hook reflects the continuing flattening of the horizontal gradients until there are small inversions in both the velocity and temperature profiles.\footnote{These have been measured experimentally, \cite{34} (both), \cite{46} (velocity), \cite{48} (temperature). \cite{60} calculates the inverted temperature profiles in a finite aspect ratio slot.} Newell and Schmidt\cite{48} point out that the inverted temperature profile means a negative horizontal gradient, and therefore a complete lack of conduction across the slot. This is the boundary layer regime. Roux et al. \cite{37}, cite \cite{83} as first using the change in the horizontal temperature gradient to classify the ‘conduction’, ‘transition’, and ‘boundary layer’ regimes.\footnote{They give an experimental curve of $\frac{dC}{dx}$ vs. $Ra$, but it is for an aspect ratio $h = 1$ and so no direct comparison is possible.}

The fields at $G = 1250$ have nearly the same forms as those shown here, and we will refer to this mode as the ‘cat’s eye’ mode. While it is no longer unique above the secondary bifurcation the cat’s eye mode is still present in essentially this form even at much higher $G$.

### 7.2.5 Basic Landau Theory

Having briefly presented the simulation results for the primary branch, let us now examine how well they are predicted by the expansion theory introduced in Section 6.2. Near $G_c$ we expect that the behavior of the perturbation fields can be described by an equation like (6.14), which, to conform to Landau’s notation, we can write as

\[
\frac{dA}{dt} = \lambda A - \alpha A^3, \tag{7.1}
\]
while we expect the kinetic energy and the Nusselt number to be described by an equation of the form of (6.8):

$$\frac{dA'}{dt} = 2\lambda A' - 2\alpha A'^2,$$  \hspace{1cm} (7.2)

where $A' = |A|^2$ and the factors of two come from the fact that

$$\frac{dA'}{dt} = \frac{1}{2} \frac{dA}{dt}.$$

Quantities obeying these equations will have solutions of the form of Equations (6.13) and (6.15) respectively. These describe an unstable equilibria at the origin and a single stable equilibria at a non-zero value of $A$, which can be found in the $t \to \infty$ limit and represent the primary branch solutions. We found that the values of $A_{asm}$ go as $\sqrt{\frac{G}{\alpha}}$ for Equation (7.1) and as $\frac{1}{\alpha}$ for Equation (7.2). Finally, we can make use of the result of Section 6.2.5 to replace $\lambda$ with $G - G_c$. This gives the $G - G_c$ dependence of quantities represented by $A$ as

\[ \sqrt{G - G_c}, \]

while those proportional to $|A|^2$ have a linear dependence and $\alpha$ is a constant independent of $G$. Looking back at Figure 7.3 we see that the kinetic energy is linear near $G_c$, while the stream function appears to go as the square
root. Here in Figure 7.8, we show that the temperature perturbation also goes as the square root, while the Nusselt number increases linearly, confirming the analysis of Section 6.4.1.

Figure 7.9: The Log-log plots vs. $G - G_c$.

To see this dependence more clearly we can make log-log graphs of these quantities against $G - G_c$. In Figure 7.9 we can see that the expected dependences hold up well to the $G = 550/600$ range.\footnote{These plots and the following plots vs. $\lambda$ extend from $G = 492$ to $G = 1250$.} In fact the stream
function perturbation seems to be in remarkable agreement even at $G = 750$.

However good the agreement is near $G_c$, it does begin to breakdown in the $G = 600 - 750$ range, if not before. There are several ways to extend the theory to try to account for this, but first recall that $A_{asm}$ depends on $G - G_c$ through $\lambda$ and that the linear dependence of $\lambda$ on $G - G_c$ itself changes in this range, going more nearly as $\sqrt{G - G_c}$ from $G = 750$ to $G = 1250/1500$. Before doing anything more complex, we might expect that looking at the dependence on $\lambda$ rather than $G - G_c$ will give us good agreement to larger $G$.

Mizushima[26] gives empirical formulas for the kinetic energy and Nusselt number as functions of $G - G_c$. These parabolic formulae give a good fit at larger $G$, in agreement with our results of a generally parabolic dependence for the range $G = 550 - 1000$, but they clearly do not match well in the range just above $G_c$ where we have seen that the dependence is linear rather than parabolic. As one would expect they pull the curve higher in $G$, in fact moving $G_c$ to $G = 499$. Experimental results can be found in Holland and Korpela[36] who worked with air ($Pr = .71$) and discuss several straight line fits to their data for the Nusselt number as a function of the Rayleigh number, $Ra = G^* Pr$. ElSherbiny et al.[7] working with inclined layers of air present general scaling laws for $Nu$ as a function of $\sqrt{Ra}$ at various finite aspect ratios. These are empirical fits to data from many authors and are carried to very high Rayleigh numbers.

Our data fit the linear formulae:

$$K.E. = .0000078(G - 491.818)$$

and

$$Nu - 1 = .000403(G - 491.818),$$

which are exact as $G_c \rightarrow 0$, but will over shoot as $G$ increases. At $G = 550$ they are off by roughly 10%. 
Figure 7.10: The Log-log plots vs. $\lambda$.

In Figure 7.10, where we plot against $\lambda$ rather than $G - G_c$, we can see that indeed the dependence given by the simple cubic Equation (7.1) and the associated Equation (7.2) is in good agreement up to $G = 1000$. Thus the break in the $G - G_c$ curves near $G = 550$ can be attributed to the change in dependence of $\lambda$ from linear to $\sqrt{G - G_c}$. Only in the range from $G = 1000 - 1350$ can we see changes in the dependence on $G - G_c$ that come from higher expansion terms or some other source rather than from the dependence of $\lambda$ itself. In this range the kinetic energy and Nusselt number vary as $\sqrt{\lambda} \propto \sqrt{G - G_c}$ and the perturbations as $\sqrt{\lambda} \propto \sqrt{G - G_c}$. 
We note here that the stream function actually grows faster than expected for larger $G$. This is consistent with the better fit for $\hat{\psi}$ in Figure 7.9 above, in which it maintains linearity longer than the other quantities. This behavior reflects the method of measuring $\hat{\psi}$, as explained above, and the fact that the spatial form of the stream function perturbation changes significantly enough to render inaccurate our assumption that the perturbation can be factored into the time-dependence (or amplitude) and a spatial form factor (see Section 6.2, Equation (6.3)). We shall come back to this below and see that this change in the form of $\hat{\psi}$ is due to the rapid growth of the $k = 0$ mode, which produces the deviations seen here.

7.2.6 Extending the Landau Theory

Empirically the Landau formulation, to third order in $A$ and second order in $A^2$, can accurately predict the size of the primary branch flows up to $G = 1000$. Analytically Stuart[61] states that this truncation holds as long as $|A|^2$ is bounded and of the order of $\lambda$. At $G = 1000$ $\lambda = .00997$ and so when the $|A|^2$ quantities change their dependence to vary as $\sqrt{\lambda} = 0.0998$ they are an order of magnitude larger than $\lambda$.

Extending the empirical approach, and ignoring the fact that the variables are solutions of a pair of coupled equations and that the magnitudes of $\hat{\psi}$ and $\hat{T}$ differ, we can add the $A^6$ term to Equation (7.1) and the $A^{13}$ term

\[\text{\footnotesize{13 Except for } \hat{\psi}, \text{ which is well predicted to } G = 750. \text{ We will see below that of the quantities we generally calculate, } \psi \text{ and its modes are generally the most difficult to match.}}\]

\[\text{\footnotesize{14 By the criterion cited in Section 6.2 } \lambda \ll (\alpha G)^{\frac{1}{3}}, \text{ the expansion is still valid at } G = 1000 \text{ where } \lambda \sim 0.01 \text{ and } (\alpha G)^{-\frac{1}{3}} \sim 0.1. \text{ For another approach see [84].}}\]
to Equation (7.2) to get:

$$\frac{dA}{dt} = \lambda A - \alpha A^3 - \beta A^5$$  \hspace{1cm} (7.3)

and

$$\frac{dA'}{dt} = 2\lambda A' - 2\alpha A'^2 - 2\beta A'^3.$$  \hspace{1cm} (7.4)

To the rough approximation that the last terms come to dominate, we can see from the pure cubic solution given in Section 6.2 that $A'$ will go as $\sqrt[3]{\lambda/\beta}$ and by the relation between $A'$ and $A$ we have $A \propto \sqrt[3]{\lambda}$. This at least qualitatively matches the $\lambda$ dependence above $G = 1000$. In fitting the dependence on $G - G_c$, however, we have to take into account that while $\alpha(G)$ is a constant, $\beta$ has a $G - G_c$ dependence inversely proportional to that of $\lambda$. This can be shown by simply requiring that all terms in the above series yield a consistent $G - G_c$ dependence for $\dot{A}$ and $\dot{A}'$.

Referring back to Equation (6.17), the solution for a general cubic, we see that these relative dependences are also necessary to yield a consistent $G - G_c$ dependence for $A'$, which is seen to be the same as that for $\dot{A}$. We are reaching the limits of the model in this regard as it fails to yield the change around $G = 1000$ to $A' \propto \sqrt[3]{G - G_c}$ and $A \propto \sqrt[3]{G - G_c}$, predicting that it would come near $G = 1500$ where the next change in the dependence of $\lambda$ on $G - G_c$ occurs.

To summarize the dependence of the various variables and derived quantities on $G - G_c$ we see one change about $G = 550$ as the dependence of $\lambda$ on $G - G_c$ changes and a second about $G = 1000$ when the dependence of $A$ on $\lambda$ changes. But while the Landau theory predicts these two changes it fails to predict the change of the dependence of $A$ on $G - G_c$ at $G = 1000$. 
7.2.7 The Expansion Coefficients

In spite of this we can still wring a bit more from the Landau theory. By letting $t \to \infty$ in Equation (7.4) and solving for $\lambda/A'$ we get

$$\lambda/A'_{asm} = \alpha + \beta A'_{asm} + \gamma A'_{asm}^2 + \cdots,$$

(7.5)

where we recall that $\alpha$ is one half the Landau constant and note that it is a constant, independent of $A'_{asm}$.

Graphing $\lambda/A'_{asm}$ against $A'_{asm}$ should give a reasonable value for $\alpha$ at low $A'_{asm}$, while the deviations for larger $A'_{asm}$ indicate the approximate importance of higher order terms in the expansions, the deviation from a constant value being due to the sum of the rest of the series. Looking first at the kinetic

![Graph of $\alpha$ for kinetic energy](image)

Figure 7.11: $\lambda/A'_{asm}$ for the change in kinetic energy.

energy, in Figure 7.11 we see that we have $\alpha = 5.659$. There is also a small linear component, such that $\beta$ is small and negative (on the order of -0.5). At larger $A'$ there are obvious contributions from higher orders in the series,

---

15The first point, equivalent to $G = 492$, is a bit off as the evolution is very slow this close to $G_c$ and was not run to so precise an asymptotic state as the other runs. The same holds for the other graphs below.
although even at $G = 1000$ these add only 4.4% to the result given by leaving them out (at $G = 750$ the deviation is only 1.1

This accuracy is achieved with a cubic approximation for the kinetic energy and will of course be better at small $t$ when $A'(t) < A'_{asm}$. We also note that to improve our results significantly we would need to include at least the next two terms in the expansion.

In Figure 7.12 we examine the case of the Nusselt number. Here we have $\alpha = 0.1094$ to within 1.5% at $G = 1250$ with a maximum deviation of 3%

![Figure 7.12: $\lambda/A'_{asm}$ for the change in Nusselt number.]

at $G = 750$. Here we also have a negative linear component (giving an estimate of -.07 for $\beta$) and then at larger $A'$ a rapid increase due to higher orders in the expansion.

Now we can turn to Equation (7.3) and with the same procedure obtain

$$\lambda/A^2_{asm} = \alpha + \beta A^2_{asm} + \gamma A^4_{asm} + \cdots \tag{7.6}$$

In Figure 7.13 we show the stream function case and from the $A_{asm} \to 0$ limit infer $\alpha = 565$, the size of $\alpha$ being inversely related to the smallness of $A_{asm}$. As
Figure 7.13: $\lambda/A'_{asm}$ for the stream function.

noted at the beginning of this chapter this ‘amplitude’ is not integrated over
the whole space (if it were it would simply be the square root of the kinetic
energy, and similarly well behaved) but rather is a measure from peak to peak.
We can see here in the large variation with $A$ the reflection of the redistribution
of energy within the mode as it grows and changes to reflect the nonlinearity of
the equations. As mentioned already, we shall see below that this reflects the
strong growth of the $k = 0$ mode of the perturbation which can be expanded
in even powers of $A$, starting with $A^2$.

Here we see strong contributions, from first a negative term and then
a higher positive term. If we assume that these are respectively the $\beta$ and
$\gamma$ terms in Equation (7.6) then we can use the fact that for $A_{asm} \to A_m$ at
$G = 1250$ we have

$$\beta A_m^2 + \gamma A_m^4 = -\frac{1}{2} \alpha$$

while

$$\frac{d}{dA_{asm}} [\beta A_{asm}^2 + \gamma A_{asm}^4] = 2\beta A_m + 4\gamma A_m^3 = 0.$$ 

Combining these we can solve for

$$\beta = -\frac{\alpha}{A_m^2} \quad \quad \gamma = \frac{\alpha}{2A_m^4}.$$ 

Inserting $A_m$ we get specifically

$$\beta \sim 1.36 \times 10^7 \quad \gamma \sim 1.63 \times 10^{11}.$$
On the other hand it is likely that the effect seen here is the result of the strong growth of the $k = 0$ mode which will add even powered terms in $A$ to the expansion for $\dot{\psi}$.

Finally in Figure 7.14 we have the temperature case. Here in the $A_{asm} \rightarrow 0$ limit we take $\alpha = .39$. This is maintained to within $\pm 3\%$ up to $G = 1000$, the variations being given by combinations of higher order terms.

7.2.8 The Perturbation Dynamics

We now examine how well the Landau expansions describe the dynamics of the growth of the perturbations as they evolve to the primary branch asymptotic solutions. In Figure 7.15 we show the time histories of a series of runs from just above $G_c$ up to $G = 1000$.\footnote{A note on times:}

Recall that we are working with dimensionless units, and thus with dimensionless time. In our graphics, however, we have made use of a time that is proportional to real time by dividing by $G/2$. Therefore to arrive at the correct dimensionless time form the values shown in the graphics one needs to multiply by $G/2$. Of course where we have calculated growth rates, etc. we have taken this into account, correcting the times to yield the correct dimensionless values.
Figure 7.1b: The perturbation time histories at $G = 492,525,600,1000$. 
nicely given by solutions to Equation (7.1) and can be written here as\textsuperscript{17}

\[
A(t) = \sqrt{-\frac{\lambda}{\alpha}} \frac{1}{1 - \left(1 + \frac{\lambda}{\alpha} \frac{1}{A_0^2} e^{-2\lambda t}\right)^\frac{1}{2}} \tag{7.7}
\]

Rewriting in terms of \(A_{asm}\) this is

\[
A(t) = A_{asm} \frac{1}{1 - \left(1 - \frac{A_{asm}^2}{A_0^2} e^{-2\lambda t}\right)^\frac{1}{2}} \tag{7.8}
\]

In the \(t \to 0\) limit this gives

\[
A(t) = A_0 e^{\lambda t},
\]

an initial exponential growth given by the linear growth rate. This is in good agreement with our results, which are discussed in greater detail in Section 7.3.3. The time taken for the initial rise of this exponential growth expands greatly as \(G_c\) is approached from above. At \(G = G_c\) as \(\lambda \to 0\) the growth timescale becomes infinite. Below \(G_c\) the same is true of the timescale for perturbations to decay away as the critical point is approached from below. This critical slow down is typical of all critical points and phase transitions.\textsuperscript{18}

Inversely as we move away from \(G_c\) (in either direction) the evolution of the system speeds up. For \(G\) above 600 the presence of the Hopf bifurcation near \(G = 1400\) starts to slow down the large \(t\) evolution but the timescale of the initial growth continues to shrink the further we move above \(G_c\). The similarities between the onset of hydrodynamical instability at \(G_c\) and the critical phenomena near a phase transition at \(T_c\) in thermodynamics or statistical mechanics suggest that the application of mean-field theories, Kadanoff’s scaling

\textsuperscript{17}Refer to Section 6.2 and Appendix A.5.

\textsuperscript{18}It is tempting to suggest that with a time-scaling at the critical point the constraint of a fixed correlation speed in the medium (set by the microscopic forces involved) requires that there also be a spatial scaling that leads to the singularity in correlation length at \(G_c\).
theory[85], and Wilson's renormalization group[86] to this basic hydrodynamic bifurcation is worth investigating. This is however beyond the scope of the present work.

For large \( t \) Equation (7.8) reduces to

\[
A(t) = A_{asm} + \frac{1}{2} \left( A_{asm} - \frac{A_{asm}^3}{A_0^2} \right) e^{-2\lambda t}
\]

and the behavior is an exponential approach to the primary branch solution, \( A(t) = A_{asm} + \beta e^{\gamma t} \). With our code we find that for \( G \) near \( G_c, \gamma \), for \( \hat{\psi} \) and \( \hat{T} \), is pure imaginary with a value of roughly twice the linear growth rate, \( \lambda \), as the basic Landau theory predicts. The measured values of \( \gamma_i \) are compared with \( \lambda \) and \( 2\lambda \) in Figure 7.16. In units of \( \lambda \), \( \gamma_i \) is 2.65\( \lambda \) at \( G = 500 \). At \( G = 510 \) it is

![Image of graph showing the imaginary frequency with \( A_{asm}, \gamma_i \) to \( G = 600 \).](image)

Figure 7.16: The exponential approach to \( A_{asm}, \gamma_i \) to \( G = 600 \).

still larger than \( 2\lambda \) (2.35\( \lambda \)) though declining, by \( G = 525 \) it is less than \( 2\lambda \), at 1.65\( \lambda \), and at \( G = 600 \) it is equal to 1.1\( \lambda \). The decline of \( \gamma_i \) towards \( \lambda \) matches the rise of the the \( k = 0 \) mode and the first harmonic which are described by even power expansions, like Equation (6.10), and may be expected to decay as \( \lambda \) rather than \( 2\lambda \).

Although the basic Landau expansion can yield a qualitative under-
standing of \( \gamma \) up to \( G = 600 \), it is not exact.\(^{19}\) More significantly, even at \( G = 510 \) \( \gamma \) is beginning to have a non-zero real component the effect of which can be seen in Figure 7.15 as the solution initially overshooting the primary branch value before decaying to it from above.\(^{20}\) As \( G \) increases this overshoot becomes more pronounced until the decay in turn overshoots and the approach to the primary branch oscillates back and forth an increasing number of times before arriving at the steady state. Similar results where found by Patterson and Imberger\(^{38}\) using a finite difference code to study the initial transients in a square cavity with a horizontal temperature gradient. They show clearly the development with increasing Rayleigh number of increasingly large and persistent oscillations in the Nusselt number as it approaches its final steady state.

This oscillatory behavior, which is the result of the growing relative importance of \( \gamma_r \), cannot be given by a single expansion for real \( A \). In order to describe this type of behavior we must have higher time derivatives than the first order derivative, \( \dot{A} \), of the Landau expansion, which result from coupling together two expansions for first order derivative. This is in fact the nature of our two equations for \( \omega \) and \( \dot{T} \). Gotoh and Mizushima\(^30\) using coupled nonlinear equations for the time dependent intensities \( (A(t)^2 \text{ and } B(t)^2) \) of single Fourier modes with the eigenvector spatial dependence\(^{21}\) and found the appearance of overshoot in the temperature perturbation amplitude with increasing \( G \). Thus one possible extension of the theory is to work with the

\(^{19}\) We note that above \( G = 600 \) \( \gamma \) starts to decline again. See Section 8.1.

\(^{20}\) A log plot of \( \psi(t) - \psi_{srn} \) at \( G = 510 \) also shows clear oscillatory behavior although the amplitude is still small. Note also the apparent break in Figure 7.16 at \( G = 510 \).

\(^{21}\) With single modes rather than full Fourier expansions these equations are of only forth order in \( A/B \), second order in the intensities.
coupled equations that we derive in Section 6.2.

Another possibility is to work in Fourier space where \( A \) and \( B \) are complex. Stuart[61] uses a full Fourier expansion (including the complex conjugate modes) and derives equations for the complex amplitudes of the form

\[
\frac{dA}{dt} = \lambda A + \alpha |A|^2 \tilde{A}
\]

(7.9)

which agrees with Equation (7.1) when \( A \) is real.\(^{22}\)

Writing the complex conjugate equation as

\[
\frac{d\tilde{A}}{dt} = \lambda \tilde{A} + \alpha^* |A|^2 \tilde{A}
\]

(7.10)

we multiply Equation (7.9) by \( \frac{1}{\tilde{A}} \) and Equation (7.10) by \( -\frac{1}{A} \) and add them to get

\[
\frac{d}{dt} \left( \frac{A}{\tilde{A}} \right) = \frac{\dot{A}}{A} + 2i\alpha|A|^2 \frac{\dot{A}}{A}
\]

(7.11)

This can be integrated to yield

\[
A(t) = |A|e^{i\alpha \int_{t_0}^{t} |A|^2 dt}
\]

(7.12)

Which for large \( t \) becomes

\[
A(t) = \sqrt{\frac{\lambda}{\alpha r}} e^{i\alpha \lambda \sqrt{r}(t-t_k)}
\]

(7.13)

where \( t_k \) is an arbitrary phase related to the choice of \( t = 0 \).

In terms of \( A_{asm} \) this is

\[
A(t) = A_{asm} e^{i\alpha A_{asm}(t-t_k)}
\]

(7.14)

If \( \alpha \) is real then

\[
\frac{d}{dt} \left( \frac{A}{\tilde{A}} \right) = 0
\]

\(^{22}\)Note that \( \alpha \) is now complex but that \( \lambda \) is still real.
and there is no oscillation, but as $G$ increases and the perturbation evolves
Stuart's results indicate that when the coefficients become complex there should
be oscillatory behavior at large $t$.

7.2.9 The Upper Primary Branch Dynamics

In Section 7.2.3 we noted that for $G$ shortly above $G_c$ the dynamics of
the perturbation growth carried the system through a series of states similar to
those that appear along the primary branch. For larger $G$ this is no longer the
case. For $G$ large enough that there is significant overshoot on the initial rise,
the evolution during the rise and subsequent decay becomes more complex.

In Figure 7.17 we have the full fields and the perturbations for early
times at $G = 1000$. In the upper left we show the perturbation amplitudes
versus time and mark the points corresponding to the field plots shown in
the rest of the figure. These progress left to right and top to bottom. The
initial perturbation is similar to that shown in Figure 5.39. It grows in a
nearly self similar fashion, as expected for linear exponential growth, until
roughly the second frame. As the rapid rise continues and the mode begins
to saturate, the form of the perturbation starts to change in the third frame.
The amplitude peaks in the fourth frame, where the increasing dominance of
the $k = 0$ mode is seen in the loss of vertical symmetry in the stream function
perturbation. Although the fundamental mode maintains its symmetry the,$k = 0$ mode drives the whole field negative, so that from a sinusoid about zero
the vertical profile becomes a sinusoid about some negative base line. Similarly
the temperature perturbation is increasingly uniform vertically with its main
variation being horizontal.

In Section 7.3.3 we find that, although the fundamental and its har-
Figure 7.17: The initial evolution of fields and perturbations for $G = 1000$. 
monics decline significantly from this peak the \( k = 0 \) does not; and so while the perturbation amplitude drops, the fields cannot return to their original form, and the perturbation itself becomes increasingly one dimensional. A good measure of this is the clockwise rotation of the major axis of the stream function’s negative vortex, which becomes increasingly vertical. The final frame shown, although not at the minimum of the perturbation’s amplitudes, does show the ‘flatest’ fields. The fields become ‘rounder’ as the amplitudes bottom out and rise again. At larger \( t \) there is an oscillatory approach to the asymptotic state (see Figure 7.7).

In the next section we will take a look at using a set of coupled expansions corresponding to the first few Fourier modes of the perturbations for both the stream function and the temperature. We will return to discuss \( \gamma_i \) and \( \gamma_r \) and the oscillations at higher \( G \) when we take up the Hopf bifurcation at \( G_h \sim 1400 \) in Chapter 8.

### 7.3 Fourier Mode Analysis of the Primary Branch

Before describing the second bifurcation at \( G_h \), we extend our analysis of the primary branch with a look at the behavior of the main Fourier modes. At the same time we make use of coupled expansions, following the actual governing equations. In doing this we shall 1) find that while most of the modes are small and can be reasonably neglected, the \( k = 0 \) mode, and to a lesser extent the first harmonic, play a very important role; 2) find the extent to which we can truncate coupled expansions and still retain the main part of our solutions; and 3) obtain a more precise picture of the mode dynamics and the limits of the single expansion approach.
7.3.1 The Main Fourier Modes

First we examine the magnitude of the fundamental mode \((k = \alpha)\) and the two modes that are most rapidly generated, the \(k = 0\) and the first harmonic \((k = 2\alpha)\). The most notable feature of these plots is the magnitude of the \(k = 0\) correction to the initial flow and the rapidity with which, as \(G\) increases, it becomes the dominant mode in the perturbations. For the stream function the \(k = 0\) mode is larger than the fundamental for \(G > 528\) while the temperature \(k = 0\) mode is largest for \(G > 570\). This \(k = 0\) dominance is especially true for the stream function and is the mode that is largely responsible for the decrease of the initial vertical flow in the slot.

We saw in Section 6.2 that each harmonic mode can be described by its own Landau-type series in powers of the fundamental amplitude. Although, as noted for undecomposed perturbations, this cannot model the coupling or the resultant oscillatory behavior we can nevertheless reproduce the asymptotic results and some of the dynamic rates for the individual modes as well. In Figure 7.19 we show the log-log plots of the time asymptotic mode amplitudes against \(\lambda\), the linear growth rate of the fundamental. The fundamental is
Figure 7.19: Log-log plots of $A_{asm}$ for the $k = 0, \alpha, 2\alpha$ modes to $G = 1250$, proportional to $\sqrt{\lambda}$ until above $G = 1000$, as expected for a quantity given by the first two terms of an odd expansion as in Equation (6.15). Similarly the $k = 0$ mode and the first harmonic are both proportional to $\lambda$ as expected from an even series with the second and fourth order terms. At the bottom we see part of the curves for the second and third harmonics indicating that they are proportional to $\lambda$ to the $\frac{3}{2}$ power and $\lambda^2$ respectively. In general we may expect that for the $n^{th}$ harmonic,

$$A_{asm}^n \propto \lambda^{n+\frac{1}{2}}$$

where the leading term is $n + 1^{st}$ order in $A$.

We note as well that the spacing of the harmonics is greater for the stream function than for the temperature. The closer spacing between temperature modes is a general feature that depends in part on the size of the Prandtl number (here 7.5 as usual), which reduces the thermal diffusion coefficient, relative to the viscous damping coefficient leading to a 1.5 order of magnitude difference in the amplitudes of the two fields fundamental modes. Since the harmonics are of increasingly high order in the fundamental, the smaller
fundamental will lead to a more rapidly decreasing set of harmonics.\textsuperscript{23}

Two final notes. First, the \( k = 0 \) mode plays a special role, due in part to reduced coupling to other modes, as the \( z \) derivatives are all zero, and in part to the reduced damping of this mode for the same reason. This mode is therefore much larger than the first harmonic even though they are both second order in the fundamental. It is also the only mode to grow larger than modes that are initially smaller. Second, there is a very noticeable deviation in the first harmonic curve above \( G = 550 \), most pronounced at \( G = 750 \) and \( G = 1000 \).

### 7.3.2 Expansion of the Coupled Equations

Here we derive the exact mode equations for the main three modes up to fourth order in the perturbations. Although in our examples above, the greater magnitude of \( T \) and the decreased spacing of its harmonics would suggest keeping more terms in the temperature than in the stream function for a consistent truncation, we note that this difference is due in part to the magnitude of the Prandtl number. Looking back at the linear eigenvectors in Section 5.6.6 we can see that for very small \( Pr \) the magnitudes become of the same order, most of the change takes place in the size of the temperature perturbation. In fact for \( Pr = 0.001 \), except at very large \( G \), the stream function eigenmode is larger than the temperature eigenmode. We shall therefore, not try to take this into account here but note that if a specific calculation is to be performed the difference may be significant.

