

Baroclinic critical layers in rotating stratified shear flows

by

Chen Wang

B.Sc., Beihang University, 2012

M.Sc., Beihang University, 2015

A THESIS SUBMITTED IN PARTIAL FULFILLMENT
OF THE REQUIREMENTS FOR THE DEGREE OF

Doctor of Philosophy

in

THE FACULTY OF GRADUATE AND POSTDOCTORAL
STUDIES

(Mathematics)

The University of British Columbia

(Vancouver)

August 2020

© Chen Wang, 2020

The following individuals certify that they have read, and recommend to the Faculty of Graduate and Postdoctoral Studies for acceptance, the dissertation entitled:

Baroclinic critical layers in rotating stratified shear flows

submitted by Chen Wang in partial fulfillment of the requirements

for the degree of Doctor of Philosophy

in Mathematics

Examining Committee:

Neil Balmforth, Mathematics
Supervisor

Gregory Lawrence, Civil Engineering
Supervisory Committee Member

Gwynn Elfring, Mechanical Engineering
Supervisory Committee Member

Michael Ward, Mathematics
University Examiner

Christian Schoof, Earth, Ocean and Atmospheric Sciences
University Examiner

Xuesong Wu, Mathematics, Imperial College London
External Examiner

Abstract

In this thesis, we study baroclinic critical layers in rotating stratified shear flows. Baroclinic critical layers are characterized by strong amplitudes surrounding the singular points of the steady inviscid wave solution, and play crucial roles in the mixing and transition to turbulence in ocean, atmosphere and astrophysical disks. This thesis studies the baroclinic critical layers in strato-rotational instability and the forced baroclinic critical layers.

The first problem we study is the baroclinic critical layer in strato-rotational instability. Strato-rotational instability (SRI) is normally interpreted as the resonant interactions between internal gravity waves or Kelvin waves. Using a combination of asymptotic analysis and numerical solution of the linear eigenvalue problem for plane Couette flow, it is shown that such resonant interactions can be destroyed by baroclinic critical layers. The critical-level coupling removes the requirement for resonance near specific wavenumbers, resulting in an extensive continuous band of unstable modes.

The second problem we study is the forced baroclinic critical layers. Linear theory predicts the baroclinic critical layer dynamics is characterized by the secular growth of flow perturbations over a region of decreasing width. Once it enters the nonlinear regime, the nonlinear dynamics filters out harmonics and the modification to the mean flow controls the evolution. At late times, we show that the vorticity begins to focus into yet smaller regions whose width decreases exponentially with time, and that the addition of dissipative effects can arrest this focussing to create a drifting coherent structure.

In the last problem, we show that the mean-flow defect generated in the forced baroclinic critical layer can make the flow unstable, and we study this ‘secondary

instability'. The instability is a horizontal shear instability with a distinct phase velocity compared to that of the forced baroclinic critical layers, and thus will excite new baroclinic critical layers. A WKB solution for the exponential growth is derived, which indicates the secondary instability grows faster than a common normal mode due to the unsteadiness of the mean-flow defect. At the later stage, the short-wave harmonics grow at extraordinarily high speeds and will finally make the linear problem ill-posed.

Lay Summary

In this thesis, we study baroclinic critical layers, which are thin layers with pronounced wave amplitudes in horizontal shear flows with density stratification. They have been found to play crucial roles in the transition to turbulence in astrophysics. We first study baroclinic critical layers in the ‘strato-rotational instability’. Such an instability is previously known to be caused by resonant interactions between two waves. Here we show that a baroclinic critical layer removes the resonant condition and induces a new kind of instability. We then study baroclinic critical layers generated by a weak forcing. We show theoretically how the critical layer will emerge, grow, focus or saturate, and at last trigger new critical layers. Our analysis can provide a theoretical foundation for the recently discovered ‘zombie vortices’ which can replicate themselves.

Preface

This thesis is the original work of the author, Chen Wang during the course of my PhD studies at UBC. The research was conducted under the guidance of Dr. Neil Balmforth.

The work of Chapter 2 has been published in the Journal of Fluid Mechanics, along with Dr. Neil Balmforth: [C. Wang and N. J Balmforth. Strato-rotational instability without resonance. *J. Fluid Mech.*, 846:815–833, 2018]. I conducted the all the research work including theoretical derivation and numerical computation. Dr. Neil Balmforth found the topic, supervised the research and edited the manuscript.

The work of Chapter 3 has been published in the Journal of Fluid Mechanics, along with Dr. Neil Balmforth: [C. Wang and N. J Balmforth. Nonlinear dynamics of forced baroclinic critical layers. *J. Fluid Mech.*, 883, 2020]. I conducted the all the research work including theoretical derivation and numerical computation. Dr. Neil Balmforth found the topic, supervised the research and edited the manuscript.

The work of Chapter 4 was conducted in collaboration with Dr. Neil Balmforth. I am the principal contributor to this research. I did most of the research, including theoretical derivation and numerical simulation in the Chapter. Dr. Neil Balmforth supervised the research and did numerical simulation for the model described in Appendix B.

Table of Contents

Abstract	iii
Lay Summary	v
Preface	vi
Table of Contents	vii
List of Figures	x
Acknowledgments	xvii
Dedication	xix
1 Introduction	1
1.1 Overview	1
1.2 Critical layers	2
1.2.1 Classical critical layers	2
1.2.2 Baroclinic critical layers	5
1.3 Protoplanetary disk, strato-rotational instability and zombie vortex instability	7
1.3.1 Protoplanetary disk	7
1.3.2 Strato-rotational instability	8
1.3.3 Zombie vortex instability	11
1.4 Scope of the present thesis	12

2	Strato-rotational instability associated with baroclinic critical layers	14
2.1	Introduction	14
2.2	Mathematical formulation	15
2.2.1	Model and governing equations	15
2.2.2	Normal modes	17
2.2.3	Solutions near the baroclinic critical levels	19
2.3	Instability of flow with stronger stratification	21
2.3.1	Resonant instabilities	24
2.3.2	Instability induced by the baroclinic critical level	27
2.4	Instability of flow with weaker stratification	31
2.5	Implications of pseudomomentum conservation	34
2.6	Concluding remarks	35
3	Nonlinear dynamics of forced baroclinic critical layers	39
3.1	Introduction	39
3.2	Model and governing equations	41
3.3	Linear theory	44
3.3.1	The steady wave response outside the baroclinic critical layers	44
3.3.2	The linear critical layers	46
3.3.3	Closure	47
3.4	The weakly nonlinear critical layer	49
3.4.1	Mean-flow response	49
3.4.2	First harmonic	50
3.4.3	Weakly nonlinear feedback	50
3.5	Nonlinear critical-layer theory	51
3.5.1	The reduction	51
3.5.2	Canonical system	53
3.5.3	Numerical solutions	55
3.5.4	Long-time focussing	58
3.6	Effects of diffusion	63
3.6.1	Modified canonical system	64
3.6.2	Dissipative coherent structures	66

3.7	Discussion	70
4	Secondary instability of forced baroclinic critical layers	73
4.1	Introduction	73
4.2	Normal mode analysis	75
4.3	An asymptotic theory for secondary instability	87
4.4	Numerical solution to the secondary instability	96
4.5	Conclusion	102
5	Summary and future directions	104
5.1	Strato-rotational instability associated with baroclinic critical layers	104
5.2	Nonlinear dynamics of forced baroclinic critical layers and their secondary instability	105
	Bibliography	108
A	The critical-layer vorticity distribution	115
B	A nonlinear model for the secondary instability	116

List of Figures

Figure 1.1	Contours of absolute vorticity of Stewartson [1] and Warn and Warn [2]’s solution for the nonlinear critical layer of Rossby waves (left) and its secondary instability (right). Source: Haynes [3].	4
Figure 1.2	Baroclinic critical layers of a tilted vortex in the experiment of Boulanger et al. [4]. Figure (a) is the density contract and figure (b) is the velocity field with the grey scale representing the vertical velocity. The pattern of tilted lines in the two figures are the baroclinic critical layers. Source: Boulanger et al. [4].	6
Figure 1.3	Protoplanetary disk surrounding the young star Elias 2-27. Source: https://www.eso.org/public/images/potw1640a/	8
Figure 1.4	Experiment of Le Bars and Le Gal [5] to demonstrate the strato-rotational instability (SRI). The Taylor-Couette flow is in the centrifugally stable regime. The left figure is for pure water, and there is no instability. The right figure is for stratified salt water, and waves of SRI develop. Source: Le Bars and Le Gal [5].	10
Figure 1.5	Zombie vortices of Marcus et al. [6]. The color scale represents the vorticity field. Vortices trigger baroclinic critical layers and replicate as time increases from figure (a) to figure (d). Source: Marcus et al. [6].	11
Figure 2.1	Sketch of the model geometry.	16

Figure 2.2	Plots of $l^2 = -\lambda^2$ against y , illustrating the two configurations of interest for the spatial arrangements of the classical critical level $y = y_c$, the turning points $y = y_{t\pm}$ and the baroclinic critical levels $y = y_{b\pm}$, which are dictated by the sign of ε_b when $\mu^2 \gg 1$	20
Figure 2.3	Phase speeds against $\mu = m/k$ for $R = 9$ and $k = 5$, with (a) $\Omega = 5/8$ (anti-cyclonic shear) and (b) $\Omega = -1/8$ (cyclonic shear). Eigenfunctions at the points I and II-IV are plotted in figures 2.5 and 2.6, respectively. The dotted line in figure (a) shows the asymptotic solution (2.28). The horizontal broken lines indicate the phase speeds for which modes develop baroclinic critical levels at the channel walls ($\text{Re}(y_{b-}) = 0$ and $\text{Re}(y_{b+}) = 1$).	22
Figure 2.4	(previous page) Unstable growth rates against $\mu = m/k$ for $R = 9$ and $k = 5$, with (a) $\Omega = 5/8$ (anti-cyclonic shear) and (b) $\Omega = -1/8$ (cyclonic shear). The insets show magnifications for higher μ and plot only the growth rates of the unstable modes generated by interaction with the baroclinic critical level. The dashed line in figure (a) shows the asymptotic solution (2.38); the dots in both panels show the asymptotic growth rates computed from (2.37). Eigenfunctions at the points I and II-IV are plotted in figures 2.5 and 2.6, respectively.	24
Figure 2.5	Eigenfunction \hat{u} of the resonant unstable mode corresponding to the coupling of IGW $_{2+}$ and IGW $_{1-}$ (labelled by the point I in figures 2.3(a) and 2.4(a); $\mu = 4.523$). Red and blue lines represent real and imaginary part of the eigenfunction, respectively. The numerical solution is shown in solid lines, and has eigenvalue $c = 0.454 + 0.00232i$. The asymptotic eigenfunction is shown in dashed lines, and is constructed using the computed asymptotic eigenvalue $c \sim c_0 + \Delta c \approx 0.453 + 0.00240i$ together with the formulae (2.22)–(2.25), which become inaccurate at the turning points $y = y_{t\pm}$	25

Figure 2.6 Eigenfunction \hat{u} of unstable modes with baroclinic critical levels for $\Omega = 5/8$ and (a) KW_- ($\mu = 5.7$; point II in figures 2.3 and 2.4), (b) $IGW1_-$ ($\mu = 9$; point III), and (c) $IGW2_-$ ($\mu = 12.5$; point IV). Red and blue lines represent real and imaginary part of the eigenfunction, respectively. The solid lines show the numerical solutions. The corresponding results from WKB theory are shown as dashed lines, which become inaccurate at the turning points $y = y_{t\pm}$ (the asymptotic eigenfunction is not shown in (a) as the comparison is poor owing to the turning point $y = y_{t-}$ close to the boundary $y = 0$). In (b), numerical computations give $c = 3.34 \times 10^{-1} + 1.54 \times 10^{-4}i$ whereas the asymptotics indicate $c = 3.35 \times 10^{-1} + 1.56 \times 10^{-4}i$. In (c), the numerical and asymptotic results are $c = 3.71 \times 10^{-1} + 3.96 \times 10^{-5}i$ and $c = 3.72 \times 10^{-1} + 3.92 \times 10^{-5}i$, respectively. Insets show magnifications of the structure near the baroclinic critical levels. 29

Figure 2.7 (a) Phase speed c_r and (b) growth rate ω_i against $\mu = m/k$ for $R = 2.5 \times 10^{-3}$, $\Omega = 5/8$ and $k = 0.4$. The horizontal broken lines indicate where baroclinic critical levels appear at the channel walls. The dots plot the asymptotic results in (2.28) and (2.43) which we do not include for the first peak in the main panel of (b) because the comparison is poor owing to the relatively small value for μ). Eigenfunctions at the points V and VI are plotted in figure 2.8. 32

Figure 2.8	Eigenfunction \hat{u} of unstable modes with baroclinic critical levels at the points marked (a) V ($\mu = 1.85$) and (b) VI ($\mu = 4.2$) in figure 2.7. Red and blue lines represent real and imaginary parts of the eigenfunction, respectively. The solid line shows the numerical solution and the dashed line in (b) is the result of asymptotics (the asymptotic eigenfunction is not shown in (a) as the comparison is poor owing to the relatively small value for μ). The insets show magnifications of the structure near the baroclinic critical levels. In (b) the numerical result for the eigenvalue is $c = 6.19 \times 10^{-2} + 8.82 \times 10^{-6}i$ and $c = 6.66 \times 10^{-2} + 5.30 \times 10^{-6}i$ for the asymptotics.	33
Figure 2.9	The pseudomomentum density \hat{p} in (2.44) for (a) the resonant SRI mode in figure 2.5, and (b,c) the KW_- and IGW_- modes destabilized by its interaction with a baroclinic critical level at $y = y_{b+}$ shown in figure 2.6 and figure 2.8. Note that the sharp negative peaks at the baroclinic critical levels have been truncated in (b,c) to help reveal the mode structure.	36
Figure 3.1	Sketch of the model. A wavemaker with wavenumber k_x and k_z is imposed at $y = 0$, and baroclinic critical levels are forced at $y = \pm N/(\Lambda k_x)$, corresponding to dimensionless locations $\pm \mathcal{N}$, where $\mathcal{N} = N\Lambda^{-1}$. The shading represents a rendering of the density perturbation based on the linear theory of Section 3.3.	41
Figure 3.2	Steady-wave solution \hat{p} under a forcing imposed at $y = 0$, with $m = 0.5$, $\mathcal{N} = 4/3$, $f = 4/3$, $\varepsilon_0 = 0.05$ (cf. Marcus <i>et al.</i> 2013). Baroclinic critical levels $y = \pm \mathcal{N}$ are indicated. . . .	48
Figure 3.3	Evolution of A and B with τ ; $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$. . .	56

Figure 3.4	<p>(a) Real and (b) imaginary parts of the critical-layer density perturbation $\gamma(\eta, \tau)$, shown as surfaces above the (η, τ)-plane. To prevent the viewing perspective from obscuring parts of the solution, we also show density maps of the solutions underneath. The insets show corresponding plots of the linear solution in (3.66). Panels (c)–(f) plot snapshots of $\gamma(\eta, \tau)$ at the times indicated; the linear result for $\gamma(\eta, \tau)$ is also included. ($m = 1/2, f = 4/3, \mathcal{N} = 4/3$.)</p>	57
Figure 3.5	<p>Snapshots of vertical vorticity ζ within the baroclinic critical layer near $y = \mathcal{N} = 4/3$, plotted as a colormap on the (x, η)-plane for (a) $\tau = 0.3$, (b) $\tau = 0.45$, (c) $\tau = 0.6$, (d) $\tau = 1$, (e) $\tau = 1.5$, and (f) $\tau = 1.8$. The domain plotted is $\eta < 25$, corresponding to $y - \mathcal{N} < 0.39$, at cross-section $z = 0$ and over one streamwise wavelengths of the forcing pattern. ($\varepsilon_0 = 0.05, \varepsilon = 0.0062, m = 1/2, f = 4/3, \mathcal{N} = 4/3$.)</p>	58
Figure 3.6	<p>Phase portraits of (γ_r, γ_i) for (a) the Hamiltonian system (3.69) with $\eta = 3/\sqrt[3]{2}$ and various H, with the thicker line indicating $H = 0$, (b) trajectories from the point $(\gamma_r, \gamma_i) = (0, 0)$ for a selection of values of η, and (c) the numerical solution of Section 3.5.3, at the five values of η indicated. The black points in (b) and the (red and blue) pairs marked (A, B) and (C, D) in (c) have the same values of $\eta - \eta_c e^{\sigma\tau}$.</p>	60
Figure 3.7	<p>(a) Evolution of ζ_0 near $\eta_c = 2.43$ at the times indicated. (b) Scaled profiles, $\zeta_0 e^{-\sigma_c(\tau-\tau_0)}$ against $\xi = (\eta - \eta_c)e^{\sigma_c(\tau-\tau_0)}$, choosing $\tau_0 = 3$; ($m = 1/2, f = 4/3, \mathcal{N} = 4/3$).</p>	62
Figure 3.8	<p>The analytical solution (3.81) for strong diffusion and $A = -1$, showing (a) $\lambda^{1/3}\gamma_r$ and (b) $\lambda^{1/3}\gamma_i$ against the scaled space and time variables $\lambda^{-1/3}\eta$ and $\lambda^{1/3}\tau$. The (red) dots show the final steady-state solution. The insets show corresponding numerical solutions to the reduced model, computed for $\lambda = 5.3$.</p>	65

Figure 3.9	Solutions of the modified canonical model, showing (a) γ_r , (b) γ_i and (c) \mathcal{U} , for $m = 1/2$, $f = 4/3$ and $\mathcal{N} = 4/3$, $c_1 = 0.238$, $c_2 = 0.219$ with the values of λ indicated (and corresponding to the three columns). The colormap is the same in the first three panels of (a) and (b), but not the rightmost panel. The quasi-steady wave amplitude $A = A_r + iA_i$ for the four computations is shown in (d).	67
Figure 3.10	A coherent structure computed from (3.83) with $\eta_* = 5.04$ and $A = -1.2$, showing (a) γ_r and γ_i , and then (c) \mathcal{U} (solid lines). The dotted lines show the numerical solution of the modified canonical model (3.79), computed for $\lambda = 0.53$ at $\tau = 17.8$ (at which moment the residual oscillations near $\eta = 0$ are less pronounced). In (b) and (d) we show G_- , c and the jump $\Delta\mathcal{U} = -[\mathcal{U}]_{-\infty}^{\infty}$ against η_* from the solutions to (3.83) for $A = -1.2$ (solid lines). The dashed lines show the limiting behaviour for $\eta_* \gg 1$ given in (3.85). The circles show data for c and $\Delta\mathcal{U}$ measured from the numerical solution of (3.79) with $\lambda = 0.53$ from $\tau = 5$ to $\tau = 15.4$	69
Figure 4.1	Mean-flow modification of streamwise velocity ΔU at $t = 20$ and $t = 30$, $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$	77
Figure 4.2	Eigenvalue c computed by the method of matrix eigenvalues (4.15). The parameters are $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $k = 1$, $\mu = 1/2$ and the mean-flow modifications are chosen at $t = 30$. The computation domain is $y \in [-7 + \mathcal{N}, 7 + \mathcal{N}]$, which is discretized by 3000 evenly distributed grid points. (4.13) is evaluated numerically using a composite trapezoidal rule.	80
Figure 4.3	Eigenvalue c at different times, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$	82
Figure 4.4	Eigenfunctions of the unstable mode at $t = 30$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$	83

Figure 4.5	\check{p} , \check{p}_+ and \check{p}_- at the onset of secondary instability at $t = 27.5$, $c = 1.425$; $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$.	84
Figure 4.6	Time of onset of instability t_s versus ε and $\varepsilon^{-1/2}$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$.	86
Figure 4.7	Secondary instability of subharmonics, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$ for $k = 1$. Solid lines represent the numerical solution of (4.47) with initial condition (4.49), dashed lines represent the WKB solution (4.48), and the dotted lines are the early-time solution (4.50).	93
Figure 4.8	Secondary instability of harmonics, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$ for $k = 1$. Solid lines represent the numerical solution of (4.47) with initial condition (4.49), dashed lines represent the WKB solution (4.48), and the dotted lines are the early-time solution (4.50).	94
Figure 4.9	Numerical solution of the secondary instability at $t = 30$; $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$. Results are obtained by solving (4.3)-(4.6) with (4.51) numerically.	98
Figure 4.10	Evolution of $\Phi(T)$, $k = 1$, $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$. Solid line is the numerical result of (4.51), and the dashed line is the WKB solution (4.48).	99
Figure 4.11	Vorticity of numerical solution (left) and the reduced model (right), $T = 9$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$.	100
Figure 4.12	Evolution of Φ for $k = 1, 2$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, Solid lines represent the numerical solution of (4.51)-(4.54), and dashed lines represent the WKB solution (4.48).	101
Figure 4.13	Vorticity of numerical solution (left) and the reduced model (right) for the harmonic $k = 2$, $\mu = m = 1/2$ at $T = 9.5$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 2.36$.	102

Acknowledgments

First of all, I would like to thank my supervisor, Professor Neil Balmforth. With his intelligent scientific instinct, he found this interesting topic for my PhD research. His supreme expertise and experience guided me through countless difficulties which might have stuck me forever, and his experienced intuition often led us to unexpected new discoveries. On top of that, I was very impressed by his dedication to science. Whenever I wanted to discuss with him, we would meet on the same day I sent him an email and often talk for hours long. Over my five years at UBC, the time I enjoyed and valued most was when I was writing on the blackboard in his office, explaining my ideas to him and then listening to his insightful comments. We have struggled together, argued with each other, and enjoyed the pleasure when the problem was finally solved elegantly.

I thank my supervisory committee members Professor Gregory Lawrence and Professor Gwynn Elfring. They have read my research proposal and draft thesis in detail and given me important suggestions. They have also attended my comprehensive exam and raised important criticism to my research. I thank Professor Michael Ward and Professor Christian Schoof for serving as my university examiner of the doctoral oral defence and raising constructive comments. I specially thank my external examiner Professor Xuesong Wu. He revised my thesis in great detail and raised a number of insightful comments and suggestions, through which I have improved my thesis significantly and learned a lot.

I'm grateful to the professors at UBC who taught me courses, which provided the fundamental and essential knowledge for my research. I thank Professors Neil Balmforth and Michael Ward for perturbation method and asymptotic analysis, Professor Colin MacDonald and Dr. Timm Treskatis for numerical methods, Pro-

fessor Stephanie Waterman for geophysical fluid dynamics, Wayne Nagata for dynamical system, and Stephen Gustafson for nonlinear wave equations.

A number of professors and colleagues have discussed with me, including Professors Philip Marcus, Stephane Le Dizès, Orkan Umurhan, Richard Kerswell, Dr. Thomas Eaves and Dr. Daniel Lecoanaet. They come from different backgrounds and therefore raised valuable constructive comments to our work. I specially thank Professor Philip Marcus for the discovery of the zombie vortex, which became the origin of my research topic. I also thank the anonymous referees to our manuscripts, and Professor Oliver Bühler serving as the associate editor of the *Journal of Fluid Mechanics* who handled our submission. Our research had significant improvements after revising the manuscripts.

During my doctoral study, I met many good friends at our Math Department: Mingfeng Qiu, Jieliang Hong, Han Hong, Elaine Wang, Yifu Zhou and Xinyu Liu. I thank them for having wonderful time with me and helping me out in my difficult moments. I thank my labmates Ivonne Medina-Lino and Marjan Zare. We talked to each other every day and share our happiness and sorrows. I thank my fellow group members Wenjing Dong, Kelsey Everard and Daniel Robb. We have group meetings together, and we enjoyed talking with each other to share our research and our life experiences. I thank my personal good friends Meinan Chen, Ruoxuan Li, Mo Yang, Hongmin Li and Sicheng Li for their concern, consolation and encouragement through my doctoral studies.

Finally, I thank my beloved parents. Their deep love has nurtured me, given me courage and helped me to get through the difficult times. I feel lucky to have this warm family that always support me whatever hardship I experience. This thesis is indebted to their love.

Dedication

To my father Desheng Wang, and my mother Yimei Jiang

Chapter 1

Introduction

1.1 Overview

A centrepiece in the theory of waves in inviscid shear flow is the critical level. In linear theory of normal modes, the critical levels become singular, demanding the inclusion of the weak effects of unsteadiness, nonlinearity or viscosity. Although these inclusions can remove the singularity of the normal mode theory, perturbations to the flow can still develop strongly in the neighbourhood of the critical levels, creating distinctive flow structures and rearrangements within the so-called critical layers that may subsequently break down to generate mixing and turbulence.

The classical critical level is the location where the phase speed c of a steady wave matches the local mean flow speed U . There has been numerous studies on this topic, and a review is given by Maslowe [7]. The critical level that this thesis focuses on is of a different kind: the baroclinic critical level. It arises in stratified fluid with horizontal shear where the Doppler-shifted phase velocity $c - U$ matches a characteristic gravity-wave speed N/k , based on the buoyancy frequency N and streamwise horizontal wavenumber k . Baroclinic critical layers have received relatively limited attention, and therefore become the topic of the present thesis. The rest of this chapter is organized as follows: in Section 1.2 we will give a literature review on classical critical layers and baroclinic critical layers. In Section 1.3, we will introduce the astrophysical protoplanetary disks, strato-rotational instability

and the self-replicating zombie vortex. They are the physical background and context for our theory of baroclinic critical layers. In Section 1.4, we will introduce the scope of our present study based on the literature review.

1.2 Critical layers

1.2.1 Classical critical layers

Since the singularity of critical levels of steady, linear waves in inviscid shear flow is unphysical, one needs to include weak effects of viscosity, unsteadiness or non-linearity to study the critical layers. But how to add these effects strongly depends on the physical context, and yields very different theories. We now briefly review some of the most typical paradigms for the classical critical layer, and then we will consider their secondary instabilities.

The first paradigm is the critical layer in unstable modes. If the flow is unstable, then the phase velocity $c = c_r + ic_i$ is complex, so $U = c$ cannot be satisfied and in this way, the singularity is avoided. For both the Rayleigh's equation and the Taylor-Goldstein equation, the phase velocity c_r of an unstable mode lies in the range of the basic flow velocity [8, 9], so for weak instability with small c_i , a critical layer will arise as long as the basic flow velocity is continuous, which makes critical levels important in instability problems.

A second paradigm is the viscous critical layer, which is primarily applied to study the neutrally stable modes in viscous flows. Representative works on this topic are Tollmien [10] and Lin [11], and a comprehensive summary has been given by Drazin and Reid [12]. It is found that a weak viscosity ν can remove the singularity, resulting in a viscous critical layer with length scale $\nu^{\frac{1}{3}}$. The critical layer matches the outer inviscid solutions on two sides but also imposes a weak viscous correction to the latter.

A third paradigm is the critical layer of wave packets. In this paradigm, researchers use the ray-tracing method to track a localized wavepacket near a critical level. Bretherton [13] investigated internal gravity waves propagating in vertical background shear, and found that when the wavepacket approaches the critical level, the vertical group velocity reduces to zero and the wavepacket will never

reach the critical level. Meanwhile, the vertical wavenumber and the wave energy density becomes larger and larger, until the linear theory finally breaks down.

A fourth paradigm is the nonlinear steady critical layer. Benney and Bergeron [14] resolved the singularity by including nonlinearity in the critical layer, and found that the streamline of the critical layer features a Kelvin cat's eye. As a result of nonlinearity, the phase velocities of the modes are related to the disturbance amplitude. Harberman [15] further considered the steady nonlinear critical layer with viscosity. As the ratio of the effects of viscosity to nonlinearity increases, Harberman [15]'s theory connects the nearly inviscid nonlinear critical layer of Benney and Bergeron [14] and the viscous linear critical layer of Lin [11] and Tollmien [10].

A fifth paradigm is the unsteady critical layer under external forcing. A typical setup for this paradigm is wind or current flowing over a wavy boundary, which continuously forces waves. This problem has been studied by Booker and Bretherton [16] and Brown and Stewartson [17] for internal gravity waves, and by Stewartson [1] and Warn and Warn [2] for Rossby waves. For both types of critical layers, the forcing establishes a steady wave response through the bulk of the flow, but the critical layer remains unsteady: the wave amplitude grows algebraically, and the critical layer becomes thinner.

The secular growth of the wave amplitude will finally make the critical layer nonlinear. For internal gravity waves, Brown and Stewartson [17], [18] and [19] have shown that over a long time, the critical layer will change from a pure wave absorber to a partial wave reflector. For Rossby waves, Stewartson [1] showed that over a very long time, the critical layer has the tendency to evolve into the nonlinear steady critical layer predicted by Benney and Bergeron [14], while Warn and Warn [2] computed the results numerically and demonstrated the pattern of a distorted Kelvin cat's eye.

The last paradigm we introduce is the nonlinear critical layer of free modes (*i.e.*, modes in absence of forcing). Representative works for this paradigm are Hickernell [20], Goldstein and Leib [21], Churilov and Shukhman [22], and a review on this topic has been recently given by Wu [23]. It is found that the bulk of the flow can be represented by a linear normal mode. The critical layer, however, is nonlinear and it controls the slow evolution of the normal mode amplitude

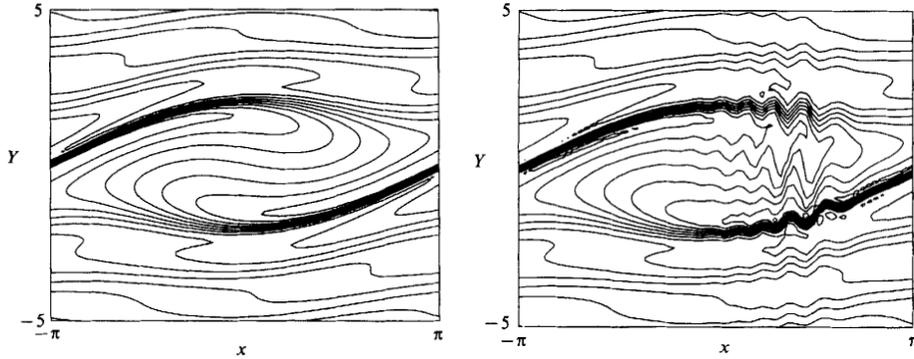


Figure 1.1: Contours of absolute vorticity of Stewartson [1] and Warn and Warn [2]’s solution for the nonlinear critical layer of Rossby waves (left) and its secondary instability (right). Source: Haynes [3].

through the solvability condition. Spatial evolution in the streamwise direction has also been considered in this paradigm.