\textsuperscript{23}This is a very general argument for small amplitude perturbations. It neglects the fact that the two fields have different structures which will also affect the harmonic mode structure.
Working from Equations (6.4) and (6.5), expanding $\psi$ and $T$ in Fourier components and equating terms of the same order in $A$ and $B$ where these are the amplitudes of the stream function and temperature fundamental modes respectively, we have a series of equations as follows:

To order $A/B$ we can write only linear equations for the fundamentals:

$$\nabla^2 \psi_k^1 = -J(\nabla^2 \psi_k^1, \psi_b) - J(\nabla^2 \psi_b, \psi_k^1) - \frac{1}{G} \partial_z T_k^1 + \frac{1}{G} \nabla^4 \psi_k^1$$

(7.15)

and

$$\dot{T}_k^1 = -J(T_k^1, \psi_b) - J(T_b, \psi_k^1) + \frac{1}{G_P r} \nabla^2 T_k^1,$$

(7.16)

where the dot represents a time derivative and $J(\ ,\ )$ is a Jacobian term. The subscript $b$ indicates the initial unperturbed background state, while other subscripts refer to the perturbation Fourier mode, and the superscripts indicate the order in $A/B$.\(^{24}\)

Moving order $A^2$ we have equations for the $k = 0$ modes and the first harmonics ($k = 2\alpha$):

$$\nabla^2 \psi_0^2 = -J(\nabla^2 \psi_0^2, \psi_b) - J(\nabla^2 \psi_b, \psi_0^2) - \frac{1}{G} \partial_z T_0^2 + \frac{1}{G} \nabla^4 \psi_0^2$$

$$- \frac{1}{2} J(\nabla^2 \psi_k^1, \psi_k^1)$$

(7.17)

$$\nabla^2 \psi_{2k} = -J(\nabla^2 \psi_{2k}^2, \psi_b) - J(\nabla^2 \psi_b, \psi_{2k}^2) - \frac{1}{G} \partial_z T_{2k} + \frac{1}{G} \nabla^4 \psi_{2k}$$

$$- \frac{1}{2} J(\nabla^2 \psi_k^1, \psi_k^1)$$

(7.18)

and

$$\dot{T}_0^2 = -J(T_0^2, \psi_b) - J(T_b, \psi_0^2) + \frac{1}{G_P r} \nabla^2 T_0^2 - \frac{1}{2} J(T_k^1, \psi_k^1)$$

(7.19)

\(^{24}\)To simplify things, we will refer only to orders in $A$ below, but this will imply $A$ and/or $B$.\)
\[ \dot{T}_{2k}^2 = -J(T_{2k}^2, \psi_b) - J(T_b, \psi_{2k}^2) + \frac{1}{GPr} \nabla^2 T_{2k}^2 - \frac{1}{2} J(T_k^1, \psi_k^1) \] (7.20)

At order \( A^3 \) we gain the first nonlinear contributions to the fundamentals, along with additions to the linear terms.

\[ \nabla^2 \psi_k^3 = -J(\nabla^2 \psi_k^3, \psi_b) - J(\nabla^2 \psi_b, \psi_k^3) - \frac{1}{G} \partial_a T_k^3 + \frac{1}{G} \nabla^4 \psi_k^3 \]
\[ - J(\nabla^2 \psi_k^2, \psi_k^3) - J(\nabla^2 \psi_0^2, \psi_k^3) - J(\nabla^2 \psi_{2k}^2, \psi_k^3) - J(\nabla^2 \psi_{2k}^2, \psi_k^1) \] (7.21)

\[ \dot{T}_k^3 = -J(T_k^2, \psi_k^3) - J(T_b, \psi_k^3) + \frac{1}{GPr} \nabla^2 T_k^3 - J(T_k^1, \psi_0^3) - J(T_0^1, \psi_k^3) \]
\[ - J(T_k^1, \psi_{2k}^3) - J(T_{2k}^2, \psi_k^3) \] (7.22)

Finally the order \( A^4 \) terms give us the last we need to compare to the expansions of Equations (7.1) and (7.2):

\[ \nabla^2 \psi_0^4 = -J(\nabla^2 \psi_0^4, \psi_b) - J(\nabla^2 \psi_b, \psi_0^4) - \frac{1}{G} \partial_a T_0^4 + \frac{1}{G} \nabla^4 \psi_0^4 - J(\nabla^2 \psi_0^2, \psi_0^2) \]
\[ - \frac{1}{2} J(\nabla^2 \psi_k^2, \psi_k^2) - \frac{1}{2} J(\nabla^2 \psi_{2k}^2, \psi_{2k}^2) \] (7.23)

\[ \nabla^2 \psi_{2k}^4 = -J(\nabla^2 \psi_{2k}^4, \psi_b) - J(\nabla^2 \psi_b, \psi_{2k}^4) - \frac{1}{G} \partial_a T_{2k}^4 + \frac{1}{G} \nabla^4 \psi_{2k}^4 - \frac{1}{2} J(\nabla^2 \psi_k^3, \psi_k^1) \]
\[ - \frac{1}{2} J(\nabla^2 \psi_{2k}^2, \psi_{2k}^2) - \frac{1}{2} J(\nabla^2 \psi_{2k}^2, \psi_{2k}^2) \] (7.24)

and

\[ \dot{T}_0^4 = -J(T_0^4, \psi_b) - J(T_b, \psi_0^4) + \frac{1}{GPr} \nabla^2 T_0^4 - J(T_0^2, \psi_0^2) - \frac{1}{2} J(T_k^3, \psi_k^1) \]
\[ - \frac{1}{2} J(T_k^1, \psi_k^3) - \frac{1}{2} J(T_{2k}^2, \psi_{2k}^2) \] (7.25)

\[ \dot{T}_{2k}^4 = -J(T_{2k}^4, \psi_b) - J(T_b, \psi_{2k}^4) + \frac{1}{GPr} \nabla^2 T_{2k}^4 - \frac{1}{2} J(T_k^3, \psi_k^1) - \frac{1}{2} J(T_k^1, \psi_k^3) \]
\[ - \frac{1}{2} J(T_{2k}^2, \psi_{2k}^2) - \frac{1}{2} J(T_{2k}^2, \psi_{2k}^2) \] (7.26)

Now since both the \( k = 0 \) and the background state are independent of \( z \) the Jacobians of two such terms vanish, while those with one such term are reduced, e.g.

\[ J(T_0, \psi_k^1) = \partial_a T_0 \partial_z \psi_k^1 \]
Taking account of this and reducing \( \nabla \) to \( \partial^2_z \) when acting on a \( z \) independent term, and recalling that the variables are \( z \) Fourier modes in order to evaluate the \( z \) derivatives, we have the reduced equations to \( A^1 \):

\[
\nabla^2 \psi^1_k = ik[\nabla^2 \psi^1_k \partial_z \psi_b - \psi^1_k \partial_z \nabla^2 \psi_b] - \frac{1}{G} \partial_z T^1_k + \frac{1}{G} \nabla^4 \psi^1_k \tag{7.27}
\]

\[
\dot{T}^1_k = ik[T^1_k \partial_z \psi_b - \psi^1_k \partial_z T_b] + \frac{1}{G P r} \nabla^2 T^1_k \tag{7.28}
\]

to \( A^2 \):

\[
\nabla^2 \psi^2_0 = -\frac{1}{G} \partial_z T^2_0 + \frac{1}{G} \partial^4_z \psi^2_0 + \frac{ik}{2} \left[ \nabla^2 \psi^1_k \partial_z \psi^1_k - \psi^1_k \partial_z \nabla^2 \psi^1_k \right] \tag{7.29}
\]

\[
\nabla^2 \psi^2_{2k} = ik[T^1_k \partial_z \psi_0 - \psi^1_k \partial_z T_b] + \frac{1}{G} \partial_z T^2_k + \frac{1}{G} \nabla^4 \psi^2_{2k} \tag{7.30}
\]

\[
\dot{T}^2_0 = \frac{1}{G P r} \partial^2_z T^2_0 + \frac{ik}{2} [T^1_k \partial_z \psi^1_k - \psi^1_k \partial_z T^1_k] \tag{7.31}
\]

\[
\dot{T}^2_{2k} = ik[T^1_k \partial_z \psi_{2k} - \psi^1_{2k} \partial_z T_b] + \frac{1}{G P r} \nabla^2 T^2_{2k} + \frac{ik}{2} [T^1_k \partial_z \psi^1_k - \psi^1_k \partial_z T^1_k] \tag{7.32}
\]

to \( A^3 \):

\[
\nabla^2 \psi^3_k = ik[\nabla^2 \psi^3_k \partial_z \psi_b - \psi^3_k \partial_z \nabla^2 \psi_b] - \frac{1}{G} \partial_z T^3_k + \frac{1}{G} \nabla^4 \psi^3_k + ik[\nabla^2 \psi^1_k \partial_z \psi^1_k - \psi^1_k \partial_z \nabla^2 \psi^1_k] \\
- \psi^3_k \partial_z \nabla^2 \psi^1_k + \nabla^2 \psi^2_k \partial_z \psi^1_k - \psi^2_k \partial_z \nabla^2 \psi^1_k + \frac{ik}{2} \left[ \nabla^2 \psi^1_k \partial_z \psi^1_{2k} - \psi^1_k \partial_z \nabla^2 \psi^1_{2k} \right] \tag{7.33}
\]

\[
\dot{T}^3_k = ik[T^3_k \partial_z \psi_b - \psi^3_k \partial_z T_b] + \frac{1}{G P r} \nabla^2 T^3_k + ik[T^1_k \partial_z \psi^1_0 - \psi^1_k \partial_z T^1_0] \\
+ T^2_k \partial_z \psi^1_k - \psi^1_{2k} \partial_z T^1_k + \frac{ik}{2} [T^1_k \partial_z \psi^1_{2k} - \psi^1_k \partial_z T^1_{2k}] \tag{7.34}
\]

to \( A^4 \):

\[
\nabla^2 \psi^4_0 = -\frac{1}{G} \partial_z T^4_0 + \frac{1}{G} \partial^4_z \psi^1_0 + \frac{ik}{2} \left[ \nabla^2 \psi^3_k \partial_z \psi^1_k - \psi^1_k \partial_z \nabla^2 \psi^3_k \right] \\
+ \nabla^2 \psi^3_k \partial_z \psi^3_k - \psi^3_k \partial_z \nabla^2 \psi^1_k + ik[\nabla^2 \psi^2_{2k} \partial_z \psi^2_{2k} - \psi^2_{2k} \partial_z \nabla^2 \psi^2_{2k}] \tag{7.35}
\]
\[ \nabla^2 \psi_{2k}^4 = i2k[\nabla^2 \psi_{2k}^{4} \partial_{z} \psi_{k} - \psi_{2k}^{4} \partial_{z} \nabla^2 \psi_{k}] - \frac{1}{G} \partial_{z} T_{2k}^4 + \frac{1}{G} \nabla^4 \psi_{2k}^4 \\
\quad + \frac{i}{2} \left[ \nabla^2 \psi_{2k}^2 \partial_{z} \psi_{k}^1 - \psi_{2k}^2 \partial_{z} \nabla^2 \psi_{k}^1 + \nabla^2 \psi_{k}^1 \partial_{z} \psi_{k}^3 - \psi_{k}^3 \partial_{z} \nabla^2 \psi_{k}^1 \right] \\
\quad + i2k[\nabla^2 \psi_{2k}^2 \partial_{z} \psi_{0}^2 - \psi_{2k}^2 \partial_{z} \nabla^2 \psi_{0}^2] \quad (7.36) \]

\[ \dot{T}_0^4 = \frac{1}{GP_{F}} \partial_{z} T_0^4 + \frac{i}{2} \left[ T_k^3 \partial_{z} \psi_{k}^1 - \psi_{k}^1 \partial_{z} T_k^3 + T_k^1 \partial_{z} \psi_{k}^3 - \psi_{k}^3 \partial_{z} T_k^3 \right] \\
\quad + i \partial_{z} \psi_{2k}^2 \partial_{z} \psi_{0}^2 \quad (7.37) \]

\[ \dot{T}_{2k}^4 = i2k[T_{2k}^4 \partial_{z} \psi_{0} - \psi_{2k}^4 \partial_{z} T_{b}] + \frac{1}{GP_{F}} \nabla^2 T_{2k}^4 + \frac{i}{2} \left[ T_k^3 \partial_{z} \psi_{k}^1 - \psi_{k}^1 \partial_{z} T_k^3 \right] \\
\quad + i \partial_{z} \psi_{2k}^2 \partial_{z} \psi_{0}^2 \quad (7.38) \]

where \( \nabla^2 \equiv \partial^2_z - k^2 \).

Finally using the definitions and relations given in Section 4.4.5 for the derivatives of the 1-D background, and combining the first- and third-order terms for the fundamental, Equations (7.27), (7.3.2), (7.33), and (7.34) give

\[ \nabla^2 \psi_{k}^{13} = -i[k + W_k \nabla^2] \psi_{k}^{13} - \frac{1}{G} \partial_{z} T_{k}^{13} + \frac{1}{G} \nabla^4 \psi_{k}^{13} + ik[\partial_{z} (\psi_{0}^2 + \frac{1}{2} \psi_{2k}) - \psi_{2k} \partial_{z}] \nabla^2 \psi_{k}^{13} \\
\quad - i \partial_{z} \psi_{2k}^2 \partial_{z} \psi_{0}^2 \quad (7.39) \]

\[ \dot{T}_{k}^{13} = -i[k \psi_{k}^{13} + W_k T_{k}^{13}] + \frac{1}{GP_{F}} \nabla^2 T_{k}^{13} + ik[\partial_{z} (\psi_{0}^2 + \psi_{2k}) - \psi_{2k} \partial_{z}] \partial_{z} T_{k}^{13} \\
\quad - i \partial_{z} (T_{0}^{2} + \frac{1}{2} T_{2k}^{2}) - T_{2k}^2 \partial_{z} \psi_{k}^{13} \quad (7.40) \]

where \( \psi_{k}^{13} = \psi_{k}^{1} + \psi_{k}^{3} \) etc.

Similarly combining the second- and fourth-order terms for the \( k = 0 \) mode, Equations (7.29), (7.31), (7.35), and (7.37) yield

\[ \nabla^2 \psi_{0}^{24} = -\frac{1}{G} \partial_{z} T_{0}^{24} + \frac{1}{G} \nabla^4 \psi_{0}^{24} + \frac{i}{2} \left[ \partial_{z} \psi_{0}^{13} - \psi_{0}^{13} \partial_{z} \right] \nabla^2 \psi_{0}^{24} - \left[ \partial_{z} \psi_{k}^{1} - \psi_{k}^{1} \partial_{z} \right] \nabla^2 \psi_{k}^{24} \]

\[ + i \partial_{z} \psi_{2k}^2 \partial_{z} \psi_{0}^2 \quad (7.41) \]

\[ \dot{T}_{0}^{24} = \frac{1}{GP_{F}} \partial_{z} T_{0}^{24} + \frac{i}{2} \left[ \partial_{z} \psi_{0}^{13} - \psi_{0}^{13} \partial_{z} \right] T_{k}^{13} + \left[ \partial_{z} \psi_{k}^{1} - \psi_{k}^{1} \partial_{z} \right] T_{k}^{3} \quad (7.41) \]
Lastly Equations (7.30), (7.36), (7.32), and (7.38) combine to give the first harmonic to fourth order:

\[
\nabla^2 \psi_{2k}^{24} = -i2k[x + W_b \nabla^2] \psi_{2k}^{24} - \frac{1}{G} \partial_\alpha T_{2k}^{24} + \frac{1}{G} \nabla^4 \psi_{2k}^{24} + \frac{i k}{2} \left\{ [\partial_\alpha \psi_k^{13} - \psi_k^{13} \partial_\alpha] \nabla^2 \psi_k^1 - [\partial_\alpha \psi_k^1 - \psi_k^1 \partial_\alpha] \nabla^2 \psi_k^1 \right\} + i2k[\partial_\alpha \psi_0^2 - \psi_0^2 \partial_\alpha] \nabla^2 \psi_{2k}^2
\]

(7.43)

\[
\dot{T}_{2k}^{24} = -i2k[\psi_{2k}^{24} + W_b T_{2k}^{24}] + \frac{1}{G \nu} \nabla^2 T_{2k}^{24} + \frac{i k}{2} \left\{ [\partial_\alpha \psi_k^{13} - \psi_k^{13} \partial_\alpha] T_k^1 + [\partial_\alpha \psi_k^1 - \psi_k^1 \partial_\alpha] T_k^2 \right\} + i2k[\partial_\alpha \psi_0^2 - \psi_0^2 \partial_\alpha] T_{2k}^2
\]

(7.44)

### 7.3.3 The Mode Dynamics

We present the mode dynamics (time evolution), while making a few comments on the mode equations derived in the last section.

As expected, Equations (7.39) to (7.44) are consistently composed of terms either even or odd in \( \psi_k^1 \) and \( T_k^1 \), thus allowing the Landau-type series expansions. In general the form of the equations is the same for each pair of modes. There are, however, several key differences. The factor of \( nk \) in front of the linear Jacobian terms (first on the right-hand sides), leads to the absence of \( z \) derivatives in Equations (7.41) and (7.42), and the 2 in Equations (7.43) and (7.44). This factor of 2, which derives from taking \( z \) derivatives, appears in some of the Jacobian terms as well. In Figure 7.20 we see the initial growth of the top three modes at \( G = 495 \). In particular we see that the \( k = 0 \) mode and the first harmonic start at the same magnitude but that the initial growth of the \( k = 0 \) mode is explosive.\(^{25}\) After a brief spurt due to the lack of \( z \)

\(^{25}\) It is probable that the initial growth of the first harmonic is actually \( 2\lambda \) but appears to be less here due to round-off error.
Figure 7.20: The initial growth of the $k = 0, \alpha, 2\alpha$ modes at $G = 495$. Derivatives in the altered linear terms, the growth rate for both terms is equal at roughly $2\lambda$, as one would expect for an even-series expansion.

In Figure 7.21 we see the full evolution of the modes, noting that not only does the $k = 0$ mode grow explosively initially but that it takes longer to saturate. These two factors combine to yield the difference that we saw in Figure 7.19 in the asymptotic values of these two modes. Comparing the right-hand-most Jacobian terms in Equations (7.41) and (7.42) with those in Equations (7.43) and (7.43), where we are specifically taking $Pr = 7.5$, we
see that the former are much smaller owing to the great difference in size of the $k = 0$ modes and the first harmonics. Thus the $k = 0$ modes saturate at a larger amplitude, as it is these nonlinear Jacobian terms that cause the growth to saturate. Figure 7.21 also shows how, near $G_c$, the modes (with the exception of the initial stage) grow exponentially and then saturate very cleanly and are probably well described by a set of series of low order.

At $G = 600$ we have checked the mode growth rates and find that they go roughly as $n\lambda$ where $nk$ is the wavenumber of the mode. Thus the fundamental grows like $\lambda$, the $k = 0$ mode and the first harmonic as $2\lambda$, and in general the $n + 1^{st}$ harmonic as $n\lambda$. In Figure 7.22 we indicate the regions

![Stream function and temperature perturbation plots](image)

**Figure 7.22:** The growth of the $k = 0, \alpha, 2\alpha$ modes at $G = 600$.

in which the growth rate is within $\sim 2\%$ for the fundamentals and within $\sim 5\%$ for the $k = 0$ mode and the first harmonic. The fundamental modes grow at the expected rate from the outset until they start to slow down as they begin to saturate. The $k = 0$ modes start out growing faster than expected, as was noted above, but rapidly approach $2\lambda$ growth, which they approximate until they too start to saturate. For the higher harmonics there is an early stage where the limited accuracy of our data obscures the real growth rate, but
once they appear out of the noise, they fit the pattern up to at least the third harmonic.

Finally in Figure 7.23 we present a sample of the mode dynamics at $G = 1000$ where the overshoot and subsequent oscillatory decay are well developed. In this figure we show the amplitudes rather than their logs as we have above, giving a good idea of the relative size of the modes.\textsuperscript{26} We can

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig7.23}
\caption{The mode amplitudes at $G = 1000$.}
\end{figure}

see the initial rise of the fundamental modes, followed shortly by the $k = 0$ and $k = 2\alpha$ modes. Note that while the peaking of harmonics follows that of the fundamental, it does so much more closely than that of the $k = 0$ mode. This mode has a generally different dynamic from the other modes. The most obvious example is that while the other modes rise far beyond their time-asymptotic values before damping down to them, the $k = 0$ modes rise throughout and exhibit a more restrained oscillation over a monotonically rising background. Thus while the initial rise of the fundamental and $k = 0$ modes is nearly the same the asymptotic values are quite different. We shall explore the

\textsuperscript{26}Refer to Figure 7.15 for the corresponding undecomposed amplitudes.
distinctions more closely in the Chapter 8 where we examine in greater detail
the oscillatory behaviors exhibited as we approach $G_h$ and go beyond.
Chapter 8

The Hopf Bifurcation and Thermal Relaxation Oscillations

In the previous chapter we discussed the primary bifurcation at $G_c = 491.818$ and applied Landau expansion theory to the VSC problem along the primary branch. up to $G = 1000$. We saw the limits of this approach as the nonlinear coupling of the two variables and many modes leads to dynamic behavior that requires a more complex description.

As shown in Figure 7.2 there is a second bifurcation at about $G = 1400$. This is a Hopf bifurcation, where, as discussed in Section 6.1, a steady solution gives way to an oscillatory one. We now investigate this second bifurcation and the oscillatory behavior associated with it, first as transients below $G_h$ and then as thermal relaxation oscillations above $G_h$.

8.1 The Approach to $G_h$

In Section 7.2.8 we defined the time dependence of the approach to the primary branch solution as $\propto e^{\gamma_t t}$. Writing the real and complex parts explicitly we have

$$A(t) - A_{asm} \propto e^{(\gamma_r + i\gamma_i)t}.$$  

The imaginary frequency being the negative of the real growth rate will drive the solution to a steady state $A_{asm}$ whenever it is positive. When there is a significant real frequency as well, $\gamma_r \neq 0$, then this approach will be oscillatory. We have found that at and just above $G_c$, $\gamma$ is purely imaginary, and directly
related to the linear growth rate $\lambda$, but that as $G$, and with it the value of $A_{asm}$, increases the nonlinearities affect the shape of the fields and $\gamma$ becomes complex with non-zero $\gamma_r$.

In Figure 7.16 we plotted $\gamma_i$ as it increased with $G$ up to 600, and indicated its relationship to $\lambda$ was in general agreement with the Landau expansion theory. Here in Figure 8.1 we plot the logarithm of $\gamma_i(G)$ above $G = 600$

![Graph showing the imaginary frequency as a function of G]

Figure 8.1: The log of the rate of decay to the asymptotic steady state.

and find that instead of continuing to grow, as $\lambda$ does, it declines. In terms of a linear stability analysis of the primary branch $\gamma_i$ forms the negative real part of an eigenvalue which passes through zero at a secondary bifurcation at $G_h \sim 1400$. Since $\gamma_r$ is non-zero, however, this next critical point is a Hopf bifurcation. We shall discuss this event in Section 8.2 after we have taken a closer look at the transient oscillations found for $G < G_h$.

In Figure 8.2 we give the perturbation dynamics at $G = 1360$, just below $G_h$. After a sharp initial rise the perturbation amplitudes peak at roughly twice their asymptotic values, followed by strong oscillatory behavior as the system relaxes towards a stationary asymptotic state. This behavior is similar
Figure 8.2: The perturbation dynamics at $G = 1360$.

to that discussed above for $G = 1000$,\textsuperscript{1} although the peak amplitudes at $G = 1360$ are roughly double those at $G = 1000$, the secondary vortex is stronger and the curl in the temperature field near peak amplitude is more pronounced. The timescale of the relaxation to the asymptotic state, is however, quite long compared to that for lower $G$ (refer to Figure 7.23 showing the dynamics at $G = 1000$). In fact the run shown in Figure 8.2 has blown up owing to accumulated numerical error before reaching the assumed final state,\textsuperscript{2} even though it has run more than 4 times longer (in real time, over 5 times in dimensionless time).\textsuperscript{3}

In Figure 8.3 we show the mode histories at $G = 1360$. We note again that 1) the temperature mode amplitudes are generally larger than those for

\textsuperscript{1}See Section 7.2.8.

\textsuperscript{2}In all cases as $G$ increases, numerical error, most likely deriving from the abrupt cutoff of mode amplitude at $NX$ and $NZ$, limits the number of timesteps that can be performed. In this case we have run slightly over 100,000ts. This problem is impervious to changing the timestep size (within bounds) but can be postponed by decreasing the grid spacing, at a cost in computation time. In the future we hope to be able to mask the high-end mode amplitudes to smooth the mode profile and avoid an abrupt cutoff. The frequency cutoff is analogous to the finite-particle techniques employed in particle simulations. See [9] and [87].

\textsuperscript{3}To convert plotted time to real time multiply by $\frac{L^2}{\alpha}$, where $2L$ is the slot width. To convert dimensionless time to real time multiply by $\frac{L^2}{2\nu}$. For water $\nu$ is of the order 0.01, so we have roughly $t_{\text{in seconds}} \sim 10(2L)^2t_{\text{dimensionless}}$. 
Figure 8.3: The mode dynamics at $G = 1360$.

the stream function; 2) the stream function harmonics are smaller relative to the $k = 0$ mode than are the temperature harmonics; and that 3) in both fields the fundamental and the harmonics experience greater relative decay from their initial peak than do the $k = 0$ modes.

Figure 8.4: A detailed study of the oscillations at $G = 1360$.

Picking a point where the oscillation amplitude is roughly constant but still fairly large, we look at in detail at a single oscillation. In Figure 8.4

---

4This oscillation is indicated by the vertical lines in Figure 8.2.
we show the three main modes for the stream function and the temperature, superimposed and normalized to the same size. The period is roughly 680 dimensionless time units, which for a 1mm wide slot converts to 1 second of real time. While not strictly sinusoidal, these oscillations are quite regular, being symmetrical in time with the maxima and minima having the same width at half height. The form of each mode is identical to that of the others and the two fundamentals are exactly in phase. In each case the fundamental leads the other modes, followed closely by the second harmonic and then, less closely, by the \( k = 0 \) mode. In Table 8.1 we list each mode and its phase lag relative to the perturbation fundamentals as a fraction of the period and in degrees.\(^5\)

<table>
<thead>
<tr>
<th>mode</th>
<th>phase lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_k ) &amp; ( T_k )</td>
<td>0.0</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>0.0645</td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>0.0968</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>0.161</td>
</tr>
<tr>
<td>( \psi_0 )</td>
<td>0.258</td>
</tr>
</tbody>
</table>

Table 8.1: The Fourier mode phase differences at \( G = 1360 \).

The coupling of the temperature modes is tighter than that of the stream function modes, reflecting the difference in magnitude of the two fields such that we expect the \( J(\hat{\omega}, \hat{\psi}) \) term to be larger than the \( J(\hat{T}, \hat{\psi}) \) term. It may also be dependent on a closer correlation between the gradients of \( \psi \) and \( T \) than between the gradients of \( \psi \) and \( \omega \). We have not yet examined this in detail but it is expected from experimental results that at larger \( G \) the vorticity is confined near the slot walls.\(^6\) This can also be seen in the high \( G \) eigenmodes. Although Figure 5.31 is for low \( Pr \), it shows a vorticity field nearly the same as those at \( Pr = 7.5 \), for comparable \( G \) and \( \alpha \).

---

\(^5\)Surely a coincidence that these are in proportion to 2,3,5,8?

\(^6\)See [3] and [34].
In Figure 8.5 we show the kinetic energy and the Nusselt number time histories for the same period as the mode analysis of Figure 8.4.\textsuperscript{7} The stream function perturbation is essentially negative, reducing the kinetic energy, and since the $k = 0$ mode is the dominant component of the perturbation, the form of the kinetic energy oscillation is very nearly the inverse of the stream function $k = 0$ mode. The form of the Nusselt number oscillations varies slightly from that of the perturbation modes and \textit{leads} the fundamentals by $\sim 23^\circ$.

During the oscillation the fields change only slightly, as shown in Figure 8.6, in which we show the fields at their 'roundest' where the the vortex is strongest relative to the $k = 0$ mode, and at their 'flattest' where it is relatively weakest. The former is on the left and the latter on the right.\textsuperscript{8} There is really very little change as the oscillations are small, and one can see, in comparison to Figure 7.7, that the asymptotic solution differs from that at

\textsuperscript{7}It is important to bear in mind that the 'kinetic energy' plotted here, and throughout our work, is the perturbation kinetic energy and that the total-flow kinetic energy varies inversely.

\textsuperscript{8}The location of these field plots is indicated by the vertical lines in Figures 8.4 and 8.5.
Figure 8.6: The fields near the minima and maxima of the oscillations in the perturbation fundamental modes at \( G = 1360 \).
\( G = 1000 \) in amplitude but very little in form.

With this, and with reference to the discussion of the dynamics at \( G = 1000 \) in Sections 7.2.8 and 7.3.3, we have characterized the upper primary branch dynamics leading up to the second bifurcation at \( G_h \).

8.2 The Second Bifurcation at \( G_h \)

As \( \gamma_i \to 0 \) at \( G_h \) the decay to \( A_{a_{em}} \) takes longer and longer, another instance of critical point time-scaling, while above \( G_h \), where \( \gamma \) is a purely real frequency, the stationary state recedes to infinity and the secondary solution is oscillatory. It would be nice to be able to say exactly what the value of \( G \) is at which this bifurcation takes place. In the case of the primary bifurcation we set up an equation (in Section 5.2) that can be solved numerically to give the neutral stability curve. Minizing the resulting solution gives \( G_c \). In practice we were able to solve for \( G_c \) to high accuracy by linearizing about the basic state, solving for \( \lambda \) as the eigenvalue of the linearized equations, and changing \( G \) until \( \lambda \) was arbitrarily small. We could also have used the full nonlinear code and
search for the critical point by working from either side, however, the critical point time-scaling makes this computationally slow and expensive.\(^9\)

In solving for \(G_h\) there is no precise analytical theory that can be solved. In theory we can linearize our equations around the primary branch solution and search for the crossing of the eigenvalue \(\gamma_i\) into the positive real side of the complex plane.\(^{10}\) The analytic form of this solution is not available, however, only the amplitudes of the spectral modes that we have solved for numerically. While it is possible to rework our linear programs to set up the eigenvalue problem for an arbitrary initial state, which could be defined by such a simulation result. Forming the necessary matrix operators from an arbitrary input field is not, however, trivial and has not yet been fully worked out.

In the meantime we can estimate \(\gamma_i\) from our simulation results and use this to give at least an approximate value for \(G_h\).\(^{11}\) Figure 8.1 for \(\gamma_i\) suggests that \(G_h\) lies above \(G = 1360\), probably near \(G = 1400\). The value of \(\gamma_i\) at \(G = 1400\) is somewhat uncertain as it is so small that when the numerical instability appears the amplitude of the oscillations is still fairly large. It is not clear whether they are shrinking to zero or will eventually stabilize at some small but finite amplitude. Based on the finite amplitude of the oscillations at \(G = 1450\) it appears that \(\gamma_i\) is negative. Since we are in effect using hugh

---

\(^9\)It is of interest that to some measure the timestep near \(G_c\) can itself be scaled, thus reducing the computation significantly.

\(^{10}\)In the eigenvalue problem for the basic 1-D state the complex eigenvalues come in complex pairs. If the same is true here then there are two new solutions rather than one. These may be related to the two 'cat's eye' modes that we will discuss in Chapter 9.

\(^{11}\)To calculate \(\gamma_i\) we estimate the asymptotic value of the perturbation amplitudes, \(|\hat{\psi}|_{asm}\) and \(|\hat{T}|_{asm}\), and then plot log of the absolute value of \(\hat{\psi}(t) - |\hat{\psi}|_{asm}\) and \(\hat{T}(t) - |\hat{T}|_{asm}\) against time. The estimates are refined until peaks on the resulting plots are all in line (if the estimate is off the peaks generated by the minima and maxima \(\hat{\psi}(t)\) and \(\hat{T}(t)\) are of different sizes). The slope of these peaks gives the value of the damping constant \(\gamma_i\).
perturbations of the primary branch solution, much larger than the secondary branch oscillatory solutions, it is impossible to estimate \( \gamma_i \) above \( G_h \) starting from the 1-D basic state. In the future we hope to be able to estimate the primary branch solution and then perturb it with small perturbations so that we can estimate \( \gamma_i \) above \( G_h \) as well.

In looking at \( \gamma_i \) we have been approaching \( G_h \) from below and have been able to give a lower bound \( G_h > 1360 \). Alternatively the results for \( G > 1400 \) can be analyzed to calculate other quantities that also change sharply at the critical point and may be used to estimate an upper bound.

In Figure 8.7 we show the real frequency \( \gamma_r \) for \( G = 1000 \) to \( G = 2000 \).\(^\text{12}\) Below \( G = 1250 \) this is roughly a constant, \( \gamma_r = .0096 \), but above

![The real frequency of oscillations around the primary branch.](image)

Figure 8.7: The real frequency of oscillations around the primary branch.

\( G = 1360 \) it declines rapidly before leveling off again near \( G \sim 2000 \) at about \( 2/3 \) of its previous value.\(^\text{13}\) Perhaps the sharpest changes can be seen in using the amplitude of the oscillations as the order parameter. In Figure 8.8 we

\(^{12}\) Note, that the point at \( G=2000 \) is a bit uncertain as there are no longer regular oscillations here. See Section 8.3 below.

\(^{13}\) The frequency is measured in inverse dimensionless time units.
see the sharp increase in the amplitude of the stream function perturbations above $G_h$. This clearly places an upper bound of $G = 1400$ on the location of $G_h$ and given enough computing time we could pin it down more closely. Although it appears that this bifurcation may not be as clean as the one at $G_c$ our data is limited by the critical point time-scaling as mentioned above. It is reasonable that given a larger grid, so that we come closer to the true asymptotic state, this bifurcation would appear more cleanly. Examination of the other diagnostics for the oscillatory regime (as discussed in Section 8.3) also suggest an upper bound of 1400 on $G_h$. We conclude therefore, that the secondary Hopf bifurcation takes place between $G = 1360$ and $G = 1400$.

In Figure 8.9 we show this bifurcation using the kinetic energy as the order parameter (see also Figure 7.2). The two curves that appear above $G_h$ represent the maximum and minimum values of the kinetic energy during each oscillation. As we will see below these are highly nonlinear oscillations with

---

14 The double values at $G = 2250, 2500$ reflect a range of amplitude for irregular oscillations found there. See also Section 9.1.

15 The values at $G = 2250, 2500$ are averages, as the oscillations are no longer regular
Figure 8.9: The kinetic energy bifurcation diagram above $G_h$. relatively large amplitudes away from $G_h$. It is not yet clear whether they represent two mutually unstable solutions or a single oscillatory solution.

As far as we know this secondary bifurcation has not been seen before. Mizushima [25] [26] has solved the VSC problem for stationary states and found the primary branch but his equations do not admit the secondary branch oscillatory states. Fujimura and Mizushima [29] using an purely kinetic model (Navier-Stokes equation only, $Pr = 0$ limit) with a small number of Fourier modes (up to 8) do find oscillatory behavior, which for $N = 8$ has similarities to the oscillations above $G_h$.\[16\] Comparison is hampered, however, by the differences in parameters. Besides taking $Pr = 0$, they chose $\alpha = .5$ at $G = 1000$ where, in a fully resolved model, the fundamental mode is linearly stable but the first and second harmonics are not.\[17\] Referring back to Figure 5.5 we see that for an 8 mode model, however, the fundamental is in fact unstable there.