A critical layer usually has locally strong shear. The shear velocity often represents an unstable basic flow, which will give rise to new instabilities. Such new instabilities are referred to as the ‘secondary instabilities’. Killworth and McIntyre [24] have shown that Stewartson [1] and Warn and Warn [2]’s solution for the nonlinear evolution of Rossby wave critical layer admits a short-wavelength instability, and they have computed the linear growth rates at typical locations of the Kelvin cat’s eye. The instability is a two-dimensional shear instability and the localized vorticity gradient bears significant similarities with the defect instability studied by Gill [25] and Balmforth et al. [26]. Haynes [3, 27] computed the subsequent nonlinear evolution of the secondary instability till the saturation of amplitude, and showed that prominent small-scale structures will appear, which cause significant rearrangement of vorticity in the critical layer. An example of the results of Haynes [3] is shown in figure 1.1.

For critical layers of internal gravity waves, the experimental research of Thorpe [28], Koop [29] and Koop and GcGee [30] revealed that an overturning secondary instability will take place, resulting in turbulent mixing in the critical layer. Lin et al. [31] have undertaken an instability analysis for the critical layer and found that the unstable modes include both the Kelvin-Helmholtz instability and the

Rayleigh-Taylor instability. However, they used idealized Gaussian defects for both velocity and density profiles instead of the actual flow of the critical layer, so their results may be less realistic compared to studies on secondary instabilities of Rossby wave critical layers.

1.2.2 Baroclinic critical layers

We next review the existing theories for baroclinic critical layers. In general, baroclinic critical layers require three fundamental elements, a density stratification, a horizontal shear and a vertical variation of the linear disturbance. The methods to study baroclinic critical layers are generally similar to the classical counterpart, but since the topic has received less attention, some of the paradigms we listed in classical critical levels are absent in literature, to our best knowledge.

For baroclinic critical layers of unstable modes, we have the study of Umurhan et al. [32]. A strong localized vorticity defect was seeded in the linear shear flow, which makes the flow unstable in a similar way to the defect instability [25, 26]. The unstable mode then admits two baroclinic critical layers due to the density stratification. Umurhan et al. [32] have revealed the detailed structure of the baroclinic critical layer, and studied how the stratification and vertical wavenumber can affect the instability. Through a weakly nonlinear analysis, they also showed that a jet may form in the baroclinic critical layer similar to what was found in the simulation of Marcus et al. [6]. However, we note that the fundamental ingredient for the instability is still the mean-flow defect rather than the baroclinic critical layers, since instability can arise in absence of the latter but not the former.

The trapping of wavepackets for baroclinic critical layers has been studied by Olbers [33], Basovich and Tsimring [34] and Badulin et al. [35]. The conclusion obtained was very similar to those of Bretherton [13], *i.e.*, the wavepacket slows down and never reaches the critical level; meanwhile, the vertical wavenumber and wave amplitude grow without bound within the linear theory, suggesting wave breaking or the commence of dissipation effects. Staquet and Huerre [36] and Edwards and Staquet [37] performed numerical simulations to study the nonlinear evolution during trapping, concluding that the trapped waves may either break into small-scale turbulence or be dissipated viscosity and diffusion.

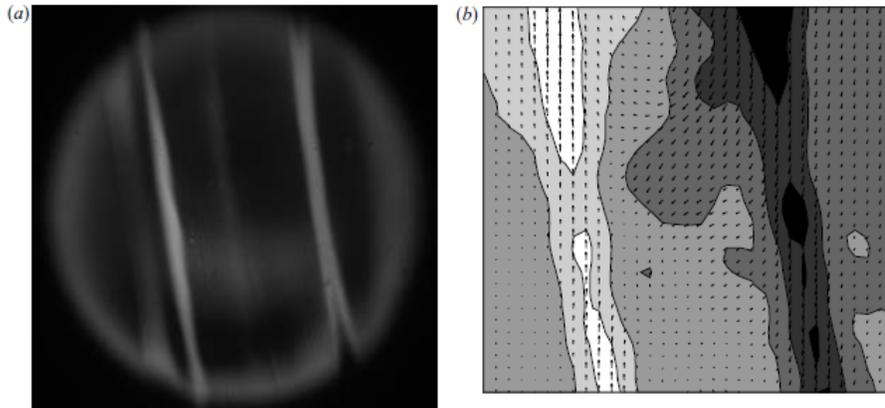


Figure 1.2: Baroclinic critical layers of a tilted vortex in the experiment of Boulanger et al. [4]. Figure (a) is the density contract and figure (b) is the velocity field with the grey scale representing the vertical velocity. The pattern of tilted lines in the two figures are the baroclinic critical layers. Source: Boulanger et al. [4].

The viscous baroclinic critical layer has been studied by Boulanger et al. [4], in the context of tilted stratified vortex. The baroclinic critical layers were first observed from their experiments. In a tank of stratified fluids, they rotated a flap and forced a tilted vortex. They observed that two light stripes in the density field appear progressively, which were then identified to be the baroclinic critical layers (figure 1.2). By analyzing the inviscid steady flow of the tilted vortex, they found the singularity of the critical level. Then they introduced viscosity and obtained the local solution for the viscous critical layer.

Boulanger et al. [38] further showed that their baroclinic critical layers of the tilted vortex also suffer from secondary instability. In their experiments, they found that when the Reynolds number is high, sinusoidal undulations develop in the critical layer, which will then evolve into vortices and create a zig-zag structure. The vortices break down at later stage and result in turbulent mixing. After a detailed analysis on the experiments, they found the instability is a Kelvin-Helmholtz instability caused by the strong shear of the axial velocity in the critical layer. Based on the observation from experiments, they made a number of assumptions and did an instability analysis. The results of the dispersion relation bear qualitative agree-

ment with experimental measurements.

1.3 Protoplanetary disk, strato-rotational instability and zombie vortex instability

The instability of a stratified shear flow may play a key role in a variety of problems from geophysics to astrophysics. The most classical stability theory considers a horizontal mean flow that is sheared in the vertical plane. The stability problem is then described by the Taylor-Goldstein equation, with the celebrated Richardson-number criterion providing an important stability condition [9, 39]. A different paradigm has arisen more recently: the stability of flow that is sheared in the horizontal plane and rotating about a vertical axis. This configuration is relevant to oceanic currents, accretion disks and stratified Taylor-Couette experiments, and it is also the setting for the baroclinic critical layers. The current thesis has a stronger relevance to the protoplanetary disks, which we will briefly introduce. Then we will introduce two potentially important instabilities in the propo-planetary disks: the strato-rotational instability (SRI) and the zombie vortex instability (ZVI).

1.3.1 Protoplanetary disk

In astrophysics, scientists found that before a new star is born, there is a very thin, and very large disk composed of gas and dust rotating around the young star at the center. Such a disk is called a protoplanetary disk, and it plays a key role in the formation of stars and planets. An observational image of the protoplanetary disk is shown in figure 1.3. Over time, the gas and dust of the protoplanetary disk will gradually accrete onto the young star on the center, and the angular momentum will be transferred outward. At the end, the disk will disappear and a new star is formed. More details of the evolution of the protoplanetary disk can be found in the review of Williams and Cieza [40].

Since the gas and dust can remain on their Keplerian orbits, there is a question of how the disk can accrete. Since the disk is very large, viscous effects are too weak to transfer the angular momentum. Researchers then argue that there must be some instability in the disk that drives the angular momentum transport. Many possible instabilities have been proposed, including the magneto-rotational insta-

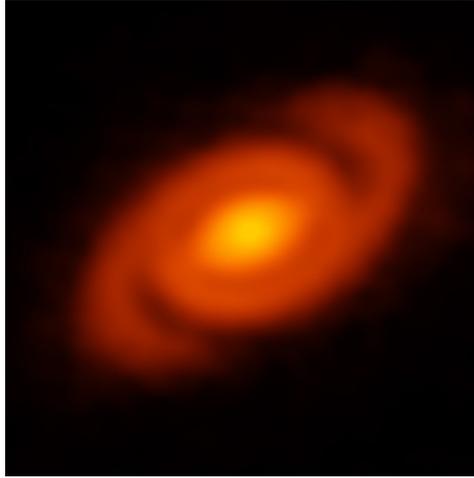


Figure 1.3: Protoplanetary disk surrounding the young star Elias 2-27.
Source: <https://www.eso.org/public/images/potw1640a/>.

bility, gravitational instability, convective over instability, horizontal shear instability, vertical shear instability, wind-driven accretion, strato-rotational instability and zombie vortex instability. A review on these instabilities and their potential roles in the accretion of protoplanetary disks is given by Lyra and Umurhan [41]. The magneto-rotational instability is the most popular candidate (*cf.* Balbus and Hawley [42]). However, some disks are too cool to be ionized [43], so the magneto-rotational instability may be inactive. Thus, what instability drives the accretion and angular momentum transport of these cool disks still remains an open question. The baroclinic critical layers we study in this thesis are closely related to the strato-rotational instability and the zombie vortex instability.

1.3.2 Strato-rotational instability

It is well known that for a non-stratified fluid, centrifugal instability may take place when the Rayleigh's instability criterion for inviscid fluid is satisfied: the square of angular momentum decreases with radius. For a stratified fluid, Bounov, Gledzer & Hopfinger (1995) have shown by both theoretical analysis and experiment that the centrifugal instability is weakened by stratification, which one may expect. However, recently researchers found that a stratified fluid can become unstable in

the centrifugally stable regime, which indicates that stratification destabilizes the flow. This new instability has been given the name ‘strato-rotational instability’, or SRI in short.

SRI was first reported in the letter of Molemaker et al. [44]. They subsequently made a comprehensive instability analysis [45], and showed that its unstable modes are non-axisymmetric and have large amplitudes near the two boundaries. These features are distinct to the Taylor vortices at the onset of centrifugal instability, which are axisymmetric and stronger in the center. The mechanism of SRI was found to be the resonance between two neutral modes localized near the two boundaries, and thus in the dispersion relation, unstable modes lie on narrow wavebands where the phase velocities of the two modes match. Molemaker et al. [44] also undertook a parameter study which indicates that instability is suppressed by increasing the ratio of rotation to shear, and instability persists for viscous flows. The studies of Yavneh *et al.* emphasized geophysical applications. Dubrulle et al. [46] studied SRI in context of astrophysics, *i.e.* considering the basic rotational shear flow approximating Keplerian rotation, and obtained similar instability behavior (with some differences in the effect of parameters). Dubrulle et al. [46] first named such instability as ‘strato-rotational instability’, and proposed that it may be the instability that drives the accretion of protoplanetary disks.

Most researchers study strato-rotational instability in Taylor-Couette flow, and much effort has been devoted to find the instability regime [44, 47–50]. From these studies, we can summarize that strato-rotational instability is present in Taylor-Couette flow as long as the inner and outer cylinders do not rotate at the same angular velocity, but it can be weak and requires strong stratification when the angular velocities of the two cylinders are similar. When the flow is centrifugally unstable, the centrifugal instability is generally stronger than SRI, but exceptions may appear when the flow has strong stratification or viscosity.

SRI has also been studied experimentally by Le Bars and Le Gal [5], Ibanez et al. [51] and Park et al. [52]. Le Bars and Le Gal [5] first observed SRI in experiments: they demonstrated that when the flow is centrifugally stable, periodic non-axisymmetric waves develop in stratified fluids, but there is no instability in non-stratified fluids. An example of their result is shown in figure 1.4. Ibanez et al. [51] broadened the parameter range to much faster rotation and much stronger

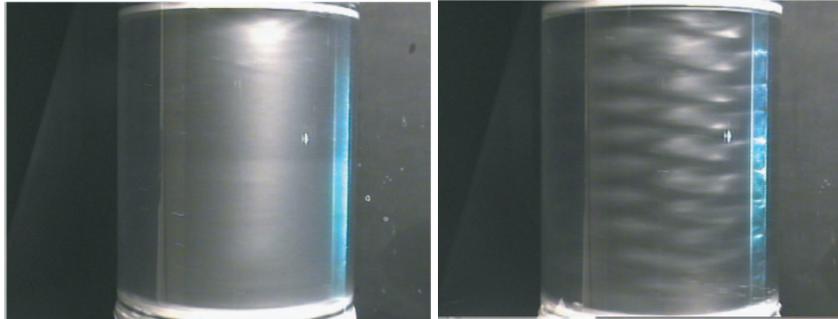


Figure 1.4: Experiment of Le Bars and Le Gal [5] to demonstrate the strato-rotational instability (SRI). The Taylor-Couette flow is in the centrifugally stable regime. The left figure is for pure water, and there is no instability. The right figure is for stratified salt water, and waves of SRI develop. Source: Le Bars and Le Gal [5].

stratification, and found that when the stratification is very strong, instability may develop when the two cylinders rotate at similar angular velocities. They also reported when the rotation is very fast, a new instability which is non-periodic and generates rapid mixing will appear. Park et al. [52] did experiments in the centrifugally unstable regime and observed strato-rotational instability and centrifugal instability simultaneously. They distinguished the two instabilities by the structures of the unstable modes and found good agreements with theoretical predictions.

In stratified plane Couette flow, instabilities similar to the strato-rotational instability have been discovered by Vanneste and Yavneh [53] and Facchini et al. [54]. Although they did not refer to such instabilities as strato-rotational instability, the hallmarks of resonance between neutral modes, strong amplitudes near the boundaries and narrow unstable wavebands are evident. As shown by Facchini et al. [54], only horizontal shear and density stratification are sufficient for instability in this configuration; planetary rotation is not an essential ingredient.

Regarding the relevance to the accretion of protoplanetary disks, the main criticism on the strato-rotational instability is that it requires two rigid boundaries to support the two neutral modes [6]. Since a protoplanetary disk has no rigid boundary (though it certainly has a free boundary), it becomes unclear whether strato-rotational instability can be present. In this context, Le Dizès and Billant [55] and

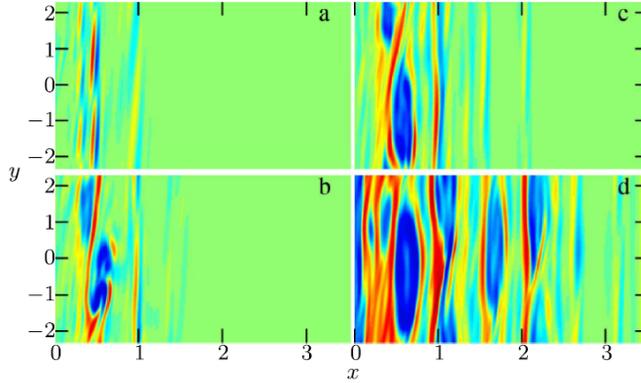


Figure 1.5: Zombie vortices of Marcus et al. [6]. The color scale represents the vorticity field. Vortices trigger baroclinic critical layers and replicate as time increases from figure (a) to figure (d). Source: Marcus et al. [6].

Billant and Le Dizès [56] have shown that instabilities may also arise in concentric vortex when the outer boundary is removed. The instability is caused by free radiation of waves to the far field, and thus such instability is named as ‘radiative instability’. Le Dizès and Riedinger [57] have shown that strato-rotational instability is connected to radiative instability in the parameter space when the outer wall is moved outward.

The configurations of both stratified Taylor-Couette flow and planar Couette could have baroclinic critical layers. However, in the studies of strato-rotational instability in the literature, their parameter sets make the baroclinic critical levels lie outside the domain, so the role of the baroclinic critical level is still unknown.

1.3.3 Zombie vortex instability

In 2013, Marcus et al. [6] reported a new type of instability in rotating stratified shear flows involving self-replicating vortices. Initially, one vortex is seeded in the periodic box, and it triggers two baroclinic critical layers. The baroclinic critical layers then roll up into new vortices, which then trigger a new generation of baroclinic critical layers. This self-replicating process is repeated until vortices finally fill the entire domain. They named the replicating vortices ‘zombie vortices’ and the instability as ‘zombie vortex instability’ (ZVI). They argued that the zom-

bie vortex instability is very possibly what drives the accretion of protoplanetary disks. An example of the zombie vortices is shown in figure 1.5.

To demonstrate the crucial role of zombie vortex instability in the protoplanetary disks, Marcus and his colleagues undertook more comprehensive investigation in the astrophysical context. Marcus et al. [58] simulated the zombie vortex instability in the Keplerian flow using the anelastic approximation. They argued that zombie vortices can emerge from random noise through an inverse cascade. The replication process and the pattern of the final turbulent flow look similar to those excited by a single vortex in Marcus et al. [6]. They also revealed that the energy of the zombie vortices is extracted from the mean flow. Marcus et al. [59] studied the threshold of the initial disturbance amplitude to trigger the zombie vortex instability. They found the threshold was determined by the initial vorticity rather than velocity, and then they determined the critical Mach number for the initial noise of the Kolmogorov spectrum. Barranco et al. [60] studied zombie vortices in nonuniform stratification which is more realistic in protoplanetary disks. They showed that the turbulence produced by the zombie vortex instability can propagate into the midplane of the disk where there is no stratification. They also argued that zombie vortices can persist in presence of radiative damping in the astrophysical environment.

1.4 Scope of the present thesis

The present thesis focuses on baroclinic critical layers in waves and instabilities of rotating stratified shear flow. In particular, we hope to fill the blanks of this topic in the existing literature that we have reviewed.

Our first objective is to study baroclinic critical layers in strato-rotational instability. Previous theoretical studies of strato-rotational instability all have their baroclinic critical levels outside the domain. Therefore, it is interesting to ask: what happens when a baroclinic critical level moves in? Will it change the resonance condition, or can it induce a different kind of instability? We will answer these questions in Chapter 2.

Our second goal is to study the inviscid unsteady baroclinic critical layer induced by external forcing. This will be undertaken in Chapter 3, where we will

study both its linear and nonlinear evolution. For classical critical layers, we have reviewed that this paradigm has been studied by Stewartson [1] and Warn and Warn [2] for Rossby wave critical layers, and Booker and Bretherton [16] and Brown and Stewartson [17] for critical layers of internal gravity waves. This paradigm is still absent for baroclinic critical layers, so it will be interesting to compare it to the classical counterpart. Specially, the baroclinic critical levels depend on wavenumbers while the classical critical level does not. This property suggests baroclinic critical layers should have important differences in the nonlinear evolution compared to classical critical layers.

Similar to other critical layers, the forced baroclinic critical layer also suffers from secondary instability, and we will consider this topic in Chapter 4. Since the baroclinic critical layer itself evolves, the problem is challenging but interesting as we will study the instability of unsteady flows. We will solve the problem taking advantage of the theory of defect instability, and obtain an asymptotic solution for the instability of an evolving basic state.

An important application of our theory of forced baroclinic critical layers is the zombie vortex instability. Marcus *et al.* discovered various interesting phenomena from their numerical simulation, but questions still remain as to the underlying mechanism of the replication process. The theoretical analysis of Umurhan *et al.* [32] has revealed that a defect in the mean flow will induce unstable modes with baroclinic critical layers, which throws some light upon the replication mechanism. But since the initial vortex of Marcus *et al.* [6] seems to have marginal mean-flow modification, one may ask where the mean-flow defect comes from. Our theory will be able to unravel part of the mystery. In Chapter 3, we will show that the baroclinic critical layer will force mean-flow responses featuring a local defect, and then in Chapter 4, we will show that the defect generates a secondary instability which may account for the replication of zombie vortices.

In Chapter 5, we will summarize the conclusions of the thesis and propose some directions of future research.

Chapter 2

Strato-rotational instability associated with baroclinic critical layers¹

2.1 Introduction

Strato-rotational instability or ‘SRI’ can arise in stratified flows with a horizontal shear and a planetary rotation. It is relevant to oceanic currents, accretion disks and stratified Taylor-Couette experiments. SRI is different from Rayleigh’s centrifugal instability (which is conventionally argued to arise when the square of the angular momentum of a rotating flow decreases with radius), and has been ascribed to resonant mode interactions [5, 44–49, 51]. For the geometry of a channel of finite width, the conditions for resonance are restrictive, demanding that instability appears only for specific narrow bands of horizontal wavelength for which the phase velocities of the two uncoupled modes are nearly the same. (If the outer boundary is removed, and waves can freely radiate, the band of unstable wavelengths becomes continuous, as shown by [55] and [56]; see also [57], who establish the connection between the resonant and radiative instabilities as the position of the outer boundary diverges.)

¹The work of this chapter has been published in the Journal of Fluid Mechanics [61].

In the existing theory of SRI, the role of baroclinic critical levels has usually been ignored. The parameter ranged examined is such that the baroclinic critical levels all lie outside the domain in previous studies, so their impact on the linear instability have not been studied in any detail. Here, our goal is to explore how the presence of baroclinic critical levels affects the resonant mode interaction of SRI. For the task, we consider the model Couette channel flow of Dubrulle et al. [46] and Vanneste and Yavneh [53] (amongst others). We use a combination of short-wavelength asymptotics and numerical solution of the linear eigenvalue problem, and examine how instability can arise through wave-mean-flow interaction from the perspective of conservation of pseudo-momentum. The model flow and governing equations, the normal-mode equation and the baroclinic critical level are described in Section 2.2; the eigenvalue problem is solved in Section 2.3 for relatively strong stratification, and then in Section 2.4 for relatively weak stratification. The instability mechanism is examined via the concept of pseudomomentum in Section 2.5, and we conclude in Section 2.6.

2.2 Mathematical formulation

2.2.1 Model and governing equations

The geometry of the model flow is described by the Cartesian coordinate system (x^*, y^*, z^*) sketched in figure 2.1. The basic Couette flow $(\Lambda y^*, 0, 0)$ is horizontal, where Λ is the shear rate, and bounded by walls at $y^* = 0$ and $y^* = L$. The channel rotates around the z^* -axis at the rate $f/2$, and the fluid is stratified with buoyancy frequency N . Here, the ‘*’ notation is used to distinguish dimensional variables; to render the equations dimensionless, we use scales for velocity, time, length, density and pressure of ΛL , $1/\Lambda$, L , $\rho_0 \Lambda^2 L/g$, and $\rho_0 \Lambda^2 L^2$, respectively, where ρ_0 is a reference density and g is gravity. Thus, we set $(x, y, z) = L^{-1}(x^*, y^*, z^*)$ and so forth, with the dimensionless variables appearing without the star decoration. The dimensionless rotation rate and Richardson number are $\Omega = f/2\Lambda$ and $R = N^2/\Lambda^2$. When $\Omega > 0$ ($\Omega < 0$), the shear is anticyclonic (cyclonic). The Rayleigh discriminant of the base flow is $\Phi = 2\Omega(2\Omega - 1)$; we consider the regime where the flow is centrifugally stable, $\Phi > 0$ ($\Omega > 1/2$ or $\Omega < 0$).

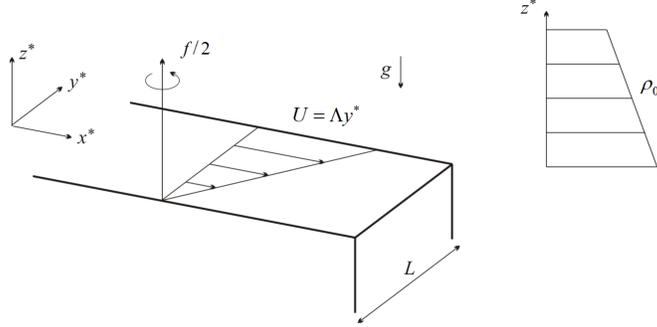


Figure 2.1: Sketch of the model geometry.

In Boussinesq approximation, the linearized governing equations for the dimensionless perturbations to the velocity field $[u(x, y, z, t), v(x, y, z, t), w(x, y, z, t)]$, pressure $p(x, y, z, t)$, and density $\rho(x, y, z, t)$ are

$$u_t + yu_x + (1 - 2\Omega)v = -p_x, \quad (2.1)$$

$$v_t + yv_x + 2\Omega u = -p_y, \quad (2.2)$$

$$w_t + yw_x + \rho = -p_z, \quad (2.3)$$

$$\rho_t + y\rho_x - R w = 0, \quad (2.4)$$

$$u_x + v_y + w_z = 0, \quad (2.5)$$

where the subscripts denote partial derivatives. The boundary conditions are that there is no flow through the walls:

$$v = 0 \quad \text{at} \quad y = 0 \quad \text{and} \quad y = 1, \quad (2.6)$$

and the disturbances are assumed to be periodic in both x and z .

The linearized potential vorticity of the Boussinesq fluid is

$$q = (2\Omega - 1)\rho_z - R(v_x - u_y), \quad (2.7)$$

and satisfies

$$q_t + yq_x = 0. \quad (2.8)$$

For normal modes, this implies $q = 0$. In this circumstance, one can then formulate the conservation law [53]:

$$\frac{d}{dt} \int_0^1 \mathfrak{p} \, dy = 0, \quad \mathfrak{p} = \frac{1}{R} \overline{\rho(w_x - u_z)}, \quad (2.9)$$

where

$$\overline{(\dots)} = \frac{1}{L_x L_z} \int_0^{L_x} \int_0^{L_z} (\dots) \, dx dz. \quad (2.10)$$

To second order in perturbation amplitude, the conserved integral in (2.9) corresponds to the net disturbance momentum in the streamwise direction, or the Eulerian pseudomomentum [62], that is, \mathfrak{p}_t represents the acceleration of the mean flow by a normal mode at the cross-stream position y , but (2.9) demands that the net effect vanishes, which places an important constraint on any unstable normal mode as we outline further in Section 2.5.

2.2.2 Normal modes

We search for normal modes with the form,

$$(u, v, w, p, \rho) = [\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{p}(y), \hat{\rho}(y)] \exp(ikx + imz - i\omega t) + \text{c.c.}, \quad (2.11)$$

where ‘c.c.’ denotes the complex conjugate, k and m are the horizontal and vertical wavenumbers, the (complex) frequency is $\omega = \omega_r + i\omega_i$, and the corresponding streamwise phase speed is $c = \omega/k = c_r + ic_i$. Substitution of this form into the linear equations followed by some algebraic manipulations leads to the eigenvalue equation [53],

$$\frac{d^2 \hat{u}}{dy^2} + h \frac{d\hat{u}}{dy} + l^2 \hat{u} = 0, \quad (2.12)$$

where

$$h = \frac{2\mu^2 k (2\Omega - 1)^2 \hat{\omega}}{(\hat{\omega}^2 - R)[\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2]}, \quad (2.13)$$

$$l^2 = -\lambda^2 = -k^2 - \frac{\mu^2 k^2}{\hat{\omega}^2 - R} \left[\hat{\omega}^2 - \Phi + \frac{2(2\Omega - 1)\hat{\omega}^2}{\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2} \right], \quad (2.14)$$

$$\hat{\omega} = \omega - ky = k(c - y) \quad \text{and} \quad \mu = m/k. \quad (2.15)$$

The boundary conditions on the channel walls become

$$\hat{v} = \frac{m^2(2\Omega - 1)(y - c)\hat{u} + (\hat{\omega}^2 - R)\frac{d\hat{u}}{dy}}{ik[\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2]} = 0 \quad \text{at} \quad y = 0 \quad \text{and} \quad y = 1. \quad (2.16)$$

Because the eigenvalue problem is real, the eigenvalues c and eigenfunctions $\hat{u}(y)$ are either real or appear as complex pairs. The equations have the symmetry, $(y, c) \leftrightarrow (1 - y, 1 - c)$, and so if we have one eigenfunction $\hat{u}(y)$ with eigenvalue c , there is another solution pair, $\hat{u}(1 - y)$ and $1 - c$.

If ω or c is real, equation (2.12) has two singular points at $\hat{\omega}^2 = R$, corresponding to the ‘‘baroclinic critical levels’’,

$$y = y_{b\pm} \equiv c \pm \frac{\sqrt{R}}{k}. \quad (2.17)$$

At these points, there is a regular Frobenius series solution $\hat{u}_R = 1 + O(y - y_{b\pm})$, and logarithmically singular one with $\hat{u}_S = \hat{u}_R \log|y - y_{b\pm}| + O(1)$. Equation (2.12) also possesses two additional singular points where $\hat{\omega}^2 = R + \mu^2(2\Omega - 1)^2$, but these can be shown to be removable [53] and merit no further discussion. Note that the classical critical level, $y = y_c \equiv c$, introduces no singularity in the current model given the constant shear of the basic flow profile.

In the following two sections, we solve the eigenvalue problem (2.12) with (2.16) both numerically and asymptotically to find unstable modes ($c_i > 0$). For the numerical computations, we use a shooting method based on the MATLAB function ODE15s, with trial guesses provided either by previous solutions at different parameter settings or the short-wavelength solutions. The calculated growth rates turn out to be relatively small (*cf.* Vanneste and Yavneh [53]), and in setting up the problem for analysis, it is advantageous to examine the situation where $c_i \rightarrow 0$.

In the limit of short vertical wavelength, $\mu^2 \gg 1$, where the character of the solutions is dictated by the sign of l^2 for $c_i \rightarrow 0$, there are two configurations to

consider depending on the sign of the factor,

$$\varepsilon_b = \frac{2\sqrt{R}(2\Omega - 1)}{k[\mu^2(2\Omega - 1)(\Phi - R) + 2R]}, \quad (2.18)$$

which determines the local behaviour near the baroclinic critical levels: $l^2 \sim \pm[\varepsilon_b(y - y_{b\pm})]^{-1}$. The two configurations are illustrated in figure 2.2; if ε_b is negative, the evanescent region ($l^2 < 0$) enclosing $y = y_c$ is bounded by two turning points $y_{t\pm}$ of the classical WKB-type where $l = 0$. The short-wavelength solutions have oscillatory character for $y_{t+} < y < y_{b+}$ and $y_{b-} < y < y_{t-}$, with the baroclinic critical levels bounding the propagation zones. In Sections 2.3 and 2.4, special attention is given to modes with $\mu^2 \gg 1$. In this situation, our first configuration therefore applies to most geophysical conditions for which $N \gg |f|$ ($R \gg \Phi$). By contrast, when $\varepsilon_b > 0$, the evanescent region is bounded by the baroclinic critical levels $y_{b\pm}$, and the propagation zones occupy $y_{b+} < y < y_{t+}$ and $y_{t-} < y < y_{b-}$. With $\mu^2 \gg 1$, this situation may apply to astrophysical disks where shear and rotation are relatively fast, so that $N \ll |f|$. We consider the case with $\varepsilon_b < 0$ (which we refer to as stronger stratification) in Sections 2.3, and that with $\varepsilon_b > 0$ (denoted weaker stratification) in Section 2.4. Note that, if μ^2 is small or Ω close to $1/2$, the spatial structure of l^2 is more complicated than indicated in figure 2.2, in which case the configurations are less easily classified.