---

\[16\] This transition is marked by the squiggly line. The vertical lines mark the changes in the phase of the 'cat's eye' modes relative to the initial perturbation. We shall discuss this Section 8.3.2 below.

\[17\] The $k = 0$ mode dominates, the period is roughly the same, and unlike the oscillations at $N = 4, 6$, all the modes have the same frequency.

\[17\] See Section 10.2 for more on the mode competition at low $\alpha$. 
It is possible therefore that they are seeing the same type of oscillations.

8.3 The Regular Relaxation Oscillations Above $G_h$

Just as above $G_c$ the value of the asymptotic perturbation amplitude increases rapidly we find that the amplitude of the new oscillations above $G_h$ grow rapidly, as shown in Figure 8.8 above. Figure 8.7 shows the decrease in frequency that mirrors this rise in the amplitude suggesting a that the period of the oscillations is proportional to their amplitude. In Figure 8.10 we plot

![THE NORMALIZED PERIOD](Image)

Figure 8.10: The normalized period.

the period divided by the oscillation amplitude. Below $G_h$, where the period is roughly constant and the amplitude is small, the normalized period appears inversely proportional to $G$, dropping rapidly. Above $G_h$ it is nearly constant as the period is directly proportional to the amplitude;

$$Period \propto 4100(\hat{\psi}_{\text{max}} - \hat{\psi}_{\text{min}}).$$

The same relationship holds for the temperature perturbation as well. It is not surprising then that as the amplitude oscillation increases with increasing $G$ we find that the form of the oscillations changes as well. We should note that
if we look again at the $G = 1360$ run, and focus the first oscillations after the initial rise, that they show many of the changes that we ascribe here to the $G > G_h$ oscillations. This suggests that the form of the oscillations depends mainly on their amplitude rather than directly on $G$.

### 8.3.1 Detailed analysis of the oscillations.

We start with a closer look at the form of the oscillations at $G = 1500$, which is far enough above $G_h$ that the oscillation amplitude is significant and we can begin to make out the accompanying changes in their form. In Figure 8.11 we present a mode decomposition at $G = 1500$ in order to show the relative scales of the three main modes. Comparing this to Figure 8.3, the initial rise is

![Figure 8.11: The mode histories at $G = 1500$.](image)

still led by the fundamental followed by the $k = 0$ mode and the first harmonic. The initial peak is proportionally higher and there is only slight damping, the asymptotic oscillations establishing themselves quite rapidly.\(^{18}\) Other points to

---

\(^{18}\)Note that the dimensionless timescales are found by multiplying by $G/2$. The given scales are proportional to real time.
note are the high average value of the $k = 0$ mode amplitudes, the nearly zero minima of the fundamental, and especially of the harmonics, and the clearly 'sawtooth' form in the $k = 0$ mode.

In Figure 8.12 we show the normalized mode amplitudes for a single oscillation. The period is now about 850 dimensionless time units (up from 680

![Diagram](image)

Figure 8.12: The detailed mode study of an oscillation at $G = 1500$.

at $G = 1360$, though the real time period is only a little more (by about 10%). The fundamental modes still have the same profile and are exactly in phase but, although their maxima and minima are still generally symmetric in time, the maxima are only 60% as wide at half-height as the minima. The profiles of the other modes are now different from that of the fundamentals. The first harmonic is time-symmetrical, but with a maxima width-at-half-height only 32% that of its minima, which really flatten out at the bottom. The harmonics are now confined to short pulses between which they have little energy. As before the temperature harmonic leads that of the stream function in its rise. We note, however, that the fundamentals and both the harmonics fall in unison.

The $k = 0$ modes show the greatest change. They have identical rise profiles, the same as that of the fundamentals (only the first harmonics are
slightly different, rising a bit more steeply), but the stream function falls more slowly than the temperature. Both \( k = 0 \) modes fall much more slowly than they rise, with a nearly constant slope most of the way and have a maxima width-at-half-height that is 75\% that of its minima. This profile which rises rapidly in explosive exponential growth but falls off slowly and roughly linearly gives rise to the term 'sawtooth'. Referring back to Sections 6.2.2 and 7.3 we noted that the \( k = 0 \) mode was special as it has no \( z \) derivatives, limiting both its dissipation and its ability to couple to other modes in the Jacobians.

In Table 8.2 we show the phase differences of the various modes, recording them separately for the rise and the fall.\(^{19}\) Comparing with Table

<table>
<thead>
<tr>
<th>mode</th>
<th>phase lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_k &amp; T_k )</td>
<td>0.0</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>0.0741</td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>0.1296</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>0.1296</td>
</tr>
<tr>
<td>( \psi_0 )</td>
<td>0.185</td>
</tr>
<tr>
<td>( \text{rising} )</td>
<td>0\°</td>
</tr>
<tr>
<td>( \psi_k &amp; T_k )</td>
<td>0.0</td>
</tr>
<tr>
<td>( \psi_1 &amp; T_1 )</td>
<td>0.0</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>0.1667</td>
</tr>
<tr>
<td>( \psi_0 )</td>
<td>0.2963</td>
</tr>
<tr>
<td>( \text{falling} )</td>
<td>0\°</td>
</tr>
</tbody>
</table>

Table 8.2: The Fourier mode phase differences, \( G = 1500 \).

8.1 for \( g = 1360 \), the \( k = 0 \) modes both have smaller phase lags which is consistent with the smaller percentage of the period taken up by the rise. On the other hand the first harmonics show a somewhat greater lag than before, as do the falling \( k = 0 \) modes. In general the stream function phases have changed considerably more than have those of the temperature.

\(^{19}\)In both cases they are measured at half-height.
In Figure 8.13 we plot the kinetic energy and Nusselt number for the full mode.\textsuperscript{20} As before the kinetic energy is inversely proportional to the stream function $k = 0$ mode owing to the dominant size of that mode. Its phase, however, lags slightly behind ($\sim 10^\circ$) having a slower rise but a steeper fall. The Nusselt number profile has changed considerably, however, apparently being composed of two parts; one a low wide roll that reaches to about half the height of the total oscillation and leads the fundamentals by $\sim 45^\circ$, and two, a tall narrow spike that coincides with the rise of the fundamentals.

Finally in Figure 8.14 we have field plots taken at the 'roundest' and 'flattest' points in the oscillation.\textsuperscript{21} There is a much greater contrast in these fields than we saw at $G = 1360$, again reflecting the greatly increased size of the oscillations. The 'cat's eye' mode, is stronger and shows a strong hook in the temperature field indicating an inverted gradient over nearly the whole wavelength. This will cause a local reversal in the sign of the buoyancy term that one would expect to have a stabilizing effect. The most pronounced

\textsuperscript{20}The timescale here is double that in Figure 8.12 but this is a numeric artifact.

\textsuperscript{21}Their location is marked in Figures 8.12 8.13 by the vertical lines.
Figure 8.14: The fields near the maxima and minima, at $G = 1500$. 'hook' coincides with the peak of the perturbations and it disappears rapidly as the $k > 0$ modes begin to fall. For its part the 'flat' mode is approaching a completely one dimensional form, occurring as the $k > 0$ modes bottom out, but the $k = 0$ mode is still relaxing (recall that its absolute magnitude is much greater at all times).

In contrast to the primary branch solutions in which the vortex axis rotated counter clockwise with increasing strength, it is just the opposite here. The observant reader will also notice that these modes are $180^\circ$ out of phase vertically from those shown in Figure 8.6. As is noted in Figure 8.9 we have here a second mode that is out of phase with the original perturbation. Whatever the overall phase of the initial perturbation this mode will be $\pi$ radians out of phase while the primary branch solutions always preserve the same phase. As we mentioned in Section 8.2 this second mode suggests that there are a pair of modes associated with the Hopf bifurcation at $G_H$. We also recall the discussion of Section 6.3.2 and restriction placed on the vertical phase by the nonlinear terms.
It is clear from Figures 8.7-8.9 that there is another change in behavior at about $G = 2000$. Although they become irregular, however, the oscillations maintain more or less the same form since, as Figure 8.8 shows, their amplitude grows slowly above $G = 1750$. Before we survey the variety of multi- and a-periodic oscillations for $G > 2000$ we briefly look at the oscillations at $G = 1750$. The period is now about 1050 dimensionless time units, though the real time period has again increased by about 10%. The fundamental mode maxima have narrowed a little more to 50% of the minima half-height, and are now slightly asymmetric in time, the rises being slower than the falls. They still have the same profile and are exactly in phase. The first harmonics are still time-symmetrical but are more sharply peaked, with maxima width-at-half-heights of 22% and 29% of the minima for the stream function and temperature respectively, and nearly zero amplitude between pulses. The harmonics though still peaking a little after the fundamentals, fall faster and at half-height lead

---

22 Although the oscillation amplitudes are roughly the same the median energy around which the oscillations occur continues to drop being down nearly 23% from that of the initial flow. See also Figure 8.9.
them slightly in phase.

The \( k = 0 \) modes now have a somewhat sharper rise profile than the fundamentals, closer to that of the harmonics and their relaxation continues to take up more of the oscillation period. The stream function falls at a slightly less than linear rate at first, then slightly faster, so that it bows out, while the inverse holds for the temperature. Consequently the maxima width-at-half-height is essentially equal to that of the minima for the stream function, while for the temperature it is 70\%. The phase lags do not change as much from \( G = 1500 \) to \( 1750 \) as they do from \( G = 1360 - 1500 \). The general trends are continued in that the rise takes up even less of the period and the fall more, so that the rise-lags are smaller, and the fall-lags larger, than at \( G = 1500 \). The harmonics now fall slightly ahead of the fundamental rather than coincidentally with them, and there is no clear general pattern of change depending on the field. The kinetic energy climbs more slowly than the \( k = 0 \) stream function mode falls and the phase lag is now 27\º. The Nusselt number profile is even more sharply defined in two components, a long slow rise starting as the fundamentals and

<table>
<thead>
<tr>
<th>mode</th>
<th>phase lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_k ) &amp; ( T_k )</td>
<td>0.0 ( \circ )</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>0.0746 ( 26.9\circ )</td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>0.1045 ( 37.6\circ )</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>0.1045 ( 37.6\circ )</td>
</tr>
<tr>
<td>( \psi_0 )</td>
<td>0.1493 ( 53.7\circ )</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>mode</th>
<th>phase lag</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \psi_k ) &amp; ( T_k )</td>
<td>0.0 ( \circ )</td>
</tr>
<tr>
<td>( T_1 )</td>
<td>-0.0298 ( -10.8\circ )</td>
</tr>
<tr>
<td>( \psi_1 )</td>
<td>-0.0373 ( -13.4\circ )</td>
</tr>
<tr>
<td>( T_0 )</td>
<td>0.209 ( 75.2\circ )</td>
</tr>
<tr>
<td>( \psi_0 )</td>
<td>0.351 ( 126.3\circ )</td>
</tr>
</tbody>
</table>

Table 8.3: The Fourier mode phase differences, \( G = 1750 \).
Figure 8.16: The kinetic energy and Nusselt number at $G = 1750$. The harmonics bottom out and apparently reaching a maximum as the $k = 0$ mode bottoms. At this point the next rise in the perturbation starts and we find a spike in the Nusselt number to twice the height of the first component.

The total perturbations are shown in Figure 8.17. We note two points.

Figure 8.17: The stream function and temperature perturbations at $G = 1750$. For future reference; first, the slight tail on the stream function as it falls is due to the slow fall of the $k = 0$ mode, and second, there is a slight asymmetry in the peak of the temperature perturbation, perhaps related to the form of the Nusselt number.
Finally in Figure 8.18 we have the 'cat's eye' and 'flat' modes at $G = 1750$. They are essentially the same as, though more extreme than, their counterparts at $G = 1500$. Note as well that they are once again in phase with the original perturbation.

### 8.3.2 A Theoretical Discussion.

Sorting through these descriptions we can propose a general scenario for the physics of these oscillations. At low perturbation amplitude there is a horizontal velocity shear and the Helmholtz instability generates vortices that reduce the both velocity shear and the thermal gradient in the center of the slot. The fundamental modes rise together and as they grow they couple in the Jacobian term, pumping energy to their harmonics. These start to rise as well, with the thermal harmonics following more closely than the kinetic harmonics, apparently due to the stronger coupling between $T$ and $\psi$. As was noted in Section 4.4.3 $\omega$ and $\psi$ are forced to be out of phase by Equation (2.17). As all the $k > 0$ modes grow they couple to themselves to feed energy to the correction to the mean fields, the $k = 0$ modes the $k = 0$ mode. At some point
the central gradients are reduced sufficiently, or even reversed. With the shear no longer unstable and the constant flow of energy into the $k = 0$ modes the $k > 0$ modes decay almost as rapidly as they grew. This decay shows a higher correlation than the initial growth.

Meanwhile the $k = 0$ mode deprived of additional energy from the self coupling of the rapidly decreasing harmonics, also starts to decay, but at a much slower rate. This decay takes place through the slow diffusion of the $k = 0$ mode itself and by the slow coupling into the other harmonics, which being stable loose it to diffusion and dissipation. In the absence of the buoyancy driven shear instability the system slowly relaxes back toward the 1-D basic state imposed by the boundary conditions. This leads to the re-establishment of the central slot gradients which at some point become unstable again.

We have, therefore, a competition between an instability that serves to relieve the stress of strong gradients in the fields, and the dissipative terms which, because of the imposed boundary conditions, seek to reestablish the basic solution and its gradients. At low $G$ these terms dominate and the basic solution is stable. Increasing $G$ represents a real increase of $\Delta T$ (given that the size of the slot and the type of fluid remain the same) and thus the stress. The magnitude of the basic shear flow is given by

$$ W_0 = \frac{G \nu}{L} $$

and therefore also increases. At some point the system becomes unstable and we find the primary branch appearing with is stationary vortices. These reduce the mean velocity and the central temperature gradient while increasing the heat flow across the slot (the Nusselt number). In practice there must be a locally definable Grashof number which the stationary flow serves to keep just
below critical. Apparently above the second bifurcation at $G_h$ this stationary solution breaks down and we see the oscillations that we have been describing in this chapter possibly mediated by the slow relaxation of the $k = 0$ mode.

The experimental evidence suggests that when $G$ is increased even farther that the instabilities will dominate and the gradients will be pushed completely out of the center of the slot and be confined entirely to boundary layers at near the walls.

In Figure 8.19 we plot the Nusselt number against the Grashof number. We can see here the conduction regime up to $G_c$, the near critical regime

![Nusselt Number vs. G](image)

**Figure 8.19: The Nusselt number and the various regimes**

where we found the Landau theory to be useful, and the oscillatory regime starting at about $G = 1000$. This is divided into a transitory part, where the oscillations damp to the primary branch, the critical point at $G_h$, the regular relaxation oscillation regime up to $G \sim 2000$, and then the irregular regime at $G/\text{geq}2250$. We have also marked points that were calculated by Mizushima [26] solving for stationary solutions numerically. These are in excellent agreement up to the point for $G = 1000$. The point at $G = 2000$ indicates the continuation of the now unstable primary branch.
The straight line above \( G = 1360 \) marks the center line for the Nusselt oscillations. It is also roughly the line indicating the separation of the two components of the Nusselt profile once they appear for \( G \geq 1500 \). An equation for this line is

\[
Nu - 1 \sim .53 \ e^{-5G}.
\]

We suspect that this may be the projection of a bifurcation in terms of a second control parameter.

Finally we indicate here, and in Figure 8.9 above, the phase of the cat's eye mode involved in these oscillations. In the case of the kinetic energy plot we see that the range in which this mode is \( \pi \) out of phase with the original perturbation coincides with a initial dip in the minimum energy just above the bifurcation. It is tempting here, and even more so in the irregular oscillatory regime to consider that we have three modes that appear in the oscillations. There is a \( k = 0 \) mode, an alteration of the mean fields, and two cat's eye modes \( (k \geq 1) \) that have phases relative to the phase set by the initial perturbation of 0 and \( \pi \). In the range from \( G_h \) to at least \( G = 2000 \) there are regular oscillations between the 'flat' mode and one of the 'cat's eye' modes superimposed on an altered mean flow. We shall see in Chapter 9 that beyond \( G = 2000 \) we find all three modes and that the oscillations become irregular or multi-periodic as a result.

8.4 Comparison with Other Thermal-relaxation modes.

It seems that the physics driving these oscillations is of a potentially very general nature. It is of interest therefore to take a quick look at some

\[\text{Footnote: For } \alpha = 1.383.\]
other physical problems were similar oscillations have been seen or where there seems to be a good potential for seeing them at some point in parameter space.

In the study of plasma physics of tokamaks there are well-known oscillations called sawtooth oscillations due to their shape. These oscillations also combine an explosive instability with a long relaxation phase. A recent paper by Aydemir et al. [88] gives a good general introduction to this problem and then discusses results of three-dimensional toroidal-geometry simulations. While the equations and physics involved are significantly more complex there are certain general similarities that are intriguing. In the tokamak the plasma in confined by a combination of magnetic and electric fields to the center of a torus. This is not however a stable situation.

A good comparison can be made between the pressure field in the tokamak and the kinetic energy in our problem. In the tokamak there is a tendency for a conical pressure field to build up in the torus, with its peak in the center dropping off rapidly towards the outside. This configuration is apparently unstable to a rapid movement of plasma towards the outside that leaves behind a flattened top on the cone. Just as in our case there is an instability that flattens central gradients in a rapid growth phase. In the VSC problem the kinetic energy is reduced by the growth of the shearing instability. In both cases there follows a long slow diffusive phase in which the pressure in the tokamak; and the kinetic energy in our case, build back up again. With this rise the central gradients are reestablished and the process happens all over again.

This parallel is not meant to be an exact one but to suggest that there

\[^{24}\text{We borrowed the name here for similar reasons.}\]
is a general form of oscillation that can occur when boundary conditions and
diffusive processes push towards the establishment of central gradients. When
these gradients become unstable to perturbations that effect a reduction in the
gradient we can get this sort of sawtooth cycle. There might be many cases
where the specific mechanisms are different and yet this general behavior can
be found.

Another system where one might look for something like this is dis-
cussed in Katanuma and Kamimura[89]. This is a slot bounded by charged
walls with a perpendicular magnetic field and a longitudinal field that is sheared
across the slot, leading to a central current sheet. This is unstable to tearing
instabilities that reduce the profile of the current sheet, flattening its peak, just
as the pressure profile peak is reduced in the sawtooth crashes.

In yet another configuration Ofman et al. [90] have been studying
tearing modes in resistive MHD with shear flow and viscosity. It seems probable
that this system can exhibit some or more forms of relaxation oscillation.

There are also problems in astrophysics, such as the bursting emis-
sions believed to come from the accretion disks of dwarf novae,\textsuperscript{25} coronal loop
dynamics and solar flares,[95] and the generation of magnetospheric substorms,[96]
to which these ideas may be applicable.

\textsuperscript{25}See for instance [91][92][93][94].
Chapter 9

Three Mode Oscillations

Up to $G = 2000$ we have seen that the perturbed VSC basic state results in regular oscillations, in which a quick rise in the perturbations caused by the instability of the central-slot gradients is followed by a slower relaxation. While the instability leads to vortex formation in the center of the slot, the relaxation phase is dominated by the $k = 0$ mode. Furthermore we have seen that the vortex can form in one of two vertical phases relative to the original perturbation, 0 and $\pi$ radians. As $G$ increases beyond 2000 we find that these oscillations become more complex and in general involve all three of these modes. In this chapter will present a sample of the possible results including one case that appears to be a chaotic attractor.

9.1 A Third Bifurcation Near $G = 2000$?

We saw in Section 8.2 that the real frequency of the oscillations appears to be leveling off approaching $G = 2000$ after dropping throughout the regular oscillation regime above $G_h$. Just as there was a qualitative change at the start of this drop, so there appears to be one at its end. There is also a parallel between the change at $G_h$ in the perturbation amplitudes and other order parameters that we used in the stationary regime and the change in the oscillation amplitudes above $G = 2000$.\(^1\) There are now oscillations of the os-

\(^1\)Compare Figures 8.9 and 8.8.
cillation amplitude, which we shall see can be either regular or not. This is also suggestive of a third bifurcation.

So far we have focused on the behavior of the VSC system once it has had time to settle into a steady or quasi-steady state. We saw that in the near critical regime the transitory behavior was not very different from the quasi-static evolution up the primary branch. We noted in passing that at higher $G$ the initial rise of the perturbation significantly overshoots the asymptotic value, or range of values and that the dynamic path from the basic state passes through far from equilibrium states that are different from those along the primary branch. As we saw in Section 8.1, in discussing the $G = 1360$ run, there is a large initial spike before the system reaches its new steady or oscillatory state. In Section 10.1 we will look in detail at the transitory behavior during this spike for high $G$, but briefly the initial perturbation grows until it saturates at the top of the spike and decays back towards its new pattern. This decay however does not pass through the same states that were transversed on the rise, but through new states that are quite distinct from the linear eigenvalues of the basic state.

What is of interest here is simply the magnitude of the initial rise in the Nusselt number and the initial drop in the kinetic energy. In Figure 9.1 we plot the minimum of the total flow.\footnote{The horizontal line at the top of the figure shows the initial total kinetic energy, and that at the bottom zero total kinetic energy. We note also that the point at $G = 2000$ is somewhat suspect. It was the last point obtained from a 16x32 mode run and should be checked with a 32x64 mode run, as where used at $G \geq 2250$. Poor resolution can be expected to enhance the natural instability of the system to this small extent.} This minimum corresponds to the initial peak in the stream function perturbation. While the minimum energy is been decreasing with increasing $G$ for all values above $G_c$ there appears to
Figure 9.1: The minimum total kinetic energy.

be a significant change in slope above $G_h$. This measure, like the oscillation frequency is changing rapidly in the regular oscillatory regime. In this case, however, there is a limit to this increase, as the kinetic energy is a positive definite quantity. Thus near $G = 2000$ we see that this limit has apparently been felt and having lost nearly 90% of its initial kinetic energy there is no room for the perturbations to grow further. Although the perturbation decays and the kinetic energy increases again to oscillate about a value of $\sim .7 e^{-2}$ at $G = 2000$ it seems reasonable that this limitation of the rise of the transient is connected with the other changes at this Grashof number.

Finally in Figure 9.2 we show the value of the initial Nusselt peak. The initial perturbation generates strong vortices which greatly increase the heat transport across the slot as measured by the increase in the Nusselt number. This dramatic change in the heat transportation during the initial transient phase is the sort of change that one expects in an order parameter at a bifurcation.

\footnote{Just as the kinetic energy at $G = 2000$ may be too low, the Nusselt number value may be a little too high.}
Figure 9.2: The height of the initial Nusselt number peak.
This completes a survey of changes near \( G = 2000 \) that suggest a possible third bifurcation.

9.2 Detailed Analysis at \( G = 2250 \)

We now turn to an examination of how the oscillations change for \( G > 2000 \). We have done runs at \( G = 2250 \) and \( G = 2500 \) obtaining results that are qualitatively the same. Here we discuss the \( G = 2250 \) results in detail.

In Figure 9.3 we show the kinetic energy, the Nusselt number, and the two perturbation amplitudes at \( G = 2250 \). It is quite apparent that these oscillations, while of a similar form to those encountered in Chapter 8 are not at all regular. They vary in amplitude and period, and in particular the Nusselt number oscillations vary in form. We had commented above that in the regular oscillatory regime it appeared that the Nusselt number form was composed of two parts. It appears here that these parts are no longer in phase with each other, but that their relationship varies from period to period. It is not very noticeable here but there are also small variations in the perturbation curves. In particular the transition from the sharp fall to the \( k = 0 \) tail at \( t = 6.9 \) (the
Figure 9.3: The kinetic energy, Nusselt number, and perturbation amplitudes at $G = 2250$, $\alpha = 1.383$

base of the large spike in the middle of the stream function frame) has a slight hitch in it. This feature is more evident at higher $G$ than it is here and again seems to indicate that the various processes involved in the oscillations are no longer precisely synchronized. In particular the triggering of the instability may vary in its timing. It is clearly not a simple function of perturbation amplitude anymore.

In Figure 9.4 we show the same diagnostics for a single oscillation.\(^4\)

\(^4\)The location of this oscillation is shown by the vertical lines in Figure 9.3.
The flattened slope at the end the fall of the perturbation amplitudes is ap-
parent as is the increasingly slow rise and sharp fall of the kinetic energy. We
note as well that while the Nusselt peaks are still clearly associated with the
instability growth it appears here that the second component correlates with
the dominance of the $k = 0$ mode. It seems likely that there are two separate
transport modes/mechanisms involved.

In Figure 9.5 we show an un-normalized mode decomposition for the
Figure 9.5: The modes for a single oscillation at $G = 2250$ same time period.\textsuperscript{5} The rise of the $k = 0$ mode now takes up barely 25\% of the oscillation cycle (compared to 30\% at $G = 1750$ and 36\% at $G = 1500$), and the maxima width-at-half-heights of the fundamental modes are an equivalent 33\% of the minima width-at-half-heights. The rise of the fundamentals starts almost shortly after they fall, are near the mid-point of their fall. The timing of the onset of the instability is likely a function of the relative size of the $k = 0$ and $k > 0$ modes. The rise of the fundamental continues to be slower as $G$ increases.

Perhaps the greatest change, however, is that the phase of the cat's eye mode is no longer constant. In fact it seems to be random. Each time the perturbation grows it has one of the two phases we have seen before, but which phase appears is unpredictable. In Figure 9.6 we illustrate the fields during a typical cycle.\textsuperscript{6} The first frame shows the peak of the cat's eye mode (0 phase). There is a complicated variation in the near zero thermal field in

\textsuperscript{5}The short vertical lines indicate roughly the transitions from one mode to another, see below.

\textsuperscript{6}The location of these frames are marked by dots in Figures 9.3 and 9.5.
Figure 9.6: Selected frames during an oscillation at $G = 2250$
the center-slot, but this pattern collapses rapidly on itself as the perturbation
begins to decay. The third frame illustrates the complete dominance of the
$k = 0$ mode, with a strongly inverted thermal gradient in the centerslot. In
the fourth frame the instability has been growing for sometime and is close to
peaking again. This time however it is the $\pi$-phase mode.

We have also run a $G = 2500$ case with similar results and continuing
evolution of the changes seen at $G = 2250$. 
9.3 Phase-space Portraits

A good way to characterize the successive solutions that we have found as we have increased the Grashof number is to look at the phase-space portraits of the perturbations. Phase-space is a tool that is widely used for conservative systems with where it can quickly and powerfully summerize much of the dynamics of a system of equations[97]. In dissipative systems there much of the analysis no longer holds but by picking the proper coordinates it is often possible to qualitatively characterize the dynamics of complex or apparently random systems with a low dimensional ‘portrait’ in which the location of the system is plotted at successive times indicating its path through the phase-space.

9.3.1 Stream Function-Temperature Phase Plots

Since we have a system of two variables the natural choice is a two dimensional space whose coordinates are given by the amplitudes of the two variables, the stream function and the temperature. In Figure 9.7 we have plotted a selection of phase portraits indicative of the evolution of the system with increasing $G$.

At $G = 1000$ where the asymptotic state is stationary we can see the system spiraling into a point attractor.⁷ A point that the system will reach from any initial point in a neighborhood around the fixed point. This evolution towards fixed points (or cycles etc.) are the mark of the dissipation in the system. If this point is the single stable solution of the VSC equations

---

⁷The points plotted here are for equal time intervals, so that their spacing gives an indication of the speed of the evolution along different parts of the curve.
Figure 9.7: Stream function-temperature phase portraits for $G = 1000, 1360, 1400, 1750, 2000,$ and 2250
in this regime then it will eventually be reached from any other point in the phase-space. In the portrait for $G = 1360$ the system is still spiraling inwards towards a point, but at such a slow rate that our run has ended before it can reach the attractor.

In the first example above $G_A$ the point has now become an oval, what is called a limit cycle. This is the mark of a Hopf bifurcation. Any initial state in a neighborhood around the oval will evolve towards it in finite time, and having reached it will follow it around forever. The width of the loop reflects the phase difference between the variables. When they are exactly in (or out) of phase they will go up and down a $45^\circ$ line together. A phase difference of $90^\circ$ results in a circle (assuming of course that the variable ranges have been scaled to be equal). Changes in the phase with time lead to bumpy loops.

In the frames for $G = 1750$ and 2000 we see an evolution in the size and the shape of the limit cycle, with the formation of a 'handle' at the outer end that slowly gets wider with $G$. Note that the low amplitude return (bottom) leg of the loop is densely dotted indicating slower evolution there. Finally in the last frame for $G = 2250$ we see that there is no longer a single line that the system follows but rather we have the 2-d projection of a band of possible paths that is at least two dimensional. If the third bifurcation were a simple period doubling bifurcation\(^8\) then we would expect to see two clear paths taken in turn, as is often found in simple mappings.[97] It is well known that a three-body, or three mode, system can be completely unpredictable[97] and here there seem to be a large number of paths, representing many frequencies and possible chaos[98].

\(^8\)See Section 9.4.1 for a period doubled solution.
9.3.2 Stream Function Mode-Mode Phase Plots

One of the objects in producing a phase portrait is to choose variables that reduce the system as simply as possible. In our case we seem to have three modes which include both the temperature and the stream function. Noticing

Figure 9.8: Stream function $k = 0$ vs. $K = \alpha$ phase-space portraits at $G = 1400, 1750, 2000$, and $2250$.

that our initial plots are nearly linear our choice of the two physical variables is not yielding a clear projection of the general system phase space. It would perhaps be more appropriate to make use of the 'modes' that we see in our
system.\textsuperscript{9} As a first approximation we try plotting the $k = 0$ and $k = \alpha$ stream function Fourier modes. The results are shown in Figure 9.8 and justify our choice as they are much clearer than before.

We can see the decay the a nice limit cycle at $G = 1400$ and 1750, although in the latter case the circle is somewhat distorted. The rotation in these frames is counter-clockwise. At $G = 2000$ it appears that the cycle is becoming less well defined (though some of this is due to the evolution into the full cycle). We note that as was the case above the path is best defined on the return swing where the variables are smaller in amplitude. At $G = 2250$ we see again the spreading of the limit cycle into a third dimension and note that, although it is hard to make out here (watching the data plotted in real time helps) the ‘turns’ at low $k = \alpha$ are of two different types depending on whether the phase is changing or remaining the same. These different ‘turn’s are related to the cycle variation noted in Section 9.2. In this run there does not seem to be any clear pattern in the phase-space depending on the phase of the cat’s eye mode for a particular cycle.

Finally we present the full phase portrait at $G = 2250$ including the whole initial spike, which clearly shows the large increase in both variables, the faster decline of the $k = \alpha$ mode and the slower oscillating approach of the $k = 0$ mode back to the attractor. This is shown in Figure 9.9.

With this we complete our description of the behavior of a vertical slot filled with water, perturbed at a wavenumber of 1.383, from $G = 0 - 2500$. There are however several interesting oscillatory states that we have found at a slightly larger $\alpha$. These are presented in the next section.

\textsuperscript{9} Especially in light of the Lorenz analogy of Section 9.4.2.
Figure 9.9: The $G = 2250$ phase portrait, including the initial transient.

9.4 Oscillatory Regimes when $\alpha = 1.6$

In brief we have found that at low $G$ increasing $\alpha$ reduces the magnitude of the initial shear flow for a given $G$, increases the stability of the system and speeds up its evolution. This is mathematically due to the increased value of all the vertical derivatives and in particular those in the dissipative terms. Physically the reduction in the initial shear flow also reduces the driving gradients, and the amount of kinetic energy available to form the perturbations. A glance at the neutral stability curve (Figure 5.3 indicates that $G_c$ larger, but not by very much. The primary branch and its stationary secondary states extends to at least $G = 1500$, so that $G_A(\alpha = 1.6) > 1500$. Above this we have done two test runs. One at $G = 1750$ with chaotic results that we will describe below and one at $G = 2000$. In general it appears that with the onset of chaos
after only a few bifurcations it would appear that we have an example of a
Ruelle and Takes [99] transition rather than a Landau type series of period
doubling bifurcations.[55]

9.4.1 The $G = 2000, \alpha = 1.6$ Case

Here we briefly present an example were the three modes are balanced
in a manner that yields a regular two period oscillation in which they each seem
to have correlated energies and the system oscillates between the flat state and,
alternately, the two cat's eye modes. The period doubled nature of this state

![Phase Portraits](image)

Figure 9.10: The $G = 2000, \alpha = 1.6$, period doubled, phase portraits.

is clearly seen in the phase portraits presented in Figure 9.10. The stream
function-temperature phase plot shows that the two fields are very nearly in-
phase, that there are two distinct cycles and that the 'turns' are clearly different
as they switch back and forth. The $k = 0$ vs. $k = \alpha$ plot presents two nice
cycles that are not quite, but nearly, exact.

In Figure 9.11 the main diagnostics are shown. The total kinetic
energy is lowest for the 'flat' mode, greatest for the $\pi$-phase mode, and halfway
inbetween for the 0-phase mode.\textsuperscript{10} We note as well that there are ‘hitches’ at the bottom of the stream function decay on every other cycle, in correspondence to the step in the $\psi - T$ phase-space curves at the same point in the cycle. Lastly we note that in this case the two components of the Nusselt number are in phase with each other again.

Finally we present for comparison the mode decomposition in Figure 9.12. Notice that the $k = 0$ modes are prominently bowed out for the stream function and are concave for the temperature. Otherwise they are essentially the same as the $G = 2250$ case for $\alpha = 1.383$.

9.4.2 Lorenz’s Attractor

In a now classic 1963 paper\cite{65} Lorenz reduced the equations for the Benard problem to three spectral equations for the three dominant components of $\psi$ and $T$. The dependence of the fields was taken as\textsuperscript{11}

$$\psi(x, z, t) = \psi_{mn}(t)\sin(mx)\sin(nx)$$ \hspace{2cm} (9.1)

and

$$T(x, z, t) = T_{mn}(t)\cos(mx)\sin(nx) - T_{0,ne}(t)\sin(2nx)$$ \hspace{2cm} (9.2)

It turns out that the behavior of the three modes is quite complex and for the right parameters will generate the now famous Lorenz attractor, which has a two lobed, butterfly-like phase-space portrait (see [65] or [97]).\textsuperscript{12}

Our contention is that the three ‘modes’ that we see in the oscillatory regime are in some sense nonlinear equivalents to Lorenz’s three Fourier modes

\textsuperscript{10}This correlates with a change of sign on the Jacobian term due to the change in phase.

\textsuperscript{11}After [100], but in our notation.

\textsuperscript{12}Note that Lorenz worked with $Pr = 10$, not far from our value of 7.5.
Figure 9.11: The $G = 2000$, $\alpha = 1.6$, diagnostics.

Figure 9.12: The $G = 2000$, $\alpha = 1.6$, mode decomposition.
components. This is based on the analogy that one in both cases there is a $k = 0$ mode (his $T_{0,\text{na}}$ component) and two $k > 0$ modes with the same form but distinct phases, such that they are linearly independent.\(^{13}\) The basis of our supposition however comes from the analysis of the oscillations at $G = 1750$ and $\alpha = 1.6$ that we present below. We should add that this is a preliminary comparison and the basis on which one can identify attractors is not an simple matter.

### 9.4.3 The Attractor at $G = 1750$, $\alpha = 1.6$

The general form of the oscillations is essentially the same as for the other parameter sets we have examined. We content ourselves with presenting a few figures that for comparison with Lorenz's attractor and leave any more detailed investigation to the future.

![Phase portraits](image)

**Figure 9.13**: The $G = 1750$, $\alpha = 1.6$, phase portraits.

In Figure 9.13 we have the phase-space portraits. They are similar to

\(^{13}\text{We note that his modes are } \frac{\pi}{3} \text{ out of phase vertically, but that the cat's eye perturbations have in effect twice the wavelength of the original perturbations and thus (relative to the original wavelength) are out of phase by a factor of } \pi.\)**
those for $G = 2000$, $\alpha = 1.6$ in that they show two types of loop. They are not however as coherent, but are now spread out into three dimensional surfaces. The largest gap in the $k = 0-k = \alpha$ bands marks the division between the two phases, the loops above are associated with the 0-phase mode while the lower larger loops describe the $\pi$-phase mode. Another diagnostic that seems to distinguish between these two modes is the potential energy (see Section 6.4.2 for the definition). In Figure 9.14 it varies to the positive side when the 0-phase mode appears and to the negative when the $\pi$-phase mode is visible. We note that there is some resemblance between this figure and Lorenz's plot.

![Potential Energy Graph](image)

Figure 9.14: The potential energy at $G = 1750$, $\alpha = 1.6$.

of his $Y$ component, the $\cos(mx)$ temperature component.

According to Tabor[97] the plot of the $n^{th}$ peak vs. the $n + 1^{st}$ peak 'contains the essence of the Lorenz attractor'. We present here in Figure 9.15 our version of this mapping, using the peaks of the temperature perturbation.
Although the number of points is not great it is large enough to clearly define the curve.\textsuperscript{14} We note that a similar plot was obtained using a smaller number of modes. It had the same form but was shifted slightly towards the upper-right. In Figure 9.16 we show similar plots for the other runs, for comparison. That at $G = 2000$ for $\alpha = 1.383$ represents regular oscillation expanding to an asymptotic size in the upper right. The $G = 2000$, $\alpha = 1.6$ case oscillates between two small areas at the lower right and the upper left. It can be noted

\textsuperscript{14}It would be nice to have more points, but as the code is being run with 32x64 modes it takes quite a bit of time to model each oscillation. This data represents over 350,000ts.
Figure 9.16: The the other peak to peak plots. that there are sets of points on the Lorenz map that will closely reproduce this oscillation, but it is unstable and if they move very much they will then wander far away. The other two frames both for $\alpha = 1.383$, the un-periodic cases at $G = 2250$, and 2500 show no particular pattern, though again it would be nice to have more points.
Chapter 10

Future Directions

In this final chapter we discuss ways in which our work can be extended in the future. We start with a brief outline of the transient behavior which is found in the large initial spikes, and suggest that we are effectively seeing higher $G$ behavior. Moving to higher $G$ being one of the main avenues to pursue. We also discuss briefly promising work that can be done by changing the other parameters of the VSC problem. Finally a review of possible applications in other branches of physics, including MHD and plasma simulation.

10.1 The Spikes

In Section 8.2 we referred to the large initial peaks that result from starting our simulations on the basic solution, far from the primary and secondary branches where they will wind up. In looking at plot of an order parameter as a function of the control parameter, $\epsilon(\mu)$ we usually think of increasing the control parameter in order to push the flow solution to higher values of $\text{epsilon}$. In these initial spikes, however, it is the order parameter that is being pushed very rapidly to a large value. If we invert the relationship, and rather than thinking of states as being a function of the control parameter we think of them as being a function of the order parameter, in a dynamic evolution, rather than a quasi-static one, then it is reasonable to associate the states seen at large $\epsilon$ with those steady or oscillatory stable solutions expected at large $\mu(\epsilon)$. We saw that this was indeed the case near the critical point $G_c$ but that
at higher $G$ the states seen in the fall are different from those seen during the rise. It turns out that the states during the fall from these large perturbation spikes resemble high $G$ states that have been observed experimentally. Therefore it seems plausible that these states 'on the other side of the hill' can be associated through an inversion of the function $\varepsilon(\mu)$, for stable states, with actual high $G$ realizable behavior.

We will look at the initial spike from a $G = 5000$ run which is pushing the limits of the $32 \times 64$ mode model. In Figure 10.1 we show the usual

![Graphs showing kinetic energy, Nusselt number, stream function perturbation, and temperature perturbation.](image)

Figure 10.1: The kinetic energy, Nusselt number, and perturbation amplitudes at $G = 5000$, $\alpha = 1.383$. 
diagnostics. We note that here, with the exception of the kinetic energy, which goes nearly to zero very rapidly and then slowly rises again, there are a whole series of peaks following the initial rise. An envelope drawn through these peaks shows a smoother relaxation, but in general, unlike the smaller spikes at smaller $G$ there is very complex behavior and continued small rises and falls during the first half of the fall. In this regard it is reminiscent of behavior Aydemir[88] with giant and complex sawteeth. It seems possible that if the oscillations seen above $G_h$ after the transients have passed are like regular sawteeth that this initial spike is like a giant sawtooth.\footnote{It is also worthwhile comparing these results to those of Ginet and Sudan [101] for 2-D compressible Benard convection.} The sample fields plots are marked by the dots in Figure 10.1. The first frame (upper-left) shows the full development of perturbation into a vortex with a large hook in the temperature field indicating extensive center-slot mixing. Notice that although they have a very large magnitude the perturbations are quite recognizable as slight modifications of the linear eigenmode. Very soon though these forms change and as they near the peak we can see they are now completely different (lower-left). There is a very strong tertiary vortex now, with opposite rotation formed by shearing between the secondary vortices, not due to any instability of the secondary. These tertiary vortices have been seen experimentally [34][43] and numerically[47], though not at quite this strength. The temperature field indicates that the mixing involves the whole width of the slot and certainly contributes to the large Nusselt number.\footnote{Compare [52] in aspect ratio 10 slots.} It is impossible here to say much about the resulting gradients, except that we can see that some of the are quite large as the hot and cold fluids are convected past each other. This phase which coincides with the large secondary spikes is illustrated in Figure 10.3 which
Figure 10.2: Sample field plots.

reads down the columns and from left to right. This mixing reminds us of the pictures produced in Libchaber and Kadanoff's work with very high Raleigh number Benard flows which have produced plumes like these and waves, such as we see below (See for instance [102]).

The third frame (upper-right) of Figure 10.2 comes as the kinetic energy bottoms out.\(^3\) This is nearly at the end of the larger after shocks at

\(^3\)Same as the middle frame right-side column in Figure 10.3.
Figure 10.3: Large scale mixing across the slot: the full temperature field.
the transition from the strong mixing phase to a quieter phase that is characterized by thermal boundary layers, a more nearly one dimensional stream function (as the $k > 0$) modes rapidly deflate, and the onset of traveling waves. The tertiary vortex has shrunk again to a size in better agreement with experiment, but both of them have rotated and reversed their pitch. This phase is shown in Figures 10.4 and 10.5 showing the temperature field and the corresponding stream function field. The small waves moving up and down the

![Figure 10.4: Waves traveling along the walls: the full temperature field.](image)

walls of the slot are probably similar to those in [102] but it is also known that at really high $G$ the VSC system changes over to boundary layers that generate traveling waves. These waves then break into the centerslot ($G$ still increasing) and contribute to turbulence there.[42] Similar waves have been seen in a numerical model for an $h = 25$ slot.[49] Meanwhile the stream function is mixing
Figure 10.5: Waves traveling along the walls: the full stream function field.

different phase modes, to in effect shift vortices up and down the slot.\textsuperscript{4}

The final frame of Figure 10.2, is well along the tail and shows the third phase, after the waves have died as well (the smaller bumps). Along the smooth tail the flow is dominated by the $k = 0$ until by this final frame the fields are almost completely one dimensional. Just off these plots ($t=2.85$) the shear instability kicks in and the whole cycle runs again. We have not yet been able to extend the run to see if it ever settles into smaller oscillations, or some other form of stable configuration. It seems possible that this is a regime of giant oscillations. In contrast to the ‘sawtooth’ oscillations the rise now accounts for only $\sim 10\%$ of the period.

One of the direction we would like extend our work in is these high $G$ behaviors. Running larger expansions, and with a contemplated rewrite to

\textsuperscript{4}Compare with similar vortex gymnastics in [101].
speed up the matrix solver we hope in the future to push on to higher $G$.

10.2 Variations in $\alpha$

In this work we have, in general, focused on the critical wavenumber in our nonlinear investigations. As we saw in the case of the attractor at $\alpha = 1.6$ though changing the wavenumber can change the solutions significantly. We have done some initial work in this regard, in particular looking at smaller $\alpha$ where the fundamental can become unstable to its harmonics.[25]. We have seen both the first harmonic and the second appearing at lower $\alpha$. Beyond the mapping of these boundaries the possibility exists of picking $\alpha$ such that two or more modes will compete. It is also possible by picking small $\alpha$ to change the effective aspect ratio of the model and look at many cells together, rather than just a single cell.

In combination with changes in the Prandtl number it is apparent that the envelope for the secondary state as a function of $\alpha$ and $G$ changes. At zero $Pr$ Mizushima has found that the envelope leans towards large $\alpha$, [25] but that at $Pr = 7.5$ it tilts towards the low $\alpha$ side. [26].

10.3 Breaking the Horizontal Symmetry

Another interesting question that has been possed by several authors involves breaking the mid-slot odd-symmetry of the initial state. This has been done by adding differential curvature[53] and by considering radiation at the walls[54]. When the axial-symmetry is broken then there are single solutions with non-zero phase speed. In one run done at $Pr$ of air (.75) near the boundary between the fundamental and the first harmonic we have seen a similar solution in an apparent case of spontaneous broken symmetry. These issues bear further
investigation.

10.4 Other Possibilities

There is another opportunity to study mode interaction at higher Prandtl number when the traveling modes also become unstable, and in general the whole issue of high $Pr$ behavior has not yet been addressed. More importantly perhaps, as suggested by our discussion at the end of Chapter 8, is the extension of the basic modeling methodology to tackle charged fluids and to investigate the new modes of behavior that are possible with the introduction of the magnetic field. This opens up numerous applications in earth and space physics that we hope to be able to explore in the future. In light of the parallels with the theories of critical phenomena and phase transitions it is worth investigating the possible use of our simulation in conjunction with more general field theories.
Chapter 11

Experimental Summary and Comparison

In this chapter we will pull together the most relevant experimental results for comparison with our work. There are unfortunately a limited number, many of which are for different types of fluids. We have left out entirely results of experiments with aspect ratios less than 10-20 for which the asymptotic solution loses its relevance. We conclude with a brief outline of the type of experimental work needed for direct comparison with the simulations we have presented in this dissertation.

11.1 A Summary of Experimental Results

Eckert and Carlson [3] working in a slot with a variable aspect ratio \( h = 10 - 50 \) performed temperature field visualizations on air \( (Pr = .71) \) using an interferometer, and assuming locally uniform pressure to convert the density data to a temperature measure. They varied the Grashof number from 50 to 12500, and delineate three regimes; a laminar regime at low \( G \) where the temperature profile is given by \( T = x \), as we expect for the basic flow, a transition regime in which the horizontal temperature gradients start to flatten beginning in the center of the slot, and a high \( G \) boundary layer regime in which the central gradients are completely flat. They estimate these transitions for a range of aspect ratios up to 100 at which they are apparently approaching asymptotic values. These are roughly \( G \sim 500 \) for the change to the transition regime, and \( G \sim 10000 \) for the shift to the boundary layer regime which shows
signs of turbulence. The lower transition is close to the analytic value of \( G_c = 495.6 \) for the onset of the shear-driven vortices which flatten the temperature gradient. The rest of our results are all in the transition regime.

In 1965 Elder performed two studies\[34][42]. The first one was done with heavy oils (\( Pr = 1000 \)) at aspect ratios from 1 to 60 and for Grashof numbers up to 6250.\(^1\) There is no mention of the low \( G \) traveling waves that the linear analysis indicates should appear first, but his estimate of the critical Grashof number \( G_c = 18.75 \pm 30\% \) is close to our rough estimate of \( G_c \sim 15.7 \), also at \( Pr = 1000 \),\(^2\) and that given by Vest and Arpaci [44] (see below). He does see the vortices and then tertiary vortices which form with reverse rotation, driven by the vertical shearing between the secondary vortices. He measures velocity profiles that clearly show the pulling back of the primary flow towards the walls as \( G \) is increased while the central gradients are reduced and even reversed slightly (\( G = 225 \)). The temperature profile is already reversed for \( G \sim 25 \) so that the heat transport is almost all by convection. Elder also measures a vertical temperature gradient, \( \beta = \frac{dT}{dz} \), which for \( G > 7 \) is given by \( \frac{1}{2h} \) where \( h \) is the aspect ratio.\(^3\) He points out that the vortices flatten this vertical gradient as well, producing a stair step effect with the average gradient remaining essentially the same.

In Elder's second study he switched to work with water (\( Pr = 7 \)) using aspect ratios of 10-30 and focusing on large \( G \) and the onset of turbulence in the central slot. These results give the most detailed picture of the breakdown

---

\(^1\)His results are given in terms of the Rayleigh number, which we have converted to a Grashof number by dividing by the Prandtl number.

\(^2\)See also Section 5.4.1.

\(^3\)This is a dimensionless relation. In dimensional form the units are \( \frac{K^2}{cm} \).

\(^4\)For most of these results \( h = 20 \).
of the vortex flows and the establishment of boundary layer flow at a Grashof number of roughly 9,000.\footnote{Elder uses orders of magnitude, saying that the vortices are steady to $G \sim 9000$, and then that they are unsteady around $G = 90,000$. The first figure agrees better with the results of other authors and would seem more likely to agree with our code as well.} At this value of $G$ the vortices become unsteady and often appear in an isolated quasi-steady form. While it might be tempting to associate these fluctuating vortices with our irregular oscillations, a proper comparison depends on obtaining more precise experimental values and on pushing our code to higher $G$. As the temperature difference is increased further, for $G$ equivalent to 13,000 traveling waves appear along the walls and generally (but not always) disrupt the vortices in the central slot. These waves travel up the hot wall and down the cold wall. As they grow with further increases in $G$, they start to curl back and break into the center of the slot as the fluid cools away from the wall and falls behind.\footnote{See for instance Figure 10.4 which shows a qualitatively similar phenomenon.} As the Grashof number is further increased and these waves grow even stronger, they perturb the central slot which has a uniform mean temperature field and the vertical gradient is zero over much of the slot. This perturbation may be the cause of the turbulent fluctuations of the velocity field.

To sum up the high $G$ state, there are very narrow thermal boundary layers near the walls, then wider velocity boundary layers and mixing layers which stir up the zero mean flow in the center of the slot. While the Grashof numbers for this behavior are much higher than anything that we have yet run, it is this type of behavior that we may begin to see dynamically in the $G = 5000$ run of Section 10.1. Again, although we see no inconsistency of behavior between these experiments and our work, no direct comparison is possible.
MacGregor and Emery [41] sum up the general picture nicely, defining the succession of regimes, conduction ($G = 0$ to 63), asymptotic flow (conduction plus convection, $G$ to 1875), laminar boundary layer flow ($G$ to 63000), a transition region, and finally turbulent boundary layer flow (above $G \sim 500,000$). These latter three regimes have little conductive heat transport due to the inversions in the temperature profiles. In terms of our results we suggest that we have seen the first two of these regimes, and that there should probably be another transition region inserted between the asymptotic and the laminar boundary layer regimes, which for the values given here, would be in the same general range as our oscillations.⁷ In the experimental portion of this paper the authors use an apparatus at aspect ratios of 10, 20, and 40 with a wide range of substances (including water) having Prandtl numbers from 1 to 20000. They give horizontal temperature profiles for these, but are mostly interested in the high $G$ dependence of the Nusselt number on the Rayleigh number.⁸

Vest and Arpaci [44] give results for air ($Pr = .71$) at an aspect ratio of 33.4. Their value for $G_c$ of 543.8 ± 10% is consistent with the linear theory value of 495.6. Their value for the critical wave number $\alpha_c = 1.37$ is a little bit lower than our value of 1.405. As we mentioned above they also ran a high $Pr$ boundary layer regime experiment with $k = 20$, yielding $G_c \sim 25$ and $\alpha_c = 1.75$, which is larger than expected from theory (they derive a theoretical value of about 0.925, while ours is $\sim 1.1$ in Section 5.4.1). As an illustration they have an excellent visualization of a cat's eye cell which compares favorably

---

⁷The use of the first appearance of temperature inversion to mark the onset of the laminar boundary flow regime places this boundary much lower than the formation of narrow layers at the walls that we see only at high $G$. It is also probable that as in Elder's case these values are meant to be general estimates rather than concrete results.

⁸The boundary conditions in this experiment are constant heat flux rather than constant temperature.
with those we see on the upper primary branch.\(^9\)

In 1971 Hart[50] carried out an extensive set of studies in a tilted box filled with water \((Pr = 7.5)\) and aspect ratios of 25 and 37. He found examples of both the stationary and the traveling wave modes and mapped their dependence on the angle of tilt. For the vertical case he found that the vortices appeared at a considerably higher Grashof number than predicted by theory. It is also of note that he seems to find traveling waves rather than stationary waves as being the first unstable mode. Hollands and Konicek, also studying general angle slots, find a critical Grashof number in the vertical case at \(687 \pm 32\), higher than expected (except for one paper by Unny [103]). It should be noted that Hart says these flows are hard to visualize and that in our work we can see that the critical point is lower than a visual inspection of the total flow would indicate. If there were an experimental technique that would allow subtraction of the mean flow from velocity data then the critical point could be determined with greater accuracy.

A third paper on tilted slots came in 1982, by ElSherbiny, Raithby, and Hollands [7]. In it they focus on the affect of varying the aspect ratio from 5 to 110. In their results for vertical slots filled with air it can be seen that \(G_c\) increases with \(h\) up to about \(h = 20\) above which it remains nearly constant. They present mostly Nusselt number curves with a focus on the large \(G\) dependence and the change from the fourthroot to the cube root with the onset of turbulence. It is consistent with numerical results as well that one needs an aspect ratio of 20 or better to approach the large \(h\) limit. In fact we saw with Eckert and Carlson that \(h = 100\) is better. A nice compromise might

\(^9\)Again the difference in parameters, theirs is in heavy oil rather than water, prevent a more direct comparison. Refer to Figure 7.7 for \(Pr = 7.5\), and \(G = 1000\).
be $h = 40 - 60$.

The last paper using water is the one by Seki et al. [43] in which they also treat oil ($Pr = 480$) and glycerin ($Pr = 12500$). Again the pictures for the high Prandtl numbers show secondary and tertiary vortices with great clarity. They give a value for $G_c$ of 23.2 at $Pr = 480$, reinforcing the values for high $Pr$ $G_c$ we have already mentioned. This concludes our review of applicable experimental results. As mentioned in the introduction there are papers with $h \sim 1$ but with such a small aspect ratio the physics is quite different. We have also omitted a few three dimensional studies as so far our work is entirely 2-D as is known to be valid in the near critical region.\(^{10}\)

In general we can summarize by saying that our results seem to be consistent with the available experimental results, but that these are so few that very little detailed or direct comparison can be made. We now turn to a brief discussion of possible future work that might enable a closer check of our numerical results with experiment.

11.2 Suggestions for Experiment and Simulation

In most cases reviewed the visualization techniques of the experiments limit the results to two dimensions even though the experiments are in three dimensions. Similarly most of the numerical work including our own have been in two dimensions. There have been a few studies more recently examining the stability of the two dimensional primary branch flows [24] [13] finding that it is limited in $\alpha$ by the Eckhaus instability and several monotone and oscillatory instabilities at higher $G$ and lower $\alpha$. The first instability with increasing $G$ is

\(^{10}\)Where Squire’s theorem shows they are two dimensional[5].
a stationary ($G \sim 530$), followed by an oscillatory one (at $G = 550$) with nearly the same linear growth rate. For our work the important point is perhaps that these instabilities lie in the third dimension, and create variation in $y$ that is a modulation of the 2-D flow. Therefore a 2-D slice, either experimental or simulated, might be expected to show the same cat's eyes before and after the 3-D stationary instability. The oscillations in $y$ would cause a movement of the cat's eyes up and down in the slot.

We should add that these are also results of periodic simulations, and long wavelength modes might be suppressed in a finite width slot. It is true that although good cat's eye visualizations have been done they are generally stationary. The general question of whether the actual flows are 2-D or 3-D seems to be in need of further study, especially experimental work focused on this question.

There are several ways in which we can attempt to make greater contact with the experimental literature. We can do diagnostics to look at the vertical temperature gradient, which has been investigated both experimentally and theoretically. We can also extend our runs both to other Prandtl numbers and to higher $G$ where much of the experimental focus has been. It will be of particular interest to see if we can reproduce the fourth root dependence of the Nusselt number on the Grashof number that is expected for boundary layer physics.

On the experimental side it would be useful to be able to measure more of the velocity field, although this appears to be more difficult than the temperature measurements. It seems possible that with the more sophisticated equipment available today, in particular the ability to computerize the data gathering perhaps using a laser doppler system, that this may be possible.
To make direct contact with the work presented here would require detailed study of the relatively low $G$ regimes at $G_c$ and above, to look for some of the transition behavior that we have seen in our simulations.

We note here that for water the Grashof number is given by

$$G \sim 20L^3\Delta T$$

so that for a four centimeter wide slot a five degree temperature difference will yield $G_c$. This would require two meters of height for an aspect ratio of 50. Doubling the dimensions of the slot will enable the reduction of the temperature gradient by a factor of 8, but require that a much larger area be kept at a uniform temperature. It is important to note that the temperature difference should be kept small if non-Boussinesq effects are to be avoided.

### 11.3 A Note on Experiment and Simulation

As is well known and oft-emphasized [9] the investigative techniques of computer simulation supplement those of experiment and theory. In fact what we have seen is a good example of this. The theory can predict very accurately the linear critical values, and the simulation has been able to reproduce them quite faithfully. Beyond this, the simulation has been able to capture the rich phenomena of the nonlinear physics, including some that the theory has not yet predicted. Simulation can delineate many effects that may be present in actual experiments and idealize them when they are considered unwanted or distracting. For example the precise measurement of the onset of instabilities in experiments is very difficult, as we pointed out above.

A perhaps more important strength of the simulations is that we can extend them to consider the case of a charged fluid with magnetic fields. While
this is relatively easy in a computer code, it would be costly and difficult to do experimentally, and impossible in many parameter regimes. It is this flexibility that originally led to undertaking this project.

On the other hand there are definite limitations to what simulations can do. In many cases, such as 3-D flows, or the high $G$ or high $Pr$ regimes we need to use many modes to resolve the fields and cannot proceed. In these and other cases the experimentalists lead the way and the theoreticians and simulationists follow behind.
Appendix A

Theoretical Appendices

The following appendices contain mathematical derivations of results that are cited in the main body of the thesis. In this chapter we have essentially theoretical results concerned with the exact formulation of the VSC problem and its analytic solutions, and with the solutions for power series expansions used in the discussion of the Landau theory. The next chapter is essentially a detailed presentation of the mathematical theory underlying our numerical methods and the last deals with the specific formulations used in our codes, along with the results of accuracy tests carried out on the numerical components.

A.1 The 1-D Solution

To obtain the 1-D solution\(^1\) we start from the unnormalized Boussinesq equations (2.4), (2.5), and (2.6)

\[
\nabla \cdot \mathbf{u} = 0
\]

\[
\frac{\partial}{\partial t} \mathbf{u} + (\mathbf{u} \cdot \nabla) \mathbf{u} = -\frac{1}{\rho_0} \nabla P + \nu \nabla^2 \mathbf{u} + (1 - \alpha T) \mathbf{g}
\]

\[
\frac{\partial}{\partial t} T + (\mathbf{u} \cdot \nabla) T = \kappa \nabla^2 T,
\]

where we have kept \(\rho_0 \mathbf{g}\) in the gravitational force term so that, as we shall see, \(P = P_0 + P_h\). \(P'\) is not a part of this solution. We have also written the

\(^1\)This is also the solution obtained in the small \(G\) small \(Pr\) limit.
derivatives explicitly, divided through by \( \rho - o \) and \( C_p \), and used the equation of state (2.7) to replace \( \rho' \) in equation (2.5).

Now we assume a solution of the form

\[
T = T(x) \\
u = (0, 0, W(x)) \\
P = P(z)
\]

Now substituting this solution first into equation (A.1), we have

\[
\nabla \cdot u \Rightarrow \frac{\partial}{\partial z} W(x) = 0,
\]

as we expect.

Next we calculate

\[
(u \cdot \nabla) \Rightarrow W(x) \frac{\partial}{\partial z}
\]

so that we have for the convective part of the derivatives

\[
(u \cdot \nabla) u = (u \cdot \nabla) T = 0,
\]

while for a static equilibrium the partials with respect to time are also zero.

This reduces equations (A.2) and (A.3) to

\[
\frac{\partial P}{\partial z} = \mu \frac{\partial^2}{\partial x^2} W(x) + (1 - \alpha T) \rho_o g
\]

and

\[
\nabla^2 T = 0
\]

This latter equation is trivially integrated to give

\[
T = c_1 x + c_2.
\]

Since at \( z = 0 \) \( T(x) = \frac{1}{2}(T_1 + T_2) \), we have \( c_2 = T_o \),
while at \( x = \pm L \quad T = \pm \Delta T \), so that \( c_1 = \Delta T \).

Thus we have for \( T \)

\[
T = T_o + \frac{\Delta T}{L} x
\]  
(A.7)

and replacing \( \delta T \) in equation (A.4);

\[
\frac{\partial P}{\partial z} = \mu \frac{\partial^2}{\partial^2 x} W(x) + (1 - \alpha \Delta T \frac{x}{L}) \rho_o g
\]  
(A.8)

Now the \( x \) independence of \( P \) and the explicit form of \( g = (0, 0, -g) \)
yield an equation for \( P \):

\[
\frac{\partial P}{\partial z} = -\rho_o g.
\]  
(A.9)

This is also trivially integrated to give

\[
P(z) = -\rho_o g z + P_o,
\]  
(A.10)

where \( P_o = \nu_o k T_o \).

Inserting this into equation (A.8), and with a little algebra, we have an equation for \( W(x) \):

\[
\frac{\nu}{\partial^2 x} W(x) = -\alpha g \frac{\Delta T}{L} x
\]  
(A.11)

Integrating this twice

\[
W(x) = \frac{-\alpha g \Delta T}{\nu} \frac{x^3}{6L} + C_1 x + C_2.
\]  
(A.12)

\( W(x) = 0 \) at \( x = \pm L \) requires that \( C_2 = 0 \),

While the boundary condition that at \( x = \pm L \) \( W = 0 \),
yields

\[
C_1 = \frac{\alpha g \Delta T L^2}{\nu 6},
\]  
(A.13)

and so

\[
W(x) = W_o \frac{x}{6L} (1 - \frac{x^2}{L^2})
\]  
(A.14)
where \( W_o \equiv \frac{\alpha g L^2 \Delta T}{\nu} \).

Note that \( W(x)_{\text{max}} = W_o \frac{\sqrt{3}}{27} \).

Finally normalizing\(^2\) by \( W_o, L, \) and \( \Delta T \) we have the 1-D solution:

\[
W(x) = \frac{x}{6} (1 - x^2) \quad \text{(A.15)}
\]

\[
T(x) = x \quad \text{(A.16)}
\]

\[
P(z) = -\frac{1}{\alpha G} z + \frac{P_o}{\rho_o W_o^2} \quad \text{(A.17)}
\]

with \( \alpha \) normalized by \( \frac{1}{\Delta T} \).

From equation (A.15) for the velocity we can derive expressions for \( \psi \) and \( \omega \) as well:

\[
\psi(x) = \frac{1}{24} (1 - x^2)^2 \quad \text{(A.18)}
\]

\[
\omega(x) = \frac{1}{6} (3x^2 - 1) \quad \text{(A.19)}
\]

Since \( P \) is eliminated from our equations when we take the curl\(^3\) this is a one-dimensional solution which depends only on \( x \) and is invariant in \( z \).

For reference we will give the Chebychev polynomial expansions of these functions:

\[
W(x) = \frac{1}{24} (T_1 - T_3) \quad \text{(A.20)}
\]

\[
\psi(x) = \frac{1}{192} (3T_0 - 4T_2 + T_4) \quad \text{(A.21)}
\]

\[
\omega(x) = \frac{1}{12} (T_0 + 3T_2) \quad \text{(A.22)}
\]

\(^2\)See Sections 2.3.