2.2.3 Solutions near the baroclinic critical levels

When we consider the eigenvalue problem in the short-wavelength limit, we take parameter settings such that some of the coefficients of (2.12) become relatively large. These choices also lead us to further analyze the solutions over small region near the critical levels in the limit $|\varepsilon_b| \ll 1$ (which is achieved below by taking $\mu^2 \gg 1$). Taking $y = y_{b+}$ as a way of illustration, we set

$$\eta = \frac{y - y_{b+}}{\varepsilon_b} \quad (2.19)$$

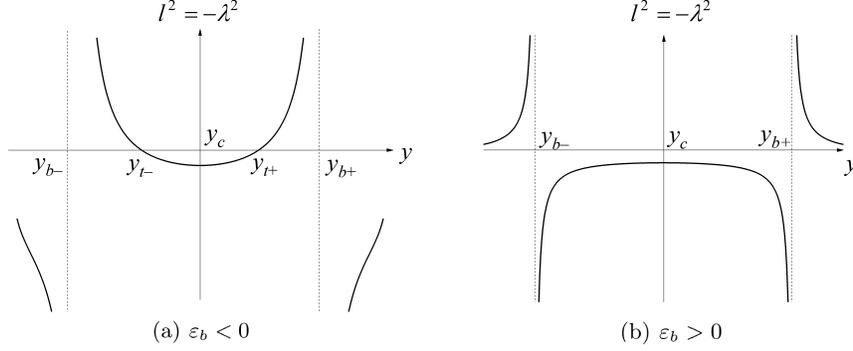


Figure 2.2: Plots of $l^2 = -\lambda^2$ against y , illustrating the two configurations of interest for the spatial arrangements of the classical critical level $y = y_c$, the turning points $y = y_{t\pm}$ and the baroclinic critical levels $y = y_{b\pm}$, which are dictated by the sign of ε_b when $\mu^2 \gg 1$.

to find the leading-order local equation,

$$\frac{d^2 \hat{u}}{d\eta^2} + \frac{1}{\eta} \frac{d\hat{u}}{d\eta} + \frac{1}{\eta} \hat{u} = 0. \quad (2.20)$$

The two independent solutions of (2.20) are Bessel functions,

$$\hat{u}_I = \begin{cases} I_0(2\sqrt{-\eta}) & \text{Re}(\eta) \leq 0, \\ J_0(2\sqrt{\eta}) & \text{Re}(\eta) \geq 0. \end{cases} \quad (2.21a)$$

$$\hat{u}_K = \begin{cases} K_0(2\sqrt{-\eta}) & \text{Re}(\eta) \leq 0, \\ -\frac{\pi}{2} [Y_0(2\sqrt{\eta}) + \text{isgn}(\varepsilon_b) J_0(2\sqrt{\eta})], & \text{Re}(\eta) \geq 0, \end{cases} \quad (2.21b)$$

where we use a particular choice for the branch cut of the square root. In particular, we take the branch cut for $\sqrt{y - y_{b+}}$ to lie along the negative real axis, which leads to the inclusion of the sign of ε_b in (2.21b), assuming that $c_i > 0$ (for $c_i < 0$ we replace this expression with the complex conjugate of \hat{u}_K). Note that \hat{u}_I

corresponds to the regular Frobenius solution \hat{u}_R , and \hat{u}_K to the singular one \hat{u}_S . The existence of c_i , however, regularizes \hat{u}_S since η cannot be zero with real y but complex y_{b+} . \hat{u}_I grows exponentially away from the critical level for $\text{Re}(\eta) < 0$, whereas \hat{u}_K decays exponentially into that region; both solutions oscillate in space for $\text{Re}(\eta) > 0$. These observations prepare the way for a match to the usual WKB solutions away from the baroclinic critical levels.

2.3 Instability of flow with stronger stratification

We first consider the case $\varepsilon_b < 0$, which we refer to as a more strongly stratified flow because the situation is achieved in the short-wavelength limit, $\mu^2 \gg 1$, when $R > \Phi$ (unless Ω approaches $\frac{1}{2}$). As illustrated in figure 2.2(a), we then expect modes to possess a wave-like character for $y_{b-} < y < y_{t-}$ and $y_{t+} < y < y_{b+}$ where $l^2 > 0$. We present results for unstable modes of the traditional SRI type that arise from a resonant coupling and have no baroclinic critical levels ($y_{b-} < 0$ and $y_{b+} > 1$), and then describe how the form of this instability changes when one of these singular levels enters the channel.

Figures 2.3 and 2.4 plot numerically calculated eigenvalues, c_r and $\omega_i = kc_i$, against $\mu = m/k$ for two typical cases with anticyclonic ($\Omega > 0$) or cyclonic ($\Omega < 0$) shear (the specific examples have $R = 9$, $k = 5$ and $\Omega = 5/8$ or $\Omega = -1/8$, implying $\Phi = 5/16$). The two horizontal broken lines in the phase-speed plots of figure 2.3 indicate where a baroclinic critical level appears on the wall at $y = 0$ or $y = 1$: the modes between the two lines do not have baroclinic levels in the domain; the modes above the line ‘ $\text{Re}(y_{b-}) = 0$ ’ have the critical level y_{b-} , and the modes below the line ‘ $\text{Re}(y_{b+}) = 1$ ’ have the critical level y_{b+} .

In addition to the normal mode solution shown in figures 2.3 and 2.4, the eigenvalue problem (2.11)-(2.11) also has ‘continuous spectrum’ $c = y$ and $c = y \pm N/k$ where $y \in [0, 1]$ varies continuously in the domain (*cf.* (4.16) and figure 4.2 in Chapter 4). As μ increases, new normal modes will separate from the continuous spectrum, and hence the number of modes increases with μ as we see in figure 2.3.

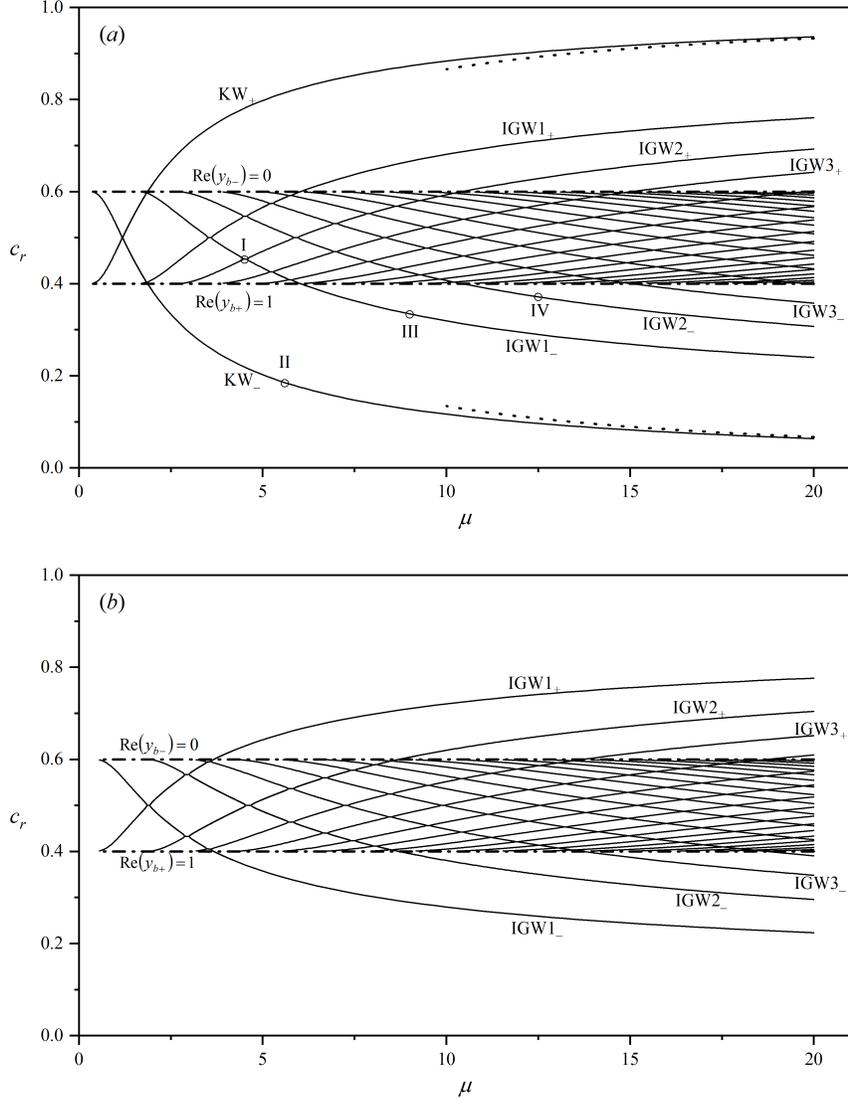


Figure 2.3: Phase speeds against $\mu = m/k$ for $R = 9$ and $k = 5$, with (a) $\Omega = 5/8$ (anti-cyclonic shear) and (b) $\Omega = -1/8$ (cyclonic shear). Eigenfunctions at the points I and II-IV are plotted in figures 2.5 and 2.6, respectively. The dotted line in figure (a) shows the asymptotic solution (2.28). The horizontal broken lines indicate the phase speeds for which modes develop baroclinic critical levels at the channel walls ($Re(y_{b-}) = 0$ and $Re(y_{b+}) = 1$).

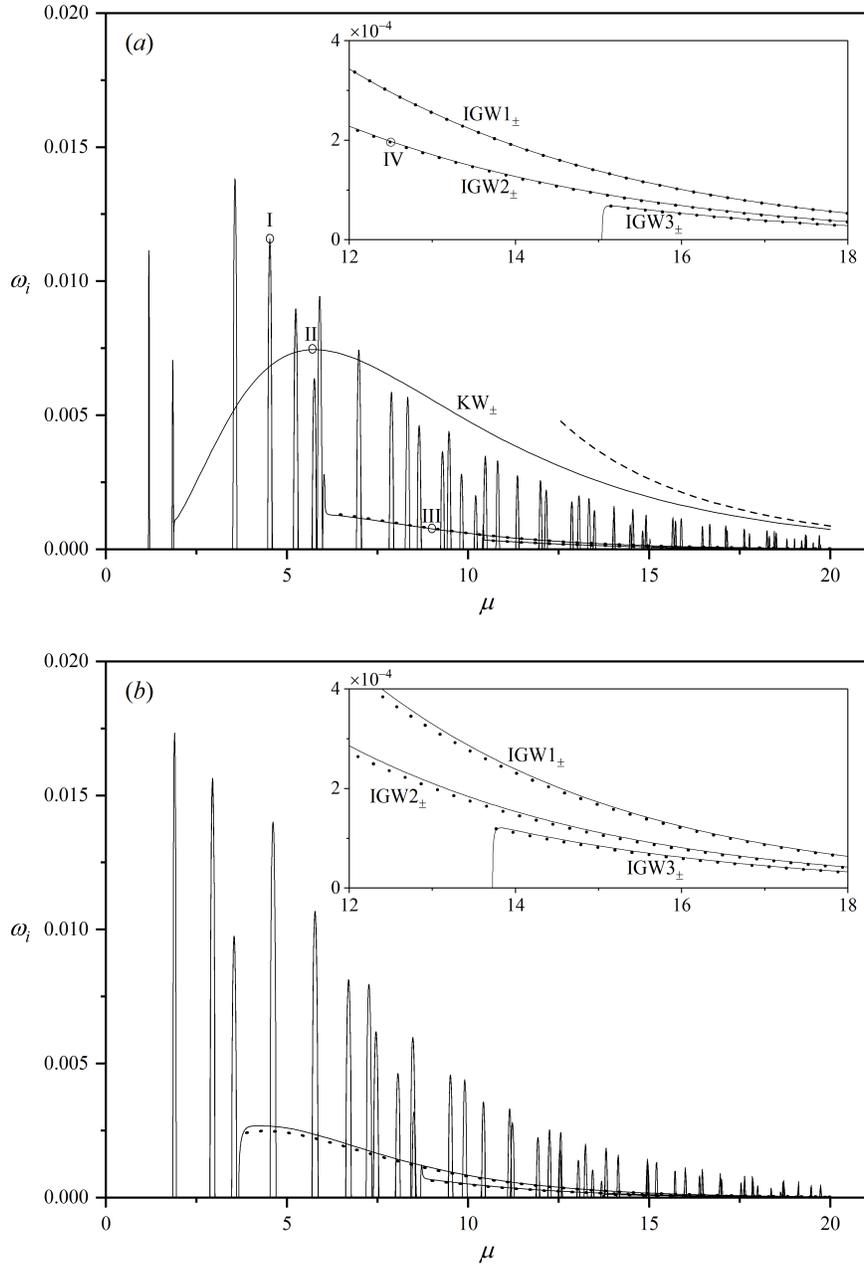


Figure 2.4: Caption on next page.

Figure 2.4: (previous page) Unstable growth rates against $\mu = m/k$ for $R = 9$ and $k = 5$, with (a) $\Omega = 5/8$ (anti-cyclonic shear) and (b) $\Omega = -1/8$ (cyclonic shear). The insets show magnifications for higher μ and plot only the growth rates of the unstable modes generated by interaction with the baroclinic critical level. The dashed line in figure (a) shows the asymptotic solution (2.38); the dots in both panels show the asymptotic growth rates computed from (2.37). Eigenfunctions at the points I and II-IV are plotted in figures 2.5 and 2.6, respectively.

2.3.1 Resonant instabilities

When there are no baroclinic critical levels, modes can be classified as either Kelvin waves (KW) or internal gravity waves (IGW); see Vanneste and Yavneh [53]. The former are localized to one of the channel walls and decay exponentially into the interior of the flow; the latter are confined primarily to one of the propagation zones within the channel. We denote the Kelvin wave localized to the wall at $y = 0$ ($y = 1$) by KW_- (KW_+). Similarly, an internal gravity wave confined to the wave cavity on the left (right) of $y = \text{Re}(y_c)$ is referred to as IGW_- (IGW_+). There is only a single Kelvin wave associated with each of the channel walls, but the wave cavities support an infinite number of internal gravity waves; we distinguish the latter by adding an integer, $IGW_{n\pm}$, corresponding to the number of nodes in $\text{Re}(\hat{u})$ within the cavity. The main difference between the cases with anti-cyclonic and cyclonic shear is the presence of Kelvin modes with phase speeds within the range of the mean flow for the former.

As summarized by Vanneste and Yavneh [53], resonant SRI arises when the phase speed of different KW_{\pm} and IGW_{\pm} modes lock together. This leads to a multitude of narrow bands of unstable wavenumbers, as illustrated by the sharp peaks in growth rate in figure 2.4. There are no unstable modes in the gaps between these bands. A sample unstable resonant mode is shown in figure 2.5, which corresponds to the interaction of IGW_{2+} and IGW_{1-} (and indicated by point I in figures 2.3 and 2.4).

The details of the neutral modes and their unstable resonant interaction can be understood using WKB theory: in the region near $y = y_c$, $-l^2 = \lambda^2 > 0$ and the

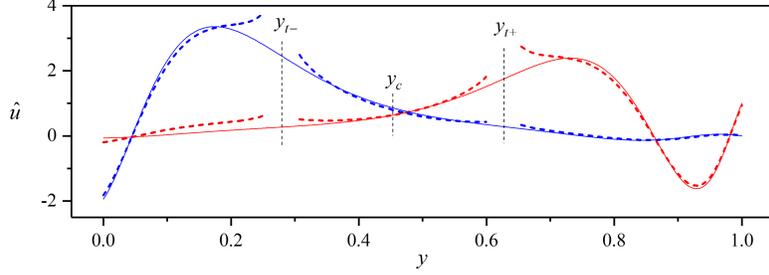


Figure 2.5: Eigenfunction \hat{u} of the resonant unstable mode corresponding to the coupling of IGW $_{2+}$ and IGW $_{1-}$ (labelled by the point I in figures 2.3(a) and 2.4(a); $\mu = 4.523$). Red and blue lines represent real and imaginary part of the eigenfunction, respectively. The numerical solution is shown in solid lines, and has eigenvalue $c = 0.454 + 0.00232i$. The asymptotic eigenfunction is shown in dashed lines, and is constructed using the computed asymptotic eigenvalue $c \sim c_0 + \Delta c \approx 0.453 + 0.00240i$ together with the formulae (2.22)–(2.25), which become inaccurate at the turning points $y = y_{t\pm}$.

decaying and growing solutions have (respectively) the exponential forms,

$$\hat{u}_- = \frac{A}{\sqrt{\lambda}} \exp \left[\int_y^{y_{t+}} \lambda(y') dy' \right], \quad \hat{u}_+ = \frac{A}{\sqrt{\lambda}} \exp \left[\int_{y_{t-}}^y \lambda(y') dy' \right] \quad \text{for } y_{t-} < y < y_{t+}, \quad (2.22)$$

where

$$A = \exp \left[-\frac{1}{2} \int h(y) dy \right] = \sqrt{\left| \frac{\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2}{\hat{\omega}^2 - R} \right|}. \quad (2.23)$$

Beyond the turning points ($y > y_{t+}$ or $y < y_{t-}$, with $l^2 > 0$), we apply the usual WKB turning-point connection formulae to find the oscillatory solutions (*cf.* Bender and Orszag [63]),

$$\hat{u}_{\pm} = \begin{cases} \frac{2A}{\Psi\sqrt{l}} \cos \left[\int_{y_{t\pm}}^y l(y') dy' \mp \frac{\pi}{4} \right], & y \gtrless y_{t\pm}, \\ \frac{A}{\sqrt{l}} \cos \left[\int_{y_{t\mp}}^y l(y') dy' \mp \frac{\pi}{4} \right], & y \lesseqgtr y_{t\mp}, \end{cases} \quad (2.24)$$

where

$$\Psi = \exp \left[- \int_{y_{r-}}^{y_{r+}} \lambda(y) dy \right]. \quad (2.25)$$

Note that both \hat{u}_- and \hat{u}_+ are real for real c , indicating that they are both standing waves over this second region.

Given their respective exponential decay (\hat{u}_- becomes exponentially small in $y > c$ and \hat{u}_+ in $y < c$), the uncoupled \pm modes therefore have phase speeds c_{\pm} that are given only by the boundary condition $\hat{v}(y, c) = 0$ at either $y = 1$ or $y = 0$, that is,

$$\hat{v}_-(0, c_-) = 0, \quad \hat{v}_+(1, c_+) = 0, \quad (2.26a, b)$$

where

$$\hat{v}_{\pm}(y, c) = \frac{m^2(2\Omega - 1)(y - c)\hat{u}_{\pm} + (\hat{\omega}^2 - R)\frac{d\hat{u}_{\pm}}{dy}}{ik[\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2]} \quad (2.27)$$

denote the WKB eigenfunctions for the cross-stream velocity component v . For KW_{\pm} , a further reduction for $\mu \gg 1$ and $\Omega > 1/2$ with $c - y = O(\mu^{-1})$ furnishes the convenient expressions,

$$c_+ \approx 1 - \frac{1}{m} \sqrt{\frac{2\Omega R}{2\Omega - 1}} \quad \text{and} \quad c_- \approx \frac{1}{m} \sqrt{\frac{2\Omega R}{2\Omega - 1}}, \quad (2.28a, b)$$

which are also plotted in figure 2.3(a).

Resonant mode coupling arises when the phase velocities of two basic modes c_- and c_+ converge. The coupled modes comprise comparable proportions of \hat{u}_- and \hat{u}_+ , and so the WKB solution for \hat{u} is expressed by the linear combination,

$$\hat{u} = C_- \hat{u}_- + C_+ \hat{u}_+, \quad (2.29)$$

for two constant C_{\pm} . The boundary conditions then become

$$C_- \hat{v}_-(0, c) + C_+ \hat{v}_+(0, c) = C_- \hat{v}_-(1, c) + C_+ \hat{v}_+(1, c) = 0. \quad (2.30)$$

implying the eigenvalue condition

$$\hat{v}_-(0, c) \hat{v}_+(1, c) = \hat{v}_+(0, c) \hat{v}_-(1, c). \quad (2.31)$$

The left-hand side of (2.31) combines the two leading-order dispersion relations in (2.26), whereas the right-hand side is exponentially small in view of the decay of \hat{v}_\pm towards $y = 0$ and $y = 1$, respectively. At resonance, the decoupled relations in (2.26) are satisfied simultaneously by $c_+ = c_- = c_0$. Treating the right-hand side of (2.31) as a small perturbation to the left-hand side (and neglecting the higher-order terms of WKB approximation (2.22) and (2.24) that also correct $\hat{v}_-(0, c)$ and $\hat{v}_+(1, c)$, but which are not expected to lead to instability as they are real), we can then estimate a correction Δc to the leading-order eigenvalue c_0 , namely

$$(\Delta c)^2 \approx \frac{\hat{v}_+(0, c_0)\hat{v}_-(1, c_0)}{\hat{v}_{-,c}(0, c_0)\hat{v}_{+,c}(1, c_0)}, \quad (2.32)$$

where the subscript ‘ c ’ denotes a derivative of the eigenfunction with respect to the eigenvalue. In figure 2.5, the leading-order WKB eigenfunction and the corrected eigenvalue are compared satisfyingly with the corresponding numerical results for the sample mode of that figure. Note that, according to (2.30),

$$\frac{C_-}{C_+} = -\frac{\hat{v}_+(0, c)}{\hat{v}_-(0, c)} \approx -\frac{\hat{v}_+(0, c_0)}{\hat{v}_{-,c}(0, c_0)\Delta c} \quad (2.33)$$

which, because Δc is purely imaginary for instability, implies a $\pi/2$ phase difference between the two coupled basic modes, as seen in figure 2.5.

2.3.2 Instability induced by the baroclinic critical level

As illustrated in figures 2.3 and 2.4, when modes develop baroclinic critical levels, the eigenvalue diagrams change dramatically in two key ways. First, the KW_- and $IGWn_-$ (KW_+ and $IGWn_+$) basic modes disappear above (below) the central band, $1 - \sqrt{R}/k < c_r < \sqrt{R}/k$, where they develop the y_{b-} (y_{b+}) critical level. This removes the possibility of any resonant mode interactions whenever there are baroclinic critical levels. Second, the same KW_- and $IGWn_-$ (KW_+ and $IGWn_+$) basic modes persist below (above) the central band, $1 - \sqrt{R}/k < c_r < \sqrt{R}/k$, even though they now develop the y_{b+} (y_{b-}) critical levels. More importantly, they become unstable over a wide band of wavenumbers. In figure 2.4(a), for the anti-cyclonic shear, three continuous lines of growth rate thereby appear for the KW_\pm , $IGW1_\pm$ and $IGW2_\pm$ modes; the growth rate of the $IGW3_\pm$ modes is also finite, but

too small to be observed in the main panels of the figure. The situation is the same for the cyclonic shear in figure 2.4(b), although the KW_{\pm} modes are not present.

These observations can be translated to stability conditions in the limit of short vertical wavelength: resonant SRI requires the existence of the central band, or $k < 2\sqrt{R}$. For the $IGWn_{\pm}$ modes, we must further demand that $0 < y_{t-}$ and $y_{t+} < 1$ in order for wavelike regions to exist. But for $\mu \gg 1$, $y_{t\pm} \approx c \pm \sqrt{\Phi}/k$. Thus, SRI between $IGWn_{\pm}$ modes, which is the only possibility for cyclonic shear, requires $2\sqrt{\Phi} < k < 2\sqrt{R}$ (cf. Park and Billant [48]). For anti-cyclonic shear, the Kelvin waves (with $c_- > 0$ or $c_+ < 1$ for $\mu \gg 1$) can participate in unstable interactions, which modifies the condition to $k < 2\sqrt{R}$. By contrast, the instability of a KW_- or $IGWn_-$ mode associated with a baroclinic critical layer requires $y_{b+} = c + \sqrt{R}/k < 1$ (or equivalently, $y_{b-} = c - \sqrt{R}/k > 0$ for a KW_+ or $IGWn_+$ mode). Hence, given that $c_- > 0$ for KW_- and $c_- > \sqrt{\Phi}/k$ for $IGWn_-$, the baroclinic critical layer instability requires $k > \sqrt{R}$ for anticyclonic shear and $k > \sqrt{R} + \sqrt{\Phi}$ for cyclonic shear.

Sample unstable modes are shown in figure 2.6, which plots the eigenfunctions of KW_- , $IGW1_-$ and $IGW2_-$ at the points denoted II, III and IV in figures 2.3(a) and 2.4(a). For each mode, we see sharp structure in the eigenfunction near the baroclinic critical level y_{b+} . To the right of these points, the mode amplitudes decay rapidly; to the left, the real part and imaginary parts have a $\pi/2$ phase difference indicative of travelling waves (propagating in the direction indicated by the arrows in figure 2.6).

In WKB theory, the solution for the ‘ \pm ’ mode is again dominated by the contribution from \hat{u}_{\pm} near either $y = 1$ or $y = 0$. To leading order, the phase velocity $c_r = c_{\pm}$ is again then given by one of (2.26). Focussing for the moment on the ‘ $-$ ’ mode, the solution is evanescent to the right of the critical level y_{b+} . Provided the right-hand wall is sufficiently far from y_{b+} , the solution is then

$$\hat{u} = \frac{AC}{\sqrt{\lambda}} \exp \left[- \int_{y_{b+}}^y \lambda(y') dy' \right], \quad y_{b+} < y \quad (2.34)$$

with C being an arbitrary constant. The exponential decay in (2.34) is fast given than $\lambda \gg 1$ near the critical level, as seen in figure 2.6.

To the left of $y = y_{b+}$, we exploit the large argument limits of Bessel function

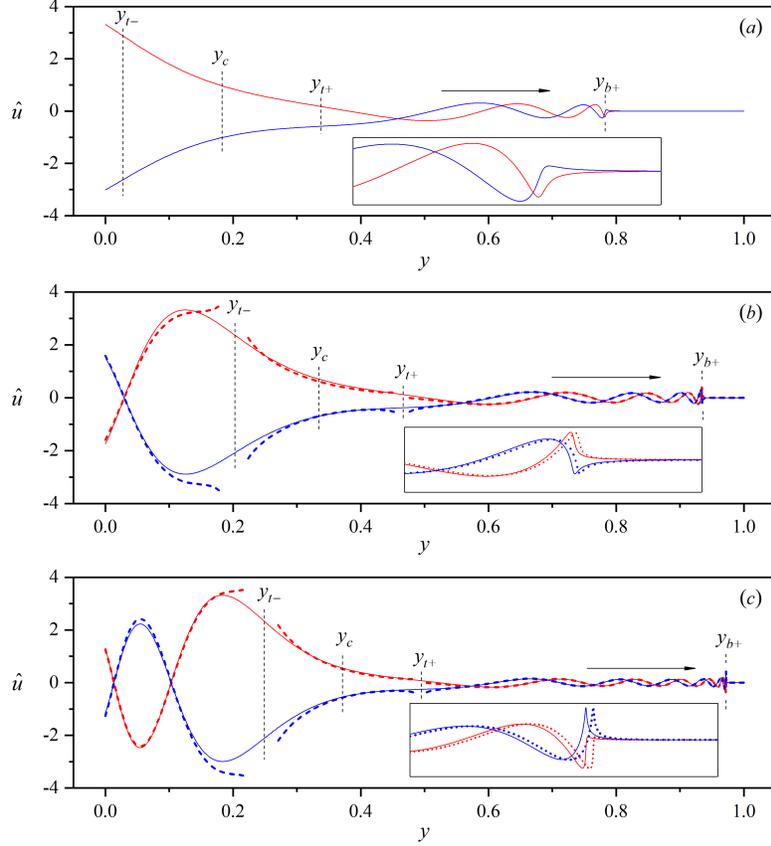


Figure 2.6: Eigenfunction \hat{u} of unstable modes with baroclinic critical levels for $\Omega = 5/8$ and (a) KW_- ($\mu = 5.7$; point II in figures 2.3 and 2.4), (b) IGW1_- ($\mu = 9$; point III), and (c) IGW2_- ($\mu = 12.5$; point IV). Red and blue lines represent real and imaginary part of the eigenfunction, respectively. The solid lines show the numerical solutions. The corresponding results from WKB theory are shown as dashed lines, which become inaccurate at the turning points $y = y_{t\pm}$ (the asymptotic eigenfunction is not shown in (a) as the comparison is poor owing to the turning point $y = y_{t-}$ close to the boundary $y = 0$). In (b), numerical computations give $c = 3.34 \times 10^{-1} + 1.54 \times 10^{-4}i$ whereas the asymptotics indicate $c = 3.35 \times 10^{-1} + 1.56 \times 10^{-4}i$. In (c), the numerical and asymptotic results are $c = 3.71 \times 10^{-1} + 3.96 \times 10^{-5}i$ and $c = 3.72 \times 10^{-1} + 3.92 \times 10^{-5}i$, respectively. Insets show magnifications of the structure near the baroclinic critical levels.

of the local solutions (2.21) to make the connection, which gives

$$\hat{u} = \frac{AC}{\sqrt{l}} \exp \left\{ i \left[\int_y^{y_{b+}} l(y') dy' + \frac{\pi}{4} \right] \right\}, \quad y_{t+} < y < y_{b+}. \quad (2.35)$$

That is, the eigenfunction now takes the form of a travelling wave over this region, unlike the standing-wave basic mode (which is exponentially small here). Given that the intrinsic frequency $\hat{\omega} = k(c - y)$ is negative in this region, the travelling waves propagate in $+y$ direction as indicated by the arrow in figure 2.6; *i.e.* they are incident on the critical level.

Over the same region, $y_{t+} < y < y_{b+}$, the solution is also given by the WKB solution (2.29), and so to match with (2.35) we must take

$$C_- = iC\Theta, \quad C_+ = \frac{1}{2}C\Psi\Theta, \quad \Theta = \exp \left[i \int_{y_{t+}}^{y_{b+}} l(y) dy \right]. \quad (2.36)$$

At $y = 0$, the leading-order boundary condition is $C_- \hat{v}_-(0, c) \sim 0$, giving $c \sim c_-$. An estimate of the correction Δc to this eigenvalue then follows on keeping the next order terms of $C_- \hat{v}_-(0, c) + C_+ \hat{v}_+(0, c) = 0$ and using (2.36):

$$\Delta c \approx - \frac{C_+ \hat{v}_+(0, c_-)}{C_- \hat{v}_{-,c}(0, c_-)} = \frac{i}{2} \Psi \frac{\hat{v}_+(0, c_-)}{\hat{v}_{-,c}(0, c_-)} \quad (2.37)$$

(ignoring higher-order corrections to the WKB approximation). The predictions computed from (2.37) for the IGW n_{\pm} modes are included in figure 2.4 and again compare well with numerical computations. The WKB eigenfunction \hat{u} and eigenvalue $c \sim c_- + \Delta c$ also match well with the numerical solutions for the two cases with higher μ^2 shown in figure 2.6(b,c) (the caption reports the numerical and asymptotic results for c).