\(^3\)See Section 2.4.
A.2 The High G Solution

Starting again from equations (A.1), (A.2), and (A.3) but this time we assume that the temperature has the form

\[ T = \theta(x) + \beta z \]  \hspace{1cm} (A.23)

so that here

\[ (u \cdot \nabla) T = \beta W(x), \]

and equation (A.5) becomes

\[ \frac{\partial^2}{\partial^2 x} \theta(x) = \beta \kappa W(x). \]  \hspace{1cm} (A.24)

Equation (A.4) now reduces to

\[ \frac{\partial P}{\partial z} = -\rho_0 g + \alpha \beta z \rho_0 g \]  \hspace{1cm} (A.25)

which, when integrated, gives

\[ P(z) = -\rho_0 g \left( \frac{1}{2} \alpha \beta z^2 - z \right) + P_o, \]  \hspace{1cm} (A.26)

where \( P_o = n_o k T_o \).

Using this result for \( P \) and definition (A.23) for \( T \) in equation (A.4) we are left with an equation for \( W(x) \):

\[ \nu \frac{\partial^2}{\partial^2 x} W(x) = -\alpha g \theta(x) \]  \hspace{1cm} (A.27)

Combining this equation with equation (A.24) we get

\[ \frac{\partial^4}{\partial^4 x} W(x) = -\frac{\alpha \beta g}{\nu \kappa} W(x) \]  \hspace{1cm} (A.28)

and an identical equation for \( \theta(x) \).
Defining \( n \equiv \frac{a_b g}{v_n} \),

and assuming an exponential form for \( W(x) \):

\[
W(x) = e^{\gamma x},
\]

so that \(
\gamma = \sqrt{-n} = \pm (1 \pm i) \sqrt{n/4},
\)

we have finally

\[
W(x) = \frac{\sinh [(1 + i)mx/L]}{\sinh [(1 + i)m]} \tag{A.29}
\]

where \( m = \frac{a_b L^4}{4 v_n} \),

which is a Rayleigh number.

For \( \theta(x) \) we take the negative sign reflecting the relationship between \( \theta(x) \) and \( W(x) \) in equations (A.27) and (A.24) to give

\[
\theta(x) = \frac{\sinh [(1 - i)mx/L]}{\sinh [(1 - i)m]} \tag{A.30}
\]

### A.3 The Boussinesq Approximation

In this section\(^4\) we will take a look at the approximations that were made in deriving the VSC equations. In this way we can estimate the parameter regimes in which the model may be expected to yield realistic results, and if not, to what degree they can be taken as qualitatively correct.

#### A.3.1 The Validity Criterion

In conducting this analysis we start with the equation of continuity written in steady-state form:

\[
u \cdot \nabla \rho + \rho \nabla \cdot u = 0 \tag{A.31}
\]

We will rewrite the variables \( P, T, \) and \( \rho \) in terms of absolute values taken at an arbitrary reference point, \( P_o, T_o, \) and \( \rho_o, \) hydrostatic variations \( P_h, \) and \( \rho_h, \)

---

\(^4\)This section follows the appendix to chp.14 of Triton[104].
which are functions of the vertical coordinate, and variations associated with the motion \( P', T', \) and \( \rho' \):

\[
P = P_o + P_h + P' \\
T = T_o + T' \\
\rho = \rho_o + \rho_h + \rho'
\]

Next inserting these in equation (A.35) and introducing the scale length \( L \), and a typical velocity \( U \), we can estimate the order of magnitude of its two terms as

\[
\frac{U}{L} (\rho_h + \rho') \quad \text{and} \quad \frac{U}{L} (\rho_h + \rho_h + \rho')
\]

We can therefore safely drop the first of these terms to yield the desired

\[
\nabla \cdot \mathbf{u} = 0 \tag{A.32}
\]

only if the conditions \( |\rho_h/\rho_o| \ll 1 \) and \( |\rho'/\rho_o| \ll 1 \) both hold. These are the incompressibility conditions.

Making use of the thermal expansion coefficient,

\[
\alpha = -\frac{1}{\rho} \left( \frac{\partial \rho}{\partial T} \right)_P,
\]

and the isothermal compressibility,

\[
\beta = \frac{1}{\rho} \left( \frac{\partial \rho}{\partial P} \right)_T,
\]

\(^{5}\) Note for reference that the speed of sound is given by

\[
c_s = \sqrt{\left( \frac{\partial \rho}{\partial P} \right)_T},
\]
we rewrite \( \rho_h \) and \( \rho' \) as

\[
\rho_h = \rho_o \beta P_h \quad \text{and} \quad \rho' = \rho_o (\alpha T' + \beta P')
\]

Inserting these in the above inequalities yields the conditions

\[
|\beta P_h| \ll 1 \quad \text{and} \quad |\alpha T' - \beta P'| \ll 1.
\]

In hydrostatic balance \( P_h \sim \rho_o g L \) so that \( |\beta P_h| \ll 1 \) will hold when

\[
B \equiv g \rho_o \beta L \ll 1. \tag{A.33}
\]

For free convection the motion induced pressure gradient and the buoyancy terms roughly balance so that \( P'/L \sim g \rho_o \alpha T' \). Thus Condition B yields \( |\beta P'| \ll |\alpha T'| \).

With this the second condition above reduces to

\[
A \equiv \alpha \Theta \ll 1 \tag{A.34}
\]

where \( \Theta \) represents the magnitude of the temperature variations.

Next we look at the kinetic equation. Assuming steady flow and making use of equation (A.32) we have

\[
u \cdot \nabla (\rho u) = -\nabla P + \mu \nabla^2 u + \rho g, \tag{A.35}
\]

where \( \mu \) has been assumed constant in deriving the viscous term. Substituting for \( \rho \) and \( P \) as above, subtracting out the hydrostatic relation \( \nabla P_h = (\rho_o + \rho_h)g \), and recalling that \( P_o \) is a constant we get

\[
\rho_o \nabla \cdot \nabla u = -\nabla P' + \rho' g + \mu \nabla^2 u \tag{A.36}
\]

Whenever Condition B holds we can drop the \( \rho_h g \) term while \( \rho' \) reduces to \( -\rho_o \alpha T' \). Finally dividing through by \( \rho_o \) we have the desired Boussinesq form

\[
u \cdot \nabla u = -\frac{1}{\rho_o} \nabla P' + g \alpha T' + \mu \nabla^2 u \tag{A.37}
\]
It is important to note that, by keeping \(-\rho_{o}aT'\) while dropping \(\rho_h g\), we have required that \(A\gg B\), allowing us to drop the pressure induced density variations while keeping those due to the temperature.\(^6\)

To derive the Boussinesq form of the heat conduction equation we start from the first law of thermodynamics:\(^7\)

\[
\frac{DE}{Dt} = \frac{DQ}{Dt} - P \frac{DV}{Dt}
\]  \(\text{(A.38)}\)

Using

\[
\frac{DQ}{Dt} = \frac{1}{\rho} (k\nabla^2 T + J + \phi)
\]  \(\text{(A.39)}\)

where \(J\) = the internal heat generation

and \(\phi = \frac{\nu^2}{2}(\partial_t u_j + \partial_i u_j)^2\) = the heat from viscous dissipation

along with

\[
\frac{DE}{Dt} = \left( C_p - \frac{\alpha P}{\rho} \right) \frac{DT}{Dt} + \frac{\beta P - \alpha T}{\rho} \frac{DP}{Dt}
\]  \(\text{(A.40)}\)

and

\[
\frac{DV}{Dt} = \frac{\alpha}{\rho} \frac{DT}{Dt} - \frac{\beta}{\rho} \frac{D\rho}{Dt}
\]  \(\text{(A.41)}\)

from thermodynamics, and canceling equal terms on the right and left sides we can write the first law in the form

\[
C_p \frac{DT}{Dt} - \frac{\alpha T}{\rho} \frac{DP}{Dt} = \frac{1}{\rho} (k\nabla^2 T + J + \phi).
\]  \(\text{(A.42)}\)

We wish to drop out the second term on the left. Assuming that \(\rho = \rho_o\), expanding \(P\) and \(T\), and making an order of magnitude comparison of the two left-hand side terms we get the conditions

\[
\frac{\alpha T u \cdot \nabla P_h}{\rho C_p u \cdot \nabla T} \sim \frac{\alpha T_o P_h}{\rho_o C_p T'} \sim \frac{g\alpha LT_o}{C_p T'}.
\]

\(^6\)See [55] sec.56.

\(^7\)Taking unit mass so that \(V = 1/\rho\).
having used the hydrostatic balance condition from above, and, with the further help of the free convection relation \( g\alpha L \sim \frac{P'}{\rho T'} \),

\[
\frac{\alpha T u \cdot \nabla P'}{\rho C_p u \cdot \nabla T} \sim \frac{\alpha T_o P'}{\rho_o C_p T'} \sim \frac{g\alpha L}{C_p} \alpha T_o.
\]

From these expressions we can form two conditions,

\[
D \equiv \frac{g\alpha L T_o}{C_p \Theta} \ll 1 \tag{A.43}
\]

and

\[
C \equiv \frac{g\alpha L}{C_p} \ll 1. \tag{A.44}
\]

The former is a stronger condition as long as \( \frac{T_o}{T'} < 0 \) and serves to insure that the whole second term is negligible as desired.

The second condition is derived assuming that \( \alpha T_o \ll 1 \), in which case it ensures that the perturbative part of the pressure term is negligible.

If we define a new thermal variable \( \theta = T - (T_o - T_o) \),

we can dispense with the stronger condition \( D \) in the following way:

First we take \( T_o \) to be the adiabatic temperature such that

\[
\frac{dT_o}{dz} = -\frac{g\alpha T_o}{C_p}.
\]

Using this relation and the earlier definition of \( T = T_o + T' \), we can transform the pressure term

\[
\frac{\alpha T}{\rho} \frac{DP_h}{Dt} = \frac{\alpha T u_z dP_h}{\rho} \frac{dz}{dt} = -\alpha g u_z T = C_p \frac{DT_o}{Dt} - g\alpha u_z T',
\]

where \( P_h \) is taken as independent of time and the horizontal coordinate, and we have used the hydrostatic relation \( \frac{dP_h}{dz} \sim \rho g \) again. Comparing these last two terms we find

\[
\frac{g\alpha u_z T'}{C_p DT / Dt} \sim \frac{g\alpha L}{C_p}.
\]
Thus when condition C holds we can drop the term \(-g\alpha u_x T'\) and arrive at

\[
\frac{\alpha T}{\rho} \frac{DP}{Dt} = C_p \frac{DT_a}{Dt}.
\] (A.45)

We can now rewrite equation (A.42). Replacing \(T\) with \(\theta\), using equation (A.45) and the time and spatial independence of \(T_a\) we have,

\[
C_p \frac{D\theta}{Dt} - \frac{\alpha T}{\rho} \frac{DP'}{Dt} = \frac{1}{\rho} (k \nabla^2 \theta + \Theta + \phi).
\] (A.46)

As we have shown above, Condition C is sufficient to knock out the \(\frac{DP}{Dt}\) term. For most fluids we can neglect \(\Theta\) as there is no internal heat generation.

Introducing a scale length for viscous effects, \(\delta\), the viscous term \(\phi\) is of the order \(\mu \frac{u^2}{\delta^2}\). Using the free convection condition, that the viscous and buoyancy forces in the kinetic equation balance, yields

\[
\mu \frac{u}{\delta^2} \sim \rho g \alpha T'.
\]

With this we can construct the condition that \(\phi\) is negligible in the heat conduction equation,

\[
\frac{\phi}{\rho C_p u \cdot \nabla T} \sim \frac{\rho g \alpha T' u}{\rho C_p u T'/L} = \frac{g \alpha L}{C_p} = C \ll 0.
\] (A.47)

Thus we have, with Condition D, assuming that \(k\) is constant,

\[
\rho C_p \frac{DT}{Dt} = k \nabla^2 T.
\] (A.48)

Similarly this equation, with \(T \rightarrow \theta\) holds for Condition C.

To summarize, we have found that Conditions A and B will yield a solenoidal velocity field and an incompressible fluid. These results on the velocity and the density, along with the invariance of \(\mu\) and \(\alpha\), and further use of Condition B, but not Condition A, such that \(A \gg B\), yield the appropriate
equation of motion. Finally Condition \(D\) or \(C\), with the constancy of \(k\), and \(C_p\), gives us a Boussinesq heat conduction equation for either \(T'\) or \(\theta\) respectively. Lastly the use of Condition \(B\), with \(A \gg B\) again, yields the appropriate equation of state. The pressure variation of the density is ignored, but in the buoyancy term we retain the temperature variations that are the key to driving convective flows.

A.3.2 The Physical Significance of These Conditions

In physical terms the ratios \(A\), \(B\), \(C\), and \(D\) represent

- \(A\): the fractional density change from expansion in the imposed temperature field.

- \(B\): the ratio of the system scale length over the isothermal-pressure scale height.

- \(C\): the ratio of the system scale length over the adiabatic temperature scale height. It can also be read as the ratio of the rate of kinetic energy generation over the rate of heat transfer.

- \(D\): the ratio of the adiabatic temperature gradient to the typical imposed temperature gradient. Usually \(D \ll C\).

In order for the Boussinesq equations to hold these must all be small. In practice, however, this will not always be the case:

In small scale laboratory systems the most common violations are due to the variation of \(\mu\) and/or \(\alpha\) with the temperature. In gases Condition \(A\) can become too large.
In large scale systems Conditions $B$ and $C$ are often violated. For large $B$ we need to include $P_h$ in the equation of motion and $\beta P'$ in the buoyancy force. For large $C$ we can no longer neglect the viscous generation of heat, $\phi$, or the other non-Boussinesq effects that we left out of the heat conduction equation, the $\frac{DP}{Dt}$ terms. In practice all four of these effects are likely to become significant at roughly the same time.

A.4 Scale Invariance

A set of equations can be said to be scale invariant if a transformation of a basic coordinate, such as length,

$$x \rightarrow \lambda x$$

leaves the equations unchanged.

The equations with which we work in these hydrodynamic flow problems can be formulated so as to bring out their inherent scale invariance. This is in fact what we did when we normalized the basic equations to obtain a dimensionless form.\(^8\) In doing this we also derived some dimensionless parameters, the Grashof and Prandtl numbers, which characterize these flows. All flows with the same values of these parameters are similar flows that differ from each other only in their dimensional scales.

Here we shall demonstrate in detail the scale invariance of the hydrodynamic equations

$$\frac{\partial u}{\partial t} + (u \cdot \nabla)u = -\frac{1}{\rho_0} \nabla P + \nu \nabla^2 u - \alpha T g$$

(A.49)

\(^8\)See Chapter 2.
\[
\frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla)T = \kappa \nabla^2 T \tag{A.50}
\]

First we need to take a look at the dimensionality of the various quantities involved in these equations.

For the coordinates:

\[
\begin{align*}
x &= L \hat{x} \propto l \\
t &= \frac{L}{W_o} \hat{t} \propto t
\end{align*} \tag{A.51}
\]

For the variables:

\[
\begin{align*}
\mathbf{u} &= W_o \mathbf{\hat{u}} \propto lt^{-1} \\
P &= W_o^2 \rho_o \Pi \propto l^2 t^{-2} m \left[l^{-3}\right] \\
T &= \Delta T \mathbf{\hat{T}} \propto \theta
\end{align*} \tag{A.52}
\]

And the dimensional parameters:

\[
\begin{align*}
\rho &= m \left[l^{-3}\right] \\
g &= lt^{-2} \\
\nu &= l^2 t^{-1} \\
\kappa &= l^2 t^{-1} \\
\alpha &= \theta^{-1} \\
C_p &= l^2 t^{-2} \theta^{-1}
\end{align*} \tag{A.53}
\]

Finally it is useful to check the consistency of all this by calculating the dimensionality of \(W_o\),

\[
W_o = \frac{\alpha g L^2 \Delta T}{\nu} \propto \frac{l}{t} \tag{A.54}
\]
A.4.1 Energy Scaling

We saw in Section 2.3 that we normalized the temperature of the system $T$ by $T_o$ so that we need only concern ourselves with the relative temperature defined by the boundary conditions. It is important to note, however, that although $T_o$ is no longer explicitly in the equations, it comes in implicitly in the parameters $\alpha$, $\rho_o$, $\nu$, and $\kappa$. Although the Boussinesq Approximation requires these to be constant throughout the region of interest, it must be remembered that they are temperature dependent quantities, and this constant value will therefore vary for different values of $T_o$.

Thus although the Equations (A.49) and (A.50) are formally invariant under scalings of the energy, we must take into account the temperature dependence of the various parameters.

A.4.2 Global Invariance

The purpose of de-dimensionalizing the hydrodynamic equations can be shown by considering what we shall call a global scale transformation. This means that we scale the overall size of the system and the temperature differential across it,

$$
L \rightarrow \lambda L \\
\Delta T \rightarrow \lambda^\theta \Delta T
$$

(A.55)

This leads to change in the scale velocity $W_o$,

$$
W_o \rightarrow \lambda^{2+\theta} W_o.
$$

(A.56)

Looking at the Definitions (A.51) and (A.52) we see that this scaling will affect the unprimed dimensional coordinates and variables,

$$
x \rightarrow x' = \lambda x
$$
\[ t \rightarrow t' = \lambda^{-1-\theta} t \]
\[ u \rightarrow u' = \lambda^{2+\theta} u \]
\[ P \rightarrow P' = \lambda^{4+2\theta} P \]
\[ T \rightarrow T = \lambda^{\theta} T \]  

(A.57)

but the dimensionless variables \( \tilde{a}, \tilde{u}, \tilde{P}, \tilde{T} \), will be unchanged. Thus in order to preserve the dimensionless equations unchanged under transformation of \( L \) and \( \Delta T \) we must only insure that the dimensionless parameters of the equations, \( G \) and \( Pr \), are likewise invariant. The Prandtl number has no dependence on either \( L \) or \( \Delta T \) and therefore remains the same. This is not surprising when we consider that this parameter is descriptive of the fluid and that we have only altered the boundary conditions to which the fluid is subject. The Grashof number, on the other hand, encodes these very boundary influences and writing it in its form

\[ G = \frac{W_o L}{\nu} \rightarrow G' = \lambda^{3+\theta} G, \]  

(A.58)

we see that the condition that \( G' = G \) requires that \( \theta = -3 \).

Thus we see that the Scaling (A.55), with \( \theta = -3 \), will define families of similar physical systems that can all be described by the same dimensionless equations. In a similar fashion a given value of \( Pr \) represents a family of fluids that all have the same ratio of \( \kappa \) to \( \nu \).

A.5 Solutions to \( \dot{A} \) Expansions

Here we work out general solutions to successive truncations of the series expansion:

\[ \frac{dA}{dt} = \gamma A + \alpha A^2 + \beta A^3 + \delta A^4 \cdots \]  

(A.59)
A.5.1 Linear

Starting with the linear case, we write

\[ \frac{dA}{dt} = \gamma A \]  \hspace{1cm} (A.60)

Writing \( \dot{A} \) for \( \frac{dA}{dt} \) we write this as

\[ \frac{\dot{A}}{A} = \gamma \]

Integrating from \( t = 0 \) to \( t \) we get \( \ln A = \gamma t + \ln A_0 \).

Finally, exponentiating we have

\[ A(t) = A_0 e^{\gamma t}. \]  \hspace{1cm} (A.61)

A.5.2 Quadratic

Keeping the quadratic term our equation is

\[ \frac{dA}{dt} = \gamma A + \alpha A^2 \]  \hspace{1cm} (A.62)

The solution proceeds as follows:

\[ \dot{A} = \alpha A \left( A + \frac{\gamma}{\alpha} \right) \]

\[ \frac{\dot{A}}{(A)(A + \frac{\gamma}{\alpha})} = \frac{\alpha}{\gamma} \left[ \frac{\dot{A}}{A} - \frac{\dot{A}}{(A + \frac{\gamma}{\alpha})} \right] = \alpha, \]

where we have used the algebraic identity

\[ \frac{1}{(A + \alpha)(A + \beta)} = \frac{1}{(\beta - \alpha)} \left[ \frac{1}{A + \alpha} - \frac{1}{A + \beta} \right]. \]

Multiplying through by \( \gamma/\alpha \) and integrating,

\[ \ln A - \ln \left( A + \frac{\gamma}{\alpha} \right) = \gamma t + \ln \left( A_0 / (A_0 + \frac{\gamma}{\alpha}) \right). \]
Exponentiating, and with a little manipulation, we arrive at

\[
\frac{1}{A} = -\frac{\alpha}{\gamma} + \left\{ \frac{\alpha}{\gamma} + \frac{1}{A_0} \right\} e^{-\gamma t}, \tag{A.63}
\]

while we can solve for

\[
A(t) = \frac{-\frac{\gamma}{\alpha}}{1 - \left\{ 1 + \frac{\gamma}{\alpha A_0} \right\} e^{-\gamma t}}. \tag{A.64}
\]

For \( t \to \infty \): \( A_{asm} = -\frac{\gamma}{\alpha} \),

while for \( t \to 0 \) this reduces to solution (A.61).

### A.5.3 Cubic

Now we add on a cubic term as well:

\[
\frac{dA}{dt} = \gamma A + \alpha A^2 + \beta A^3 \tag{A.65}
\]

The solution follows as before:

\[
\dot{A} = \beta A \left( A^2 + \frac{\alpha}{\beta} A + \frac{\gamma}{\beta} \right)
\]

Define \( \delta_1 = a + b \) and \( \delta_2 = a - b \),

where \( a = \frac{\alpha}{2\beta} \) and \( b = \frac{1}{2} \sqrt{\frac{\alpha^2}{\beta^2} - 4\frac{\gamma}{\beta}} \).

For \( \alpha^2 > 4\beta\gamma \) the \( \delta \)s are real and we can write

\[
\frac{\dot{A}}{(A)(A + \delta_1)(A + \delta_2)} = \beta.
\]

We can derive the required algebraic identity

\[
\frac{1}{(A + \alpha)(A + \beta)(A + \gamma)} = \frac{1}{\beta\gamma + \alpha\gamma + \alpha\beta} \left[ \frac{c_1}{A + \alpha} + \frac{c_2}{A + \beta} + \frac{c_3}{A + \gamma} \right]
\]

by solving

\[(c_1 + c_2 + c_3)A^2 = 0 \quad \text{and} \quad c_1(\beta + \gamma)A + c_2(\alpha + \gamma)A + c_3(\alpha + \beta)A = 0\]
to yield

\[ c_1 = \frac{(\beta - \gamma)}{(\beta - \gamma)} \quad c_2 = \frac{(\gamma - \alpha)}{(\beta - \gamma)} \quad c_3 = \frac{(\alpha - \beta)}{(\beta - \gamma)}. \]

Making use of this we have

\[
\frac{\dot{A}}{(A)(A + \delta_1)(A + \delta_2)} = \frac{1}{\delta_1\delta_2} \left[ \frac{\dot{A}}{A} + \frac{\delta_2}{\delta_1 - \delta_2} \frac{\dot{A}}{(A + \delta_1)} - \frac{\delta_1}{\delta_1 - \delta_2} \frac{\dot{A}}{(A + \delta_2)} \right].
\]

Integration gives

\[
\frac{1}{a^2 - b^2} \left[ \ln A + \frac{\delta_2}{2b} \ln(A + \delta_1) - \frac{\delta_1}{2b} \ln(A + \delta_2), \right]_0^t = \beta t
\]

and exponentiation

\[
\left[ \frac{A(A + \delta_1)^{\frac{\delta_2}{2\gamma} - \frac{\alpha}{\beta}}}{(A + \delta_2)^{\frac{\delta_1}{2\gamma} + \frac{1}{2}}} \right] = \left[ \frac{A^2}{(A + \delta_1)(A + \delta_2)} \right]^{\frac{\alpha}{\beta}} \left[ \frac{A + \delta_2}{A + \delta_1} \right]^{-\frac{\beta}{\delta_1}} = \text{constant } e^{\beta t},
\]

where the constant is of the same form as the central expression but with \(A(t)\) replaced by \(A_0\). Substituting back for the \(\delta s\) and shifting exponents we get

\[
\frac{A^2 + \alpha \beta A + \gamma}{A^2} \left[ \frac{A + \frac{\alpha}{\beta} - b}{A + \frac{\alpha}{\beta} + b} \right]^{-\frac{\alpha}{\beta}} = \text{constant } e^{-2\gamma t}. \quad (A.66)
\]

Defining

\[
\left[ \right] = \left[ \frac{A + \frac{\alpha}{\beta} - b}{A + \frac{\alpha}{\beta} + b} \right]^{-\frac{\alpha}{\beta}} = \left[ 1 - \frac{2b}{A + \frac{\alpha}{\beta} + b} \right]^{-\frac{\alpha}{\beta}}
\]

we can write a polynomial for \(A\):

\[
A^2 + \alpha \beta A \left\{ \left[ \right] - \text{cnst } e^{-2\gamma t} \right\} + \gamma \left\{ \left[ \right] - \text{cnst } e^{-2\gamma t} \right\} = 0
\]

Finally defining

\[
\left\{ \right\} \equiv \left\{ \frac{\left[ \right]}{\left[ \right] - \text{cnst } e^{-2\gamma t}} \right\},
\]

we have a solution for \(A\)

\[
A(t) = -\frac{\alpha}{2\beta} \left\{ \right\} \pm \frac{1}{2} \sqrt{\frac{\alpha^2}{\beta^2} \left\{ \right\}^2 - 4 \frac{\gamma}{\beta} \left\{ \right\}}. \quad (A.67)
\]
For $t \to \infty$ we have simply

$$A_{sm} = \frac{-\alpha}{2\beta} \pm \frac{1}{2} \sqrt{\frac{\alpha^2}{\beta^2} - \frac{4\gamma}{\beta}}$$

(A.68)

We note that the solutions are equivalent to $-\delta_1$ and $-\delta_2$ as defined above. This can also be had by simply letting $e^{-2\gamma t}$ go to zero in Equation (A.66). Note that these two solutions are also found by dividing out the polynomial in $A$ leaving $[\ ] = 0$. Which of the two we get depends on the sign of the exponent $\frac{\alpha}{2\beta \beta}$.

For a pure cubic $\alpha = 0$, Equation (A.66) can be reduced to

$$A(t)^2 = \frac{-\frac{\gamma}{\beta}}{1 - \left\{ 1 + \frac{\gamma}{\beta} \frac{1}{A^2} \right\} e^{-2\gamma t}}.$$

(A.69)

For $t \to \infty$, $A = \sqrt{-\frac{\gamma}{\beta}}$.

The $t \to 0$ limit is the same as before.

For $\alpha^2 > 4\beta \gamma$ we would have complex roots if we tried to factor the quadratic as we have done above. From the integral tables we have

$$\int \frac{\dot{A}}{A(A^2 + \frac{\alpha}{\beta} A + \frac{\gamma}{\beta})} = \frac{\beta}{2\gamma} \ln \left( \frac{A^2 + \frac{\alpha}{\beta} A + \frac{\gamma}{\beta}}{A^2 + \frac{\alpha}{\beta} A + \frac{\gamma}{\beta}} \right) - \frac{\alpha}{2\gamma b'} \tan^{-1} \left( \frac{2A + \frac{\alpha}{\beta}}{2b'} \right),$$

where $b' = \frac{1}{2} \sqrt{4\frac{\gamma}{\beta} - \frac{\alpha^2}{\beta^2}}$.

### A.5.4 Higher Orders

We can see that in general we can follow the same general procedure and for $t \to \infty$ get the solution for $A$ as a series of roots of an increasingly higher order polynomial, one order less than the equation for $\dot{A}$. Care must be taken for complex roots.
Appendix B

Numerical Appendices

These appendices outline some of the mathematical background for the spectral use of Fourier and Chebychev series in approximating general functions.\footnote{A good general reference for these subjects with an emphasis on computational applications is Tajima\cite{9}.} The exact transformations used in the numerical treatment are also derived. We pay special attention to the Chebychev series, discussing its advantages in approximating a general smooth function, and deriving various properties and relations that we use in the numerical treatment of the linear and nonlinear equations. Finally we present a brief discussion of convolutions and aliasing, which enter into the treatment and testing of the nonlinear terms.

B.1 Approximation Theory

In working with spectral methods, we are using a finite series of functions to approximate a known or unknown function $f(x)$. To understand some of the basic theory\footnote{This section and the following sections on Fourier and Chebychev Series follow closely discussions in Fox and Parker\cite{105}.} involved and to demonstrate in particular some of the properties of the series representations that we use, it is best to start with the basic idea of using an $n^{th}$ degree polynomial $p_n(x)$ to approximate a function $f(x)$. In this case there will be an error

$$e_n(x) = f(x) - p_n(x)$$
We are interested in minimizing this error. To do this there are two commonly used criteria used to judge the 'minimum' error when considering a range of $x$.

### B.1.1 The Minimax Criterion

The first method is to minimize the maximum value of $|e_n(x)|$ in the range of $x$ of interest, usually taken here to be $-1 \leq x \leq 1$. We now look at two particular forms for $f(x)$.

First when $f(x) \equiv 0$, it can be shown that the polynomial $p_n(x)$ that satisfies the minimax condition in $-1 \leq x \leq 1$ must have $n + 1$ alternating maxima and minima with values $+M$ and $-M$ respectively, including the end points $\pm 1$. It consequently also has $n$ zeros in this interval.

Second, by considering the Lagrangian interpolation formula we can find a polynomial approximation for a function $f(x)$ having bounded derivatives of order $n + 1$ on the interval $(-1, 1)$ of the form

$$p_n(x) = f(x) - \frac{\Pi(x)f^{(n+1)}(\xi)}{(n + 1)!} \quad \xi = [0, x]$$  \hspace{1cm} (B.1)

$$\Pi(x) = (x - x_0)(x - x_1)\cdots(x - x_n)$$

This polynomial is of degree $n$ and matches $f(x)$ at $n + 1$ points given by the zeros of $\Pi(x)$. If we write $M_{n+1}$ for the maximum value of $f^{(n+1)}$, the error (the second term on the right) can be minimized according to the minimax criterion by minimizing

$$\max \left| \frac{e_n(x)}{M_{n+1}} \right| = \frac{\max|\Pi(x)|}{(n + 1)!}$$  \hspace{1cm} (B.2)

It is clear that these two cases actually have the same solutions to within a multiplicative constant.
B.1.2 The Least Squares Technique

The second method for defining a minimum error is to reduce the average error over the whole range $-1 \leq x \leq 1$. In the general case we can give different weights over the range and arrive at the least squares criteria,

$$S = \int_{-1}^{1} w(x)e_n^2(x)dx = \text{minimum}, \quad \text{(B.3)}$$

or in the corresponding discrete case,

$$S_n = \sum_{k=0}^{N} w(x_k)e_n^2(x_k) = \text{minimum}. \quad \text{(B.4)}$$

In the following we work with the continuous case. Replacing $e_n(x)$ in Equation (B.3), assuming a general form for the polynomial

$$p_n(x) = \sum_{k=0}^{n} c_k \phi_k(x), \quad \text{(B.5)}$$

where the $\phi_k(x)$ are polynomials of degree $k$, and using the condition for a minimum that $\partial S/\partial c_j = 0$ for $j = 0, \ldots, n$, we can derive the set of equations

$$\int_{-1}^{1} w(x)\phi_j(x) \{f(x) - (c_0 + \ldots + c_k \phi_k(x))\} dx = 0 \quad \text{(B.6)}$$

for $k = 0, \ldots, n$.

From these one can derive a set of linear simultaneous equations for the $c_k$, called the normal equations, taking the form $A\mathbf{c} = \mathbf{b}$. If we now take a set for the $\phi_k(x)$ which is mutually orthogonal with respect to $w(x)$ we can diagonalize $A$ and reduce its components to the form

$$a_{kk} = \int_{-1}^{1} w(x)\phi_k^2(x)dx, \quad \text{(B.7)}$$

while for the components of $\mathbf{b}$ we have

$$b_k = \int_{-1}^{1} w(x)\phi_k(x)f(x)dx, \quad \text{(B.8)}$$
and thus for the coefficients \( c_k \),

\[
    c_k = \frac{\int_{-1}^{1} w(x) \phi_k(x) f(x) \, dx}{\int_{-1}^{1} w(x) \phi_k^2(x) \, dx}.
\]

(B.9)

When \( w(x) \) does not change sign in the interval \((-1,1)\), and the \( \phi_k(x) \) do not vanish identically, the \( c_k \) are finite, and we have a solution for \( p_n \).

We can take any normalization that is convenient and choose one such that the coefficient of \( x^k \) in \( \phi_k(x) \) is \( A_k \) and that

\[
    l_k = A_k \int_{-1}^{1} w(x) x^k \phi_k(x) \, dx.
\]

(B.10)

There are various ways\(^3\) to construct suitable polynomials \( \phi_k(x) \), but in any case they must have the following properties:

- \( \phi_k(x) \) is orthogonal to all \( \phi_j(x) \), \( j < k \),

- which implies as well that the set is linearly independent,

- and that \( \phi_k(x) \) has \( k \) real and distinct zeros in the interval \((-1,1)\).

It can also be shown that they satisfy a recurrence relation:

\[
    \phi_{k+1}(x) = (\alpha_k x + \beta_k) \phi_k(x) + \gamma_{k-1} \phi_{k-1}(x).
\]

(B.11)

From this and making use of Normalization (B.10), we can solve for the \( \alpha_k \), \( \beta_k \), and \( \gamma_k \) as follows:

\[
    \alpha_k = \frac{A_{k+1}}{A_k}, \quad \beta_k = \frac{\alpha_k}{l_k} \int_{-1}^{1} w(x) x \phi_k^2(x) \, dx
\]

(B.12)

\[
    \gamma_{k-1} = -\alpha_k \left( \frac{A_{k-1}}{A_k} \right) \left( \frac{l_k}{l_{k-1}} \right).
\]

\(^3\)For instance the Gram-Schmidt technique, see [105]
By reconfiguring the problem we can derive a differential generating formula for the $\phi_k(x)$. First we write

$$w(x)\phi_k(x) = V_k^{(k)} \equiv \frac{d^k V_k(x)}{dx^k}. \quad (B.13)$$

Integrating by parts $k$ times we get a system of equations

$$\frac{d^{k+1}}{dx^{k+1}} \left[ \frac{1}{w(x)} \frac{d^k V_k(x)}{dx^k} \right] = 0 \quad (B.14)$$

with boundary conditions

$$V_k(\pm 1) = V_k'(\pm 1) = \ldots = V_k^{(k-1)}(\pm 1) = 0.$$ 

Finally this gives a simpler relation between $l_k$ and $A_k$,

$$l_k = (-1)^k(k!) A_k \int_{-1}^{1} V_k(x) dx \quad (B.15)$$

For a general weighting of $w(x) = (1 - x^2)^\alpha$, we now have a class of polynomials called the ultra spherical polynomials that for specific values of $\alpha$ yield such well-known sets as the Legendre ($\alpha = 0$), the Chebychev ($\alpha = -\frac{1}{2}$), and in the limit ($\alpha \to \infty$) the Taylor series of polynomials.

In studying the convergence rates of these series of polynomials, the important issue is the form of $A_k$ for large $k$. For a large value of $A_k$ the convergence will be rapid and vice versa. It can be shown that the minimax criteria in Appendix B.1.1 are equivalent to maximizing $A_k$ here, and thus maximizing the convergence. We will see below that the series which satisfies these criteria is none other than the Chebychev series.

### B.2 Fourier Series

In the previous section on approximation theory we discussed the use of orthogonal polynomials for approximating functions, but one can also use
wider classes of orthogonal functions. Of these the trigonometric functions have been found to lend themselves very well to computational purposes, and they form the basis of both the Fourier series and the derivative Chebychev series. The Fourier series is especially suited for functions periodic with period $2\pi$.

The basic definition of a Fourier series is

$$f(x) = \frac{1}{2} c_0 + \sum_{k=1}^{\infty} \{ c_k \cos(kx) + s_k \sin(kx) \}, \quad (B.16)$$

where we can use the orthogonality of the sine and cosine functions to solve for the real coefficients:

$$c_k = \frac{1}{N} \int_{-N}^{N} f(x) \cos(kx) \, dx \quad \text{and} \quad s_k = \frac{1}{N} \int_{-N}^{N} f(x) \sin(kx) \, dx \quad (B.17)$$

for $-N \leq x \leq N$,

where $N$ is usually taken to be $\pi$.

If we limit $k$ to a maximum value of $N$ we have a 'trigonometric polynomial' that is a least-squares approximation to $f(x)$ with unit weight function in the interval $-\pi \leq x \leq \pi$.\footnote{See [105] chp.2}

$$f(x) = \frac{1}{2} c_0 + \sum_{k=1}^{N} \{ c_k \cos(kx) + s_k \sin(kx) \}. \quad (B.18)$$

We may also rewrite Definition (B.16) in terms of complex coefficients by substituting the exponential expressions for the sine and cosine, giving

$$f(x) = \frac{c_0}{2} + \frac{1}{2} \sum_{k=1}^{\infty} (c_k - is_k) e^{ikx} + \frac{1}{2} \sum_{k=1}^{\infty} (c_k + is_k) e^{-ikx}. \quad (B.19)$$

Now using the conditions $c_k = c_{-k}$, $s_k = -s_{-k}$, and $b_0 = 0$ can derive,

$$f(x) = \frac{1}{2} \sum_{k=-\infty}^{\infty} (c_k - is_k) e^{ikx}. \quad (B.20)$$
Thus we can now write the exponential form of the Fourier series,

\[ f(x) = \sum_{k=-\infty}^{\infty} f_k e^{ikx}, \]  

(B.21)

where \( f_k \equiv \frac{1}{2}(c_k - i\sigma_k) \).

The rate of convergence over an interval \(-\pi \leq x \leq \pi\), irrespective of periodicity, depends on the smoothness of \( f(x) \), the number of derivatives that can be taken before a discontinuity is encountered. Beyond the smoothness of \( f(x) \) on its interval there may also be ‘terminal’ discontinuities when the function or its derivatives at one end of the interval don’t match those at the other. In that a function with period \( 2\pi \) will never have such discontinuities it is especially suited for approximation by a Fourier series.

We can see that this series is made up of odd(sine) and even(cosine) parts. For an arbitrary \( f(x) \) the terminal discontinuity will be in \( f(x) \) itself for the sine series, while for the cosine series it will be in the first derivative. Thus in general the cosine series will converge more readily.

It will be noted that a function which has fixed boundaries with differing conditions imposed on each one, as is the case of the horizontal modes in our convection problems, is unsuited for Fourier representation.

### B.2.1 The Fourier Transforms

In order to do Fourier transforms computationally, we need a discrete form of Equation (B.21) above. The usual form consistent with our previous discussion is\(^5\)

\[ f(x_n) = \sum_{k=-N/2}^{N/2-1} f_k e^{ikx_n}, \]

(B.22)

\(^5\)This is given by [106] and [87].
where \( x_n = 2\pi n/N \) and \( n = 0, 1, \ldots, N - 1 \).

The normalized inverse transform is then\(^6\)

\[
f_k = \frac{1}{N} \sum_{n=0}^{N-1} f(x_n) e^{-ikx_n},
\]

\[\text{where } k = -N/2, \ldots, 0, \ldots, N/2 - 1.\]

The normalization is derived using the orthogonality relations:\(^7\)

\[
\frac{1}{N} \sum_{k=0}^{N-1} e^{i2\pi kn/N} e^{-i2\pi km/N} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases}
\]

\[\text{(B.24)}\]

Now the FFTs used in our code implement the transforms

\[
f_k = \frac{1}{N} \sum_{n=0}^{N-1} f_n e^{-i2\pi kn/N} \quad \text{and} \quad f_n = \sum_{k=0}^{N-1} f_k e^{i2\pi kn/N}.
\]

\[\text{(B.25)}\]

We must examine how changing the bounds on the summation in Equation (B.22) affects the value of the sum. We can do this by rewriting this equation as

\[
f_n = \sum_{k=0}^{N/2-1} f_k e^{i2\pi kn/N} + \sum_{k=-N/2}^{-1} f_k e^{i2\pi kn/N},
\]

\[\text{(B.26)}\]

where having replaced \( x_n \), we have changed notation from \( f(x_n) \) to \( f_n \).

Next, letting \( k \rightarrow k' - N \) in the second term on the right, we have

\[
\sum_{k=-N/2}^{-1} f_k e^{i2\pi kn/N} = \sum_{k'=-N/2}^{N-1} f_{k' - N} e^{i2\pi (k' + N)n/N}.
\]

\[\text{(B.27)}\]

In the case where \( f_k \) is periodic with period \( N \),\(^8\) noting that \( e^{\pm i2\pi} = 1 \) gives the desired result,

\[
f_n = \sum_{k=0}^{N-1} f_k e^{i2\pi kn/N} = \sum_{k=-N/2}^{N/2-1} f_k e^{i2\pi kn/N}.
\]

\[\text{(B.28)}\]

\(^6\)Note that in some cases the choice of which transform carries the normalization is reversed, see for instance Brigham [107].

\(^7\)Ibid, equ.6-19, p99.

\(^8\)We will look at the non-periodic case in Appendix B.3.4.
In the case where \( f_k \), or \( f_n \), is not periodic we cannot simply re-arrange the summation order in this way. Instead we must perform an overall phase shift by letting the index \( k \to k' = k - N/2 \):

\[
f_n = \sum_{k=0}^{N-1} f_k e^{i2\pi kn/N} = \sum_{k=-N/2}^{N/2-1} f_{k'} e^{i2\pi (k'+N/2)n/N} = \sum_{k=-N/2}^{N/2-1} f_{k'} e^{-i\pi n} e^{i2\pi k'n/N}.
\]

Finally noting that \( e^{\pm i\pi n} = (-)^n \), we have the relation

\[
(-)^n f_n = \sum_{k=-N/2}^{N/2-1} f_{k'} e^{i2\pi k'n/N}
\]  \((B.29)\)

This is of course equally true for a calculation of \( f_k \) from \( f_n \).

### B.3 Chebychev Polynomials

In Appendix B.1 above we discussed approximation theory but did not go so far as to derive specific solutions under the various criteria presented. In each case application of the criteria leads to a set of polynomials that may be expressed in either an algebraic or trigonometric form. These are the Chebychev Polynomials. Our interest in them stems from a combination of analytical and computational properties that make them an excellent basis for a series representation of a non-periodic function \( f(x) \) defined on a discrete grid.

In their trigonometric form these Chebychev polynomials are cosine polynomials

\[
T_n(x) = \cos(n\theta) \quad \cos(\theta) = x \quad -1 \leq x \leq 1
\]  \((B.31)\)

Their algebraic form can be found via the transform \( \cos(\theta) = x \) or from recursion relations given in Appendix B.3.5 below.
B.3.1 As Solutions that Minimize Error

Returning now to approximation theory, we find that the solution to the minimax criterion for a function $f(x) = 0$ is given at $n^{th}$ order by $2^{1-n}T_n(x)$ which has $n$ zeros (and thus matches exactly at $n$ points) located at $x_k = \cos\left(\frac{2k+1}{n} \pi\right)$ for $k = 0, \ldots, n - 1$.

For all $n^{th}$ degree polynomials with a leading coefficient of unity this one has the smallest maximum deviation from zero in the interval $(-1, 1)$. Applying this to the Lagrangian interpolation formulation for a general $f(x)$ leads to a solution for $p_n(x)$ with

$$\Pi(x) = 2^{-n}T_{n+1}(x),$$

which matches $f(x)$ at $n + 1$ points $x_k = \cos\left(\frac{2k+1}{n+1} \pi\right)$, $k = 0, \ldots, n$.

The location of the matching points is significant as in this case if

$$M_{n+1} \equiv \max|f^{(n+1)}(\xi)|$$

is finite for all $n$ then the error $e_n \to 0$ as $n \to \infty$. This is not true for all choices of points.

We should mention here that a least-squares minimization of $\Pi(x)$ also leads to the Chebychev polynomials.

Turning now to our discussion of Least-Squares Theory, we recall that we found a class of solutions called ultra spherical polynomials which include the Legendre and Taylor series as well as the Chebychev series. These are polynomial series found by using a weight function $w(x) = (1 - x^2)^\alpha$ in either the recurrence relation (B.11) or the differential generating function (B.14).

\footnote{See Appendix B.1.2.}
The Legendre series comes from choosing $\alpha = 0, \ w(x) = 1$, while the Chebychev result when $\alpha = -\frac{1}{2}, \ w(x) = (1 - x^2)^{-\frac{1}{2}}$.

For the Chebychev case we have a normalization

$$l_k = \int_0^\pi \cos^2(k\theta)d\theta = \begin{cases} \pi & \text{for } k = 0 \\ \frac{\pi}{2} & \text{for } k \neq 0 \end{cases} \quad (B.32)$$

and for the coefficients of the $\phi_k(x)$,

$$c_k = \frac{2}{\pi} \int_{-1}^1 (1 - x^2)^{-\frac{1}{2}} T_k(x)f(x)dx, \quad (B.33)$$

yielding a Chebychev least-squares polynomial for $f(x)$\(^{10}\)

$$p_n = \sum_{k=0}^{n} c_k T_k(x) \quad (B.34)$$

As noted in Appendix B.1.2 the convergence of the ultra-spherical series is inversely related to the magnitude of the coefficient $A_k$ of the leading term $x^k$. More precisely one can derive a relationship for large $k$\(^{11}\)

$$c_k \sim \frac{f^{(k)}(0)}{k!A_k}, \quad (B.35)$$

showing that the increase of $A_k$ with increasing $k$ is directly proportional to the decrease of $c_k$ and thus the convergence rate. For the series we have mentioned (plus an estimate for the Fourier series),

- Legendre $\ c_k \sim k! \frac{\sqrt{\pi k}}{2^k}$
- Chebychev $\ c_k \sim k! \frac{1}{2^{k-1}}$
- Taylor $\ c_k \sim k!$
- Fourier $\ c_k \leq k^{-3}$

\(^{10}\)We will often write ' or '' with a summation sign

$$\sum_{k=0}^{n}, \quad \text{or} \quad \sum_{k=0}^{n}''$$

where ' indicates a factor of $\frac{1}{2}$ in front of the $k = 0$ term of the summation and '' a factor of $\frac{1}{2}$ in front of both the $k = 0$ and $k = n$ terms of the summation.

\(^{11}\)This discussion is predicated on a general function $f(x)$ which is not specially conditioned in a way that favors any particular representation.
Of these the Chebychev value is the smallest possible for an ultra-spherical polynomial, while that for the Taylor series is largest. Note that the values for the Legendre and Fourier series are also much worse than for the Chebychev. One can in fact recast the statement about minimizing the maximum error in the range \((-1, 1)\) above to say that of all polynomials of degree \(n\), normalized to have a maximum value of unity in the interval \((-1, 1)\), the Chebychev polynomials have the largest possible coefficient of \(x^k\) and thus the best convergence characteristics.

**B.3.2 Chebychev Series Representation**

We can now proceed to derive the Chebychev series in one of several ways. Most trivially we can let \(n \rightarrow \infty\) in the Chebychev polynomial of Equation (B.34) above, thus yielding the Chebychev series

\[
f(x) = \lim_{n \rightarrow \infty} p_n = \sum_{k=0}^{\infty} c_k T_k(x)
\]

On the other hand, it is useful to derive it from the Fourier series, to illustrate the close connection between the two, and to clarify as well its superior convergence properties, as its use allows us to avoid the terminal discontinuities for non-periodic functions that were discussed in Appendix B.2 above. The Chebychev series is in fact a modified form of a Fourier series designed to handle non-periodic functions.

Take the function \(f(x)\) in the range \(-1 \leq x \leq 1\) and make a change of variable \(x = \cos(\theta)\):

\[
f(x) = f(\cos(\theta)) = g(\theta), \quad \text{for } 0 \leq \theta \leq \pi.
\]

This new function \(g(\theta)\) is both even and has period \(2\pi\), properties inherent in the cosine function. Furthermore as long as \(f(x)\) has smooth derivatives in
\(-1 \leq x \leq 1\), \(g(\theta)\) will likewise in \(0 \leq x \leq \pi\). Thus \(g(\theta)\) is well suited for representation by a Fourier cosine series.

\[
g(\theta) = \frac{1}{2} c_0 + \sum_{k=1}^{n} c_k \cos(k\theta) \quad c_k = \frac{2}{\pi} \int_{0}^{\pi} g(\theta) \cos(k\theta) d\theta. \tag{B.38}
\]

Transforming back to \(x\) we have

\[
f(x) = \frac{1}{2} c_0 + \sum_{k=1}^{n} c_k T_k(x) \quad c_k = \frac{2}{\pi} \int_{-1}^{1} (1 - x^2)^{-\frac{1}{2}} f(x) T_k(x) dx. \tag{B.39}
\]

We now have a series without any terminal discontinuities.

### B.3.3 The Discrete Representation

For numerical work we have discrete functions \(f(x_k)\) rather than continuous functions \(f(x)\) and must therefore derive a discrete representation. Theoretically most of the development of Appendix B.1 can be redone for the discrete case, but here we want to work specifically to derive a Chebychev series. This is easiest if we work with the function \(g(\theta)\) on the interval \((0, \pi)\). Start with the trigonometric identity

\[
\frac{1}{2} + \cos(\theta) + \cdots + \cos((N - 1)\theta) + \frac{1}{2} \cos(N\theta) = \frac{1}{2} \sin(N\theta) \cot(\frac{1}{2}\theta) \tag{B.40}
\]

which vanishes at \(\theta_k = k\pi/N\) for all integer values of \(k\). Combining this with the identity

\[
\cos(\theta) \cos(n\theta) = \frac{1}{2} \left[ \cos((n + 1)\theta) + \cos((n - 1)\theta), \right] \tag{B.41}
\]

one can show that the functions \(\phi_r(\theta) = \cos(r\theta)\) are orthogonal under \(^n\) summation with a normalization given by

\[
\sum_{k=0}^{N} \phi_r^2(\theta_k) = \left\{ \begin{array}{ll}
\frac{1}{2} N & \text{for } r \neq 0 \\
\frac{N}{2} & \text{for } r = 0, N.
\end{array} \right. \tag{B.42}
\]
Using these functions to form a least-squares fit to \( g(\theta) \) and then transforming with \( x = \cos(\theta) \), we have a Chebychev discrete least-squares fit:

\[
p_n(x) = \sum_{r=0}^{n} c_r T_r(x)
\]

\[
c_r = 2 \frac{N}{N} \sum_{k=0}^{N} f(x_k) T_r(x_k), \quad x_k = \cos\left(\frac{k\pi}{N}\right)
\]

(B.43)

where \( p_n(x) \) is the approximation to \( f(x) \) caused by using \( N < \infty \), and \( x_k \) represents our grid.

It is important to note that in the case that \( n = N \) we need to replace \( ' \) with \( '' \) in the sum over the \( T_r \)'s and that this gives an exact fit at the points of definition.

\[
p_N(x) = \sum_{r=0}^{N} c_r T_r(x)
\]

\[
c_r = 2 \frac{N}{N} \sum_{k=0}^{N} f(x_k) T_r(x_k), \quad x_k = \cos\left(\frac{k\pi}{N}\right)
\]

(B.44)

In summary we see that there are several features of the Chebychev polynomials that make them attractive for numerical work:

- Using the cosine transformation for our grid provides the best points for matching the function \( f(x) \) over the interval \((-1,1)\). It also gives us increased accuracy near the boundaries.

- These polynomials give us rapid convergence allowing better results with fewer modes.

- Compared to the Fourier series they eliminate the problem of terminal discontinuities and are therefore especially suited for non-periodic functions, i.e., physically bounded dimensions.
Finally, they can be expressed as a cosine series, and thus we may make use of existing FFT's to transform into mode space by forcing the FFT to mimic a cosine transformation.

**B.3.4 The Chebychev Transforms**

The basic Chebychev transform follows directly from the discrete representation given above:

\[
  f_n = \sum_{k=0}^{N} T_k(x_n) \\
  f_k = \frac{2}{N} \sum_{n=0}^{N} f_n T_k(x_n), \quad x_n = \cos\left(\frac{n\pi}{N}\right). \tag{B.45}
\]

Making the substitution for \( x_n \) we can write these as

\[
  f_n = \sum_{k=0}^{N} T_k(\pi kn/N) \\
  f_k = \frac{2}{N} \sum_{n=0}^{N} f_n \cos(\pi kn/N). \tag{B.46}
\]

In order to implement these transforms computationally, we take advantage of the fact that they are essentially cosine transformations and that by extending our data we can use the FFTs to perform them. First starting with a function \( f(x_n) \) with \( x_n = \cos(2\pi n/N) \), we extend the data symmetrically about \( N \), such that \( f_{2N-n} = f_n \). We can now write down an inverse Fourier transformation with period \( 2N \),\(^{12}\)

\[
  f_k = \frac{1}{2N} \sum_{n=0}^{2N-1} f_n e^{-i\pi kn/2N} \tag{B.47}
\]

\(^{12}\)See Equation (B.25).
Now splitting the sum in two and letting the index on the second change, $n \to m = 2N - n$,

$$f_k = \frac{1}{2N} \sum_{n=0}^{N-1} f_n e^{-i\pi kn/N} + \frac{1}{2N} \sum_{m=1}^{N} f_{2N-m} e^{-i\pi k(2N-m)/N}. \quad (B.48)$$

Using the even parity of our extension, $f_{2N-m} = f_m$, and the fact that $\exp(i2\pi n) = 1$ we have

$$f_k = \frac{1}{2N} \sum_{n=0}^{N-1} f_n e^{-i\pi kn/N} + \frac{1}{2N} \sum_{m=1}^{N} f_m e^{i\pi km/N}. \quad (B.49)$$

Now we extend both summations to the range $[0,N]$ but in doing so we have to multiply the $f_0$ and $f_N$ terms by $\frac{1}{2}$, yielding

$$f_k = \frac{1}{2N} \sum_{n=0}^{N} \quad f_n \left\{ e^{i\pi kn/N} + e^{-i\pi kn/N} \right\} \quad (B.50)$$

Finally combining the exponentials we have\footnote{We can also simply shift the index on the second sum, letting $n = m - 1$ and, making use of the periodicity of the original $f_n = f_{n+N}$ to equate $f_0$ to $f_N$, we have a standard cosine transform:

$$f_k = \frac{1}{N} \sum_{n=0}^{N-1} \quad f_n \cos(\pi kn/N). \quad (B.51)$$

This is to within a factor of 2 the desired Chebychev transform. Thus we can use the FFT, with a post factor of 2, to implement the forward transform (the second of the pair (B.46)) while the backward transformation is to be had by using the FFT with a pre-factor of $\frac{1}{2}$ to compensate for the extra factor of 2 introduced by using the inverse Fourier transform in place of the inverse cosine transform.}
There is, however, one last complication. Our functions in real space are defined over the interval [-1,1] rather than [0,2] and thus the summation in the forward transform should be from -N/2 to N/2-1 rather than 0 to N. In Appendix B.2.1 we looked at this problem, but only for the case where the function to be transformed is periodic. Here our functions in x are non-periodic, one of the main reasons for using the Chebychev decompositions. This results in an extra factor of \((-1)^k\) as a post-factor on the forward transform and as a pre-factor on the inverse transform.

### B.3.5 Chebychev Properties

We present here various analytical properties of the Chebychev polynomials that have been of use in our work.

Making use of the trigonometric identity (B.41) from above we can easily derive a recursion relation

\[ T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad (B.53) \]

which along with \(T_0(x) = 1\) and \(T_1(x) = x\) can be used to find the explicit form of the polynomials as functions of \(x\). The Chebychev polynomials are alternately even and odd functions of \(x\) (as \(n\) is even or odd), and at the ends of their range,

\[
T_n(1) = \begin{cases} 
+1 & \text{for even } n \\
-1 & \text{for odd } n.
\end{cases} \quad (B.54)
\]

\[
T_n(-1) = \begin{cases} 
+1 & \text{for even } n \\
-1 & \text{for odd } n.
\end{cases} \quad (B.55)
\]

\[14\]This relation holds only for \(n \geq 1\). For \(n = 0\) we have \(T_1 = 2xT_2\).
For a product of Chebychev polynomials we have

\[ T_n(x)T_r(x) = \frac{1}{2} [T_{n+r}(x) + T_{n-r}(x)]. \]  \hspace{1cm} (B.56)

This follows from the identity giving the product of two cosines and is related to the result of the example given in Appendix B.5.

To express a power of \( x \) we have

\[ x^r = (T_1)^r = \frac{1}{2^{r-1}} \left\{ T_r(x) \left[ \binom{r}{i} T_{r-i}(x) + \frac{r!}{(r-i)!i!} \right] + \cdots \right\}, \]  \hspace{1cm} (B.57)

where \( \binom{r}{i} = \frac{r!}{(r-i)!i!} \).

This expression can also be derived from \( x^r = \cos^r(\theta) \), or

\[ x^r = \left( \frac{e^{i\theta} + e^{-i\theta}}{2} \right)^r. \]  \hspace{1cm} (B.58)

Using Equation (B.57) we can now work out the result of multiplying a Chebychev polynomial by a power of \( x \):

\[ x^rT_n(x) = \frac{1}{2^{r-1}} \left\{ T_r(x)T_n(x) \left[ \binom{r}{i} T_{r-i}(x)T_n(x) + \frac{r!}{(r-i)!i!} \right] + \cdots \right\} \]  \hspace{1cm} (B.59)

\[ = \frac{1}{2^r} \sum_{i=0}^{r} \binom{r}{i} T_{r-i+2i}(x), \]

where we have used the identity

\[ \binom{r}{i} = \binom{r}{r-i}. \]  \hspace{1cm} (B.60)

To calculate the derivatives of the Chebychev polynomials start with Definition (B.31),

\[ T_n = \cos(n\theta) \quad x = \cos(\theta). \]  \hspace{1cm} (B.61)
Making use of the relation

\[ dx = -\sin(\theta)d\theta, \]  \hspace{1cm} (B.62)

we have

\[ \frac{d}{dx} T_n = \frac{n \sin(n\theta)}{\sin(\theta)}. \]  \hspace{1cm} (B.63)

Multiplying by \( \frac{\sin(\theta)}{\sin(\theta)} \) and converting to cosines we get

\[ \frac{d}{dx} T_n = \frac{n}{2} \left[ \frac{\cos((n-1)\theta) - \cos((n+1)\theta)}{\sin^3(\theta)} \right]. \]  \hspace{1cm} (B.64)

Adding and subtracting \( \frac{1}{2} \cos((n-1)\theta) \) and making use of the cosine Identity (B.41),

\[ \frac{d}{dx} T_n = \frac{n}{\sin^3(\theta)} \left[ \cos((n-1)\theta) - \cos(\theta) \cos(n\theta) \right]. \]  \hspace{1cm} (B.65)

Finally making use of Definitions (B.61) to write this expression in terms of Chebychev polynomials, we have

\[ \frac{d}{dx} T_n = \frac{n(T_{n-1} - xT_n)}{(1 - x^2)}. \]  \hspace{1cm} (B.66)

Likewise for the second derivative,

\[ \frac{d^2}{dx^2} T_n = \frac{d}{dx} \left[ \frac{n \sin(n\theta)}{\sin(\theta)} \right] \frac{d\theta}{dx}, \]  \hspace{1cm} (B.67)

we have

\[ \frac{d^2}{dx^2} T_n = \frac{-1}{\sin(\theta)} \left[ \frac{n^2 \cos(n\theta)}{\sin(\theta)} - \frac{n \sin(n\theta) \cos(\theta)}{\sin^2(\theta)} \right]. \]  \hspace{1cm} (B.68)

Using Equation (B.63) and the Definitions (B.61) again,

\[ \frac{d^2}{dx^2} T_n = \frac{(xT'_n - n^2T_n)}{(1 - x^2)}. \]  \hspace{1cm} (B.69)

The calculation of higher derivatives by this method is increasingly cumbersome; we will stop here as this is as much as we need for our work.
To calculate the integral of a Chebychev polynomial we make use of the Definitions (B.61) and Relation (B.62) to write

$$\int T_r(x)dx = -\int \cos(r\theta)\sin(\theta)d\theta. \quad (B.70)$$

Using a trigonometric identity

$$\int T_r(x)dx = -\frac{1}{2} \int \{\sin((r+1)\theta) - \sin((r-1)\theta)\} d\theta \quad (B.71)$$

and finally carrying out the integration and writing the result in terms of the $T_n$s, we have

$$\int T_r(x)dx = \frac{1}{2} \left[ \frac{1}{r+1}T_{r+1}(x) - \frac{1}{r-1}T_{r-1}(x) \right]. \quad (B.72)$$

The special cases for $r = 0$ and $r = 1$ yield

$$\int T_0(x) = T_1(x), \quad \int T_1(x) = \frac{1}{4}T_2(x). \quad (B.73)$$

### B.4 Convolutions

In the non-linear code we work mostly in mode space. The major exception is the calculation of the Jacobian terms which require real-space multiplications. As mentioned, above the main usage of the FFTs comes from transforming the fields to real space for these multiplications and then transforming the result back to mode space every time step. This may seem to be a lot of work, but whereas the other operations (calculating powers of $x$, or $x$ and $z$ derivatives) can be handled most conveniently in mode space, the equivalent of a real-space multiplication would be a convolution in mode space. It is computationally advantageous to avoid calculating a convolution which requires essentially a fourfold summation in 2-D. In fact, one of the main uses of FFTs
is to enable mode-space calculation of real space convolutions encountered in
data-analysis problems.\(^{16}\)

In order, however, to test the accuracy of our non-linear terms, and
especially to try out various de-aliasing methods,\(^{16}\) we need to be able to ac-
tually calculate the mode-space convolutions to use as a reference. We will
also need to sort out some normalization issues involved in making use of an
alternate form of the Discrete Convolution Theorem.\(^{17}\)

Given two periodic mode space functions: \(G(\frac{n}{NT})\) and \(H(\frac{n}{NT})\),
where \(N\) is the number of data points in a period, \(T\) is the sampling interval,
and \(n = 0, N - 1\), one writes the discrete convolution of \(G\) and \(H\):

\[
y(kT) = \sum_{n=0}^{N-1} G(nT)H((k - n)T). \quad (B.74)
\]

In our case we start with two mode-space functions \(G\) and \(H\) and do
inverse Fourier transforms\(^{18}\) to get two real-space functions \(g\) and \(h\):

\[
g(kT) = \sum_{n=0}^{N-1} G \left( \frac{n}{NT} \right) \exp^{i2\pi nk/N} \quad (B.75)
\]

\[
h(kT) = \sum_{m=0}^{N-1} H \left( \frac{m}{NT} \right) \exp^{i2\pi mk/N}. \quad (B.76)
\]

Now multiplying these we get

\[
g(kT)h(kT) = \sum_{n=0}^{N-1} G \left( \frac{n}{NT} \right) \sum_{m=0}^{N-1} H \left( \frac{m}{NT} \right) \exp^{i2\pi (n+m)k/N}. \quad (B.77)
\]

\(^{16}\)For a good general reference on FFTs, convolutions, etc. see [107].

\(^{16}\)See Appendices B.5 and C.4.4 for discussion of aliasing and de-aliasing.

\(^{17}\)Since we are working computationally this discussion deals with discrete functions, transforms etc. We will also work in one dimension to make the notation simpler. The extension
to two dimensions is straightforward.