In the limit $\mu \gg 1$, we can again derive an explicit expression for KW $_-$:

$$c_i \approx c_- \exp \left(-2 \int_0^{y_{t+}} \lambda(y) dy \right). \quad (2.38)$$

The prediction (2.38) is compared with the numerical results in figure 2.4(a).

2.4 Instability of flow with weaker stratification

If $\varepsilon_b > 0$, we expect that the evanescent region around $y = \text{Re}(y_c)$ is bounded by the baroclinic critical levels $y_{b\pm}$ (see figure 2.2(b)); for $\mu^2 \gg 1$, this case is realized for $\Phi > R$, and so corresponds to a more weakly stratified flow. In this situation, the exponential tail of a Kelvin wave riding on one of the channel walls can become transformed into internal waves beyond the adjacent critical level. If these internal waves can satisfy the boundary conditions at the other channel wall, a new type of normal mode is generated. Moreover, the interaction with that critical level destabilizes the mode *via* a mechanism that we shed light on later. The phase speed and growth rate of such solutions are plotted against μ in figure 2.7, with two sample eigenfunctions shown in figure 2.8.

In figure 2.7(a), the line ‘ $\text{Re}(y_{b-}) = 0$ ’ lies below that for ‘ $\text{Re}(y_{b+}) = 1$ ’ because $\sqrt{R}/k < 0.5$ and the two Kelvin wavespeeds never intersect. Moreover, the critical levels removes any neutral internal waves with speeds inside the range of the base flow (which can be understood from the connection conditions across $y = y_{b\pm}$ implied by the local solutions derived in Section 2.2.3: \hat{u}_K possesses an exponentially decaying amplitude that corresponds to a travelling wave, which cannot satisfy the boundary condition). Thus, resonances do not occur, precluding traditional SRI. Instead, instability is generated purely through the interaction of the Kelvin waves with the baroclinic critical level. As seen in figure 2.8, to the right of y_{b+} the unstable mode KW_- takes the form of a standing wave, and there is an abrupt phase change across $y = y_{b+}$.

For a short-wavelength description of the unstable KW_\pm modes, we once more express the eigenfunction as a linear combination of decaying and growing solutions in the evanescent region, $\hat{u} = C_- \hat{u}_- + C_+ \hat{u}_+$, with constants C_- and C_+ . Focussing again on KW_- , in the evanescent region $0 \leq y < y_{b+}$, \hat{u}_- and \hat{u}_+ are expressed by

$$\hat{u}_- = \frac{A}{\sqrt{\lambda}} \exp \left[\int_y^{y_{b+}} \lambda(y') dy' \right], \quad \hat{u}_+ = \frac{A}{\sqrt{\lambda}} \exp \left[- \int_y^{y_{b+}} \lambda(y') dy' \right], \quad 0 \leq y < y_{b+}, \quad (2.39)$$

with A given by (2.23). Matching (2.39) to the the local solution in (2.21), we can

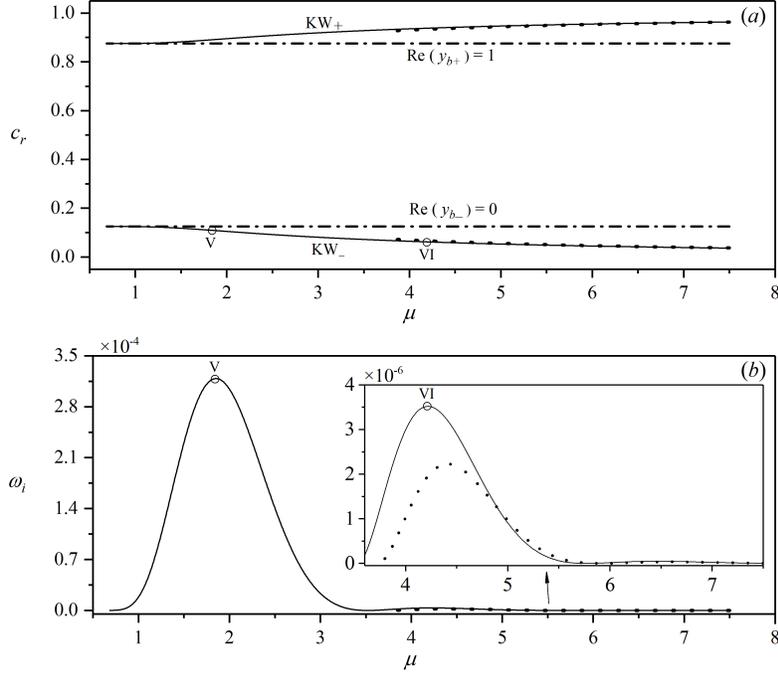


Figure 2.7: (a) Phase speed c_r and (b) growth rate ω_i against $\mu = m/k$ for $R = 2.5 \times 10^{-3}$, $\Omega = 5/8$ and $k = 0.4$. The horizontal broken lines indicate where baroclinic critical levels appear at the channel walls. The dots plot the asymptotic results in (2.28) and (2.43) which we do not include for the first peak in the main panel of (b) because the comparison is poor owing to the relatively small value for μ . Eigenfunctions at the points V and VI are plotted in figure 2.8.

derive their corresponding expressions in $y > y_{b+}$:

$$\hat{u}_- = \frac{2A}{\sqrt{l}} \cos \left[\int_{y_{b+}}^y l(y') dy' - \frac{\pi}{4} \right], \quad \hat{u}_+ = \frac{A}{\sqrt{l}} \exp \left[-i \int_{y_{b+}}^y l(y') dy' - \frac{i\pi}{4} \right],$$

$y_{b+} < y \leq 1. \quad (2.40)$

The phase jump of \hat{u}_+ across the critical layer rationalizes that seen in the eigenfunction in figure 2.8.

For $\mu \gg 1$, the boundary condition (2.16) on $y = 1$ implies $\hat{u} = 0$ to leading

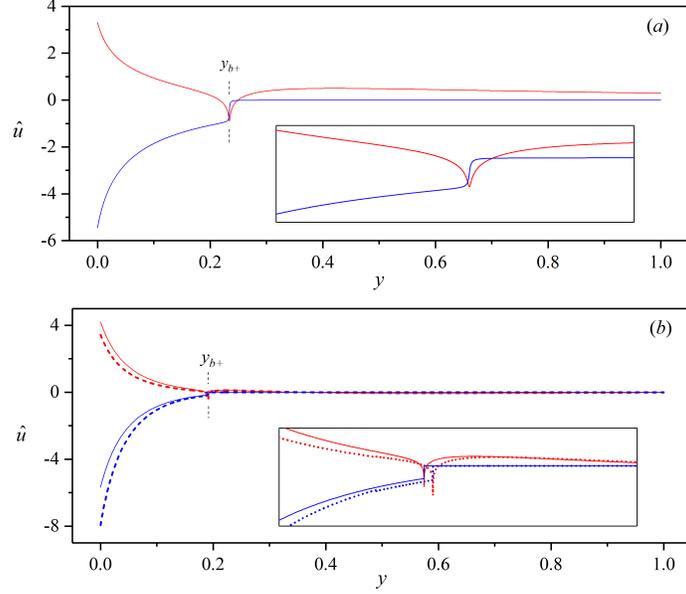


Figure 2.8: Eigenfunction \hat{u} of unstable modes with baroclinic critical levels at the points marked (a) V ($\mu = 1.85$) and (b) VI ($\mu = 4.2$) in figure 2.7. Red and blue lines represent real and imaginary parts of the eigenfunction, respectively. The solid line shows the numerical solution and the dashed line in (b) is the result of asymptotics (the asymptotic eigenfunction is not shown in (a) as the comparison is poor owing to the relatively small value for μ). The insets show magnifications of the structure near the baroclinic critical levels. In (b) the numerical result for the eigenvalue is $c = 6.19 \times 10^{-2} + 8.82 \times 10^{-6}i$ and $c = 6.66 \times 10^{-2} + 5.30 \times 10^{-6}i$ for the asymptotics.

order, so the WKB solution in the oscillatory region $y_{b+} < y \leq 1$ is the standing gravity wave,

$$\hat{u} = \frac{AC}{\sqrt{l}} \sin \int_y^1 l(y') dy' \quad \text{for } y_{b+} < y \leq 1, \quad (2.41)$$

where C is a constant. Reconciling (2.41) with the superposition of \hat{u}_{\pm} in (2.40) implies

$$C_- = -\frac{C}{2} \exp(-i\theta), \quad C_+ = C \sin \theta, \quad \theta = \int_{y_{b+}}^1 l(y) dy + \frac{\pi}{4}. \quad (2.42)$$

At the left boundary $y = 0$, the solution is dominated by \hat{u}_- , leading to a basic phase speed c_- which is again expressed by (2.28b). The contribution of \hat{u}_+ is exponentially small, and keeping the next-order correction Δc to the eigenvalue that it induces (again neglecting the potentially higher-order WKB approximations), provides the estimate,

$$c_i \approx -\text{Im} \left[\frac{C_+ \hat{v}_+(0, c_-)}{C_- \hat{v}_-, c(0, c_-)} \right] \approx 4c_- \cos^2 \left[\int_{y_{b+}}^1 l(y) dy - \frac{\pi}{4} \right] \exp \left[-2 \int_0^{y_{b+}} \lambda(y) dy \right]. \quad (2.43)$$

The prediction (2.43) is included in figure 2.7. Note the sinusoidal dependence of c_i in (2.43) which is also evident in this figure.

2.5 Implications of pseudomomentum conservation

For the normal-mode form (2.11), the pseudomomentum in (2.9) becomes

$$\mathbf{p} = \hat{\mathbf{p}}(y) e^{2\omega_i t}, \quad \hat{\mathbf{p}}(y) = \frac{(\mu \hat{u}^* - \hat{w}^*) \hat{w}}{y - c} + \text{c.c.}, \quad \hat{w} = -\frac{\mu \hat{\omega} \left[(2\Omega - 1) \frac{d\hat{u}}{dy} - k \hat{\omega} \hat{u} \right]}{k[\hat{\omega}^2 - R - \mu^2(2\Omega - 1)^2]}, \quad (2.44)$$

where the $*$ superscript denotes complex conjugation. Thus the conservation law (2.9) reduces to

$$2\omega_i \int_0^1 \hat{\mathbf{p}}(y) dy = 0. \quad (2.45)$$

This condition places no constraint on the spatial structure of neutral modes with $\omega_i = 0$. However, unstable modes with $\omega_i \neq 0$ must have zero total pseudomomentum $\int_0^1 \hat{\mathbf{p}}(y) dy = 0$. This demands that the spatial structure of the modes, as characterized by the density $\hat{\mathbf{p}}(y)$, must break down into cancelling contributions from regions of opposite sign of pseudomomentum.

The pseudomomentum density $\hat{\mathbf{p}}$ is plotted against y in figure 2.9(a) for a mode destabilized by the resonant coupling of traditional SRI, corresponding to the eigenfunction of figure 2.5. As pointed out by Vanneste and Yavneh [53], the $\text{IGW}n_{\pm}$ basic modes have a \mp sign for their pseudomomentum, whereas the signature of KW_{\pm} depends on Ω : the KW_{\pm} modes have a \mp signature if $\Omega > 1/2$, and a \pm one when $\Omega < 0$. The coupled mode in figure 2.5 and 2.9(a) consists of an IGW_- component that is concentrated near the left wall with positive pseudomomentum

(implying the local mean flow is accelerated in the positive x -direction), and an IGW_+ component confined near the right wall with negative pseudomomentum (accelerating the local mean flow in the negative x -direction). Unstable resonance is achieved when their net contributions are balanced.

Figures 2.9(b,c) show the density $\hat{\rho}$ for KW_- and IGW_- modes destabilized by their interaction with baroclinic critical levels. The modes have very different distribution of pseudomomentum: $\hat{\rho}$ is again locally large and positive near the left wall reflecting the Kelvin wave or internal gravity wave localized there. The net contribution of this wave is balanced by that from a sharp negative peak around the baroclinic critical level.

The pseudomomentum contribution of the baroclinic critical level can be understood by considering the local behaviour of the solutions in the vicinity of $y = y_{b\pm}$. As outlined in Section 2.2.2, $d\hat{u}/dy \sim a(y - y_{b\pm})^{-1}$ for the singular Frobenius solution, with a a constant. Thus, $\hat{\rho} \sim O(y - y_{b\pm})^{-2}$ near $y = y_{b\pm}$ according to (2.44). An estimate for the contribution of the critical level $y_{b\pm}$ therefore follows as a sharply localized acceleration of the mean flow:

$$\mp \frac{2\sqrt{R}k|a|^2}{m^2(2\Omega - 1)^2} \int_{\text{Re}(y_{b+})-\Delta}^{\text{Re}(y_{b+})+\Delta} \frac{dy}{|y - y_{b\pm}|^2} \approx \mp \frac{2\pi\sqrt{R}k|a|^2}{m^2(2\Omega - 1)^2|c_i|}, \quad (2.46)$$

where the range of integration identifies a local region with $\Delta \ll 1$ but $\Delta/c_i \gg 1$ over which the pseudomomentum is large, and a direct evaluation of the Lorentzian integrand furnishes the right-hand side of (2.46).

It is now apparent how the basic modes KW_- and IGW_{n-} can become unstable by coupling with the y_{b+} critical level (or the KW_+ and IGW_{n+} by interacting with y_{b-}). The signatures of the basic modes and the critical-level contributions also rationalize how other types of modal instabilities cannot appear.

2.6 Concluding remarks

In this chapter, we have studied the linear instability of three-dimensional rotating stratified shear flow, with a focus on how baroclinic critical levels impact the strato-rotational instability (SRI). We have found that the baroclinic critical levels destroy the resonance mode interactions between the internal gravity and Kelvin waves that

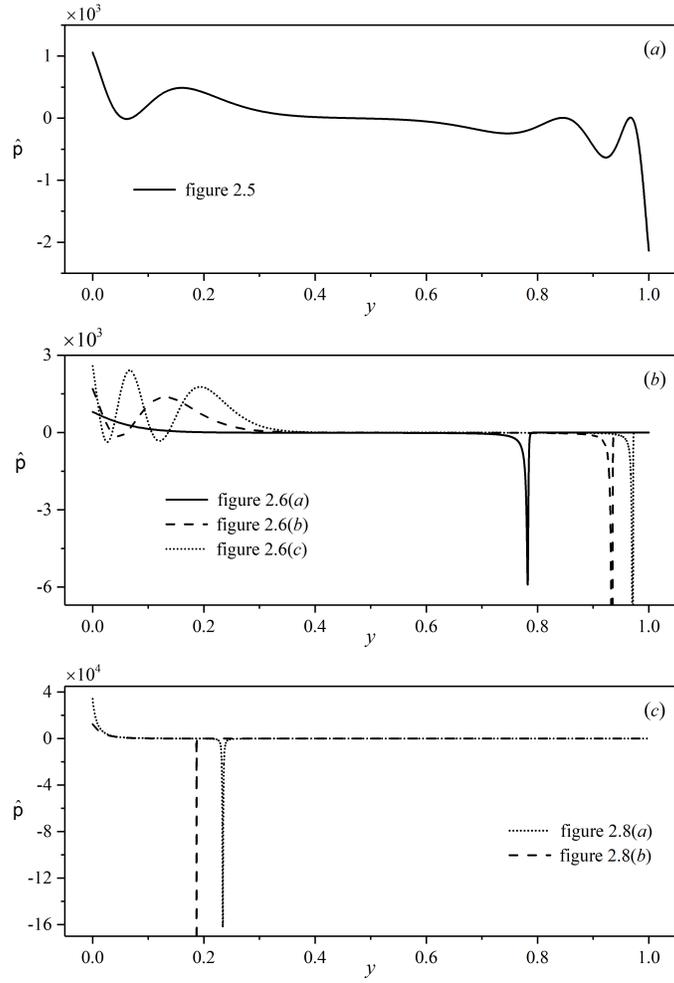


Figure 2.9: The pseudomomentum density \hat{p} in (2.44) for (a) the resonant SRI mode in figure 2.5, and (b,c) the KW_- and IGW_- modes destabilized by its interaction with a baroclinic critical level at $y = y_{b+}$ shown in figure 2.6 and figure 2.8. Note that the sharp negative peaks at the baroclinic critical levels have been truncated in (b,c) to help reveal the mode structure.

underpin traditional SRI. Instead, these waves can become unstable by interacting directly with one of the critical levels. An immediate consequence is that the instability no longer requires any resonance conditions on the modal wavenumber, but occupies an extensive continuous waveband. The situation is somewhat like how the classical critical level can affect the linear instabilities of shear flows in shallow water when there are potential vorticity gradients [64, 65].

To establish these results, we combined numerical computations of the linear eigenvalue problem with short-wavelength asymptotics. We further examined the pseudomomentum of the unstable modes to shed further light on how the mode interactions could lead to instability. This quantity must exactly vanish for unstable modes in the stability problem we have studied, and traditional SRI comes about through the resonant interaction of modes with different signatures of pseudomomentum. For modes destabilized by their interaction with a baroclinic critical level, on the other hand, a thin layer surrounding this level provides a sharp source of pseudomomentum to balance that of the mode. In other words, the critical level acts like an emitter of the relevantly signed pseudomomentum.

A key limitation of the present analysis is the assumption of a linear velocity profile, which is unlikely in any physical setting. Once the background vorticity gradient is not zero, the classical critical layer can directly affect the dynamics of the normal modes, whereas here this position plays a more subtle role (our unstable modes all require classical critical levels within the domain to set the stage for destabilizing interactions). In the common configuration of Taylor-Couette flow, classical critical levels do not appear to play a significant role. For example, in the study of Park and Billant [48], there is no observable singular behavior near the classical critical level, nor does it render instability in absence of resonance. This may be because the vorticity gradient of the Taylor-Couette flow is relatively weak. But in other flows with stronger vorticity gradients and more complicated shear, for example, in the Lamb–Oseen vortex of the radiative instability of Le Dizès and Billant [55], the classical critical level plays a crucial role in the instability through an ‘over-reflection’ phenomenon. In the present thesis, we will also encounter the classical critical level of secondary instability in Chapter 4, where we have strong mean-flow defect forced by the baroclinic critical layer.

Previous studies on classical critical levels often proceed beyond inviscid lin-

ear theory and incorporate nonlinearity and viscosity in view of the relatively high amplitudes and sharp gradients that develop nearby. Here too there are large velocity gradients inside the thin region around the baroclinic critical level, demanding that both viscosity and nonlinearity are important ingredients in “baroclinic critical layer” theory. We leave such a theory for SRI for future work, but in Chapter 3, we will study in detail the effects of nonlinearity and dissipation on forced baroclinic critical layers.

Chapter 3

Nonlinear dynamics of forced baroclinic critical layers¹

3.1 Introduction

In Chapter 2, we studied the baroclinic critical layers in unstable modes, which are the eigen modes of a homogeneous system. In this chapter, we embark on a different paradigm: we will study the baroclinic critical layers in waves generated by external forcing. This paradigm is related to wind or current following over topography, waves generated by an oscillating source, or waves forced by localized vortices as in the zombie vortex problem that we are most interested in.

For classical critical layers, Stewartson [1] and Warn and Warn [2, 67] studied the nonlinear dynamics of the critical layers of forced Rossby waves. They found that steady waves developed over the bulk of the shear flow, but that the critical layer remained unsteady, exciting mean-flow corrections and all the harmonics of the original wavenumber, and twisting up the background vorticity into Kelvin cat's eye pattern. A similar scenario exists for the critical layers of internal gravity waves travelling vertically through stratified shear flow, with important repercussions on wave breaking, momentum transport and mixing in the atmosphere [16–19].

Existing literature on the baroclinic critical layers has mainly focused on the

¹The work of this chapter has been published in the Journal of Fluid Mechanics [66]

propagation of linear wave packets. Using ray-tracing theory, Olbers [33], Basovich and Tsimring [34] and Badulin et al. [35] found that wave packets slow down as they approach the baroclinic critical level, never reaching it. Simultaneously, the wave amplitude and cross-stream wavenumber grow indefinitely, indicating that linear theory eventually fails in a wave-trapping process like that found earlier for classical critical levels [13]. Staquet and Huerre [36] and Edwards and Staquet [37] performed numerical simulations to study the nonlinear evolution during trapping, concluding that the trapped waves may either break into small-scale turbulence or be dissipated by dispersion, viscosity and diffusion. More related to the current work is the study by [4], who explored the analogues of baroclinic critical levels in stratified, tilted vortices, and resolved the singularities by introducing viscosity.

The aim of the present chapter is to theoretically study the evolution of forced baroclinic critical layers, following the paradigm of Stewartson [1] and Warn and Warn [2, 67] for Rossby waves, or Booker and Bretherton [16] and Brown and Stewartson [17, 18, 19] for internal waves in stratified shear flow. The linear dynamics of a forced baroclinic critical layer is expected to be similar to that of a classical critical layer, owing to the similarity of the singularities in the linear wave equations. However, the subsequent nonlinear evolution is likely to be very different because the location of the baroclinic critical level itself is dependant of the streamwise wave number, which is different among all the harmonics of the original wave. This suggests that they cannot feature in the nonlinear dynamics within the baroclinic critical layer, unlike in classical critical layer theory.

The layout of the chapter is as follows: in Section 3.2, we give the model and governing equations of the problem. In Section 3.3, we solve the linear problem explicitly and draw out structure that first develops within the baroclinic critical level. In Section 3.4, we extend the analysis by considering weakly nonlinear perturbations, which allows us to determine the time and length scales that characterize the nonlinear critical layer. This leads us, in Section 3.5, to derive a reduced model of nonlinear dynamics via a matched asymptotic expansion. We then present numerical solutions of the reduced model and a further asymptotic analysis of them. We explore the effects of dissipation in the baroclinic critical layer in Section 3.6, and then discuss the implications of the results and the relation to previous and future work in Section 3.7.

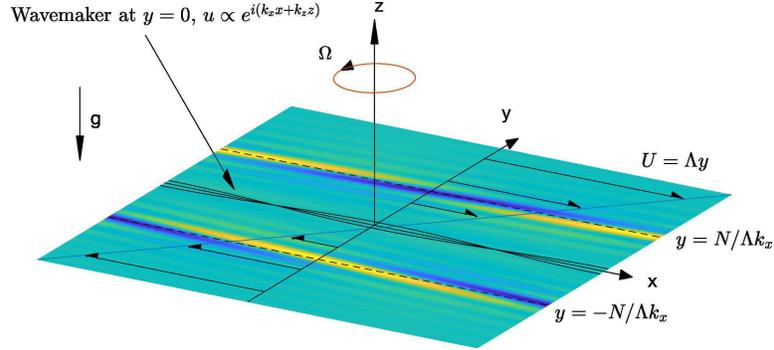


Figure 3.1: Sketch of the model. A wavemaker with wavenumber k_x and k_z is imposed at $y = 0$, and baroclinic critical levels are forced at $y = \pm N/(\Lambda k_x)$, corresponding to dimensionless locations $\pm \mathcal{N}$, where $\mathcal{N} = N\Lambda^{-1}$. The shading represents a rendering of the density perturbation based on the linear theory of Section 3.3.

3.2 Model and governing equations

We consider forced disturbances to an unbounded horizontal shear flow, orientated in the x -direction with a constant shear rate $\Lambda > 0$ in the y -direction. The domain rotates around the vertical axis at angular velocity Ω , and the fluid is stratified in z with constant buoyancy frequency N . Waves are driven into the shear flow by a wavemaker that we locate along $y = 0$. This forcing has the streamwise and vertical wavenumbers, k_x and k_z , respectively. The baroclinic critical levels are located at $y = \pm N/(\Lambda k_x)$. The sketch of the model is shown in figure 3.1.

We work with a dimensionless version of the governing fluid equations in which length, time, velocity, pressure and density perturbations are scaled by k_x^{-1} , Λ^{-1} , Λk_x^{-1} , $\rho_0 \Lambda^2 k_x^{-2}$ and $\rho_0 \Lambda^2 / (k_x g)$, respectively. Here, ρ_0 is a reference density and g is gravity. We employ the Boussinesq approximation and, for the most part of our study, neglect viscosity and diffusion in view of the large spatial scales that characterize geophysical and astrophysical flows. At the end of the work, we briefly explore the effect of diffusion. The perturbations to the velocity (u, v, w) ,

pressure p and perturbation density ρ then satisfy

$$u_t + yu_x + (1 - f)v + uu_x + vu_y + wu_z = -p_x, \quad (3.1)$$

$$v_t + yv_x + fu + uv_x + vv_y + wv_z = -p_y, \quad (3.2)$$

$$w_t + yw_x + uw_x + vw_y + ww_z = -p_z - \rho, \quad (3.3)$$

$$\rho_t + y\rho_x - \mathcal{N}^2 w + u\rho_x + v\rho_y + w\rho_z = 0, \quad (3.4)$$

$$u_x + v_y + w_z = 0, \quad (3.5)$$

where subscripts represent partial derivatives and we have introduced the dimensionless Coriolis parameter $f = 2\Omega/\Lambda$ and buoyancy frequency $\mathcal{N} = N\Lambda^{-1}$. Because our interest lies in the forcing of the baroclinic critical layers of an internal wave, we consider basic flows that are linearly stable to prevent unstable modes from dominating the dynamics. Centrifugal instabilities arise when $0 < f < 1$ [68], so we set $f > 1$ or $f < 0$ to eliminate them; strato-rotational instability is not present because it requires reflective boundaries [45, 61] which are absent here.

Initially, there is no disturbance, implying $u = v = w = \rho = p = 0$ at $t = 0$. The wavemaker is then switched on to excite waves with baroclinic critical levels. To idealize the forcing and formulate a concise mathematical problem, we assume that the wavemaker introduces a time-independent jump in the tangential horizontal velocity at $y = 0$, but not in the normal velocity. That is, we impose the jump conditions,

$$u|_{y=0+} - u|_{y=0-} = \varepsilon_0 \exp(ix + imz) + \text{c.c.}, \quad v|_{y=0+} = v|_{y=0-}, \quad (3.6)$$

where ε_0 represents the strength, $m = k_z/k_x$, c.c. represents the complex conjugate, and the \pm superscripts indicate the limits from either side. This forcing approximates a thin, spatially periodic vortex sheet. In the numerical simulation of Marcus et al. [6], waves were forced by a periodic array of localized Gaussian vortices. Our forcing therefore represents an idealization of their model in that we consider the leading-order Fourier component while neglecting the evolution and cross-stream thickness of the forcing. The configuration is slightly different to that in the studies of Stewartson [1] and Booker and Bretherton [16], where a wavy boundary forced the normal velocity. The current configuration implies that waves are gen-

erated at $y = 0$ and develop with baroclinic critical levels to either side (although simplifications are afforded by the symmetry described presently). Had we placed the wavemaker along a boundary at $y = 0$, only one critical level would have featured, but the wall may also make the basic flow linearly unstable [61]. Other idealizations include wavemakers that gradually switch on [69], that generates disturbances with finite phase speed (displacing the baroclinic critical levels), or that with finite thickness (as for the vortices of Marcus *et al.*). Nevertheless, the precise form of forcing of the wave is not expected to affect the qualitative dynamics of the baroclinic critical layers, a feature on which we elaborate further later.

Note that the system in (3.1)-(3.6) is invariant under the transformation,

$$(u, v, w, \rho) \rightarrow -(u, v, w, \rho) \quad \text{and} \quad p \rightarrow p, \quad \text{for} \quad (x, y, z) \rightarrow -(x, y, z). \quad (3.7)$$

This observation permits us to solve the problem only in $y > 0$, and therefore consider only one baroclinic critical layer, then generate the solution in $y < 0$ using the implied symmetry conditions.

Also, combining (3.1)-(3.5), we may derive an equation for the vertical component of vorticity:

$$\frac{D}{Dt}(v_x - u_y) - \mathcal{N}^{-2}(f - 1 + v_x - u_y) \frac{\partial}{\partial z} \frac{D\rho}{Dt} + w_x v_z - w_y u_z = 0, \quad (3.8)$$

where

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + (y + u) \frac{\partial}{\partial x} + v \frac{\partial}{\partial y} + w \frac{\partial}{\partial z}. \quad (3.9)$$

3.3 Linear theory

The linearized governing equations are

$$u_t + yu_x + (1 - f)v = -p_x, \quad (3.10)$$

$$v_t + yv_x + fu = -p_y, \quad (3.11)$$

$$w_t + yw_x + \rho = -p_z, \quad (3.12)$$

$$\rho_t + y\rho_x - \mathcal{N}^2 w = 0, \quad (3.13)$$

$$u_x + v_y + w_z = 0. \quad (3.14)$$

The linearized version of (3.8) reduces to a conservation law of potential vorticity, $q_t + yq_x = 0$, or, given that $q = 0$ everywhere at $t = 0$,

$$q = (f - 1)\rho_z - \mathcal{N}^2(v_x - u_y) = 0. \quad (3.15)$$

In the absence of linear instability, the forcing (3.6) drives a steady wave response throughout the bulk of the flow (as can be established by solving the initial-value problem using Laplace transforms, and then performing a large-time asymptotic analysis, following Warn and Warn [67] and Booker and Bretherton [16]). Near the baroclinic critical levels, however, the flow remains unsteady, requiring a finer analysis of those regions similar to that used by Stewartson [1].

3.3.1 The steady wave response outside the baroclinic critical layers

The steady wave solution outside the critical layers takes the form:

$$(u, v, w, p, \rho) = [\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{p}(y), \hat{\rho}(y)] \exp(ix + imz) + \text{c.c.} \quad (3.16)$$

Substituting (3.16) into (3.10)-(3.14), one can derive an equation for $\hat{p}(y)$,

$$\hat{p}'' - \frac{2y}{y^2 - f(f-1)}\hat{p}' - \left[\frac{y^2 - f(f+1)}{y^2 - f(f-1)} + m^2 \frac{y^2 - f(f-1)}{y^2 - \mathcal{N}^2} \right] \hat{p} = 0, \quad (3.17)$$

with

$$\hat{u} = \frac{(f-1)\hat{p}' - y\hat{p}}{y^2 - f(f-1)}, \quad \hat{v} = \frac{i(y\hat{p}' - f\hat{p})}{y^2 - f(f-1)}, \quad \hat{w} = -\frac{my\hat{p}}{y^2 - \mathcal{N}^2}, \quad \hat{\rho} = \frac{im\mathcal{N}^2\hat{p}}{y^2 - \mathcal{N}^2}. \quad (3.18a,b,c,d)$$

(cf. Vanneste & Yavneh, 2007). Note that the singularities at $y^2 = f(f-1)$ in (3.17) and (3.18) are removable. The baroclinic critical levels $y = \pm\mathcal{N}$, however, are true singular points. The Frobenius solutions near $y = \mathcal{N}$ are,

$$\hat{p}_A = 1 - \frac{m^2[\mathcal{N}^2 - f(f-1)]}{2\mathcal{N}}(\mathcal{N} - y)\log|\mathcal{N} - y| - \alpha(\mathcal{N} - y) + \dots \quad (3.19)$$

$$\hat{p}_B = y - \mathcal{N} + \dots \quad (3.20)$$

where α is determined by the condition that $\hat{p}_A \rightarrow 0$ as $y \rightarrow \infty$. In terms of these Frobenius solutions, we express \hat{p} for $y > 0$ by

$$\hat{p} = \begin{cases} A_L \hat{p}_A, & y > \mathcal{N}, \\ A_L \hat{p}_A + B_L \hat{p}_B, & 0 < y < \mathcal{N}, \end{cases} \quad (3.22)$$

where A_L and B_L are constants.