\(^{18}\)For the various properties of Fourier series and transforms see Appendix B.2.
Performing an inverse transform

\[ Y \left( \frac{l}{NT} \right) = \frac{1}{N} \sum_{k=0}^{N-1} g(kT)h(kT) \exp^{-i2\pi k/l} \]  \hspace{1cm} \text{(B.78)}

\[ Y \left( \frac{l}{NT} \right) = \frac{1}{N} \sum_{k=0}^{N-1} \sum_{n=0}^{N-1} G \left( \frac{n}{NT} \right) \sum_{m=0}^{N-1} H \left( \frac{m}{NT} \right) \exp^{i2\pi (n+m-l)/N} . \]  \hspace{1cm} \text{(B.79)}

Rearranging, we have

\[ Y \left( \frac{l}{NT} \right) = \sum_{n=0}^{N-1} G \left( \frac{n}{NT} \right) \sum_{m=0}^{N-1} H \left( \frac{m}{NT} \right) \frac{1}{N} \sum_{k=0}^{N-1} \exp^{i2\pi (n+m-l)/N} . \]  \hspace{1cm} \text{(B.80)}

Next we make use of the orthogonality relationship for the exponentials, Equation (B.24),

\[ \frac{1}{N} \sum_{k=0}^{N-1} \exp^{i2\pi nk/N} \exp^{-i2\pi mh/N} = \begin{cases} 1 & \text{if } n = m \\ 0 & \text{otherwise} \end{cases} \]  \hspace{1cm} \text{(B.81)}

to reduce Equation (B.24) to

\[ Y \left( \frac{l}{NT} \right) = \sum_{n=0}^{N-1} G \left( \frac{n}{NT} \right) H \left( \frac{l-n}{NT} \right) . \]  \hspace{1cm} \text{(B.82)}

Comparing with Equation (B.74) above, we see this is indeed the mode-space convolution.

We can now write the Convolution Theorem:

\[ g(kt)h(kt) \iff G \left( \frac{n}{NT} \right) * H \left( \frac{n}{NT} \right) \]  \hspace{1cm} \text{(B.83)}

where * is the symbol for a convolution operation.

This theorem can be easily extended to more than one dimension.

We note that since our transforms are actually a mixture of Fourier with Chebychev transforms implemented by the forced conversion of the Fourier transform into a modified Cosine transform there is an extra factor of \( \frac{1}{2} \) that must be inserted as well. This is owing to the factor of 2 difference in the
magnitude of the Chebychev and Cosine transformations. As a result the straightforward calculation of the convolution in mode space gives twice the field obtained by transforming both fields to physical space, multiplying, and transforming the product back again.

Having worked all this out we can test the accuracy of our real space calculation. When the test fields $G$ and $H$ consist of single modes (one non-zero value in each field) and the mode numbers are small enough that there is no aliasing, then the error is on the order of $1e-16$ compared to a convolution amplitude of $1e-2$, which is essentially the same as the FFT error we saw above.

## B.5 Aliasing

Aliasing is an error which occurs whenever a continuous function is discretized with a sampling frequency that is less than twice the highest frequency component of the sampled function.

For a given sampling interval $\Delta$ there is a corresponding frequency called the Nyquist critical frequency, given by

$$f_c \equiv \frac{1}{2\Delta}.$$  \hspace{1cm} (B.84)

This is a cutoff frequency. Any frequency components in the sampled function cannot be properly resolved if they are greater than the Nyquist critical frequency. These frequency components, however, do not simply drop out of the resulting discretized function, as they would for a low-pass filter with a cutoff at $f_c$; they are folded back onto the frequencies below the Nyquist frequency.

---

19 See [87] chp.12, and Appendix B.3.4.
20 See Appendix B.5.
Figure B.1: Illustration of Aliasing

This is what we call aliasing. Thus when the sampled function has an amplitude in frequency space $A_\delta$ at the frequency $f_c + \delta f$, we find after sampling that the amplitude at frequency $f_c - \delta f$ is greater by $A_\delta$ than it ought to be. This is illustrated in Figure B.5.

Since any time we make use of discrete FFTs in numerical simulation we are performing this type of discretation, we must examine possible aliasing errors. In most instances where we have a function defined on a grid and we do a transformation with an FFT, there is no problem so long as we use as many samples as we have grid points. In this case the original function is limited in frequency by the discreteness of the grid, and the sampling for the FFT does not introduce new error.

In the case of our nonlinear terms, however, we have to be more careful. These involve the transformation and multiplication of two functions before inversely transforming the result back to mode-space. Although each function is separately frequency limited (they have in fact been transformed
to real-space from mode-space representations and are therefore always appropriately limited), when we multiply them we couple the frequencies producing the sums and differences of the original components. This process can generate frequency components up to twice the critical frequency which are then folded back into the lower frequency components when the result of the multiplication is transformed back into mode-space.

To illustrate this very simply we take two delta functions in mode space:

\[ G\left(\frac{n}{kT}\right) = \delta(k - p) + \delta(k + p) \quad \text{and} \quad H\left(\frac{n}{kT}\right) = \delta(k - q) + \delta(k + q). \]

Transforming these to physical space (see Equations (B.75) and (B.76)) we have

\[ g(kT) = \left(\exp^{i2\pi pk/N} + \exp^{-i2\pi pk/N}\right) = \cos(2\pi pk/N) \quad \text{(B.85)} \]

\[ h(kT) = \left(\exp^{i2\pi qk/N} + \exp^{-i2\pi qk/N}\right) = \cos(2\pi qk/N). \quad \text{(B.86)} \]

Multiplying these and forward transforming, yields

\[ g(kT)h(kT) = \cos(2\pi pk/N)\cos(2\pi qk/N) \quad \text{(B.87)} \]

\[ y\left(\frac{l}{NT}\right) = \frac{1}{N} \sum_{k=0}^{N-1} \cos(2\pi pk/N)\cos(2\pi qk/N)\exp^{i2\pi lk/N}, \quad \text{(B.88)} \]

which, after converting the cosines back to exponentials, rearranging, and making use of Relation (B.81), we have the result

\[ y\left(\frac{l}{NT}\right) = \{\delta(k - p - q) + \delta(k - p + q)\}. \quad \text{(B.89)} \]

---

21 Recall that the reality of the physical space functions leads to mode-space parities equivalent to making ± modes complex conjugates.
We can see here that as long as \( p \) and \( q \) are small there is no problem, but if they are large enough that their sum is greater than \( N \) then there will be aliasing.
Appendix C

Computational Appendices

In these appendices we outline the details of the numerical treatment of the linear and non-linear problems. We also present some results of tests on the various components, the FFT's, the matrix operators, and the non-linear terms, so as to gauge the accuracy of our codes.

C.1 The Linear Eigenvalue Problem

We want to take the linearized VSC equations derived in Section 4 and formulate them as a generalized eigenvalue problem of the form

\[ A\phi = \lambda B\phi \]  \hspace{1cm} (C.1)

where \( \phi \) is a vector containing the Chebychev components of \( \omega, \psi, \) and \( T, A \) is a 3x3 block differential matrix operator, \( B \) is a conditioning matrix that allows us to include the various constraints on the main pair of equations, and \( \lambda \) is an eigen value. These are all complex quantities, with \( A \) and \( B \) being given and \( \lambda \) and \( \phi \) being solved for.

More specifically we have, for \( \phi \),

\[ \phi = (\omega_0, \ldots, \omega_{n-2}, 0, 0, \psi_0, \ldots, \psi_{n-2}, 0, 0, T_0, \ldots, T_{n-2}, 0, 0) \]

where the zero's replace the \( N-1 \), and \( N \)th modes in order to make room for the boundary conditions.
To derive $A$ we take the two main equations with the condition linking $\psi$ and $\omega$ and write them in terms of operators:

\[
\begin{align*}
A(x)\omega + B(x)\psi + C(x)T &= \lambda\omega \\
+ B'(x)\psi + D(x)T &= \lambda T \\
E(x)\omega + F(x)\psi &= 0.
\end{align*}
\]

We can now formulate the problem fully, showing the block nature of $A$,

\[
\begin{bmatrix}
F(x) & E(x) \\
B(x) & A(x) & C(x) \\
B'(x) & D(x)
\end{bmatrix}
\begin{pmatrix}
\psi \\
\omega \\
T
\end{pmatrix}
= \lambda
\begin{bmatrix}
\psi \\
\omega \\
T
\end{bmatrix}
\tag{C.2}
\]

where the operators are, from Equations (4.9), (4.7), and (4.8),

\[
\begin{align*}
A(x) &= \frac{1}{G}D^2_x - \frac{1}{G}\alpha^2 + \frac{i}{6}\alpha X^3 - \frac{i}{6}\alpha X \\
B(x) &= -i\alpha X \\
C(x) &= -\frac{1}{G}D_x \\
D(x) &= \frac{1}{PrG}D^2_x - \frac{1}{PrG}\alpha^2 + \frac{i}{6}\alpha X^3 - \frac{i}{6}\alpha X \\
B'(x) &= -i\alpha \\
E(x) &= -1 \\
F(x) &= \Delta_2^2 - \alpha^2.
\end{align*}
\]

$X$, $X^3$, $D_x$, and $D^2_x$ are matrix operators,\(^1\) and $\alpha$ is a diagonal matrix with all non-zero values equal to the wave number in the $z$ direction, resulting from taking derivatives of the Fourier expansion in $z$.

Using what is known as the Tau method,\(^2\) the last two rows of the blocks $F(x)$, $B(x)$, and $D(x)$ are replaced by the boundary conditions on $\psi$, $\omega$, and $T$ respectively. Similarly the last two rows of blocks $E(x)$, $A(x)$, $C(x)$, and $B'(x)$ will be filled with zeros.

\(^1\)See Appendices C.2.1 and C.2.2 below.

\(^2\)See for instance [108].
Finally the matrix $B$ is unit diagonal with zeros replacing ones in positions corresponding to the boundary conditions and to the conditioning equation. Thus diagonal components are

$$B_{ii} = (0, 0, \ldots, 0, 0; 1, 1, \ldots, 1, 0, 0; 1, 1, \ldots, 1, 1, 0, 0)$$

and all the rest are zero.

This selectively enforces zeros on the right-hand side of our equations, so that we may include the boundary conditions and the constraint relating $\psi$ and $\omega$ as part of the eigenvalue problem.

The result of setting this problem up and feeding it through a complex generalized eigen solver is a list of eigenvalues and (if so desired) their corresponding eigen functions. We must be careful to expect that for each of the zeros in $B$ we will have a spurious infinite eigenvalue. Also we will get as many solutions as there are Chebychev modes in our decomposition, as each mode gives its own contributions to the set of linear equations. In general we will only be interested in the largest few values, which can possibly be considered physical.

C.1.1 The Two-Variable Formulation

We can recast this problem in just two variables by eliminating the vorticity, leaving a fourth-order equation for $\psi$ and the same second-order equation for $T$. The number and nature of the boundary conditions would remain unchanged.

The matrix equation now takes the form

$$\begin{bmatrix} A(x) & C(x) \\ B'(x) & D(x) \end{bmatrix} \begin{pmatrix} \psi \\ T \end{pmatrix} = \lambda \begin{bmatrix} F(x) \\ I \end{bmatrix} \begin{pmatrix} \psi \\ T \end{pmatrix}; \quad (C.3)$$
with the operators defined as

\[
A(x) = \frac{1}{G} (D_x^2 - \alpha^2v^2 + \frac{i}{6}\alpha(X^3 - X)(D_x^2 - \alpha^2) - i\alpha X
\]

\[
C(x) = -\frac{1}{G} D_x
\]

\[
D(x) = \frac{1}{PrG} D_x^2 - \frac{1}{PrG} \alpha^2 + \frac{i}{6}\alpha X^3 - \frac{i}{6}\alpha X
\]

\[
B'(x) = -i\alpha
\]

\[
F(x) = D_x^2 - \alpha^2
\]

\[I = \text{the identity.}\]

The primary interest in such a reformulation is that we can reduce by a factor of \(\frac{3}{4}\) the number of matrix elements involved and thus gain roughly a factor of two in computation time.

But the transformation

\[
\lambda \omega \to \lambda \nabla^2 \psi
\]

requires a more complicated form for \(B\) with the use of operator group \(F(x)\), and the numerical stability of the problem may be greatly reduced. In the future we plan to investigate the practicality of this approach.

### C.2 The Operators

In both the linear and non-linear problems we are interested in doing all linear calculations in mode space. This requires operator equivalents for two types of operations: [1] Multiplication by known analytic functions (e.g. the background fields in the linear problem) and [2] Differentiation. By operators

---

3Rubel and Landis[46] used a two-variable formulation as the basis for a non-linear finite difference simulation of the VSC problem, but the spectral case is quite different.
we mean matrix operators that when applied to the matrix representations of the mode-space fields will yield new fields equal to those resulting from performing the same operation on the real-space fields and then converting to mode space.

C.2.1 The Algebraic Operators

Since our background fields are only functions of $x$ and are independent of $z$ we only need operator equivalents for powers of $X$. The decomposition in the $x$ direction is in Chebychev polynomials, and we therefore need to know how to represent the product of a power of $X$ and a Chebychev polynomial in terms of other Chebychev polynomials.

Starting from the formula for the product of a Chebychev polynomial and a power of $X$,\(^5\)

\[
X^r T_x(x) = \frac{1}{2^r} \sum_{i=0}^{r} \binom{r}{i} T_{x-r+2i}(x).
\]  
(C.4)

For the first three powers of $X$ we get specifically

\[
XT_n(x) = \frac{1}{2}(T_{n-1} + T_{n+1})
\]  
(C.5)

\[
X^2 T_n(x) = \frac{1}{4}(T_{n-2} + 2T_n + T_{n+2})
\]  
(C.6)

\[
X^3 T_n(x) = \frac{1}{8}(T_{n-3} + 3T_{n-1} + 3T_{n+1} + T_{n+3}).
\]  
(C.7)

Due to the special case for $n = 0$ such that

\[
XT_0 = \frac{1}{2} T_1 \neq \frac{1}{2} T_1,
\]

\(^4\)See Appendices C.1 and B.3 for details on the decomposition and the Chebychev polynomials, respectively.

\(^5\)See Appendix B.3.5.
we need to modify the recursion relation slightly by introducing the notation

\[ c_n = \begin{cases} 
2 & \text{if } n = 0 \\
1 & \text{otherwise} 
\end{cases} \]

to get

\[ X T_n(x) = \frac{1}{2} [T_{n-1} + c_n T_{n+1}], \quad (C.8) \]

where we are assuming implicitly that \( T_n = 0 \) for \( n < 0 \).

For the higher powers of \( X \) we can use this relation on itself to yield successively

\[ X^2 T_n(x) = \frac{1}{4} [T_{n-2} + (c_{n-1} + c_n) T_n + c_n T_{n+2}] \tag{C.9} \]

and

\[ X^3 T_n(x) = \frac{1}{8} [T_{n-3} + (c_{n-2} + c_n) T_{n-1} + (c_{n-1} + c_n^2 + c_n) T_{n+1} + c_n T_{n+3}] \tag{C.10} \]

To derive the explicit coefficient recursion relations needed to make the matrix operators for the linear code, we start from

\[ f(x) = \frac{1}{2} a_0 + \sum_{n=1}^{N-1} a_n T_n + \frac{1}{2} a_N T_N \equiv \sum_{n=0}^{N} "a_n T_n, \tag{C.11} \]

where as usual " signifies a factor of \( \frac{1}{2} \) for \( n = 0, N \).

Now, writing

\[ x f(x) = \frac{1}{2} b_0 + \sum_{n=1}^{N-1} b_n T_n + \frac{1}{2} b_N T_N \equiv \sum_{n=0}^{N} "b_n T_n, \tag{C.12} \]

equating

\[ \sum_{n=0}^{N} "b_n T_n = \sum_{n=0}^{N} "a_n x T_n, \tag{C.13} \]

and making use of Relation (C.8) on the right-hand side, we get

\[ \frac{1}{2} b_0 + \sum_{n=1}^{N-1} b_n T_n + \frac{1}{2} b_N T_N = \frac{1}{2} a_0 T_1 + \sum_{n=1}^{N-1} a_n \left( T_{n-1} + T_{n+1} \right) + \frac{a_N}{4} \left( T_{N-1} + T_{N+1} \right). \tag{C.14} \]
Equating coefficients of the $T_n$ we can write in a general form

$$b_n = \frac{1}{2}(d_na_{n-1} + \frac{c_n}{d_{n+1}}a_{n+1}) \quad (C.15)$$

where $c_n$ is defined as above and

$$d_n = \begin{cases} 
2 & \text{if } n = N \\
1 & \text{otherwise}
\end{cases}$$

The factors of $c_n$ and $d_n$ derive from the $b_0$ and $b_N$ terms and the factors of $\frac{1}{d_n}$ from the $a_N$ term.

By recursive substitution we can derive the relations for the coefficients of the series expansions of $X^2f(x)$ and $X^3f(x)$:

$$b_n = \frac{1}{4}(d_na_{n-2} + (c_{n-1} + c_n)a_n + \frac{c_n}{d_{n+2}}a_{n+2}) \quad (C.16)$$

and

$$b_n = \frac{1}{8}[d_na_{n-3} + d_n(c_{n-2} + c_{n-1} + c_n)a_{n-1} + \frac{c_{n-1} + c_n}{d_{n+1}}a_{n+1} + \frac{c_n}{d_{n+3}}a_{n+3}] \quad (C.17)$$

Note that for a representation

$$f(x) = \sum_{n=0}^{n} a_n x T_n,$$

we can set all the $d_n$s to unity and these equations are then exactly parallel to ones for the $T_n$s given above.

These relations define the components of the matrix operators used to implement multiplication by powers of $X$ in Chebychev mode space. Using these on the mode representation of a field $\psi(x)$ yields the mode representation of the product $x\psi(x)$ (see Appendix C.1).\(^6\)

---

\(^6\)Note that Gottlieb and Orzag [57] use an expansion

$$f(x) \equiv \sum_{n} a'_n T_n$$
C.2.2 The Derivative Operators

Since doing a derivative with respect to $z$ in Fourier space is a trivial matter of multiplying by the proper wave number, we again concentrate on deriving methods to calculate the $z$ derivatives in Chebychev space.

For a Chebychev representation of a function

$$f(x) = \sum_{n=0}^{N} ''a_n T_n$$

(C.18)

there are two possible ways to proceed. We can either keep the functions $T_n$ and calculate the new coefficients $b_n$ such that

$$f'(x) = \sum_{n=0}^{N-1} 'b_n T_n,$$

(C.19)

or we can differentiate the Chebychev polynomials themselves while keeping the coefficients the same:

$$f'(x) = \sum_{n=0}^{N} ''a_n T_n'$$

(C.20)

The first of these methods results in recursion relations giving the new coefficients $b_n$ in terms of the $a_n$ and is suited for forming a matrix operator that operates on a mode-space representation of a field $\psi(x)$ to yield the mode-space representation of the real field $\psi'(x)$.

so that $a_0 = \frac{1}{2} a_0$. With this we have

$$\sum_n b_n' T_n = \sum_n a_n' x T_n = \sum_n a_n' (T_{n-1} + c_n T_{n+1})$$

with $c_n$ defined as above. This leads to their coefficient relations based on

$$b_n' = \frac{1}{2} (c_{n-1} a_{n-1}' + a_{n+1}')$$. 

To derive these relations we start from the series representations of \( f(x) \) and its derivatives,

\[
f(x) = \sum_{n=0}^{N} a_n T_n \quad f'(x) = \sum_{n=0}^{N-1} b_n T_n \quad f''(x) = \sum_{n=0}^{N-2} c_n T_n,
\]

where we notice that each differentiation eliminates the constant term, \( a_0, b_0, \) etc., reducing the number of terms in the representation of the derivative by one each time.

In order to evaluate the coefficients \( b_n \) and \( c_n \) in terms of the \( a_n \) we make use of the relationship

\[
f(x) = \int f'(x) dx = \int \int f''(x) dx.
\]

Using Definitions (B.72) and (B.73) to integrate term by term we can write

\[
\sum_{n=0}^{N-1} b_n T_n = \frac{1}{2} b_0 T_0(x) + \frac{1}{2} c_0 T_1(x) + \frac{1}{4} c_1 T_2(x) + \frac{1}{2} \sum_{n=2}^{N-2} c_n \left( \frac{T_{n+1}(x)}{n+1} - \frac{T_{n-1}(x)}{n-1} \right)
\]

and

\[
\sum_{n=0}^{N} a_n T_n = \frac{1}{2} a_0 T_0(x) + \frac{1}{2} b_0 T_1(x) + \frac{1}{4} b_1 T_2(x) + \frac{1}{2} \sum_{n=2}^{N-1} b_n \left( \frac{T_{n+1}(x)}{n+1} - \frac{T_{n-1}(x)}{n-1} \right).
\]

Equating coefficients of each \( T_n \) in Equation (C.23) separately, we find relations

\[
b_n = \frac{1}{2n} (c_{n-1} - c_{n+1}) \quad n = 1, 2, \ldots, N - 3
\]

\[
b_{N-2} = \frac{1}{2(N - 2)} c_{N-3} \quad b_{N-1} = \frac{1}{2(N - 1)} c_{N-2}.
\]

Likewise for the coefficients in Equation (C.24) we have

\[
a_n = \frac{d_n}{2n} (b_{n-1} - b_{n+1}) \quad n = 1, 2, \ldots, N - 2
\]

\[
a_{N-1} = \frac{1}{2(N - 1)} b_{N-2} \quad a_N = \frac{1}{N} b_{N-1},
\]

\[
(\text{C.25})
\]

\[
(\text{C.26})
\]
where as above

\[ d_n = \begin{cases} 
2 & \text{if } n = N \\
1 & \text{otherwise}
\end{cases} \]

Substituting Relations (C.26) in Equation (C.23) gives

\[ \sum_{n=0}^{N} a_n T_n = \frac{1}{2} a_0 T_0(x) + \frac{1}{2} b_0 T_1(x) + \frac{1}{8} (c_0 - c_2) T_2(x) + \frac{1}{4} \sum_{n=0}^{N-1} \frac{1}{n} (c_{n-1} - c_{n+1}) \left\{ \frac{T_{n+1}(x)}{n+1} - \frac{T_{n-1}(x)}{n-1} \right\}, \quad (C.27) \]

and we can calculate a general relation for the \( a_n \) and \( c_n \):\(^7\)

\[ a_n = \frac{d_n}{4n} \left( \frac{1}{n-1} c_{n-2} - \frac{2n}{n^2-1} c_n + \frac{1}{n+1} c_{n+2} \right) \quad n = 1, 2, \ldots, N - 4 \]

\[ a_{N-3} = \frac{1}{4(N-3)} \left( \frac{1}{N-4} c_{N-5} - \frac{2(N-3)}{(N-3)^2-1} c_{N-3} \right) \]

\[ a_{N-2} = \frac{1}{4(N-2)} \left( \frac{1}{N-3} c_{N-4} - \frac{2(N-2)}{(N-2)^2-1} c_{N-2} \right) \quad (C.28) \]

\[ a_{N-1} = \frac{1}{4(N-1)(N-2)} c_{N-3} \quad a_N = \frac{1}{2N(N-1)} c_{N-2} \]

Now by inverting Relations (C.26) and (C.28) we get top-down recursions for the \( b_n \) and \( c_n \) in terms of the \( a_n \) and the larger \( b_n \) and \( c_n \), respectively:

\[ b_{n-1} = b_{n+1} + \frac{2n}{d_n} a_n \quad n = 1, 2, \ldots, N - 2 \]

\[ b_{N-1} = N a_N \quad (C.29) \]

\[ b_n = 0 \quad \text{for } n \geq N \]

and

\[ c_{n-2} = \frac{2n}{n+1} c_n - \frac{n-1}{n+1} c_{n+2} + \frac{4n(n-1)}{d_n} a_n \quad n = 1, 2, \ldots, N - 4 \]

\[ c_{N-3} = 4(N-1)(N-2) a_{N-1} \quad c_{N-2} = 2(N)(N-1) a_N \]

\[ c_n = 0 \quad \text{for } n \geq N - 1 \quad (C.30) \]

These give top-down recursion relations for the \( c_n \)s and \( b_n \)s in terms of the \( a_n \)s.

\(^7\)Alternatively, we can simply combine Relations (C.25) and (C.26).
Alternatively we can use Relation (C.29) in itself to get

\[ b_{n-3} = b_{n+1} + \frac{2n}{d_n} a_n + 2(n - 2)a_{n-2}. \]  \hspace{1cm} (C.31)

Continuing this process and using the condition on \( b_n \) for \( n \geq N \), we see that in general we can rewrite these recursion relations in a bottom-up form:

\[ b_n = 2 \sum_{\substack{p = n + 1 \\ p + n \text{ odd}}}^{N} \frac{p}{d_p} a_p. \]  \hspace{1cm} (C.32)

Similarly from Relation (C.30) we can derive a bottom-up expression for \( c_n \):

\[ c_n = \sum_{\substack{p = n + 2 \\ p + n \text{ even}}}^{N} \frac{p(p^2 + n^2)}{d_p} a_p. \]  \hspace{1cm} (C.33)

As above, these new relations must be corrected for indexing from 1 to \( N+1 \) rather than 0 to \( N \). This gives us a bottom-up method suitable for modifying with a cutoff to neglect small \( a_n \).\(^8\)

Finally the method based on Equation (C.20) above allows us to calculate the real-space derivative field \( \psi'(x) \) given the mode representation of the original field \( \psi(x) \), in other words the \( a_n \), and the values of the \( T_n \). Thus we can use Equations (B.66) and (B.69) to evaluate \( \frac{d}{dx} T_n \) and \( \frac{d^2}{dx^2} T_n \).

We thus have three different options for derivative operators, and in Appendix (C.4.3) below we shall examine their accuracy and speed to determine which ones are best to use in our codes.

\(^8\)See Appendix C.4.3 below for more details.
C.3 The Nonlinear Numerical Methods

In this section we will outline the numerical techniques used to set up the nonlinear problem. While they resemble those used for the linear case in many respects, especially in the spectral analysis, there are differences as well due to the time differencing of the nonlinear case.

C.3.1 The Differencing Techniques.

In performing time differencing of a set of differential equation there are various methods that may be employed. In our case in order to stabilize the numerical evolution we use a mixture of explicit technique, where all the quantities used to calculate something at the $N+1^{st}$ timestep are from the $N^{th}$ timestep or before, and implicit technique, in which the values at the $N+1^{st}$ step appear on both sides of the equation. In our case the nonlinear terms in the Jacobians are handled with an explicit method and are simply calculated outright once each timestep. The rest of the terms are treated implicitly and must be solved for by setting up a series of linear equations and inverting the matrix of the coefficients in a fashion not unlike the method for the linear eigen problems.

In a very general sense, time differencing is the method of discretizing the time derivatives in our equations wherein we have, schematically,

$$\frac{\partial \xi}{\partial t} = \{ \} \Rightarrow \xi_{i}^{n+1} = \xi_{i}^{n} \frac{\Delta t}{2} \{ \}$$

Adams-Bashforth Differencing.

The advection terms, the Jacobians, are handled using the Adams-Bashforth method. Roache [109] (pp. 74-75) describes this is a one-step, three
time level, forward time scheme with an error of the order \((\Delta t^2, \Delta x^2)\).\(^9\)

The derivation of this technique is as follows. In order to find the value of a variable \(\xi\) at the \(i^{th}\) grid location and the \(n + 1^{st}\) time step, being \(\xi^{n+1}_i\), we Taylor expand it in time:

\[
\xi^{n+1}_i = \xi^n_i + \frac{\partial \xi}{\partial t} \bigg|_i^n \Delta t + \frac{1}{2} \frac{\partial^2 \xi}{\partial t^2} \bigg|_i^n \Delta t^2 + O(\Delta t^2).
\]

Next, using a one-sided finite differencing in time to replace the second order term,

\[
\frac{\partial^2 \xi}{\partial t^2} \bigg|_i^n = \frac{\partial}{\partial t} \left[ \frac{\partial \xi}{\partial t} \bigg|_i^n \right] \Rightarrow \frac{1}{\Delta t} \left\{ \frac{\partial \xi}{\partial t} \bigg|_i^n - \frac{\partial \xi}{\partial t} \bigg|_i^{n-1} \right\} + O(\Delta t)
\]

which upon substitution yields

\[
\xi^{n+1}_i = \xi^n_i + \frac{\Delta t}{2} \left\{ 3 \frac{\partial \xi}{\partial t} \bigg|_i^n - \frac{\partial \xi}{\partial t} \bigg|_i^{n-1} \right\} + O(\Delta t^3)
\]

Finally for \(\frac{\partial \xi}{\partial t}\) we insert the various terms that define the time evolution of \(\xi\).

This differencing technique is unconditionally, but weakly, unstable, with an amplification factor \(G = 1 + O(\Delta t^2)\).\(^{10}\) The addition of viscous terms, however, stabilizes it, giving a usable \(\Delta t\) as a function of the order parameter. Lilly [110] found it to be more accurate than the Lax-Wendroff method.

**Crank-Nicolson Differencing.**

The rest of the terms, all the linear terms, are treated with the implicit Crank-Nicolson method (see [109] pp. 83-85). This has no time-splitting and an error of order \((\Delta t^2, \Delta x^2)\), the same as for the Adams-Bashforth technique.

---

\(^9\) See Roache [109] for a thorough review and discussion of differencing techniques, upon which these sections are based. For an analysis of the stability and accuracy of various commonly used techniques see Lilly [110].

\(^{10}\) The amplification factor \(G\) arises in von Neumann stability analysis by solving \(V^{n+1} = GV^n\) for the amplification of variables from one time step to the next. For stability we must have \(|G| \leq 1\). See [109] pp. 42-45.
This method is a straightforward calculation for spatial derivatives:

\[
\xi_i^{n+1} = \xi_i^n + \alpha \frac{\Delta t}{2} \left[ \frac{\partial^k \xi_i^n}{\partial x^k} + \frac{\partial^k \xi_i^{n+1}}{\partial x^k} \right].
\] (C.37)

The stability of this technique depends on the order of the derivative, \( k \). For a first order derivative (also referred to as 'Euler's modified method') |\( G \) = 1. For \( k = 2 \) this method is unconditionally stable, as for \( \Delta t \to 0 \) we have \( \alpha \Delta t/\Delta x^2 \to \infty \) and the amplification factor \( G \to -1 \). We should note, however, that when we use this method in combination with spectral decomposition that for large \( \Delta t \), \( \Delta t > \frac{1}{2} \Delta x^2/\alpha \), the Fourier modes with wavelength \( \lambda \geq 2\Delta x \) will overshoot leading to a rapid loss of stability.