Although \hat{p} is bounded for $y \rightarrow \mathcal{N}$, the amplitudes of the velocity, $(\hat{u}, \hat{v}, \hat{w})$, and density, $\hat{\rho}$, all diverge, signifying that the steady wave solution fails at the critical levels. In particular, we observe that

$$\hat{p} \rightarrow A_L, \quad \hat{\rho} \rightarrow \frac{im\mathcal{N}A_L}{2(y - \mathcal{N})} \quad (3.23)$$

and

$$\hat{u} \rightarrow \left[\frac{m^2(f-1)}{2\mathcal{N}}(\log|\mathcal{N} - y| + 1) + \frac{\alpha(f-1) - \mathcal{N}}{\mathcal{N}^2 - f(f-1)} \right] A_L + \begin{cases} 0 & y > \mathcal{N}, \\ \frac{f-1}{\mathcal{N}^2 - f(f-1)} B_L & y < \mathcal{N}, \end{cases} \quad (3.24)$$

for $y \rightarrow \mathcal{N}$.

3.3.2 The linear critical layers

We now focus on the baroclinic critical layer at $y = \mathcal{N}$. Here, we search for an unsteady solution depending on the long timescale $T = \delta t$ and with the short spatial scale $Y = (y - \mathcal{N})/\delta$, where $\delta \ll 1$ is a small parameter organizing an asymptotic expansion. We then set

$$(u, v, w, p, \rho) = [\tilde{u}(Y, T), \tilde{v}(Y, T), \delta^{-1}\tilde{w}(Y, T), A_L, \delta^{-1}\tilde{\rho}(Y, T)] \exp(ix + imz) + \text{c.c.}, \quad (3.25)$$

in view of the limits in (3.23)-(3.24).

Combining (3.12) and (3.13) to eliminate w , then substituting in (3.25) now gives, to leading order in δ ,

$$\left(\frac{\partial}{\partial T} + iY\right)\tilde{\rho} = -\frac{1}{2}m\mathcal{N}A_L. \quad (3.26)$$

In the early stage of linear evolution, $t \sim O(1)$, $\rho \sim O(1)$, so we have the initial condition $\tilde{\rho} \rightarrow 0$ as $T \rightarrow 0$, which yields

$$\tilde{\rho} = -\frac{1}{2}im\mathcal{N}A_L \frac{e^{-iYT} - 1}{Y}, \quad (3.27)$$

Hence

$$\rho = -\frac{1}{2}im\mathcal{N}A_L t \left[\frac{e^{-i(y-\mathcal{N})t} - 1}{(y-\mathcal{N})t} \right] e^{ix+imz} + \text{c.c.} \quad (3.28)$$

This solution has a spatial structure dependent on the self-similar combination $t(y - \mathcal{N})$. Hence, the amplitude grows linearly and the width of the critical layer shrinks with time.

Next, the main balance in (3.15) implies that $\tilde{u}_Y \sim -im(f-1)\mathcal{N}^{-2}\tilde{\rho}$, or

$$\tilde{u}_Y = -\frac{m^2(f-1)A_L}{2\mathcal{N}} \frac{e^{-iYT} - 1}{Y}. \quad (3.29)$$

But the limits of the steady wave response in (3.24) imply that \tilde{u} jumps by an

amount $(f-1)B_L/[\mathcal{N}^2 - f(f-1)]$ across the baroclinic critical layer. Hence,

$$B_L = -\frac{m^2 A_L [f(f-1) - \mathcal{N}^2]}{2\mathcal{N}} \lim_{L \rightarrow \infty} \int_{-L}^L (e^{-iYT} - 1) \frac{dY}{Y} = i\pi \frac{m^2 [f(f-1) - \mathcal{N}^2]}{2\mathcal{N}} A_L. \quad (3.30)$$

(cf. Stewartson 1978).

3.3.3 Closure

We can now apply the forcing condition to close the problem. The symmetry property (3.7) applied to the steady wave (3.16) indicates that

$$[\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{\rho}(y)] = -[\hat{u}(-y), \hat{v}(-y), \hat{w}(-y), \hat{\rho}(-y)]^*, \quad \hat{p}(y) = \hat{p}(-y)^*, \quad (3.31)$$

where the superscript $*$ represents the complex conjugate. Hence, substituting the steady wave solution into the jump condition (3.6) representing the forcing, we arrive at

$$\begin{aligned} (A_L - A_L^*)\hat{p}_A(0) + (B_L - B_L^*)\hat{p}_B(0) &= 0, \\ (A_L + A_L^*)\hat{p}'_A(0) + (B_L + B_L^*)\hat{p}'_B(0) &= -f\varepsilon_0. \end{aligned} \quad (3.32a, b)$$

Exploiting (3.30), we obtain

$$A_L = -\frac{f\varepsilon_0(\hat{p}_A - i\beta\hat{p}_B)}{2(\hat{p}_A\hat{p}'_A + \beta^2\hat{p}_B\hat{p}'_B)} \Big|_{y=0}, \quad \beta = \frac{\pi m^2 [f(f-1) - \mathcal{N}^2]}{2\mathcal{N}}. \quad (3.33)$$

The amplitude of the pressure perturbation at the critical layer is therefore

$$\varepsilon = |A_L| = \frac{|f\varepsilon_0| \sqrt{\hat{p}_A^2 + \beta^2\hat{p}_B^2}}{2|\hat{p}_A\hat{p}'_A + \beta^2\hat{p}_B\hat{p}'_B|} \Big|_{y=0}. \quad (3.34)$$

A sample steady wave solution is plotted in figure 3.2.

Note that equations (3.32)-(3.34) appear to become trivial if $f = 0$, suggesting that rotation is essential to the forcing of the baroclinic critical layer. In fact, a deeper analysis of the power series solution about $y = 0$ demonstrates that this is not the case, because $\hat{p}'_A(0)$ and $\hat{p}'_B(0)$ become $O(f)$ in this limit, and the closure

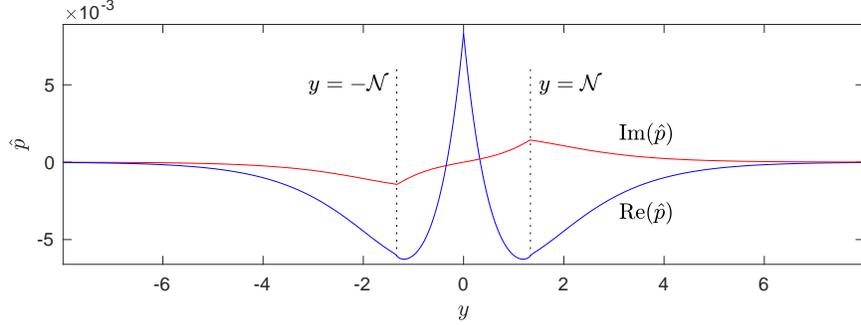


Figure 3.2: Steady-wave solution \hat{p} under a forcing imposed at $y = 0$, with $m = 0.5$, $\mathcal{N} = 4/3$, $f = 4/3$, $\varepsilon_0 = 0.05$ (cf. Marcus *et al.* 2013). Baroclinic critical levels $y = \pm\mathcal{N}$ are indicated.

relation in (3.32) remains non-trivial. Consequently, in the model, we may take the limit $f \rightarrow 0$, highlighting how rotation is not an essential ingredient to the dynamics of forced linear waves.

The same feature does not apply to the vertical wavenumber or stratification, which control the secular growth inside the critical layer, as seen in (3.27) and (3.29); without either a vertical dependence in the forcing or stratification, there is no baroclinic critical-layer dynamics. Note that, despite appearances, the limit $\mathcal{N} \rightarrow 0$ in (3.29) is not problematic: further analysis of \hat{p}_A and \hat{p}_B indicates that $|A_L| \sim \mathcal{N} / \log \mathcal{N}$ for $\mathcal{N} \rightarrow 0$, and so the secular growth in the critical layer is eliminated in this limit.

It is also noteworthy that, in the limit that any of the parameters m , f , or \mathcal{N} are large, the disturbance decays exponentially from the forcing to the baroclinic critical levels (cf. Vanneste and Yavneh [53] and Wang and Balmforth [61]). The amplitude ratio $\varepsilon/\varepsilon_0$ then becomes exponentially small, and the secular growth in the critical layer is much weakened.

3.4 The weakly nonlinear critical layer

We now advance beyond linear theory and perform a weakly nonlinear expansion by setting

$$\begin{aligned}
& (u, v, w, \rho, p) \\
&= \varepsilon \left\{ [u_1(Y, T), v_1(Y, T), \delta^{-1} w_1(Y, T), \delta^{-1} \rho_1(Y, T), p_1(Y, T)] e^{ix+imz} + \text{c.c.} \right\} \\
&+ \varepsilon^2 [u_0(Y, T), v_0(Y, T), w_0(Y, T), \rho_0(Y, T), p_0(Y, T)] \quad (3.35) \\
&+ \varepsilon^2 \left\{ [u_2(Y, T), v_2(Y, T), w_2(Y, T), \rho_2(Y, T), p_2(Y, T)] e^{2(ix+imz)} + \text{c.c.} \right\},
\end{aligned}$$

focussing upon the critical layer with $y = \mathcal{N} + \delta Y$. The scaling of the fundamental Fourier component follows the linear critical layer theory outlined above, and we have $\varepsilon[u_1, v_1, w_1, \rho_1, p_1] \rightarrow [\tilde{u}, \tilde{v}, \tilde{w}, \tilde{\rho}, A_L]$ at early times ($T \ll 1$). The goal of the current section is to identify the timescale and width of the critical layer (as denoted by the small parameter δ) for which the mean flow correction and first harmonic reach sufficient strength to modify the evolution of fundamental mode. This connects δ to the amplitude parameter ε , establishing the scalings of the nonlinear critical layer.

3.4.1 Mean-flow response

The mean-flow component of (3.5) gives $v_{0Y} = 0$, which implies $v_0 = 0$ since the mean flow response decays outside the critical layer. The streamwise mean-flow velocity u_0 is described by the $j = 0$ component of (3.1), which is

$$\frac{\partial u_0}{\partial T} = \delta^{-2} (imw_1 u_1^* - v_1^* u_{1Y}) + \text{c.c.} \quad (3.36)$$

To leading order in δ , the mean-flow components of (3.3) and (3.4) are,

$$\rho_0 = -\delta^{-2} v_1^* w_{1Y} + \text{c.c.}, \quad (3.37)$$

$$-\mathcal{N}^2 w_0 = -\delta^{-2} (v_1^* \rho_{1Y} + imw_1^* \rho_1) + \text{c.c.} \quad (3.38)$$

Thus, u_0 , w_0 , and ρ_0 are all $O(\delta^{-2})$.

3.4.2 First harmonic

The largest first harmonic components of (3.3), (3.4), (3.5) and (3.8) indicate that

$$2i\mathcal{N}w_2 + \rho_2 + 2imp_2 = -\delta^{-2}(v_1w_{1Y} + imw_1^2), \quad (3.39)$$

$$2i\mathcal{N}\rho_2 - \mathcal{N}^2w_2 = -\delta^{-2}(v_1\rho_{1Y} - imw_1\rho_1). \quad (3.40)$$

$$\delta^{-1}v_{2Y} + 2imw_2 = 0. \quad (3.41)$$

$$\frac{u_{2Y}}{\delta} + \frac{2im(f-1)}{\mathcal{N}^2}\rho_2 = \frac{i\delta^{-2}}{2\mathcal{N}} \left[\frac{2im(f-1)}{\mathcal{N}^2}(v_1\rho_{1Y} + imw_1\rho_1) + (v_1u_{1Y} + imw_1u_1)_Y \right]. \quad (3.42)$$

However, (3.2) demands that $p_2 = O(\delta u_2, \delta v_2)$ and so p_2 is much smaller than w_2 or ρ_2 . Hence,

$$w_2 = \frac{\delta^{-2}}{3\mathcal{N}^2} [2i\mathcal{N}(v_1w_{1Y} + imw_1^2) - v_1\rho_{1Y} - imw_1\rho_1], \quad (3.43)$$

$$\rho_2 = \frac{\delta^{-2}}{3\mathcal{N}} [\mathcal{N}(v_1w_{1Y} + imw_1^2) + 2i(v_1\rho_{1Y} + imw_1\rho_1)], \quad (3.44)$$

which are $O(\delta^{-2})$, whereas u_2 and v_2 are $O(\delta^{-1})$.

3.4.3 Weakly nonlinear feedback

On again combining (3.3) and (3.4), we find the fundamental components,

$$\left(\frac{\partial}{\partial T} + iY \right) \rho_1 + \frac{1}{2}m\mathcal{N}p_1 = -\varepsilon^2\delta^{-1}iu_0\rho_1, \quad (3.45)$$

with the leading-order nonlinear terms included on the right, and after a considerable number of cancellations stemming from the use of (3.37), (3.38), (3.43) and (3.44) and the leading-order relations $\rho_1 = -i\mathcal{N}w_1$ and $v_{1Y} = -imw_1$. Note that the nonlinear terms generated by the first harmonic and mean-flow components w_0 and ρ_0 completely cancel out at this stage, leaving only the effect of the modification to the streamwise mean flow u_0 . But the scaling established for the mean flow correction implies that the right-hand side of (3.45) is $O(\delta^{-3}\varepsilon^2)$. Thus, the mean

flow feedbacks on the fundamental mode when $\delta = \varepsilon^{2/3}$. That is, for

$$t = O(\varepsilon^{-\frac{2}{3}}), \quad y = \mathcal{N} + O(\varepsilon^{\frac{2}{3}}). \quad (3.46)$$

These are the scalings for the nonlinear critical layer theory in the next section.

Note that we may extend the analysis to consider the higher harmonics. One finds that when $\delta = \varepsilon^{2/3}$, the Fourier component $e^{ij(x+mz)}$ with $j > 1$ is $O(\varepsilon^{j/3})$, which signifies that the higher-order harmonics $j \geq 3$ are still weak when the mean flow correction begins to feedback on the fundamental. Thus, they play no role in the nonlinear theory.

3.5 Nonlinear critical-layer theory

3.5.1 The reduction

Motivated by the weakly nonlinear analysis, we now introduce the rescalings,

$$T = \varepsilon^{\frac{2}{3}}t, \quad Y = \frac{y - \mathcal{N}}{\varepsilon^{\frac{2}{3}}}. \quad (3.47)$$

The outer solution for the pressure is

$$p = \varepsilon p_1 e^{i(x+mz)} + \text{c.c.}, \quad p_1 = \begin{cases} A(T)\hat{p}_A(y), & y > \mathcal{N}, \\ A(T)\hat{p}_A(y) + B(T)\hat{p}_B(y), & 0 < y < \mathcal{N}, \end{cases} \quad (3.48)$$

which is a single dominant Fourier mode characterized by the steady wave solution. However, the amplitudes A and B now evolve with the slow time T , because the nonlinear evolution of critical layer can affect the outer flow. Initially, A and B are given by the linear analysis:

$$A(0) = \frac{A_L}{\varepsilon}, \quad B(0) = \frac{B_L}{\varepsilon}. \quad (3.49)$$

Inside the critical layers, we set

$$\begin{aligned}
p &= \varepsilon A(T) e^{i(x+mz)} + \text{c.c.} + \dots, \\
[w, \rho] &= \varepsilon^{\frac{1}{3}} [w_1(Y, T), \rho_1(Y, T)] e^{i(x+mz)} + \text{c.c.} + \dots \\
[u, v] &= \varepsilon [u_1(Y, T), v_1(Y, T)] e^{i(x+mz)} + \text{c.c.} + \varepsilon^{\frac{2}{3}} [U_0(Y, T), 0] + \dots \quad (3.50)
\end{aligned}$$

Equation (3.45) and the leading-order fundamental-mode components of (3.1), (3.3)-(3.5) and (3.8) now become

$$\frac{\partial \rho_1}{\partial T} + iY \rho_1 + \frac{m\mathcal{N}}{2} A = -iU_0 \rho_1. \quad (3.51)$$

$$i\mathcal{N} u_1 - (f-1)v_1 + iA = -v_1 U_{0Y}, \quad (3.52)$$

$$w_1 = \frac{i}{\mathcal{N}} \rho_1, \quad v_{1Y} = -imw_1, \quad (3.53)$$

$$\mathcal{N}^2 u_{1Y} + im(f-1-U_{0Y})\rho_1 = i\mathcal{N} v_1 U_{0Y}. \quad (3.54)$$

The initial condition of ρ_1 is given by the linear result

$$\rho_1 \rightarrow -\frac{im\mathcal{N}A(0)}{2} \frac{e^{-iYT} - 1}{Y}, \quad T \rightarrow 0. \quad (3.55)$$

Similar to (3.36), the mean-flow velocity U_0 is governed by

$$\frac{\partial U_0}{\partial T} = -v_1^* u_{1Y} + imw_1 u_1^* + \text{c.c.} \quad (3.56)$$

The initial condition is $U_0 \rightarrow 0$ as $T \rightarrow 0$, as in early linear evolution the mean-flow modification is minimal.

It is possible to algebraically manipulate (3.51)-(3.54) and then integrate in T to show that

$$U_0 = -\frac{2}{\mathcal{N}^3} |\rho_1|^2, \quad (3.57)$$

a result that can be traced back to the fact that the change to the mean flow is given by the Eulerian pseudo-momentum [62], which is the right-hand side of (3.57) to

leading order in the critical layer. Hence

$$\frac{\partial \rho_1}{\partial T} + iY\rho_1 + \frac{1}{2}m\mathcal{N}A = i\frac{2}{\mathcal{N}^3}|\rho_1|^2\rho_1. \quad (3.58)$$

To match the inner and outer solutions, we first note, from (3.53), that $v_{1Y} = m\mathcal{N}^{-1}\rho_1$. Integrating this relation in Y over the critical layer then provides the jump of the outer solution $v_1 = i(y\rho_{1,y} - f\rho_1)/[y^2 - f(f-1)]$ for the limit of $y \rightarrow \mathcal{N}$ (cf. (3.18b)), which yields

$$B = -im\frac{f(f-1) - \mathcal{N}^2}{\mathcal{N}^2} \int_{-\infty}^{\infty} \rho_1 dY, \quad (3.59)$$

in a similar manner to Section 3.3.2 and (3.30).

Last, we again use the forcing condition at $y = 0$ to close the problem:

$$\begin{aligned} (A - A^*)\hat{p}_A(0) + (B - B^*)\hat{p}_B(0) &= 0, \\ (A + A^*)\hat{p}'_A(0) + (B + B^*)\hat{p}'_B(0) &= -f\frac{\varepsilon_0}{\varepsilon} \end{aligned} \quad (3.60)$$

(cf. Section 3.3.3 and (3.32)). Note that the form of the forcing impacts the reduced model only through the closure relations in (3.60). Had we used a different idealization of the forcing here, there would be a different algebraic relation between A , B and $\varepsilon_0/\varepsilon$. However, this relation still connects A with the forcing amplitude and the integral of ρ_1 over the critical layer, and in the scaled, canonical system presented below, all that would change would be how the parameters of that system (denoted c_0 , c_1 and c_2 in Section 3.5.2) depend on the original physical constants. In this sense, the reduced model is independent of the choice of forcing.

3.5.2 Canonical system

The final rescalings

$$\rho_1 = \left(\frac{m\mathcal{N}^4}{4}\right)^{\frac{1}{3}} \gamma(\eta, \tau), \quad T = \left(\frac{2\mathcal{N}}{m^2}\right)^{\frac{1}{3}} \tau, \quad Y = \left(\frac{m^2}{2\mathcal{N}}\right)^{\frac{1}{3}} \eta, \quad (3.61)$$

lead to the canonical form,

$$\frac{\partial \gamma}{\partial \tau} + i\eta\gamma + A = i|\gamma|^2\gamma, \quad (3.62)$$

$$A(\tau) = c_0 + \frac{ic_1}{\pi} \int_{-\infty}^{\infty} \gamma_r d\eta - \frac{c_2}{\pi} \int_{-\infty}^{\infty} \gamma_i d\eta, \quad (3.63)$$

where $\gamma = \gamma_r + i\gamma_i$,

$$c_0 = -\text{sgn}\left(\frac{f}{\hat{p}'_A(0)}\right) \frac{|1 + c_1 c_2|}{\sqrt{1 + c_1^2}}, \quad (3.64)$$

and

$$\begin{Bmatrix} c_1 \\ c_2 \end{Bmatrix} = \frac{\pi m^2 [f(f-1) - \mathcal{N}^2]}{2\mathcal{N}} \begin{Bmatrix} \hat{p}_B(0)/\hat{p}_A(0) \\ \hat{p}'_B(0)/\hat{p}'_A(0) \end{Bmatrix}. \quad (3.65)$$

For $\tau \ll 1$, we must match $\gamma(\eta, \tau)$ to the corresponding solution of the linear problem, given by

$$\gamma = iA \frac{1 - e^{-i\eta\tau}}{\eta}, \quad A = \frac{c_0(1 - ic_1)}{1 + c_1 c_2}, \quad (3.66)$$

which provides the initial condition for (3.62).

The reduced model equations in (3.62)-(3.66) are solved numerically in the next section. The system is integro-differential in the sense that (3.62) is an equation of motion in time, solved at each level of η , with the integral constraint in (3.63). There is no dependence on either x or z , because the leading-order dynamics involves only the fundamental mode of the forcing wave pattern and the mean-flow response (which is then prescribed by the pseudo-momentum). The only nonlinearity is the cubic term on the right of (3.62), which is generic in weakly nonlinear theories of systems with few degrees of freedom. The model is therefore rather different from those that emerge for classical forced critical layers, which usually take the form of partial differential equations in all the spatial variables. The reduced model has the two parameters, c_1 and c_2 , and the choice of sign for $f\hat{p}'_A(0)$ in c_0 . In most situations \hat{p}_A and \hat{p}_B are characterized by a similar exponential away from $y = 0$, implying $c_1 \approx c_2$.

From (3.62)-(3.63), one can establish that the quantity,

$$\mathcal{H} = \int_{-\infty}^{\infty} \left[\frac{1}{2} |\gamma|^4 - \eta |\gamma|^2 + 2\text{Im}(A^* \gamma) \right] d\eta + \frac{c_1}{\pi} \left[\int_{-\infty}^{\infty} \gamma_r d\eta \right]^2 + \frac{c_2}{\pi} \left[\int_{-\infty}^{\infty} \gamma_i d\eta \right]^2, \quad (3.67)$$

must be conserved, and therefore equal to $\pi c_1(1 + c_1 c_2)/(1 + c_1^2)$ in view of the initial conditions. This constraint implies that the linear-in-time growth of $\gamma(\eta, \tau)$ predicted by linear theory must eventually become arrested, as otherwise the quartic first term in (3.67) cannot be counter balanced by the remaining quadratic and constant terms. To determine the manner in which the arrest takes place, we turn to a numerical solution of the reduced model.

3.5.3 Numerical solutions

To solve the canonical system of equations numerically, we first select a grid in η spanning a finite domain (we use 1501 equally spaced gridpoints over the interval $1.5 < \eta < 3$ where γ has large gradients, then 1544 gridpoints distributed evenly over $-25 < \eta < 1.5$ and $3 < \eta < 25$). We then integrate (3.62) forward in time numerically using a 4th-order Runge-Kutta method at each of the grid points. To evaluate the integrals in (3.63), we use an approach similar to Warn & Warn (1978) to extrapolate the limits to infinity. We use parameter settings guided by the computations of Marcus *et al.* (2013): $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, which yield $c_1 = 0.238$, $c_2 = 0.219$.

Figure 3.3 displays the evolution in τ of the forced wave amplitudes, A and B , which is relatively mild with $\text{Re}(A) \approx c_0 \approx -1$ and $\text{Im}(A)$, $\text{Re}(B)$ and $\text{Im}(B)$ all remaining small. This mild behaviour results because, in (3.63), $|c_1|$ and $|c_2|$ are fairly small. Thus, the forced wave evolves slowly over the bulk of the shear flow (*i.e.* the outer region), maintaining a profile similar to the linear distribution in figure 3.2.

The density perturbation $\gamma(\eta, \tau)$, shown in figure 3.4, exhibits a richer behaviour: for $\tau < 1$, the numerical solution follows the linear prediction in (3.66), with its characteristically developing undulations and linear growth near $\eta = 0$ (see figure 3.4(a,b)). Once $|\gamma|$ reaches order-one values there, however, the growth of the numerical solution saturates, as demanded by the constraint in (3.67). De-

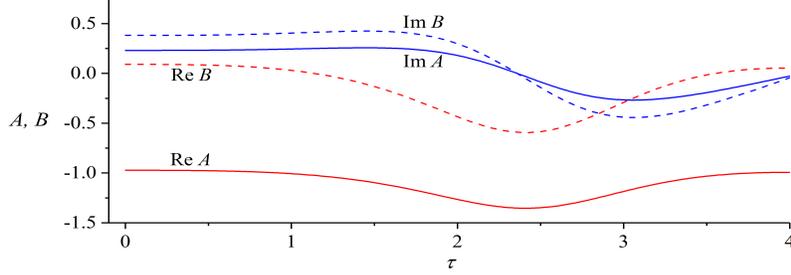


Figure 3.3: Evolution of A and B with τ ; $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$.

spite this, the solution continues to undulate over increasingly shorter spatial scales. Moreover, nonlinear effects distort the density profile further, shifting the maximum magnitude from $\eta = 0$ to a small, positive level in η and generating pronounced fine structure over a narrow region nearby.

The rapid spatial variation in $\gamma(\eta, \tau)$ significantly impacts the critical-layer vorticity, which depends on the η -derivatives of $\gamma(\eta, \tau)$. In particular, the leading-order vertical vorticity is given by the mean-flow vorticity ζ_0 :

$$\zeta \sim \zeta_0 \equiv \frac{\partial}{\partial \eta} |\gamma|^2. \quad (3.68)$$

However, from the matched asymptotic expansion, we may reconstruct $\zeta(x, \eta, z, t)$ to higher orders, incorporating the fundamental Fourier mode ζ_1 and first harmonic ζ_2 , as summarized in Appendix A. The evolution of the reconstructed vertical vorticity field is plotted in figure 3.5. For early times, $\zeta_0 \ll 1$, and the vertical vorticity is actually given by the higher-order linear solution (as in figure 3.5a, *cf.* (3.29)). With the increase of τ , the vorticity distribution tilts over and ζ_0 grows to dominate ζ , as seen in figure 3.5b,c. This growth leads to the distinctive dipolar stripe seen in figure 3.5d. In the later stages of evolution (figure 3.5e,f), the stripe becomes stronger and more focussed, shifting slightly above $\eta = 0$, and corresponding to the sharpening oscillations in γ seen in figure 3.4.

The behaviour of the numerical solution seen in figures 3.3-3.5 is generic for most parameter settings; for moderate m , f (either $f > 1$ or $f < 0$) and \mathcal{N} , the parameters c_1 and c_2 of the reduced model are relatively small in magnitude, prompt-

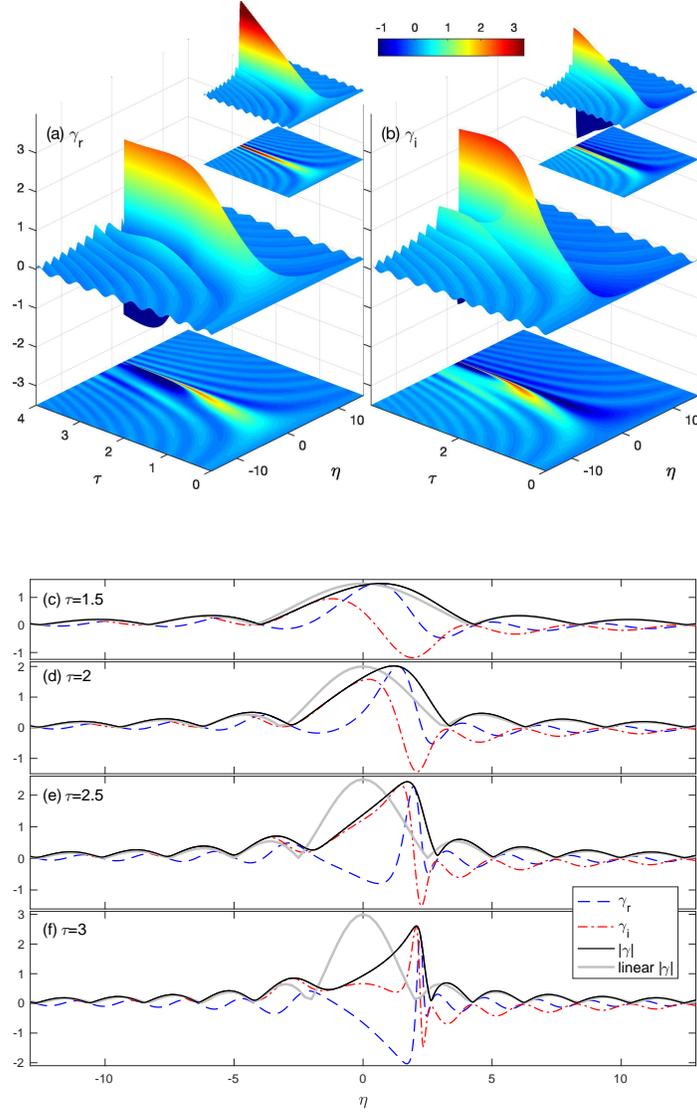


Figure 3.4: (a) Real and (b) imaginary parts of the critical-layer density perturbation $\gamma(\eta, \tau)$, shown as surfaces above the (η, τ) -plane. To prevent the viewing perspective from obscuring parts of the solution, we also show density maps of the solutions underneath. The insets show corresponding plots of the linear solution in (3.66). Panels (c)–(f) plot snapshots of $\gamma(\eta, \tau)$ at the times indicated; the linear result for $|\gamma(\eta, \tau)|$ is also included. ($m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$.)

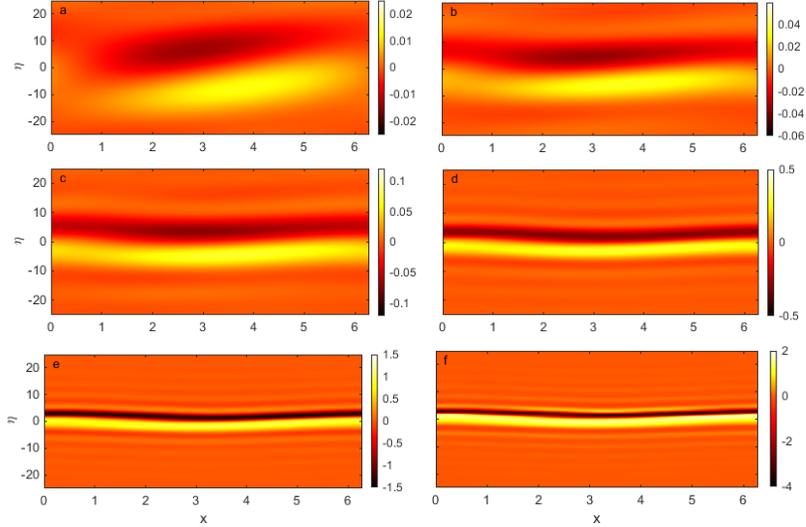


Figure 3.5: Snapshots of vertical vorticity ζ within the baroclinic critical layer near $y = \mathcal{N} = 4/3$, plotted as a colormap on the (x, η) -plane for (a) $\tau = 0.3$, (b) $\tau = 0.45$, (c) $\tau = 0.6$, (d) $\tau = 1$, (e) $\tau = 1.5$, and (f) $\tau = 1.8$. The domain plotted is $|\eta| < 25$, corresponding to $|y - \mathcal{N}| < 0.39$, at cross-section $z = 0$ and over one streamwise wavelengths of the forcing pattern. ($\varepsilon_0 = 0.05$, $\varepsilon = 0.0062$, $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$.)

ing similar dynamics. Even when $|c_1|$ and $|c_2|$ become order one, the evolution still bears qualitative similarities. However, more complicated behaviour can occur in the reduced model when these parameters take higher values. Such parameter settings can be achieved at special combinations of m , f and \mathcal{N} for which $\hat{p}_A(0)$ becomes small, or perhaps for other types of forcing. We avoid consideration of special situations of this sort, and instead turn to a deeper analysis of the focussing dynamics observed in the reduced model.

3.5.4 Long-time focussing

In view of the result that A changes slowly, we now use the approximation of $A = \text{constant}$ to gain further analytical insights to the focussing phenomenon. This device was used previously by Stewartson (1978) to obtain an analytical solution

to the nonlinear evolution of Rossby wave critical layers. In our model, constant A in (3.63) requires $c_1 = c_2 = 0$, hence $A = -\text{sgn}(f\hat{p}'_A(0))$, which is -1 for the current parameter setting. The evolution equation (3.62) can then be written as the one-degree-of-freedom Hamiltonian system,

$$\begin{aligned}\frac{\partial\gamma_r}{\partial\tau} &= \frac{\partial H}{\partial\gamma_i} = 1 + \eta\gamma_i - \gamma_r^2\gamma_i - \gamma_i^3, \\ \frac{\partial\gamma_i}{\partial\tau} &= -\frac{\partial H}{\partial\gamma_r} = -\eta\gamma_r + \gamma_r^3 + \gamma_r\gamma_i^2,\end{aligned}\tag{3.69}$$

with Hamiltonian,

$$H = -\frac{1}{4}(\gamma_r^2 + \gamma_i^2)^2 + \frac{1}{2}\eta(\gamma_r^2 + \gamma_i^2) + \gamma_i\tag{3.70}$$

(the point-wise version of the conserved quantity \mathcal{H} in (3.67) for $c_1 = c_2 = 0$). For the specific initial condition of our critical-layer problem, we have $H = 0$ for all values of η .

Figure 3.6(a) illustrates the phase portrait of the system (3.69) for the special choice $\eta = \eta_c = 3/\sqrt[3]{2}$. In this case, the orbit from $(\gamma_r, \gamma_i) = (0, 0)$ lies along a separatrix that converges to a saddle point at $(\gamma_r, \gamma_i) = (0, -\gamma_e)$, for $\tau \rightarrow \infty$, with $\gamma_e = \sqrt[3]{2} \approx 1.26$. Trajectories from $(\gamma_r, \gamma_i) = (0, 0)$ for a spread of values of η around η_c are illustrated in figure 3.6(b); the presence of the separatrix at $\eta = \eta_c$ implies that these trajectories bifurcate in direction on the phase plane on passing through that special level. Thus, a small variation in η about η_c can result in a large change of γ at later times, implying high values of γ_η to feed into ζ .

For the numerical solutions of Section 3.5.3, although c_1 and c_2 do not vanish, the forced-wave amplitude does remain slowly varying in τ , leading to a qualitatively similar dynamics: figure 3.6(c) plots the phase portrait of γ for five values of η within the region where the dipolar stripe is focussed. As η varies from 2.38 to 2.48, the trajectories for different levels abruptly switch in direction near the point $(\gamma_r, \gamma_i) = (0, -1.2)$. Although the slow variation of $A(\tau)$ precludes any trajectory from reaching a steady value, the numerical solution for $\eta = 2.43$ slows down, lingers and hesitates before selecting one of the two possible directions, much like the orbits for $c_1 = c_2 = 0$ near the separatrix in figure 3.6(a,b). The level of this

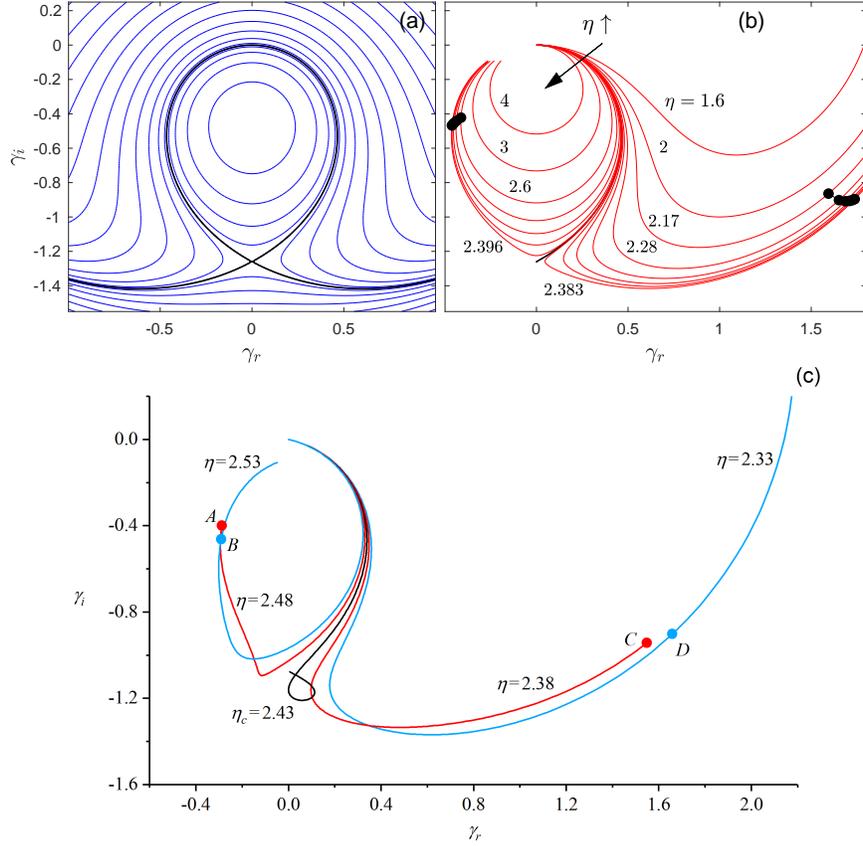


Figure 3.6: Phase portraits of (γ_r, γ_i) for (a) the Hamiltonian system (3.69) with $\eta = 3/\sqrt[3]{2}$ and various H , with the thicker line indicating $H = 0$, (b) trajectories from the point $(\gamma_r, \gamma_i) = (0,0)$ for a selection of values of η , and (c) the numerical solution of Section 3.5.3, at the five values of η indicated. The black points in (b) and the (red and blue) pairs marked (A, B) and (C, D) in (c) have the same values of $|\eta - \eta_c|e^{\sigma\tau}$.

trajectory is slightly shifted from $3/\sqrt[3]{2} \approx 2.38$ because c_1 and c_2 are non-zero and $A(\tau) \neq -1$. Nevertheless, we conclude that the close passage to an effective saddle point on the (γ_r, γ_i) phase plane is responsible for the focussing effect. For the numerical solution, we therefore define $\eta = \eta_c \approx 2.43$ to be the level for which γ evolves slowest near the effective saddle, and refer to this location as the nonlinear critical level.

Continuing the analysis for $c_1 = c_2 = 0$, we may linearize the system (3.69) about $\gamma = -i\gamma_e$ to find that

$$\frac{\partial}{\partial \tau} \begin{pmatrix} \gamma_r \\ \gamma_i + \gamma_e \end{pmatrix} = \begin{pmatrix} 0 & \eta - 3\gamma_e^2 \\ -\eta + \gamma_e^2 & 0 \end{pmatrix} \begin{pmatrix} \gamma_r \\ \gamma_i + \gamma_e \end{pmatrix}. \quad (3.71)$$

The two eigenvalues of the matrix are $\pm\sigma$, with corresponding eigenvectors \mathbf{v}_+ and \mathbf{v}_- , where

$$\sigma = \sqrt{[(\gamma_e^2 - \eta)][\eta - 3\gamma_e^2]} \approx \sigma_c = \frac{\sqrt{3}}{\sqrt[3]{2}} \quad \text{if } \eta \approx \eta_c. \quad (3.72)$$

The solution of (3.71) is then

$$\begin{pmatrix} \gamma_r \\ \gamma_i + \gamma_e \end{pmatrix} = r_+ \mathbf{v}_+ e^{\sigma(\tau - \tau_0)} + r_- \mathbf{v}_- e^{-\sigma(\tau - \tau_0)}, \quad (3.73)$$

for some constants r_{\pm} and a time constant τ_0 indicating when the orbit reaches the neighbourhood of the saddle point.

Now, along the separatrix converging to $\gamma = -i\gamma_e$ for $\eta = \eta_c$, the constant r_+ must vanish. But when η is close to, but not at η_c , this factor is small but finite, hence a local linearization of $r_+(\eta)$ near $\eta = \eta_c$ leads us to set $r_+ \approx C(\eta - \eta_c)$, for some constant C . Therefore,

$$\begin{pmatrix} \gamma_r \\ \gamma_i + \gamma_e \end{pmatrix} \sim C(\eta - \eta_c) \mathbf{v}_+ e^{\sigma_c(\tau - \tau_0)}, \quad (3.74)$$

at large times. That is, for η near η_c , those pairs of (η, τ) with the same $(\eta - \eta_c)e^{\sigma_c\tau}$ should have the same γ . Although this property is derived from the local linearization about the fixed point, it still holds when trajectories have progressed further along the unstable manifolds of that saddle because the trajectories shadow that curve. This is illustrated in figure 3.6 for both the Hamiltonian system and the numerical solution, where the pairs of points plotted along sample orbits have the same values for $(\eta - \eta_c)e^{\sigma_c\tau}$, and therefore similar γ , even though they correspond to different choices of (η, τ) . We can express the property mathematically

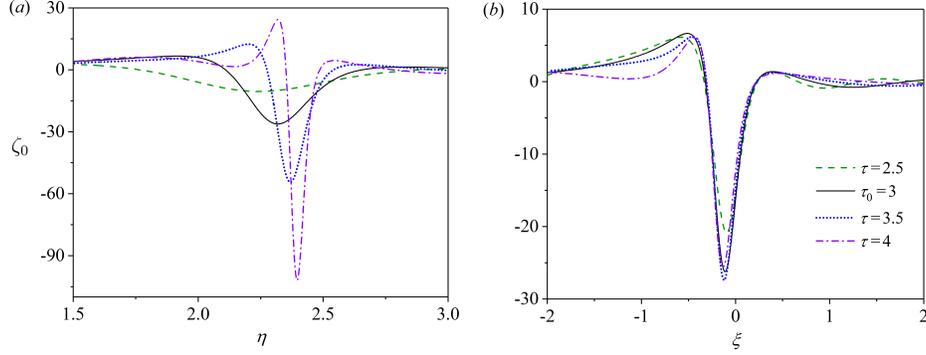


Figure 3.7: (a) Evolution of ζ_0 near $\eta_c = 2.43$ at the times indicated. (b) Scaled profiles, $\zeta_0 e^{-\sigma_c(\tau-\tau_0)}$ against $\xi = (\eta - \eta_c) e^{\sigma_c(\tau-\tau_0)}$, choosing $\tau_0 = 3$; ($m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$).

by writing the solutions in the self-similar form,

$$\gamma \approx F(\xi) \quad \text{and} \quad \zeta_0 \approx e^{\sigma_c(\tau-\tau_0)} \frac{d}{d\xi} |F(\xi)|^2, \quad \text{with} \quad \xi = (\eta - \eta_c) e^{\sigma_c(\tau-\tau_0)}, \quad (3.75)$$

for some function $F(\xi)$ related to the shape of the unstable manifolds of the saddle point. Thus, the lengthscale of the nonlinear critical layer at $\eta = \eta_c$ decreases exponentially in time, accounting for the relatively rapid focussing of sharp spatial variations in γ at later times in figure 3.4, and the amplitude of the vertical vorticity grows exponentially. Figure 3.7 presents four snapshots of $\zeta_0(\eta, \tau)$ for the numerical solution, then replots them against ξ and scaled by $e^{\sigma_c(\tau-\tau_0)}$, adopting $\tau_0 = 3$; while the profile of ζ_0 keeps sharpening and strengthening, the rescaled profile remains nearly unchanged, confirming the self-similar structure in (3.75).

The exponential focussing towards the nonlinear critical level is problematic as it implies that the higher-order harmonics of the forcing pattern, which are neglected in our nonlinear critical layer model, grow faster than the re-arrangements of the mean flow. In particular, one can deduce that the vertical vorticity of the j^{th} Fourier component, $\exp[j(ix + imz)]$, grows like $e^{(j+1)\sigma_c\tau}$. The model therefore fails once the solution becomes overly focussed, heralding the onset of a further, more complicated, stage of evolution.

It is noted that in an even longer time, the variation of A also yields some significant effects. When the near-homoclinic trajectories in figure 3.6 return after one period, the variation of A will render a different saddle point. Consequently, in the long-time behavior, there could be multiple levels of focussing, which can be observed in the left panels of figure 3.9(a),(b),(c) given subsequently. A deeper analysis of this phenomenon concerns the topic of slow passage through the homoclinic orbits of slowly varying hamiltonian systems. We refer readers to Haberman and Ho [70] and Diminnie and Haberman [71] for a detailed analysis on this topic. We will not further discuss it here, as we expect that our model already fails when the first focussing takes place.

3.6 Effects of diffusion

The increasingly fine scales encountered in the critical layer due to the exponential focussing suggest that dissipation may also become prominent over later times, even if small initially. To explore this possibility in more detail, we return to the governing equations and include the viscous terms $\nu \nabla^2(u, v, w)$ in (3.1)–(3.3) and diffusive term $\kappa \nabla^2 \rho$ in (3.4). We then take the distinguished limit $(\nu, \kappa) = O(\varepsilon^2)$, which corresponds to the order when dissipation first becomes important. In particular, with this scaling of ν and κ , the dissipative terms are too small to affect the quasi-steady wave in the bulk of the flow, but enter the analysis of the baroclinic critical layers owing to the reduced spatial scale in y . Equation (3.51) is now replaced with

$$\frac{\partial \rho_1}{\partial T} + iY \rho_1 + \frac{m\mathcal{N}}{2} A = -iU_0 \rho_1 + \frac{(\nu + \kappa)}{2\varepsilon^2} \frac{\partial^2 \rho_1}{\partial Y^2}. \quad (3.76)$$

The Eulerian pseudomomentum is no longer equal to the mean-flow response, as in (3.57), and we have to return to the mean-flow evolution equation:

$$\frac{\partial U_0}{\partial T} = \frac{m}{\mathcal{N}^2} (A^* \rho_1 + A \rho_1^*) + \frac{\nu}{\varepsilon^2} \frac{\partial^2 U_0}{\partial Y^2}, \quad (3.77)$$

(which follows from the substitution of (3.52)–(3.54) into the modified version of (3.56)). The initial condition is still given by (5.20), the dissipative terms being negligible at early times when the spatial scales are larger. The closure relations

given by the match to the outer solution remain (3.59) and (3.60). Equations (3.76) and (3.77) can be combined to furnish the integral relation,

$$\frac{d}{dT} \int_{-\infty}^{\infty} (|\rho_1|^2 + \frac{1}{2} \mathcal{N}^3 U_0) dY = -\frac{(\nu + \kappa)}{\varepsilon^2} \int_{-\infty}^{\infty} \left| \frac{\partial \rho_1}{\partial Y} \right|^2 dY, \quad (3.78)$$

provided that ρ_1 and U_0 decay sufficiently quickly for $|Y| \rightarrow \infty$. We now briefly discuss the dynamics captured by this dissipative version of the model, focussing on the astrophysically relevant limit $\nu \ll \kappa$, which allows the viscous term in (3.77) to be neglected.

3.6.1 Modified canonical system

A scaling similar to that in Section 3.5.2, now furnishes the modified canonical system,

$$\frac{\partial \gamma}{\partial \tau} + i\eta\gamma + A = -i\gamma\mathcal{U} + \lambda \frac{\partial^2 \gamma}{\partial \eta^2}, \quad \frac{\partial \mathcal{U}}{\partial \tau} = A^* \gamma + A\gamma^* \quad (3.79)$$

and (3.63), where

$$\mathcal{U}(\eta, \tau) = \left(\frac{2\mathcal{N}}{m^2} \right)^{\frac{1}{3}} U_0 \quad \text{and} \quad \lambda = \frac{\kappa \mathcal{N}}{m^2 \varepsilon^2}. \quad (3.80)$$

This system may be solved numerically. For the task, we now use a Crank-Nicolson method to evolve the system in time and centred finite differences method to evaluate spatial derivatives, exploiting Newton iteration at each time step to solve the nonlinear equations.

Before characterizing the features of the numerical solutions, we first pause to examine the dynamics in the limit that diffusion is relatively strong, $\lambda \gg 1$. In this limit, the large diffusive term $\lambda \gamma_{\eta\eta}$ in (3.79) must be balanced by introducing the rescalings, $(\gamma, \tau) = O(\lambda^{-1/3})$, $\eta = O(\lambda^{1/3})$ and $\mathcal{U} = O(\lambda^{-2/3})$. The advection of the density perturbation by the mean-flow correction, $i\gamma\mathcal{U}$, is then small in the first equation in (3.79), and if we again make the approximation of constant A , we find

$$\gamma \approx -A \int_0^\tau e^{-\lambda q^3/3 - iq\eta} dq, \quad (3.81)$$

which is plotted in figure 3.8. At $\tau \ll 1$, (3.81) recovers the secular growth of the

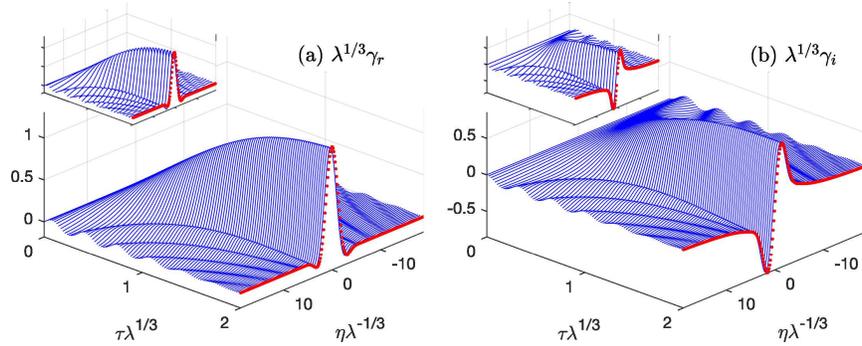


Figure 3.8: The analytical solution (3.81) for strong diffusion and $A = -1$, showing (a) $\lambda^{1/3}\gamma_r$ and (b) $\lambda^{1/3}\gamma_i$ against the scaled space and time variables $\lambda^{-1/3}\eta$ and $\lambda^{1/3}\tau$. The (red) dots show the final steady-state solution. The insets show corresponding numerical solutions to the reduced model, computed for $\lambda = 5.3$.

linear non-dissipative critical layer (*cf.* (3.66)), but over longer times, this solution approaches a steady state, illustrating how diffusion is able to saturate that growth before nonlinearity (and the advective term $i\gamma\mathcal{U}$) enters the fray. Figure 3.8 also illustrates how this dynamics does indeed characterize the full modified model for larger values of the diffusivity, demonstrating how the analytical solution in (3.81) agrees satisfyingly with numerical results computed with $\lambda = 5.3$. The steady state prediction from (3.81) corresponds to the result of viscous critical-layer theory presented by [4] for stratified tilted vortices (in which case, $\tau \rightarrow \infty$ in (3.81) and the solution can be related to the Scorer function).

Nevertheless, the establishment of a steady state with spatial structure in the density perturbation is inconsistent with the integral relation in (3.78). Indeed, if γ approaches a steady state, \mathcal{U} continues to grow linearly with τ , and for times of order $\lambda^{1/3}$, the advective term $i\gamma\mathcal{U}$ can no longer be neglected in (3.79), heralding the onset of a different, more complicated phase of evolution. Figure 3.9 shows a suite of numerical solutions, illustrating this later evolutionary stage for cases with stronger diffusion (right-hand panels), and other examples with smaller λ (left-hand panels). For the latter, diffusion is too weak to arrest the linear growth in the critical layer and nonlinear focussing begin to occur; only when the spatial scale

has reduced sufficiently does the dissipative effect take hold to limit the exponential amplification found for $\lambda = 0$. At that stage, a new phase of evolution again emerges, much like that found for stronger diffusion. In particular, the oscillations of the non-dissipative dynamics begin to fade with time, and a localized coherent structure emerges that drifts to larger η under the advective effect of the mean-flow correction. The structure leaves in its wake an increasingly strong deficit in \mathcal{U} , which is permitted by the constraint in (3.78) because diffusion may continually lower \mathcal{U} as long as the gradients of γ remain finite.

3.6.2 Dissipative coherent structures

The drifting coherent structure can be analyzed further owing to its fine spatial scale and the relatively slow timescale over which the system develops once the larger-scale transients have subsided: assuming that $\lambda \ll 1$ and A is real and constant, we search for a quasi-steady travelling wave solution in which

$$\gamma \approx \gamma(\xi) \quad \text{and} \quad \xi = \frac{\eta}{\sqrt{\lambda}} - \int c \, d\tau, \quad (3.82)$$

which characterizes a coherent structure with a length scale of $\sqrt{\lambda} \ll 1$ and a drift velocity given by c . Hence,

$$-c\gamma' + i\eta_*\gamma + A \approx -i\gamma\mathcal{U}' + \gamma'' \quad \text{and} \quad -c\mathcal{U}' \approx A^*\gamma + A\gamma^* \approx 2A\gamma_r, \quad (3.83)$$

where $\xi = 0$, or $\eta_* = \sqrt{\lambda} \int c \, d\tau$, prescribes the center of the coherent structure. From (3.79) to (3.82), we have assumed $\eta = \eta_*$ to obtain the travelling wave solution. The justification of this approximation is that in figure 3.9, the coherent structure drifts very slowly and η_* changes little over a very long time period.

The fifth-order system of (3.83) may be solved subject to the far-field constraints that γ and \mathcal{U} approach constant values as $|\xi| \rightarrow \infty$. In particular, since the coherent structure invades a region to the right in which $\gamma_r = \mathcal{U} = 0$, but \mathcal{U} remains finite to the left (see figure 3.9), we demand the limits

$$(\gamma_r, \gamma_i, \mathcal{U}) \rightarrow \begin{cases} (0, G_+, 0) & \text{for } \xi \rightarrow \infty, \\ (0, G_-, \Delta\mathcal{U}) & \text{for } \xi \rightarrow -\infty, \end{cases} \quad (3.84)$$

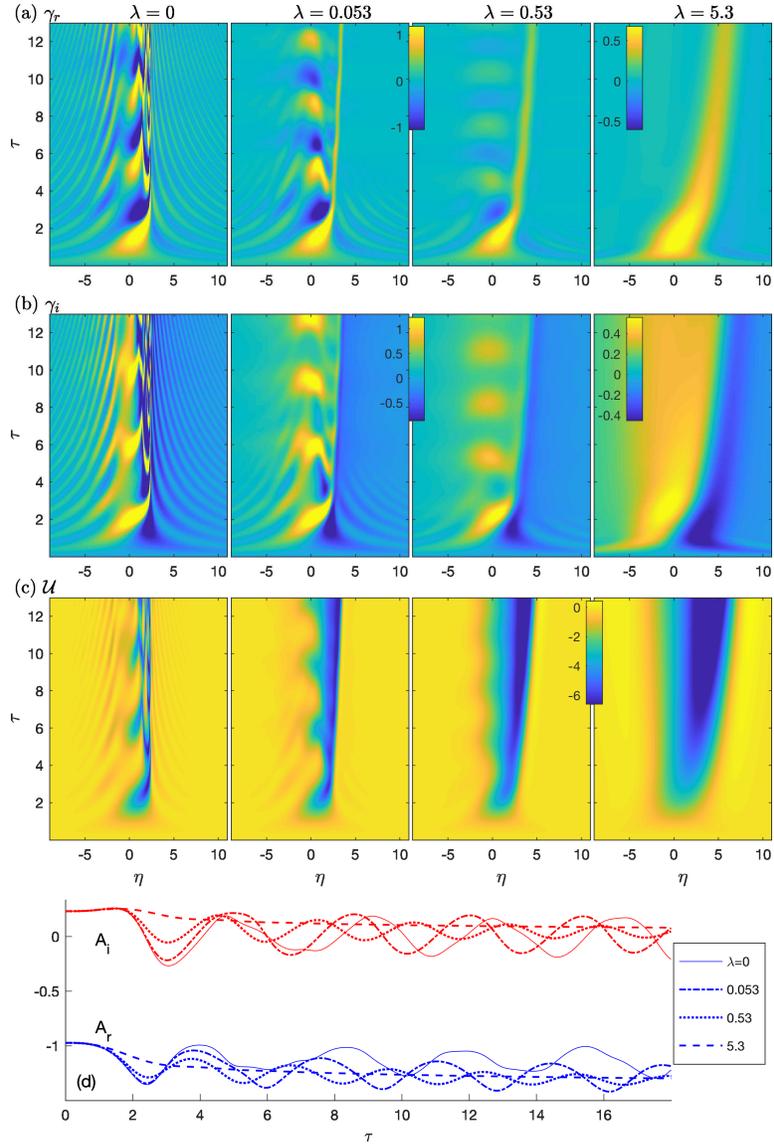


Figure 3.9: Solutions of the modified canonical model, showing (a) γ_r , (b) γ_i and (c) \mathcal{U} , for $m = 1/2$, $f = 4/3$ and $\mathcal{N} = 4/3$, $c_1 = 0.238$, $c_2 = 0.219$ with the values of λ indicated (and corresponding to the three columns). The colormap is the same in the first three panels of (a) and (b), but not the rightmost panel. The quasi-steady wave amplitude $A = A_r + iA_i$ for the four computations is shown in (d).

where $G_+ = A\eta_*^{-1}$, $G_- = A(\eta_* + \Delta\mathcal{U})^{-1}$ and $\Delta\mathcal{U}$ is the jump in the mean flow across the structure. (3.84) imposes six boundary conditions to (3.83). One must also remove the translational invariance of the system by imposing an additional constraint. Thus, given η_* , we solve (3.83) subject to those seven conditions, treating G_- and c as unknown parameters (eigenvalues). This furnishes localized structures taking the form of “pulses” in γ_r and “fronts” in γ_i and \mathcal{U} . Note that, as the coherent structure drifts to the right, η_* increases, corresponding to an evolution of the coherent structure, which is treated parametrically in the quasi-steady approximation of (3.82) and (3.83).

Figure 3.10 shows a sample solution to (3.83) for $(A, \eta_*) = (-1.2, 5.04)$, giving $G_+ = -0.24$. These choices for A and η_* correspond to the numerical solution of the modified canonical model for $\lambda = 0.53$ shown in figure 3.9 at $\tau \approx 18$, and they are also plotted in figure 3.10. The solution to (3.83) compares satisfyingly with the snapshot of the simulations near the core of the coherent structure, although there are discrepancies further away arising from the influence of the far-field flow.

Figure 3.10 also includes data computed from (3.83) for G_- , c and $\Delta\mathcal{U}$, as functions of η_* . In the limit of large η_* , a simple rescaling of (3.83) and (3.84) indicates the limiting behaviour,

$$G_- \rightarrow G_+ = O(\eta_*^{-1}), \quad c = O(\eta_*^{-5/2}), \quad \Delta\mathcal{U} = O(\eta_*). \quad (3.85)$$

The solution of (3.83) is compared to (3.85) together with measurements from the numerical simulation in the figure. Similarly, the characteristic strength and width of the structure are $\gamma = O(\eta_*^{-1})$, $\mathcal{U} = O(\eta_*)$ and $\xi = O(\eta_*^{-1/2})$. Thus, as the coherent structure drifts to the right, and η_* slowly increases, the drift velocity declines, and the peak in γ_r and jump in γ_i must decrease and narrow. However, the jump in $\Delta\mathcal{U}$ continues to build up, predicting that the deficit in the mean flow grows linearly with η for $\eta < \eta_*$.