### C.3.2 The Coefficient Equations

In order to derive the explicit equations for the spectral coefficients we start from the vorticity-stream function formulation of the VSC equations.\(^{11}\) Making use of the Adams-Bashforth technique for the nonlinear terms and the Crank-Nicolson technique for the linear terms, we can derive a set of time differenced equations:

\[
\omega^{n+1} = \omega^n - \frac{\Delta t}{2} \left\{ 3 \frac{\partial (\omega, \psi)^n}{\partial (x, z)} - \frac{\partial (\omega, \psi)^{n-1}}{\partial (x, z)} \right\}
- \frac{\Delta t}{2G} \left\{ \frac{\partial T^n}{\partial x} + \frac{\partial T^{n+1}}{\partial x} - \nabla^2 \omega^n - \nabla^2 \omega^{n+1} \right\}
\] (C.38)

\[
T^{n+1} = T^n - \frac{\Delta t}{2} \left\{ 3 \frac{\partial (T, \psi)^n}{\partial (x, z)} - \frac{\partial (T, \psi)^{n-1}}{\partial (x, z)} \right\}
- \frac{\Delta t}{2PrG} \left\{ -\nabla^2 T^n - \nabla^2 T^{n+1} \right\}.
\] (C.39)

\(^{11}\)See Equations (2.12) and (2.16) in Section 2.4.
Now we define $\beta = 2G/\Delta t$, and collect all the $n + 1^{st}$ terms on the left-hand side,

\[
(\beta - \nabla^2)\omega^{n+1} + \frac{\partial T^{n+1}}{\partial x} = (\beta + \nabla^2)\omega^n - \frac{\partial T^n}{\partial x}
\]

\[
= \frac{\beta \Delta t}{2} \left\{ 3 \frac{\partial (\omega, \psi)^n}{\partial (x, z)} - \frac{\partial (\omega, \psi)^{n-1}}{\partial (x, z)} \right\} \quad \text{(C.40)}
\]

\[
(Pr\beta - \nabla^2)T^{n+1} = (Pr\beta + \nabla^2)T^n - \frac{\beta \Delta t}{2} \left\{ 3 \frac{\partial (T, \psi)^n}{\partial (x, z)} - \frac{\partial (T, \psi)^{n-1}}{\partial (x, z)} \right\} \quad \text{(C.41)}
\]

We proceed to carry out a spectral decomposition. First, in the linear terms we replace the variables $\psi, \omega, \text{and} \ T$ with Fourier series in $z$, as we did in Section 4.3 for the linear problem:

\[
\psi(x, z) = \sum_{k_z} \psi_{k_z}(x) \exp ik_z z \quad \text{(C.42)}
\]

so that all the $z$ derivatives bring down a factor of $ik_z$. This yields equations for each $k_z$:

\[
(\beta + k_z^2)\omega^{n+1} - \frac{\partial^2}{\partial x^2} \omega^{n+1} - \frac{\partial T^{n+1}}{\partial x} = (\beta - k_z)\omega^n - \frac{\partial^2}{\partial x^2} \omega^n - \frac{\partial T^n}{\partial x}
\]

\[
- \left[ \frac{\beta \Delta t}{2} \left\{ 3 \frac{\partial (\omega, \psi)^n}{\partial (x, z)} - \frac{\partial (\omega, \psi)^{n-1}}{\partial (x, z)} \right\} \right]_{k_z} \quad \text{(C.43)}
\]

\[
(Pr\beta + k_z^2 T^{n+1} - \frac{\partial^2}{\partial x^2} T^{n+1} = (Pr\beta - k_z)T^n - \frac{\partial^2}{\partial x^2} T^n
\]

\[
- \left[ \frac{\beta \Delta t}{2} \left\{ 3 \frac{\partial (T, \psi)^n}{\partial (x, z)} - \frac{\partial (T, \psi)^{n-1}}{\partial (x, z)} \right\} \right]_{k_z} \quad \text{(C.44)}
\]

where $[ \cdot ]_{k_z}$ indicates the $k_z^{th}$ component of the Fourier transform of these terms.

Next we expand the remaining functions of $x$ (in the linear terms only) in terms of the following Chebychev series:

\[
\psi_{k_z}(x) = \sum_{m=0}^{N} \psi_{m, k_z} T_n(x)
\]

\[
\omega_{k_z}(x) = \sum_{m=0}^{N} \omega_{m, k_z} T_n(x) \quad \text{(C.45)}
\]

\[
T_{k_z}(x) = \sum_{m=0}^{N} T_{m, k_z} T_n(x).
\]
Inserting these we get equations for the \((m, k_z)^{th}\) spectral coefficients at the \(n + 1^{st}\) time step,

\[
(\beta + k_z^2)B_{m,k_z}^{n+1} - B_{m,k_z}^{n+1}'' + C_{m,k_z}^{n+1}' = Q\omega_{m,k_z}
\]  

\(\)  

\[
(\beta + k_z^2)C_{m,k_z}^{n+1} - C_{m,k_z}^{n+1}'' = QT_{m,k_z}
\]  

along with a similar representation of the definition of the vorticity (Relation (2.17)),

\[
k_z^2A_{m,k_z}^{n+1} - A_{m,k_z}^{n+1}'' + B_{m,k_z}^{n+1} = 0
\]  

Here the terms \(Q\omega_{m,k_z}\) and \(QT_{m,k_z}\) contain terms from the \(n^{th}\) and \(n - 1^{st}\) timesteps as follows,

\[
Q\omega_{m,k_z} = (\beta - k_z^2)B_{m,k_z}^{n} + B_{m,k_z}^{n}'' - C_{m,k_z}^{n}'
\]

\[
- \left[ \frac{\beta\Delta t}{2} \left\{ 3 \frac{\partial \omega}{\partial (x,z)} - \frac{\partial \psi}{\partial (x,z)} \right\} \right]_{m,k_z}
\]  

\[
QT_{m,k_z} = (Pr\beta - k_z^2)C_{m,k_z}^{n} + C_{m,k_z}^{n}''
\]

\[
- \left[ \frac{Pr\beta\Delta t}{2} \left\{ 3 \frac{\partial T}{\partial (x,z)} - \frac{\partial \psi}{\partial (x,z)} \right\} \right]_{m,k_z}
\]  

where \([\ ]_{m,k_z}\) indicates the \(k_z^{th}, m^{th}\) component of the Fourier-Chebychev transform of these terms, and in all cases the 's are indicative of an \(x\) derivative.

These are terms that we will calculate in real space, where the multiplications are simple, and then transform back to spectral space to calculate the \(Q\)'s and set up the equations for the linear terms.

We still need to eliminate the \(x\) derivatives in Equations (C.46), (C.47), and (C.48), and this can be done by means of the recursion relations for the Chebychev components. These can be derived using Equation (C.26), and for the coefficients defined above, take the form

\[
2mB_m = B_{m-1}' - B_{m+1}'
\]

\[
2mB_m' = B_{m-1}'' - B_{m+1}''
\]

\[
2mC_m = C_{m-1}' - C_{m+1}'
\]

\[
2mC_m' = C_{m-1}'' - C_{m+1}''
\]
Subtracting Equation (C.46) with \( m \to m - 1 \) from itself with \( m \to m + 1 \) and eliminating the second derivatives with the recursion relations, we get

\[
2mB'_m - (\beta + k_z^2)(B_{m-1} - B_{m+1}) - 2mC_m = Q\omega_{m+1} - Q\omega_{m-1}, \tag{C.51}
\]

where we have dropped the indices for the timestep (as it is uniformly \( n+1 \)) and the Fourier wavenumber (likewise uniformly \( k_z \)). Note also that the summation over \( m \) now starts at \( m = 1 \). A similar treatment of Equation (C.47) yields

\[
2mC'_m - (Pr\beta + k_z^2)(C_{m-1} - C_{m+1}) = QT_{m+1} - QT_{m-1} \tag{C.52}
\]

We repeat the process again, this time subtracting \( \frac{1}{2(m+1)} \) times Equation (C.51) with \( m \to m + 1 \) from \( \frac{1}{2(m-1)} \) times itself with \( m \to m - 1 \). Doing the same for Equation (C.52) yields a new pair of equations, now without derivatives:

\[
2mB_m - \frac{\beta'}{2(m-1)}B_{m-2} + \frac{\beta'm}{m^2 - 1}B_m - \frac{\beta'}{2(m+1)}B_{m+2} - C_{m-1} + C_{m+1} =
\]

\[
-\frac{Q\omega_{m-2} + Q\omega_m}{2(m-1)} + \frac{Q\omega_m - Q\omega_{m+2}}{2(m-1)} \tag{C.53}
\]

and

\[
2mC_m - \frac{\beta''}{2(m-1)}C_{m-2} + \frac{\beta''m}{m^2 - 1}C_m - \frac{\beta''}{2(m+1)}C_{m+2} =
\]

\[
-\frac{QT_{m-2} + QT_m}{2(m-1)} + \frac{QT_m - QT_{m+2}}{2(m-1)}, \tag{C.54}
\]

where we have defined two new quantities \( \beta' \equiv (\beta + k_z^2) \) and \( \beta'' \equiv (Pr\beta + k_z^2) \), and the index on \( m \) now starts at \( m = 2 \).

Finally, the same process applied to Equation (C.48) yields the third set of equations:

\[
2mA_m - \frac{k_z^2}{2(m-1)}A_{m-2} + \frac{k_z^2m}{m^2 - 1}A_m - \frac{k_z^2}{2(m+1)}A_{m+2} - \frac{k_z^2}{2(m-1)}B_{m-2}
\]
\[ + \frac{k^2 m}{m^2 - 1} B_m - \frac{k^2}{2(m+1)} B_{m+2} = 0. \tag{C.55} \]

These three sets of equations for the three sets of coefficients can be written as a matrix equation of the form

\[ A x = c \tag{C.56} \]

where \( A \) is block matrix of the factors multiplying the coefficients in Equations (C.53), (C.54), and (C.55), \( x \) is a vector of the coefficients we are solving for, and \( c \) is a vector of constants. The matrix \( A \) is a 9-block matrix where the blocks are either empty, bi-diagonal, or tri-diagonal. The vector \( x \) contains all of the \( A_m \)'s, \( B_m \)'s, and \( C_m \)'s, and the vector \( c \), in corresponding order the appropriate right-hand sides; the \( Q\omega \) expressions, the \( QT \) expressions, and the zeros for the third equation.

The boundary conditions\(^{12}\) take the form

\[
\begin{align*}
\sum_{m=0}^{N} "B_m(x = 1) &= 0 \\
\sum_{m=0}^{N} "(-1)^m B_m(x = -1) &= 0 \\
\sum_{m=0}^{N} "B'_m(x = 1) &= 0 \\
\sum_{m=0}^{N} "(-1)^m B'_m(x = -1) &= 0 \\
\sum_{m=0}^{N} "C_m(x = 1) &= 1 \\
\sum_{m=0}^{N} "(-1)^m C_m(x = -1) &= -1
\end{align*} \tag{C.57}
\]

and are inserted into the first rows of the appropriate blocks. This procedure is opposite to the one used in the linear problem, where the boundary conditions were inserted into the last rows of the blocks,\(^{13}\) since the process of ‘taking the derivatives’ has emptied the first two rows already.

Once the matrices are set up they are solved by a lower-upper decomposition and back substitution technique.\(^{14}\)

\(^{12}\)Refer to Section 2.5 for their derivation.

\(^{13}\)See Appendix C.1.

\(^{14}\)See the appendix in [48] or [87] for details.
Lastly we note that the possibility of a two variable implementation for the non-linear equations has not yet been tried, although it should be possible, in particular in combination with the use of matrix operators to set up the linear equations in a fashion similar to the formulation of the linear problem. These techniques will be investigated in the future especially in relation to addition of the magnetic potential to the equations.

C.4 Accuracy Tests

In order to gain a quantitative idea as to the accuracy of the codes we have carried out a series of tests on the various component routines. This process has resulted in a much better understanding of how these components function and in several cases has enabled us to redesign them to reduce the numerical error that they introduce into the calculations.\footnote{These tests were run on a Cray X-MP and supplemented and checked later on a Cray Y-MP, where we found some slight improvements.}

C.4.1 The FFTs

One of the key elements of the nonlinear code is the process of transformation from physical space to mode space. These transformations are carried out every timestep for the calculation of the nonlinear terms, and therefore any significant error might be expected to compound rapidly. Although we use a Fourier transformation in \( z \) and a Chebychev transformation in \( \varpi \) both are carried out using Fast Fourier Transforms. This is done by extending the data to be transformed in \( \varpi \) and forcing even parity, thus converting the Fourier transform into a Cosine transform suitable for the cosine-based Chebychev
series.\textsuperscript{16}

To test that our handlers were properly set up, in particular that the various normalizations had been implemented properly, and to ascertain the accuracy of the FFTs themselves we wrote a simple routine to do test transforms and display the results.

The test fields were usually on an \( x \) by \( z \) grid of \( (32,31) \), as this is the grid most often run. This produces 16 independent fourier modes and 32 Chebychev modes in mode space.


- The errors for a single forward back transform in \( x \) or \( z \) or both are of the order of \( 1e^{-13} \) times the amplitude of the test field.

- Overall, the Fourier error is less than the Chebychev, as one might expect, since there are extra computations due to special normalizations, etc. This difference is usually 10-20\%. Note however that some or all of this is due to normally having twice the length of data in this direction. This error also changes if we use exact Chebychev modes. In general the greater number of modes needed to resolve a field, the higher the error.

- The combined error is always 10-20\% less than the sum of the errors of independent transforms.

- Reducing grid size in one direction only affects the error in that direction, and of course the combined error.

\textsuperscript{16}See also Appendix B.3.
Finally for the first three of the test fields below we ran multiple transform tests to see how the error behaved 'over time'. In all cases it grew at a rate just over linear in the first 10 iterations but by 100 iterations the overall growth was slightly under linear. We ran to 10,000 iterations (100,000 for the random fill case), and growth continued to be slightly under linear. At 100,000 iterations the errors had grown roughly $1e+5$ times and were of order $1e-8$. This is the order of the smallest perturbations we have used and seems consistent with empirical evidence that numerical instabilities overwhelm the simulation after several hundred thousand time steps.

2. Random Fill.

- The error is linearly proportional to the range, or amplitude of the test field data ($-r, +r$).
- Shifting the range to $(0, 2r)$ increases the error not more than 50%.
- Reducing the grid to $(16, 15)$ reduces the error slightly, $\sim 20\%$.
- Truncating during transformation produces drastic errors. Even for 32x16 modes truncated to 31x15 modes, the error produced is a factor of 2 larger than the test fields. Truncating to 20x10 modes increases this error by a factor of 5.

3. Cosine by Cosine.

- The error is roughly 60% that for a random field of the same magnitude.
- Reducing the grid to $(16, 15)$ reduces the error slightly: $\sim 20\%$. 
- Truncation to 20x10 modes (from 32x16) seems to leave the error unaffected.


- In this case the Sine function is transformed with difficulty by the Cosine based Chebychev transform. While the Fourier error is only slightly larger than the Cosine case, the Chebychev error is nearly double, in fact being 25% larger than the random case. The combined error is about 90% that for the Random Fill case.

- Reducing the grid to (16,15) reduces the Fourier error exactly as for the Cosine case, but the Chebychev error although reduced by about 40%, as is the combined error, is still larger than in the Cosine case.

- Truncation to 31x15 produces no change in the error, while truncating to 20x10 seems to improve the Fourier error slightly while worsening the Chebychev error a bit. The combined error is about 10% larger.

5. Summation over Sines: \( \sum_k (\text{ran}(x)/N) \sin(\pi kx) \sin(\pi kx) \)

- There are two significant features of this test field: 1) it has a very small amplitude, on the order of \(10^{-13}\), and 2) it includes all scales.

- The errors are still roughly \(10^{-13}\) times the test fields and scale linearly with the test field amplitude.

- Reducing the grid size to (16,15) actually increases the error in the Chebychev transform by a factor of 5, while the Fourier error stays the same. The combined error in this case is actually even larger, being about 7 times larger.
C.4.2 The Algebraic Operators

Tests were run for operators $X, X^2, \text{ and } X^3$ as well as comparison tests where two operators were used in succession. The test fields were 32 by 8.\(^{17}\) In all cases the error is on the order of $0.1 - 0.2e^{-13}$, comparable to that for the FFTs. There was no appreciable difference between the errors from a single operator and those due to an equivalent combination, and no consistent pattern as to which is more accurate.

C.4.3 The Derivative Operators

In Appendix C.2.2 we discussed three different methods for implementing the derivative operators that we need in both the linear and nonlinear codes. First there are the Chebychev based derivatives which also transform back to physical space in $X$. Secondly, there are those based on recursive formulas for the mode space coefficients, including the top-down method and the related bottom-up technique. Here we test each of these for accuracy and speed to decide which is the best to use.

*Varying $NX$: the need for a cutoff.*

The first series of tests run was for a given test function but varying the number of modes, $NX$, used to represent the fields. We can see that for a low number of modes there is a very large error as the fields are poorly resolved, but that as $NX$ is increased the error drops off very rapidly. This trend, however, does not hold for all $NX$; the error bottoms out and then rises slowly again for large $NX$. In the case shown in Figure C.1, for $\sin(x) + \cos(x)$, the error is minimum at about 16 modes; it increases rapidly to 32 modes and

\(^{17}\)Since they are tests of a $X$ operator, the number of $x$ modes used is irrelevant.
Figure C.1: Error vs. number of modes for a $\sin(x) + \cos(x)$ initial field. at a slower rate beyond that. It is disturbing that by increasing the number of modes used we should actually increase the error, especially by several orders of magnitude.\textsuperscript{18}

The high $Nx$ increase in error comes from using more modes than are necessary to resolve the fields involved. If we run these routines with a long tail of modes that should be zero (or less than the level of machine-noise) but have in fact machine noise values, we will add in a lot of machine noise and thus significantly increase our error.

We can already see that the errors involved in these routines are much larger than those involved in any of the other operations; remembering that they are mostly used while calculating the Jacobian terms, which are crucial for the nature of the results, it will be important to try to reduce them as much

\textsuperscript{18}We should point out that all these results depend on the nature of the test function, but the broad trends should hold for most reasonable functions. See also our discussion of the series for $\sin(\alpha x)$ below.
as possible. This is especially true of the larger second-derivative errors, high by themselves are large enough to destabilize unstable equilibria. Therefore, in order to improve accuracy we can add a test to each routine, eliminating these machine noise values from the calculations.

The results of these modified routines are also shown in Figure C.1, and we can see that they help considerably at high $NX$. Note that the top-down recursion method does not lend itself well to this modification, and thus we will be interested mostly in the modified versions of the Chebychev based routines and those using the bottom-up recursion.\footnote{In our tests we compared the unmodified top-down recursion method, the modified and unmodified versions of the Chebychev routines, and the modified versions of the bottom-up routines, using the other routines and versions for comparison.} It seems that this later method is especially good for high $NX$.

In running this series of tests we used cutoffs of $1.0e^{-14}$ and $3.33e^{-14}$. These gave nearly identical results for low $NX$ and for high $NX$, but in the middle range the results for $1.0e^{-14}$ are considerably better.

Although it would seem from these results that setting our cutoff at $1.0e^{-14}$ and using the bottom-up recursion relations will generally limit the first derivatives to an error of the order of $.5e^{-13}$ in the first derivative, and of order $.25e^{-11}$ in the second, we would like to check more carefully the effect of varying cutoff and of changing the test function.

\textit{Varying the Cutoff.}

In Figure C.2, we show a series of results for the same test field, $\sin(x) + \cos(x)$, run for $NX$ at 32 and 64, while we vary the cutoff. It is clear that a cutoff of the order of $3.33e^{-15}$ or less is too small to catch much of the noise, and therefore the modified routines give results essentially the same as
Figure C.2: Error vs. cutoff at $NX = 32$ and 64 for a $\sin(x) + \cos(x)$ initial field.

the unmodified routines. Although not shown here, if we take a cutoff that is too high (perhaps of order $1.0e^{-12}$ or more), we start to lose parts of the field along with the noise, and the modified routines have errors that are actually larger than those of the unmodified routines.

Thus there is a window in which this technique will be most effective, which will lie just above the value of the machine noise, so that we lose as little signal as possible and yet most of the noise. In the case shown here we can see that for a cutoff of $1.0e^{-14}$ or $3.33e^{-14}$ we have the best results and in fact can reduce the errors by 1.5-2 orders for the first derivative and by 2-3 orders for the second derivative.

In order to check that these results also hold for other test fields, we ran a similar series using $\sin(5x) + \cos(5x)$ as a test function. Here the results are more closely bunched for the different routines; because this function peaks at a higher mode and thus needs more modes to be resolved, there seems to be
a much smaller window for an effective cutoff. The results for the derivatives at $NX = 32$ show improvement when using a cutoff in the same range as before, but for $NX = 64$ they seems to be a little bit worse.

Varying $\alpha$.

In order to check the dependence of the improvement due to testing we ran a series, see Figure C.3, using test fields of the form $\cos(\alpha x) + \sin(\alpha x)$

![Figure C.3: Error vs. wavenumber (1-8) for a sin(x) + cos(x) initial field.](image)

where we let $\alpha$ range from 1 to 8. We used cutoffs of $1.0e^{-14}$ and $3.33e^{-14}$.

In general the modified versions are usually better than or at worst only slightly worse than the unmodified routines. They perform best for small $\alpha$, where we expect a rapid drop in mode amplitude with increasing wavenumber and thus a cleaner separation of meaningful values from noise. The spread of values seems to converge as $\alpha$ increases (at least up to $\alpha = 5$). At $\alpha = 8$ there seems to be spreading again with the $NX = 64$ results being worse and the $NX = 32$ results being better for the tested routines, as opposed to the
untested ones.$^{20}$

Note that here for $\alpha$ equaling 1 or 3 the results are significantly better with a cutoff at $1.0e^{-14}$ rather than at $3.33e^{-14}$.

Since increasing $\alpha$ represents increasing harmonics in physical space we might expect that we will have a mixture of $\alpha$'s with the higher $\alpha$ components having smaller amplitudes. To see the results of this increase, we used a test field $\frac{1}{100}\cos(4x) + \frac{1}{100}\sin(4x)$. Using a small amplitude like this reduces all the errors by up to 2 orders leaving them all closely bunched near $4e^{-13}$ for the first derivative and $5e^{-10}$ for the second derivative. Here the modified and unmodified versions perform similarly.

In one final test we combined this last test field with the earlier $\sin(x) + \cos(x)$ field. The results are in most cases slightly better than for the $\sin(x) + \cos(x)$ field alone, though not far off. The notable exception was for the modified bottom-up routine for the second derivative, where the error was an order greater, being close to that for the $\frac{1}{100}\cos(4x) + \frac{1}{100}\sin(4x)$ alone. These results were identical for both a $1.0e^{-14}$ and a $3.33e^{-14}$ cutoff.

Thus overall, one might expect that for a field dominated by low modes the errors are similar to those for low $\alpha$ in this series. In other words, it is more important to minimize the error for low $\alpha$ than for high, for which purpose using a 'tested' routine with a $1.0e^{-14}$ cutoff is most effective.

In summary we can make several general comments (usually but not always true), based on these series of tests:

$^{20}$We also did some runs with the function $x^3$. The results were closely bunched for $NX = 32$, and a bit more spread out for $NX = 64$. The tested bottom-up routines are again best followed by the tested Chebychev routines, as before. Most important, the tested routines are all better than the untested ones.
1. The best routines for accuracy are those based on the bottom-up recursions, and the difference is sometimes substantial. In most cases the 'tested' single-use Chebychev routines are equally good, with the exception of the $NX = 64$ use of the double derivatives for low-peaking modes.

2. A best value for the cutoff can be taken as $1.0e^{-14}$.

3. All the unmodified routines give nearly the same results.

4. It seems that using a single routine for a double derivative is more accurate than using a single derivative twice, although the second derivative routines are always less accurate than the single derivative routines, owing to the nature of taking derivatives of a series representation.

5. In situations where the low-order harmonics in a harmonic decomposition of the fields are not dominant, then we may expect increased error in the tested derivative routines (especially at high $NX$), but overall these routines are better than or equal to the untested ones.

*Time Tests: How fast are they?*

Finally we come to the consideration of time as well as accuracy. Since these routines are used many times in both the linear and nonlinear codes it is important computationally to use the fastest routines we can. In reading the results of timing the routines, it is important to keep in mind where we are using FFT routines along with the derivative routines, as this must be included along with the routines' own time for a valid comparison with other routines that don't need the FFT. We can tell even before we see the actual figures that it is faster to use the single second derivative routines rather than a single derivative twice.
<table>
<thead>
<tr>
<th>subroutine</th>
<th>time in sec.</th>
<th>descriptor</th>
</tr>
</thead>
<tbody>
<tr>
<td>D1X</td>
<td>0.00111</td>
<td>unmodified Chebychev</td>
</tr>
<tr>
<td>D2X</td>
<td>0.00112</td>
<td>unmodified Chebychev</td>
</tr>
<tr>
<td>F1DIR</td>
<td>0.00026</td>
<td>unmodified top down</td>
</tr>
<tr>
<td>F2DIR</td>
<td>0.00046</td>
<td>unmodified top down</td>
</tr>
<tr>
<td>ODIR</td>
<td>0.00297</td>
<td>unmodified bottom up</td>
</tr>
<tr>
<td>O2DIR</td>
<td>0.00301</td>
<td>unmodified bottom up</td>
</tr>
<tr>
<td>DT1X</td>
<td>0.00173</td>
<td>modified Chebychev</td>
</tr>
<tr>
<td>DT2X</td>
<td>0.00141</td>
<td>modified Chebychev</td>
</tr>
<tr>
<td>ODIR</td>
<td>0.01329</td>
<td>modified bottom up</td>
</tr>
<tr>
<td>O2DIR</td>
<td>0.01287</td>
<td>modified bottom up</td>
</tr>
<tr>
<td>XFFT</td>
<td>0.00116</td>
<td>the FFT routine</td>
</tr>
</tbody>
</table>

We have included the time for the FFT as in either code it would have to be used with the Chebychev-based routines to bring a result back to mode space, while in the nonlinear code we would have to use it with the recursive routines to transform the result into real space where that is called for.

We see that by far the fastest routines are the unmodified top-down recursives. Even when used with the FFT, they are close to the times for the Chebychev routines standing alone. The unmodified bottom-up recursives are roughly equal to the Chebychev routines used with the FFT. On the other hand the modified recursive routines are up to an order of magnitude slower than the Chebychev-based routines. It thus seems we must choose between an order of magnitude in speed or in accuracy.

Looking carefully again at Figure C.3, we see that for low $\alpha$ the modified Chebychev routines are between a half and a full order of magnitude less accurate than the bottom-up recursions for the first derivative, while for the second derivative they are slightly more than an order of magnitude less accurate. By the fifth harmonic they are equal and for higher $\alpha$ possibly a little better (at least for the first derivative).

Thus in general we can probably settle with the modified Chebychev
routines. For best low mode accuracy we use the bottom-up modified routines and if we want pure speed and can give up a bit of accuracy then the unmodified top-down routines are best.

C.4.4 The Nonlinear Terms and De-aliasing

The de-aliasing methods examined here are based on the idea that if the high frequency modes in the two fields to be multiplied have zero amplitude then there will be no aliasing in the result. Specifically, if the two initial fields are limited to non-zero mode amplitudes for $\frac{2}{3}$ (or less) of the total number of modes employed, then the result will also contain valid results in the first $\frac{2}{3}$ of its modes.

There are essentially three ways to produce this result. First, if the mode amplitudes drop off with increasing frequency at a sufficiently rapid rate (and/or a large enough number of modes is employed so that the upper $\frac{1}{3}$ are negligible), then there will be very little aliasing, and in general it may be ignored. When these conditions are not met, however, as when the field contains a steep gradient so that the mode amplitude versus frequency curve is relatively shallow, or when computational constraints place an effective limit on the number of modes employed, then one must contrive to limit artificially the number of non-zero modes. The other two methods are of this sort.

We should note that the linear calculations for our code can be done in mode-space and that we can make full use of all modes with no problems.\(^{21}\) It is when we wish to calculate the nonlinear terms that we would like the fields to be band limited. To do this there are two options. First, we can zero out

\(^{21}\)Of course, it is best if we have enough modes to resolve the flows that are generated.
the high-end modes before transforming the fields to real-space, keeping the total number of modes the same but using only \( \frac{2}{3} \) of them. Second, we can, for this operation only, extend the size of the fields by padding them with zeros before transforming to real-space and then discard the extra modes when we transform the result back. This second choice will minimize the loss of signal but at a potential computational cost.

In our tests we have examined these three options, comparing them to an analytic calculation of the convolution in mode space as discussed in Appendix B.4. These options were 1) to do no de-aliasing, in order to gauge the error involved, 2) to throw out modes at the start, and 3) to pad with extra modes. If the amplitudes tail off fast enough, then these methods all give the same results, but since the VSC problem often develops a steep gradient in the temperature perturbation at the center of the slot, which is exacerbated by the fact that \( x = \cos(\theta) \) grid is sparsest in the slot center, it seems likely that this is often not the case for our work.

In initial tests we have used test fields that are of the form

\[
\psi(k_x, k_z) = \exp^{-\alpha k_x - \beta k_z}
\]

(C.58)

varying \( \alpha \) and \( \beta \) from 1 to 10. In most cases we used \( \alpha = \beta \), but we also performed tests varying only one or the other.

Results:

- We looked at two functions to determine the errors, one being the simple difference between the results by a given method and the convolution results, which we will call the difference, and the other being this difference normalized by the convolution value at the same point in the field, which we will call the error.
• The difference as well as the error are largest where the convolution is smallest, along the \( k_x = N k_{xo} \) and \( k_z = N k_{zo} \) boundaries of the mode-space fields. The maximum error is always the same and is surely some numerical effect, as the value is rational.

• The values of the convolution drop with increasing \( \alpha \) and \( \beta \), the minima very steeply and exponentially, the maximum value more slowly and less than exponentially. It may even have an absolute lowest upper bound.

• For all measures the error decreases rapidly as the values of \( \alpha \) and \( \beta \) increase, dropping exponentially with a logarithmic slope of the same order as that of the convolution minima.

• The differences are smallest for [3], as might be expected, and largest for [2]. Apparently the error introduced by truncation is greater than that avoided by de-aliasing.

• At \( \alpha = \beta = 1 \) the maximum differences for [1] and [2] are nearly the same while that for [3] is slightly more than an order of magnitude lower. For \( \alpha = \beta = 10 \) [3] is still better than [1] by about the same amount, but [2] is worse by an order of magnitude. Even this case, however, is 2.5 orders lower than at \( \alpha = \beta = 1 \).

For \( \alpha = \beta < 4 \) all differences for [3] are at most of the same order as the convolution minima.

• The minimum differences for [1] and [3] are essentially the same and drop at the same rate as the minima of the test fields. This line is always more than an order of magnitude lower than the minimum value of the convolution. Since this represents the error where the convolution is
largest, we can see that unless there is some special sensitivity to the small high-frequency values of the mode-space field the aliasing errors are probably inconsequential. For test fields with even $\alpha = \beta = 1$ the error is barely larger than 1% while for $\alpha = \beta = 5$ it is on the order of .0025%.

- For $\alpha$ or $\beta$ held at 1 while the other varies, we find that the differences level off very quickly. For $\beta = 1$ this is at the 1% level, for $\alpha = 1$ it is at the .5% level. This reflects the faster Chebychev convergence. It also shows that the errors may be much worse when one of the fields drops off slowly with increasing frequency (i.e. the Chebychev decomposition of the temperature fields).

Overall we can see that if the fields are resolved with a few modes there is probably little error. When this is not the case, however, chopping modes introduces more error than it saves,\textsuperscript{22} and the preferred de-aliasing technique must be to extend the fields and pad them with zeros, discarding the extra modes in the result.

\textsuperscript{22}This indicates that attempts to use a number of modes, $N$, that is other than a power of 2 by means of truncation may be expected to cause problems. This has indeed been found to be true empirically, and we have abandoned this technique.
BIBLIOGRAPHY


VITA

Alan Holland McAllister was born in Philadelphia, Pennsylvania, on August 3, 1957, the son of Louis Erhardt McAllister jr. and Marian Holland McAllister. After graduating from Germantown Friends School in Philadelphia in 1975, he entered Haverford College, Haverford, Pennsylvania. After a leave of absence from January 1978 to August 1979 during which time he worked at The Philadelphia Savings Fund Society, he returned to Haverford College and received the degree of Batchelor of Science in May, 1981. Moving to Mountain View, and then to Tokyo in the fall of 1982 he worked as a systems engineer(M.V.), programmer(M.V. & Tokyo) and technical proofreader (Tokyo) until the end of 1984. In January 1985 he entered the University of Texas at Austin, where he has been employed as a grader, a research assistant, and at present a research engineer/scientist assistant. On August 17th 1985 he married Chiwa Higashi and they have two children, Madoka age 4, and Takumi age 9 months.

Permanent address: 5106 N. Lamar #247
    Austin, Texas 78751

This dissertation was typeset\textsuperscript{23} with \LaTeX\ by the author.

\textsuperscript{23}\LaTeX\ document preparation system was developed by Leslie Lamport as a special version of Donald Knuth's \TeX\ program for computer typesetting. \TeX\ is a trademark of the American Mathematical Society. The \LaTeX\ macro package for The University of Texas at Austin dissertation format was written by Khe-Sing The.