This behaviour of the coherent structure rationalizes the dynamics of the modified canonical model seen in figure 3.9: once the linear dynamics and nonlinear focussing have subsided, the two features that remain are the decaying oscillations near $\eta = 0$ and the drifting coherent structure. The structure leaves in its wake a slowly diffusing density perturbation $\gamma \approx iG_-$ (see the right-hand plots in fig-

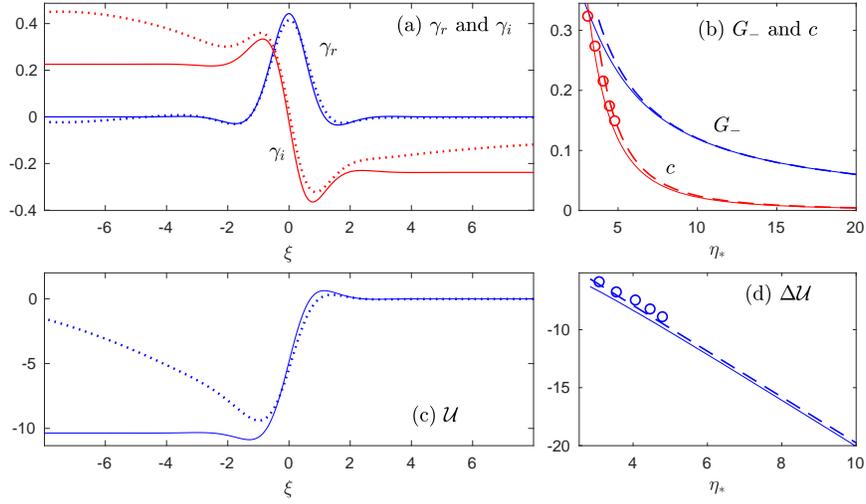


Figure 3.10: A coherent structure computed from (3.83) with $\eta_* = 5.04$ and $A = -1.2$, showing (a) γ_r and γ_i , and then (c) \mathcal{U} (solid lines). The dotted lines show the numerical solution of the modified canonical model (3.79), computed for $\lambda = 0.53$ at $\tau = 17.8$ (at which moment the residual oscillations near $\eta = 0$ are less pronounced). In (b) and (d) we show G_- , c and the jump $\Delta\mathcal{U} = -[\mathcal{U}]_{-\infty}^{\infty}$ against η_* from the solutions to (3.83) for $A = -1.2$ (solid lines). The dashed lines show the limiting behaviour for $\eta_* \gg 1$ given in (3.85). The circles show data for c and $\Delta\mathcal{U}$ measured from the numerical solution of (3.79) with $\lambda = 0.53$ from $\tau = 5$ to $\tau = 15.4$.

ure 3.9(b)) and a gradually strengthening mean flow correction $\Delta\mathcal{U}$, as seen on the right of figure 3.9(c). Thus, with diffusion, all growth in the density perturbation becomes arrested, leaving a widening and strengthening, jet-like defect in the streamwise mean flow.

One final concern is the impact of viscosity on the dynamics of the coherent structure: it is clear from (3.77) that the growth of the mean flow correction may be halted when $\nu = O(\varepsilon^2)$. Indeed, in the limit of stronger diffusion, the viscous term may allow \mathcal{U} to also reach a steady state within the critical layer. However, as for the classical critical layers of Rossby waves [72] and clear from the constraint (3.78), a genuine steady state is not possible with dissipation. Instead, the mean-

flow correction must inevitably spread viscously out of the critical layer, even if a quasi-steady state is reached locally. Such considerations suggest that viscosity, if sufficiently strong, may prevent the creation of the drifting coherent structure, although a widening jet-like defect might still appear in the mean flow.

3.7 Discussion

In this chapter, we have studied the non-dissipative, nonlinear dynamics of forced baroclinic critical layers using matched asymptotic expansion. In the linear regime, the forcing establishes a steady wave response outside the critical layers, but disturbances grow secularly inside the critical layer, which thins with time. The behavior is very similar to the forced critical layers of both Rossby and internal gravity waves [1, 2, 16, 17, 67]. Continuing the analysis, we then studied the weakly nonlinear dynamics of the critical layer, finding that the adjustment of the mean flow provides the most important feedback on the growing disturbance there. Guided by the critical-layer scalings exposed by the weakly nonlinear analysis, we then derived a reduced model for the nonlinear critical layer. The numerical solution of the reduced model reveals a continued growth of the vertical vorticity as the disturbance is focussed exponentially quickly into a finer region within the critical layer. The focussing progresses uninterrupted until the reduced model breaks down.

Such pathological behaviour is quite different to that of the forced critical layer of a Rossby wave, where nonlinearity halts the secular linear growth and the mean vorticity distribution overturns into a distinctive cat's eye structure [1, 2, 24]. In that process, all the harmonics of the forcing pattern are excited to the same strength of the fundamental component. By contrast, in our nonlinear theory of the forced baroclinic critical layer, the adjustment to the mean flow arrests the linear growth and prompts the focussing of the vorticity before any of the higher harmonics becomes important. It is only once the strength and lengthscale of the focussed vorticity pass out of the asymptotic regime of our theory that the harmonics will appear. One important contributor to this feature is that the position of the baroclinic critical level itself is dependant on the streamwise wavenumber. The critical level of the forcing does not therefore coincide with those of the harmonics. This filtering action weakens the impact of those harmonics within the baroclinic critical layer,

leaving the adjustment of the mean flow as the main nonlinearity.

The nonlinear structures developed in our forced baroclinic critical layers (jet-like defects in the mean velocity and dipolar stripes in the vorticity) may well be the analogues of the features seen in the simulations of Marcus *et al.* (2013) and Wang [73]. Unlike in the reduced model, however, where these structures continue to focus, the mean flow structures spawned in the simulations roll up into new vortices, providing part of the chain of events leading to self replication. Thus, our model likely misses important secondary instabilities. Indeed, Killworth and McIntyre [24] and Haynes [3] have shown that the nonlinear evolution of a forced Rossby wave can be susceptible to shorter-wavelength shear instabilities and generate “critical layer turbulence” along the filaments of vorticity wrapped around the main cat’s eye (see also [74]). We will study the secondary instability caused by the mean-flow defect in Chapter 4, which may explain the rollup and self-replication of zombie vortices.

Marcus *et al.* [59] further argued that self replication is a finite-amplitude instability, requiring the amplitude of the initial disturbance to exceed a certain threshold. By contrast, the secular growth and nonlinear focussing of the disturbance inside the critical layer is triggered for an arbitrarily small forcing amplitude in our analysis. Nevertheless, we have idealized the driving as a steady wavemaker, and ignored any possible evolution of that forcing. If the wavemaker cannot be sustained indefinitely, a threshold likely emerges that demands that the forcing act for sufficient time and strength to drive the baroclinic critical layers to the point where secondary instability can arise.

The continued focussing of the mean vorticity layer also indicates that dissipative effects are likely to become important in the later stages of evolution inside the baroclinic critical layer. Including the diffusion of density (*i.e.* heat or salt) in the theory leads to a modification of the reduced model, which now takes a partial differential form. A brief exploration of the modified model demonstrates that weak diffusion can arrest the focussing to the nonlinear critical level. Interestingly, a drifting solitary-wave like object then emerges, with a structure that can be analyzed analytically. The solitary wave leaves in its wake another jet-like defect in the mean flow, but this time the defect gradually widens and deepens as the object slowly drifts.

In summary, when a steady forcing drives waves with baroclinic critical levels into a horizontally sheared flow with vertical stratification, the growing density perturbations predicted by linear theory become saturated by nonlinear effects. Although this saturation is demanded by the conservation laws of the governing equations, those constraints still permit the density perturbation to develop finer spatial structure over a region within the baroclinic critical layer. This nonlinear focussing effect takes place exponentially quickly, developing sharp jet-like defects in the mean flow, which can survive even in the presence of weak dissipation. This dynamics of the baroclinic critical layers is more destructive than that for the classical critical layers of Rossby waves and internal gravity waves, and plausibly rationalizes part of the cycle of vortex self replication observed by Marcus et al. [6] in numerical simulations.

Chapter 4

Secondary instability of forced baroclinic critical layers

4.1 Introduction

In Chapter 3, we have proposed that a secondary instability could take place in the baroclinic critical layer, which may account for the rollup of critical layers into new zombie vortices as in the simulation of Marcus et al. [6]. In this chapter, we show that such a secondary instability does exist, and we will study its linear evolution in detail.

For Rossby wave critical layers, Killworth and McIntyre [24] have shown that Stewartson [1] and Warn and Warn [2]'s solution for the critical layer's nonlinear evolution is unstable to short-wave disturbances. The instability is a two-dimensional shear instability caused by the local reversal of vorticity gradient. Haynes [27] and Haynes [3] performed numerical computation for the later nonlinear evolution of the instability. It was shown that the instability could bring about small structures in the critical layer, which can cause considerable rearrangement of vorticity and change the critical layer's momentum absorptivity.

One important property of the critical-layer instability is that it is caused by the strong vorticity gradient localized in the small critical-layer region. This structure allows for a matched asymptotic analysis which leads to a simple dispersion relation represented by an integral [24]. This type of instability has been studied

earlier by Gill [25]. He showed that a localized vorticity defect seeded in a Couette flow is sufficient to make the flow unstable, and derived an integral relation for the stability boundary. Balmforth et al. [26] further exploited the mathematical simplification of the defect dynamics and made a more comprehensive analysis. They studied instabilities of typical defect profiles, provided the general conditions for instability and solved the initial value problems.

Another important aspect of the critical-layer instability is that the basic flow is unsteady as the critical layer itself evolves. This renders difficulties to the analysis since a normal-mode solution is not valid in general for unsteady basic flows. The analysis of Killworth and McIntyre [24] has circumvented this difficulty by taking advantage of the scale separation of Stewartson [1] and Warn and Warn [2]: under the condition that the streamwise wavelength of the forcing is much longer than the cross-stream length scale, the instability grows much faster than the evolution of the critical layer itself, so we may parameterize the time of the latter and simply study the normal mode at different times. When such a scale separation is not available, the normal mode solution does not yield valid results and one may have to study the temporal evolution via numerical simulation [3]. More generally, so far, only several types of unsteady basic flows have been found to allow for analytical insights to their instabilities. As summarized by Drazin and Reid [12], the few examples include periodic basic flows [75], basic flows whose time and spatial variables are separable [76] and the solitary wave solution of KdV equation [77].

For our current problem, we have shown in Chapter 3 that the baroclinic critical layer evolves into a pattern of a dipolar vortex stripe, which possesses local maxima and minima of vorticity. We therefore expect the baroclinic critical layer admits a secondary instability similar to that of Rossby-wave critical layers and vorticity defects. We can then take advantage of the matched asymptotic analysis in those problems. Surprisingly, this method helps us to derive an asymptotic solution for the temporal evolution of the secondary instability. We therefore find another example where we may study the instability of unsteady flows with analytical insights.

Our secondary instability also shares similarity with the defect instability of Umurhan et al. [32]. They also considered stratified flow with horizontal shear, and seeded a defect in the basic flow. The defect then induces linear instability, but

different from previous defect instabilities of two-dimensional flows, the unstable mode possess two baroclinic critical levels, which may represent the replication of zombie vortices. Our picture is very similar, but instead of using simple idealized defect profiles as they did, our mean-flow defect is the actual unsteady mean-flow response we obtained in Chapter 3. Hence our results may be more realistic in view of explaining the zombie vortex instability.

Boulanger et al. [38] have shown that the baroclinic critical layers in stratified, tilted vortices also suffer secondary instability. The instability is induced by the strong shear of vertical velocity of the linear viscous critical layer. But their critical layer has specific properties endowed by the tilted vortices which may not apply to our critical layer, hence their instability is a little more different from ours, and we will not further comment.

The layout of this chapter is as follows. In Section 4.2, in view that the baroclinic critical layer evolves in a slow time scale, we first parameterize time and undertake a normal mode analysis. This allows us to gain important insights to the properties of the secondary instability. We will identify similarities to defect instabilities in previous studies. Then in Section 4.3, we take advantage of these properties to build an asymptotic theory for the secondary instability. The analysis is again similar to that for defect dynamics, but we may include the effect of unsteady basic flow and derive the actual temporal evolution of the secondary instability. We will also consider the instability of various wavenumbers, bringing the special behavior of higher harmonics into attention. We confirm the asymptotic theory in Section 4.4, where we solve the equations of the secondary instability numerically. Finally, we conclude in Section 4.5.

4.2 Normal mode analysis

To study the secondary instability, we first need to decide the basic flow which small disturbances should be added to. In the analysis of Killworth and McIntyre [24] for Rossby wave critical layers, the basic flow was the nonlinear solution of Stewartson [1] and Warn and Warn [2] which include all wave components. But in our problem, in Chapter 3 we have shown that higher-order harmonics remain small until the nonlinear model breaks down. For the fundamental mode, we do

not think it will give rise to secondary instability of Rossby waves by itself, as the instability in Killworth and McIntyre [24] appeared after the critical layer becomes nonlinear. Boulanger et al. [38] have shown that the linear baroclinic critical layer itself does admit a secondary instability, but the instability is strongly dependent on their tilted-vortex structure, for example, a small tilting angle which allows for a vertical scale separation between the linear critical layer and its instability. Such properties are not available in our problem, so we think the fundamental mode is unlikely to induce a secondary instability. The mean-flow response in the baroclinic critical layer, however, has strong similarities with the basic-flow defects of Gill [25] and Balmforth et al. [26], and is therefore very likely to be unstable. So we will use the mean-flow response as the basic flow for the secondary instability. The benefit of this approach is that we may consider the instability of any wavenumber, otherwise the wave of the disturbance must be much shorter than the streamwise length scale of the basic flow, as in Killworth and McIntyre [24].

We continue with the model in Chapter 3, and consider the basic flow of velocities $[U(y,t), 0, W(y,t)]$ and density $G(y,t)$, which is the time-dependent mean flow in the weakly nonlinear analysis we studied in Section 3.4:

$$U = y + \Delta U, \quad \Delta U = \varepsilon^2 u_0, \quad W = \varepsilon^2 w_0, \quad G = \varepsilon^2 \rho_0, \quad (4.1)$$

where u_0 , w_0 and ρ_0 are defined in (3.35). Unless specially indicated, we use the same notations as in Chapter 3. We consider U , W and G to be forced by linear disturbances (*i.e.*, we substitute the solution of linear critical layers in Section 3.3.2 to the mean-flow responses (3.36)-(3.38)), because we will see that the secondary instability first appears when the baroclinic critical layer is still linear. In this regime, ΔU , W , and G are all of order $O(\varepsilon^2 t^2)$ (up to $\log t$ corrections), and their length scales shrink with time in the form of $|y - \mathcal{N}| \sim t^{-1}$.

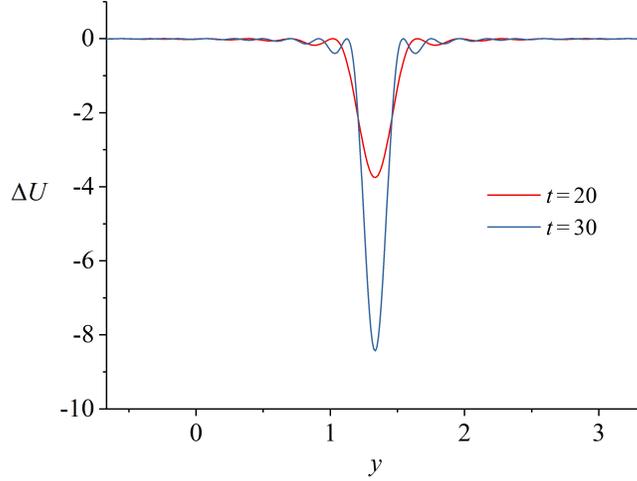


Figure 4.1: Mean-flow modification of streamwise velocity ΔU at $t = 20$ and $t = 30$, $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$.

We will see that ΔU plays the key role in the secondary instability, and we have a simple expression for it:

$$\Delta U = \frac{\varepsilon^2 m^2 \cos(y - \mathcal{N})t - 1}{\mathcal{N} (y - \mathcal{N})^2}. \quad (4.2)$$

In figure 4.1, we have plotted two sample profiles of ΔU . We see that ΔU features a jet-like profile localized around the baroclinic critical level $y = \mathcal{N}$. Such a localized mean-flow modification added to the linear shear flow is similar to the structure of defect dynamics as in Gill [25], Balmforth et al. [26] and Umurhan et al. [32]. The profile admits inflectional points, so it is likely to be unstable. Another important property of ΔU is that it is unsteady: it keeps strengthening and sharpening. As a result, its instability could admit significant differences to that of steady flows.

We consider small disturbances (u', v', w', p', ρ') added to such unsteady mean flows, then we linearize the governing equations and obtain:

$$u'_t + Uu'_x + Wu'_z + (U_y - f)v' = -p'_x, \quad (4.3)$$

$$v'_t + Uv'_x + Wv'_z + fu' = -p'_y, \quad (4.4)$$

$$w'_t + Uw'_x + Ww'_z + v'W_y + \rho' = -p'_z, \quad (4.5)$$

$$\rho'_t + U\rho'_x + W\rho'_z + v'G_y - \mathcal{N}^2 w' = 0, \quad (4.6)$$

$$u'_x + v'_y + w'_z = 0. \quad (4.7)$$

We have again neglected dissipation effects as the first step to understand the secondary instability. As before, we may derive an equation for the vertical component of vorticity:

$$\left(\frac{\partial}{\partial t} + U \frac{\partial}{\partial x} + W \frac{\partial}{\partial z} \right) (v'_x - u'_y) - v'U_{yy} + (U_y - f)w'_z - W_y u'_z = 0. \quad (4.8)$$

In this section, we consider the normal mode solution

$$(u', v', w', \rho', p') = \varepsilon[\check{u}(y), \check{v}(y), \check{w}(y), \check{\rho}(y), \check{p}(y)]e^{ik(x+\mu z-ct)} + \text{c.c.}, \quad (4.9)$$

where k is the wavenumber in the horizontal streamwise direction (scaled by the wavenumber of the forcing k_x), and μ is the ratio of vertical wavenumber to horizontal wavenumber. We search for the phase velocity c such that it enables all disturbances to decay to 0 as $y \rightarrow \pm\infty$ and therefore serves as an eigenvalue to the linear system.

Since U , W and G are time-dependent, the solution of \check{u} , \check{v} , \check{w} , $\check{\rho}$, \check{p} and c will be functions of time, too. But in the previous chapter, we have shown that U , W and G evolve in a slow time scale, so we could expect the time-dependence of \check{u} , \check{v} , \check{w} , $\check{\rho}$, \check{p} and c are also weak, and therefore, as a leading order approximation, we parameterize time (*i.e.*, we neglect ∂_t except on e^{-ikct}). This approach will enable us to gain essential insights to the secondary instability, which helps us to further probe the actual time-dependence later on.

We may solve the eigenvalue problem first by a shooting method. After some

algebraic manipulation, we can obtain an ODE for \check{p} :

$$\begin{aligned} \check{p}_{yy} + \left[-\frac{2D_c D_{c,y} + f U_{yy}}{D_f} + k\mu \frac{(D_{\mathcal{N}} - D_c^2)W_y - iG_y D_c}{D_{\mathcal{N}} D_c} \right] \check{p}_y \\ + \left[kf \frac{2D_c D_{c,y} + f U_{yy}}{D_c D_f} - k^2 + \frac{k\mu}{D_{\mathcal{N}}} \left(-k\mu D_f + kf \frac{D_c W_y + iG_y}{D_c} \right) \right] \check{p} = 0, \end{aligned} \quad (4.10)$$

where

$$\begin{aligned} D_c &= k(U - c + \mu W), \quad D_f = k^2(U - c + \mu W)^2 - f(f - U_y), \\ D_{\mathcal{N}} &= k^2(U - c + \mu W)^2 - \mathcal{N}^2. \end{aligned} \quad (4.11)$$

We require $\check{p} \rightarrow 0$ for both $y \rightarrow \infty$ and $y \rightarrow -\infty$, and this poses an eigenvalue problem for c . In practice, it is helpful to exploit the asymptotic limit of \check{p} at large $|y|$:

$$\check{p} \approx C_+ e^{-k\sqrt{1+\mu^2}y} \quad y \gg 1, \quad \check{p} \approx C_- e^{k\sqrt{1+\mu^2}y} \quad y \ll -1, \quad (4.12)$$

where C_- are C_+ constants. The problem is then translated to finding c and C_+/C_- , which may be solved by shooting from $|y| \gg 1$ to the central region.

To confirm the results, we also solve the problem by a matrix method. We take the divergence of (4.3), (4.4) and (4.5) and use (4.7) to eliminate u'_i , v'_i and w'_i , then we can derive an integral representation of \check{p} without c :

$$\begin{aligned} \check{p} = & -\frac{e^{-k\sqrt{1+\mu^2}y}}{2\sqrt{1+\mu^2}} \int_{-\infty}^y e^{k\sqrt{1+\mu^2}y'} \left[(f - 2U_y - 2\mu W_y) i\check{v} + f\sqrt{1+\mu^2}\check{u} - i\mu\check{p} \right] dy' \\ & -\frac{e^{k\sqrt{1+\mu^2}y}}{2\sqrt{1+\mu^2}} \int_y^{\infty} e^{-k\sqrt{1+\mu^2}y'} \left[(f - 2U_y - 2\mu W_y) i\check{v} - f\sqrt{1+\mu^2}\check{u} - i\mu\check{p} \right] dy'. \end{aligned} \quad (4.13)$$

We then discretize the domain into a finite number of grid points: $[y_1, y_2, \dots, y_n]$, and we define a vector to represent the values of \check{u} , \check{v} , \check{w} and \check{p} on these points:

$$X = [\check{u}_1, \check{u}_2 \dots \check{u}_n, \check{v}_1, \check{v}_2 \dots \check{v}_n, \check{w}_1, \check{w}_2 \dots \check{w}_n, \check{\rho}_1, \check{\rho}_2 \dots \check{\rho}_n]^T. \quad (4.14)$$

We evaluate the integrals in (4.13) numerically, using the values of integrands on the grid points, then substitution into (4.3)-(4.6) yields a discretized linear system in form of

$$MX = ikcX, \quad (4.15)$$

where M is a large matrix of size $(4n \times 4n)$. Finally, we solve the eigenvalues of M to obtain c . This method is relatively expensive and less precise, but can give us all the possible eigenvalues.

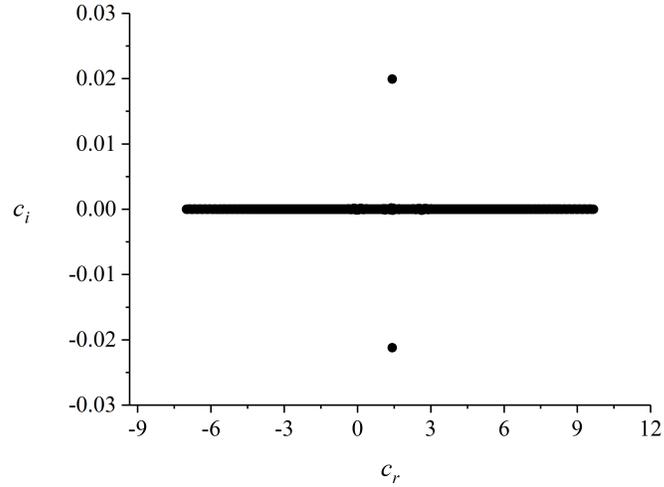


Figure 4.2: Eigenvalue c computed by the method of matrix eigenvalues (4.15). The parameters are $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $k = 1$, $\mu = 1/2$ and the mean-flow modifications are chosen at $t = 30$. The computation domain is $y \in [-7 + \mathcal{N}, 7 + \mathcal{N}]$, which is discretized by 3000 evenly distributed grid points. (4.13) is evaluated numerically using a composite trapezoidal rule.

Figure 4.2 shows an example for the numerical results of c computed by the method of matrix eigenvalues. For each eigenvalue, we plot a dot for its real part c_r versus imaginary part c_i . We consider the parameter set of $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$ as before. The parameterized time which we used to evaluate the mean flow is chosen at $t = 30$. The computation domain is $y \in [-7 + \mathcal{N}, 7 + \mathcal{N}]$ where we put 3000 evenly distributed grid points to evaluate (4.13) numerically

using a composite trapezoidal rule. For the disturbance, we select the Fourier component of the forcing $k = 1$, $\mu = m$ as a typical example. This is the dominant Fourier component observed in the rollup of the simulation of Marcus et al. [6] and Wang [73], which may be because the forcing and the secondary instability share the same periodic domain, and the fundamental mode has much stronger excitation compared to the harmonics.

We see a horizontal thick line lying on the real axis. These eigenvalues correspond to $D_c = 0$ and $D_{\mathcal{N}} = 0$ in (4.11), which are

$$c_V = U + \mu W \quad c_G = U + \mu W \pm \mathcal{N}. \quad (4.16)$$

c_V and c_G are the Doppler-shifted phase velocities of vortical modes and internal gravity waves, respectively. Since U varies continuously with y due to the shear, we get a ‘continuous spectrum’ of eigenvalues.

We also see a pair of dots lying above and below the continuous spectrum. These discrete eigenvalues are the normal modes. In particular, the one with a positive imaginary part is an unstable mode, signifying the existence of the secondary instability. Here we only have one unstable mode, and in fact, it is associated with one of the two inflectional points of the major peak of ΔU in figure 4.1. From a Nyquist method, Balmforth et al. [26] have shown that if the basic flow has N inflectional points, there can be at most $(N + 1)/2$ unstable modes. At later times, we may have more unstable modes as the inflectional points of adjacent small peaks come into effect, but their instabilities are much weaker and we will not consider them.

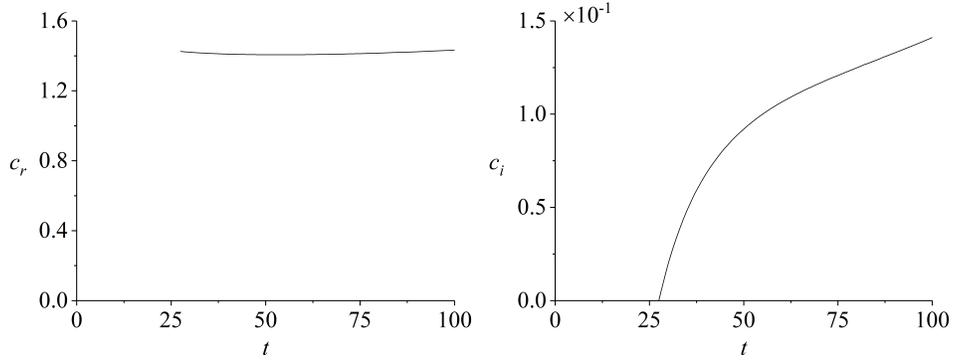


Figure 4.3: Eigenvalue c at different times, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$.

Figure 4.3 shows the eigenvalue c computed by basic flows at different times. Before $t = 27.5$, the mean-flow modification is too weak to support a normal mode, and we only have the continuous spectrum in figure 4.2. At $t = 27.5$, a mode first emerges with $c_i = 0$; it then becomes more unstable due to a strengthening mean-flow modification. The real part of the phase velocity c_r , however, remains nearly a constant. Its value is always close to $\mathcal{N} = 4/3$, and importantly, it is distinct to $c = 0$ of the fundamental mode generated directly by the forcing. The baroclinic critical layer of the fundamental mode around $y = \mathcal{N}$ is therefore the classical critical layer for the secondary instability.

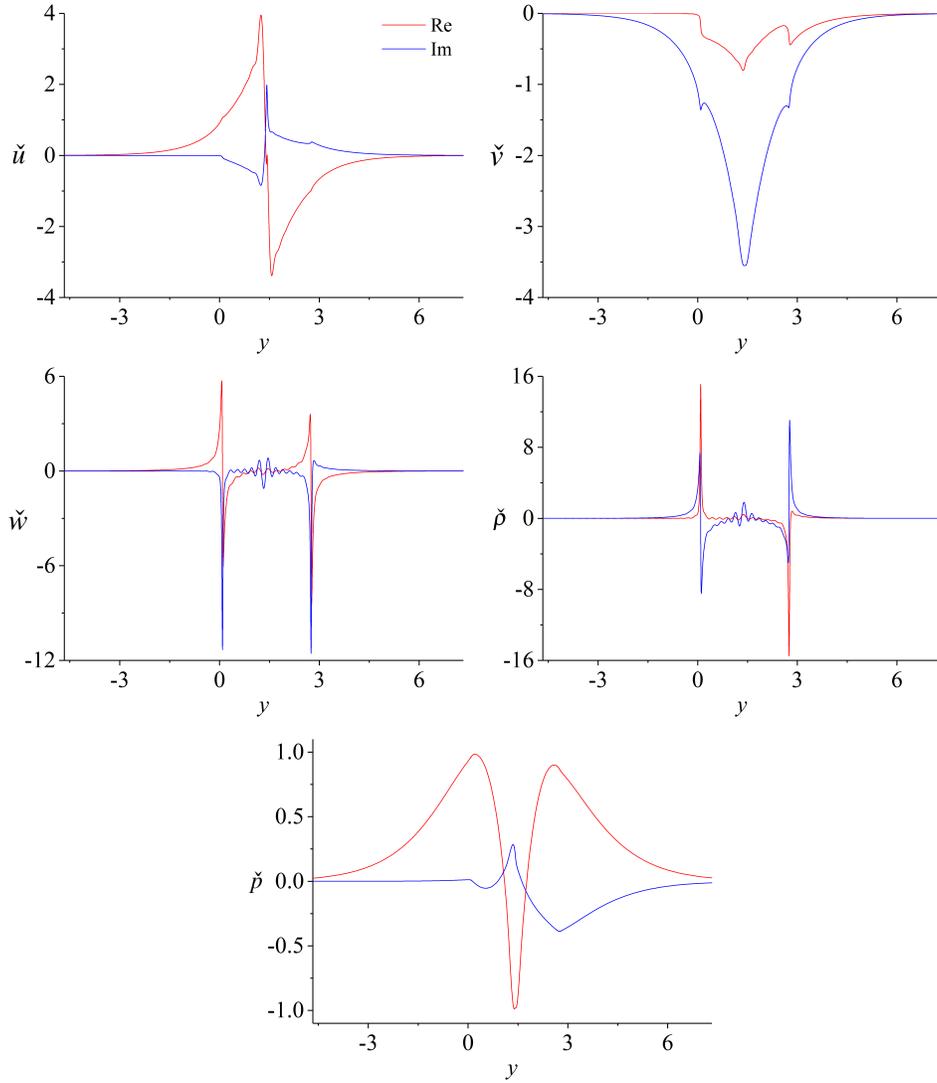


Figure 4.4: Eigenfunctions of the unstable mode at $t = 30$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$.

The eigenfunctions of the secondary instability are shown in figure 4.4. Their distributions are centered around their classical critical level near $y = \mathcal{N}$. Here we have \tilde{u} , \tilde{v} much stronger than \tilde{w} , and \tilde{u} features a jump across the critical level. These characteristics are the hallmark of 2-dimensional horizontal shear instability. We also observe that $\tilde{\rho}$ and \tilde{w} have large amplitudes near their baroclinic critical

levels $y = c \pm \mathcal{N}$. $y = 2\mathcal{N}$ is essentially a new critical level and we think its formation explains how the zombie vortex can replicate itself. Our eigenfunctions share qualitative similarities with Umurhan et al. [32]'s results: they also have strong amplitudes of density and vertical velocity near the new baroclinic critical levels. But one of their baroclinic critical level is significantly strong than the other, which we guess may be because they used a different defect from ours.

With the results we obtained, we now propose a mechanism for the secondary instability, from which we may obtain its characteristic scales. One way to understand the secondary instability is to look at the situation without the mean-flow modification where such unstable modes do not exist, and then consider how a localized mean-flow modification makes them possible. We define \check{p}_- and \check{p}_+ as the solution of (4.10) with the same c as \check{p} but with $U = y, V = W = G = 0$. Clearly \check{p}_- and \check{p}_+ cannot decay to zero for both $y \rightarrow \infty$ and $y \rightarrow -\infty$, but they can do so in one of the limits, which we define as:

$$\check{p}_- \rightarrow 0 \quad \text{as } y \rightarrow -\infty, \quad \check{p}_+ \rightarrow 0 \quad \text{as } y \rightarrow +\infty. \quad (4.17)$$

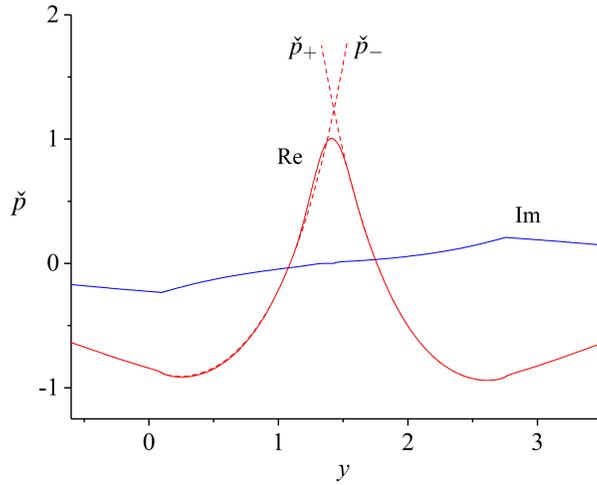


Figure 4.5: \check{p} , \check{p}_+ and \check{p}_- at the onset of secondary instability at $t = 27.5$, $c = 1.425$; $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$.

Figure 4.5 shows \check{p} and \check{p}_\pm at the onset of instability $t = 27.5$. In this situation

$c = 1.425$, which is close to $\mathcal{N} = 4/3$. Away from the critical level, the mean-flow modification is negligible, so \check{p} is the same as \check{p}_\pm for $y \gtrless \mathcal{N}$. In the critical layer, \check{p}_y gradually reverts its direction and connects $\check{p}_{+,y}$ and $\check{p}_{-,y}$ on two sides, making it possible that $\check{p} \rightarrow 0$ for both $y \rightarrow \infty$ and $y \rightarrow -\infty$. Hence the essential condition for a normal mode to form is a sufficiently strong \check{p}_{yy} to revert the direction of \check{p}_y across the critical layer. We may use this information to obtain the scales of the critical layer at the onset of secondary instability.

Let us assume that the onset of secondary instability appears at $t = t_s$, and the baroclinic critical layer has length scale $|y - \mathcal{N}| \sim \delta_s$ at this time. From the dynamics of linear baroclinic critical layers, we have $\delta_s \sim t_s^{-1}$. \check{p} changes little in the critical layer, so we may normalize

$$\check{p} \approx 1, \quad y \approx \mathcal{N}. \quad (4.18)$$

\check{p}_y changes continuously from $\check{p}_{-,y}$ to $\check{p}_{+,y}$ which are both $O(1)$ determined by the outer flow. Hence to revert the direction of \check{p}_y , we need

$$\check{p}_{yy} \sim \frac{\check{p}_{+,y} - \check{p}_{-,y}}{\delta_s} \sim \delta_s^{-1} \quad (4.19)$$

From (4.10), the leading-order of \check{p}_{yy} is given by

$$\check{p}_{yy} \sim k \frac{U_{yy}}{D_c} \check{p} \sim \frac{U_{yy}}{\delta_s}. \quad (4.20)$$

where we have used $D_c = k(U + \mu W - c) \approx k(y - \mathcal{N}) \sim k\delta_s$ in the critical layer. Comparing with (4.19), we therefore need

$$U_{yy} \sim 1 \quad (4.21)$$

for the secondary instability. From the mean-flow response of baroclinic critical layer, (*cf.* (4.2)),

$$U_{yy} \sim \frac{\varepsilon^2}{\delta_s^4} \quad (4.22)$$

Thus we finally obtain

$$\delta_s \sim \varepsilon^{\frac{1}{2}}, \quad t_s \sim \varepsilon^{-\frac{1}{2}}. \quad (4.23)$$

To confirm this conclusion, we have computed the values of t_s at various ε , shown in figure 4.6. We see that the relation $t_s \sim \varepsilon^{-\frac{1}{2}}$ is very well satisfied. Since $\varepsilon^{-\frac{1}{2}} < \varepsilon^{-\frac{2}{3}}$, the secondary instability first appears when the baroclinic critical layer is still linear. This justifies our analysis of using weakly nonlinear mean-flow responses as the basic flow, but it could also mean that the nonlinear dynamics of baroclinic critical layers we studied in Chapter 3 may be affected by the secondary instability.

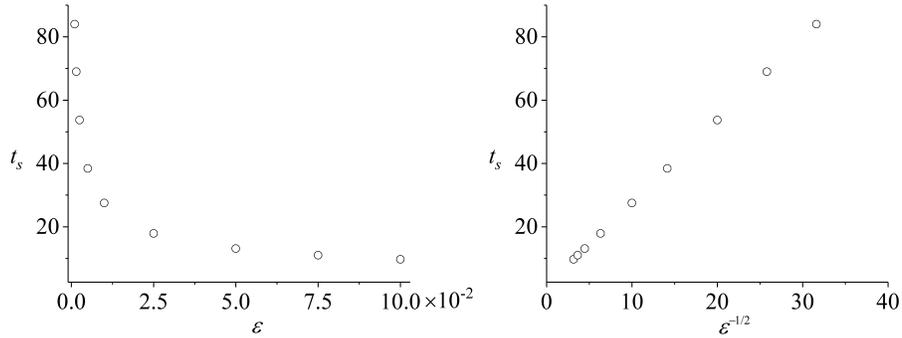


Figure 4.6: Time of onset of instability t_s versus ε and $\varepsilon^{-\frac{1}{2}}$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$.

We identify significant similarities of our current unstable modes to those in the defect instability of Gill [25] and Balmforth et al. [26]. That c is close to the basic flow velocity in the critical layer, that \check{p} is nearly constant but \check{p}_y has a jump, that U_{yy} plays the crucial role and that the length scale is in order of square root of the defect amplitude, are all essential properties of the defect instability. Moreover, although our system is more complicated with density stratification, vertical velocity and densities disturbances, they do not come into the leading order dynamics of the secondary instability. All these similarities suggest that the secondary instability shares the fundamental mechanism with the defect instability, so we may adopt their tool of matched asymptotic analysis to gain further analytical insights, which will be undertaken in the next section.

At the end of this section, we comment on the validity of the normal mode analysis. The assumption of the analysis is that we may parameterize time, and it is justified if the variation of the mean flow is slow compared to that of the mode e^{-ikct} itself. We have found $c \approx \mathcal{N}$ to leading order, so the mode does oscillate

much faster than the mean flow. Hence we expect the normal mode $\sim e^{-ik\mathcal{N}t}$ as well as the profiles of the eigenfunctions are good approximations to the actual solution. However, the growth rate c_i shown in figure 4.3 is not big. One can find that c_i is in the same order of δ_s which is the order of mean flow's rate of change. As a result, the value of c_i may not be the actual growth rate of the secondary instability; all we can say is that the flow will become unstable and grow over a time scale of δ_s^{-1} . Similarly, the precise value of t_s does not make sense, either; it only tells us the order of time when the instability takes place. We will reveal the actual evolution of the secondary instability in the next section.

4.3 An asymptotic theory for secondary instability

In this section, we build an asymptotic theory for the secondary instability, taking advantage of the small scale of the critical layer and the properties of the secondary instability obtained in the normal mode analysis. The matched asymptotic analysis we use is very similar to that of Gill [25] and Balmforth et al. [26], but we may take into account the unsteadiness of the basic flow and derive the actual temporal evolution of the secondary instability.

The secondary instability has phase velocity close to \mathcal{N} , and outside the critical layer, the mean-flow modification is negligible. So we may construct a quasi-steady wave solution for the outer region:

$$(u', v', w', p', \rho') = \varepsilon \Phi(T) [\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{p}(y), \hat{\rho}(y)] e^{ik(x+\mu z - \mathcal{N}t)} + \text{c.c.}, \quad (4.24)$$

$\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{p}(y), \hat{\rho}(y)$ are the steady wave solution without mean-flow modification, and the amplitude Φ varies with the slow time $T = \varepsilon^{\frac{1}{2}}t$ of the critical layer. The exponential growth of the secondary instability is included in $\Phi(T)$. $\hat{u}(y), \hat{v}(y), \hat{w}(y), \hat{p}(y), \hat{\rho}(y)$ are governed by

$$\begin{aligned} & \hat{p}_{yy} - \frac{2k^2(y - \mathcal{N})}{k^2(y - \mathcal{N})^2 - f(f-1)} \hat{p}_y \\ & - \left[k^2 \frac{k^2(y - \mathcal{N})^2 - f(f+1)}{k^2(y - \mathcal{N})^2 - f(f-1)} + k^2 \mu^2 \frac{k^2(y - \mathcal{N})^2 - f(f-1)}{k^2(y - \mathcal{N})^2 - \mathcal{N}^2} \right] \hat{p} = 0, \end{aligned} \quad (4.25)$$

and

$$\begin{aligned}\hat{u} &= \frac{(f-1)\hat{p}_y - k^2(y-\mathcal{N})\hat{p}}{k^2(y-\mathcal{N})^2 - f(f-1)}, & \hat{v} &= \frac{ik[(y-\mathcal{N})\hat{p}_y - f\hat{p}]}{k^2(y-\mathcal{N})^2 - f(f-1)}, \\ \hat{w} &= \frac{-k^2\mu(y-\mathcal{N})\hat{p}}{k^2(y-\mathcal{N})^2 - \mathcal{N}^2}, & \hat{\rho} &= \frac{ik\mu\mathcal{N}^2\hat{p}}{k^2(y-\mathcal{N})^2 - \mathcal{N}^2}.\end{aligned}\quad (4.26a,b,c,d)$$

Note that they are different from the $\hat{\cdot}$ variables in Chapter 3. Motivated by figure 4.5, we demand \hat{p} is continuous across $y = \mathcal{N}$ but \hat{p}_y has a jump. Hence we normalize the series solution of \hat{p} near $y = \mathcal{N}$ to be

$$\hat{p} = \begin{cases} (f-1)[1 + \alpha_+(y-\mathcal{N}) + \dots] & y > \mathcal{N}, \\ (f-1)[1 + \alpha_-(y-\mathcal{N}) + \dots] & y < \mathcal{N}, \end{cases}\quad (4.27)$$

where α_+ and α_- are determined by the condition that $\hat{p} \rightarrow 0$ as $y \rightarrow \infty$ and $y \rightarrow -\infty$. In this setting, the classical critical level of the secondary instability, $y = \mathcal{N}$ is not a singularity of (4.25) and hence there is no singular terms in the series solution (4.27). The discontinuity of \hat{p}_y is induced by the local mean-flow anomaly which we will subsequently analyze. It is also noted that the steady-wave solutions have their baroclinic critical levels $y = 0$ and $y = 2\mathcal{N}$, which are their singularities. We expect that the flow here behaves similarly to the forced linear baroclinic critical layers which we studied in Chapter 3. We will defer this issue to future studies on the ‘replication of zombie vortices’; here we only use the fact that the baroclinic critical layers render a constant jump of \hat{u} , \hat{v} and \hat{p}_y as long as they remain linear so that α_+ and α_- remain constant.

In the limit of $y \rightarrow \mathcal{N}$, (4.26) indicates

$$\hat{u} \rightarrow -\frac{1}{f}\hat{p}_y, \quad \hat{v} \rightarrow ik, \quad \hat{w} \rightarrow 0, \quad \hat{\rho} \rightarrow -i\mu k(f-1),\quad (4.28a,b,c,d)$$

so \hat{u} has a jump across $y = \mathcal{N}$:

$$(\hat{u})_{y=\mathcal{N}+} - (\hat{u})_{y=\mathcal{N}-} = -\frac{f-1}{f}(\alpha_+ - \alpha_-).\quad (4.29)$$

These relations imply the properties of flow in the critical layer: we expect p , v and ρ are nearly uniform to leading order, w is small, and u will vary.

Next, we study the critical layer around $y = \mathcal{N}$. We consider the time and length scales at the onset of instability which are obtained from the normal mode analysis in the previous section:

$$T = \varepsilon^{\frac{1}{2}} t, \quad Y = \frac{(y - \mathcal{N})}{\varepsilon^{\frac{1}{2}}}, \quad (4.30)$$

so the disturbances are expressed by

$$(u', v', w', p', \rho') = \varepsilon [\check{u}(Y, T), \check{v}(Y, T), \check{w}(Y, T), \check{p}(Y, T), \check{\rho}(Y, T)] e^{ik(x + \mu z - \mathcal{N}t)} + \text{c.c.} \quad (4.31)$$

Note that T and Y in this chapter have different definitions with those in Chapter 3. The mean-flow modifications ΔU , W and G grow to order ε at $t = O(\varepsilon^{-\frac{1}{2}})$, hence we rescale them by

$$\Delta U = \varepsilon \mathcal{U}, \quad W = \varepsilon \mathcal{W}, \quad G = \varepsilon \mathcal{G} \quad (4.32)$$

For \mathcal{U} , we again have the simple expression

$$\mathcal{U} = \frac{m^2 \cos(YT) - 1}{\mathcal{N} Y^2}. \quad (4.33)$$

One may derive the leading-order of (4.5) and (4.6), which are

$$\check{\rho} = -ik\mu\check{p}, \quad \check{w} = \varepsilon^{\frac{1}{2}} \frac{\check{\rho}_T + ikY\check{\rho} + \check{v}\mathcal{G}_Y}{\mathcal{N}^2}. \quad (4.34)$$

This indicates the vertical velocity \check{w} is at a higher order compared to other disturbances, hence we may assume the flow is two-dimensional to leading order and define a stream function ψ such that

$$u' = -\psi'_y, \quad v' = \psi'_x. \quad (4.35)$$

To realize the property that to leading order, v' is a constant identical to (4.28b) while u' has a jump across the critical layer, ψ' has the form

$$\psi' = \varepsilon \left[\Phi(T) + \varepsilon^{\frac{1}{2}} \Psi(Y, T) \right] e^{ik(x + \mu z - \mathcal{N}t)} + \text{c.c.} \quad (4.36)$$

Substituting (4.30)-(4.36) into the vorticity equation (4.8), we obtain

$$\left(\frac{\partial}{\partial T} + ikY\right)\Psi_{YY} = ik\Phi\mathcal{U}_{YY}. \quad (4.37)$$

The last two terms of (4.8) related to w' and W_y are at higher orders, so (4.37) only concerns the horizontal flow. We define the vorticity $\mathcal{Z} = \Psi_{YY}$, then applying the jump condition of u' given by (4.29), we obtain the governing equations of the reduced model for $\Phi(T)$ and $\mathcal{Z}(Y, T)$:

$$\left(\frac{\partial}{\partial T} + ikY\right)\mathcal{Z} = ik\Phi\mathcal{U}_{YY}, \quad (4.38)$$

$$\int_{-\infty}^{\infty} \mathcal{Z} dY = \frac{\Phi(f-1)}{f}(\alpha_+ - \alpha_-). \quad (4.39)$$

This reduced model is very similar to that of the defect instability of Gill [25] and Balmforth et al. [26], though it involves f due to the planetary rotation and α_{\pm} due to a more complicated outer flow. But unlike their steady defect, our mean-flow modification \mathcal{U} itself evolves with T , so the secondary instability may not be in form of a normal mode.

The solution of \mathcal{Z} from (4.38) is

$$\mathcal{Z} = \mathcal{Z}_0(Y)e^{-ikYT} + \int_0^T ik\Phi(T')\mathcal{U}_{YY}(Y, T')e^{ikY(T'-T)}dT', \quad (4.40)$$

where \mathcal{Z}_0 is the initial condition of \mathcal{Z} at $T = 0$. Substituting (4.40) into (4.39) yields

$$\int_{-\infty}^{\infty} \mathcal{Z}_0(Y)e^{-ikYT} + \int_0^T ik\Phi(T')I(T', T)dT' = \frac{\Phi(f-1)}{f}k(\alpha_+ - \alpha_-), \quad (4.41)$$

where

$$I = \int_{-\infty}^{\infty} \mathcal{U}_{YY}(Y, T')e^{ik(T'-T)Y}dY = -k^2(T'-T)^2 \int_{-\infty}^{\infty} \mathcal{U}(Y, T')e^{ik(T'-T)Y}dY. \quad (4.42)$$

With the expression of \mathcal{U} given in (4.33), we may evaluate this integral analyti-

cally:

$$I = \begin{cases} 0 & 0 \leq T' \leq \frac{k}{k+1}T, \\ \frac{\pi m^2 k^2}{\mathcal{N}} (T' - T)^2 [(k+1)T' - kT] & T' > \frac{k}{k+1}T. \end{cases} \quad (4.43)$$

The results may be derived taking advantage of the Fourier-transform relation

$$\int_{-\infty}^{\infty} \frac{\sin^2 at}{t^2} e^{-i\omega t} dt = \begin{cases} \pi \left(a - \frac{|\omega|}{2} \right) & |\omega| \leq 2a, \\ 0 & |\omega| > 2a. \end{cases} \quad (4.44)$$

Substituting (4.43) into (4.41) and (4.42), we obtain an integral equation for Φ :

$$\Phi(T) = -\frac{i}{2} \sigma \int_{\frac{k}{k+1}T}^T \Phi(T') (T' - T)^2 [-kT + (k+1)T'] dT' + tilt, \quad (4.45)$$

with

$$\sigma = -\frac{2\pi m^2 k^3 f}{\mathcal{N}(f-1)(\alpha_+ - \alpha_-)}, \quad tilt = \frac{f}{k(f-1)(\alpha_+ - \alpha_-)} \int_{-\infty}^{\infty} \mathcal{L}_0 e^{-ikYT} dT. \quad (4.46a, b)$$

Here σ is a real number (α_+ and α_- have the same imaginary parts due to symmetry about $y = \mathcal{N}$ of the outer solution). *tilt* represents the effect of vorticity tilting caused by the shear. It decays in $O(T^{-2})$ for large T (cf. Brown and Stewartson [78]), so we may neglect it if the focus is on the instability. We may then differentiate (4.45) to obtain a differential equation:

$$\Phi_{TTTT} + i\sigma [(T\Phi)_T - 3k\Phi] = \frac{-i\sigma k^2}{(k+1)^3} \left[\frac{(T^2\phi)_{TT}}{2} + 2(k+1)(T\phi)_T + 3(k+1)^2\phi \right], \quad (4.47)$$

where $\phi = \Phi\left(\frac{k}{k+1}T\right)$ is Φ evaluated an earlier time, but when the solution grows exponentially, we should have $\phi \ll \Phi$ for $T \gg k+1$, so we may neglect the terms on the righthand side of (4.47) and obtain a WKB solution

$$\Phi = \frac{\Phi_0}{T^{k+\frac{1}{3}}} \exp \left\{ \frac{3}{8} \left[\sqrt{3} - i \operatorname{sgn}(\sigma) \right] |\sigma|^{\frac{1}{3}} T^{\frac{4}{3}} \right\}. \quad (4.48)$$

Here Φ_0 is a constant dependent on the wavenumber k . It is determine by the

earlier evolution before the WKB solution becomes valid, which is difficult to study analytically. In practice, we fix it by matching to the numerical solutions. (4.48) reveals the essential property of the secondary instability: due to the unsteadiness of the basic flow, the exponential growth is different from the usual normal mode instability: it grows exponentially with $T^{\frac{4}{3}}$ and there is a prefactor $T^{-\frac{1}{3}-k}$ in the slowly varying amplitude.

Next, we explore the effect of wavenumber k on the secondary instability. The mean-flow modification is forced by the linear waves with a specific wavenumber; as a result, the instability depends strongly on whether the disturbance has shorter waves ($k > 1$, which we refer to as harmonics) or longer waves ($k < 1$, which we refer to as subharmonics) compared to the forcing. The WKB solution (4.48) indicates if we wait long enough, all Fourier components will grow exponentially and thus they are all unstable. From the outer solution discussed in (4.25)-(4.27), a simple rescaling of $\xi = k(y - \mathcal{N})$ yields $\alpha_+, \alpha_- \sim k$, so $\sigma \sim k^2$ according to (4.46a), indicating much stronger growth rates for higher-order harmonics.

In addition to the exponential growth at later times, we also hope to consider the early evolution before the exponential growth is turned on. A natural setup for the reduced model (4.38)-(4.39) is to give an initial condition \mathcal{L} at $T = 0$, and then solve the initial value problem. Such a problem has been studied by Balmforth et al. [26] in detail: using Laplace transform, they have derived explicit solutions to the initial value problem for a Lorentzian defect. In addition to unstable modes, the solution also includes other disturbances representing the effect of Landau damping and possibly a transient growth (they should be in *tilt* in (4.45) in our analysis). However, these phenomena may not be the case in our model of forced baroclinic critical layers, because we have no initial disturbance at $T = 0$; secondary instabilities are generated by weakly nonlinear forcing in a manner similar to Section 3.4. While we have not yet considered details of these forcing, we hope to simply study the evolution of secondary instability and the effect of k on it. For this task, our basic assumption is that the solution to the homogeneous ODE (4.47) represents the exponentially growing part of the solution, or the ‘the unstable mode’. This is obviously true for large T , and we assume that we may trace back to small T as well, because (4.47) dictates how disturbances grow from initially small amplitudes to later large ones. We then consider the ‘unstable modes’ with a simple

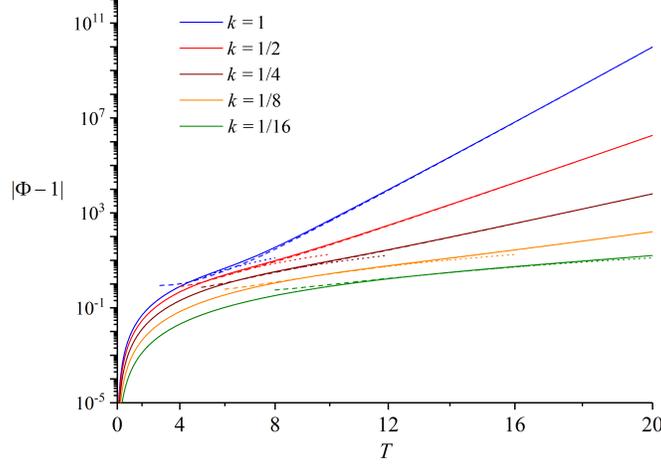


Figure 4.7: Secondary instability of subharmonics, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$ for $k = 1$. Solid lines represent the numerical solution of (4.47) with initial condition (4.49), dashed lines represent the WKB solution (4.48), and the dotted lines are the early-time solution (4.50).

initial condition

$$\Phi = 1, \quad \Phi_T = \Phi_{TT} = \Phi_{TTT} = 0, \quad T = 0, \quad (4.49)$$

which is equivalent to setting $tilt = 1$ in the integral equation (4.45). We solve the problem numerically and plot the results in solid lines in figures 4.7 and 4.8, compared to the WKB solution in dashed lines at large T . For small T , we may derive an asymptotic solution:

$$\Phi = -\frac{i}{24(k+1)^3} \sigma T^4 + 1. \quad (4.50)$$

The solutions of (4.50) are plotted in dotted lines in figures 4.7 and 4.8, which demonstrate good agreements with numerical solutions for small T . Hence $\Phi - 1$ features a T^4 power growth in the early stage. There is a transition from the power growth to the exponential growth of WKB solution at some time later.

For subharmonics, $(k + 1)$ in the denominator of (4.50) plays a minor role,

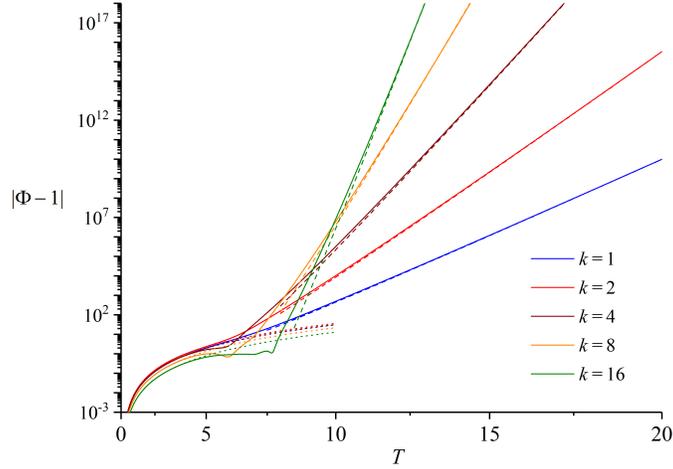


Figure 4.8: Secondary instability of harmonics, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$ for $k = 1$. Solid lines represent the numerical solution of (4.47) with initial condition (4.49), dashed lines represent the WKB solution (4.48), and the dotted lines are the early-time solution (4.50).

hence $\Phi - 1 \sim ik^2 T^4$, which means all harmonics are weaker than the fundamental mode at the early stage. When the exponential growth commences, $\sigma \sim k^2$ represents the exponential growth rate, so the subharmonics become even weaker compared to the fundamental mode. We also observe that as k becomes smaller, the transition from the power growth to the exponential growth is more and more delayed. For $k = 1/16$, (4.50) very well describes the actual solution through $T = 0 \sim 20$. In conclusion, the instability of subharmonics is weak compared to the fundamental mode, and becomes weaker as the wave becomes longer.

The behaviour of harmonics is more complicated. With $\sigma \sim k^2$, (4.50) indicates that harmonics are weaker than the fundamental mode at early times, which we may confirm from figure 4.8. After some time, transitions to exponential growth take place. As k increases, this transition is delayed, *i.e.*, the exponential growth of higher harmonics are turned on later. But this still happens much earlier than the condition $T \gg k + 1$ required by the WKB solution: for example, the harmonic $k = 16$ starts to grow exponentially at about $T = 8$, and it soon dominates over all

the other lower modes at $T \approx 10$. The divergence of exponential growth rates as wavenumbers increase is evidently demonstrated.

The explosive behaviour of harmonics is in sharp contrast to common types of hydrodynamic instabilities. For the latter, there is usually a finite band of unstable wavenumber; harmonics higher than the cutoff wavenumber are stable. The unusual behavior of the secondary instability must be attributed to the unsteadiness of the basic flow. To give an intuitive explanation, as time increases, the scale of ΔU becomes smaller and its profile becomes sharper, so we expect it will favour shorter waves and make their instabilities stronger.

An important consequence of the fast-growing harmonics is that the problem will become ill-posed. Theoretically, ill-posedness happens when the exponential growth rate goes to infinity as the wavenumber k goes to infinity, and this will make the problem lose stability to perturbations with infinitesimal length scales¹. We cannot compute $k \rightarrow \infty$, but for a very high harmonic $k = 100$, we found it starts to grow exponentially at about $T = 22$ with a huge growth rate comparable to the WKB solution. So we think at least in practice, one will encounter the problem of ill-posedness, which we actually did. In a further analysis, we made an attempt to study the nonlinear evolution of the secondary instability by adding the leading-order nonlinear term to the vorticity equation. We then need to solve a PDE in Y, T and phase $\theta = x + \mu z - \mathcal{N}t$ in a manner similar to Stewartson [1] and Warn and Warn [2]. Details of this model is shown in Appendix B. When solving the problem, we found that a grid-size instability invariably appears after some time. We believe this is because when the problem becomes ill-posed, the numerical error with the scale of grid size has extremely large growth rate, so it will quickly pollute the results.

The divergent growth of the harmonics is unphysical and it must be controlled by other effects. In our nonlinear model, we tried adding the leading-order nonlinear term corresponding to $-v'u'_{yy}$ to the vorticity equation (4.8), similar to Brown and Stewartson [72], but this did not stop the numerical instability. Since the main reason for the unusually strong exponential growth of higher harmonics is

¹Here we are referring to the stability in view of posing the problem: in the limit of the amplitude of the perturbation going to zero, its response should also go to zero. Otherwise the problem is unstable and ill-posed.

the mean-flow defect \mathcal{U} which keeps growing and sharpening, we might consider how \mathcal{U} might be weakened. Obviously, this may be accomplished by viscosity or thermal diffusion. In the presence of thermal diffusion, we have shown in Section 3.6.1 that the mean-flow defect will still grow secularly but stop shrinking in size. According to the analysis of Shen [76], such a flow may grow very fast in order $\exp(\sigma t^2)$, but the unstable waveband where $\text{Re}(\sigma) > 0$ will become finite and hence the ill-posedness will be removed. We have been neglecting viscosity based on the astrophysical contexts, but if viscosity is added, then according to (3.76) and (3.77), the mean-flow defect will saturate (though the defect will spread to outer flows in a manner to Brown and Stewartson [72]), and the instability will be expected to be similar to a normal mode instability.

Even though the ill-posedness should be eliminated finally, the explosive growth of the higher harmonics suggests the secondary instability should be very strong. Indeed, in the numerical simulations, the secondary instability of baroclinic critical layers in Marcus et al. [6] is much stronger than that of Rossby wave critical layers in Haynes [27] and Haynes [3]: in the former, the dipolar stripe rolls up and then evolves violently into much larger vortices, but in the latter, the general shape of the Kelvin cat's eye is still maintained. We have not yet studied the nonlinear evolution of secondary instability, but from our current results, we may infer a possible reason: in the Rossby wave case, the nonlinear critical layer itself can saturate, and then merely oscillates in a pattern of Kelvin cat's eye. But for the baroclinic critical layers, the mean-flow defect keeps growing. This causes ill-posedness in inviscid linear instability and may generate strong repercussions in the further nonlinear evolution.

4.4 Numerical solution to the secondary instability

To verify the asymptotic theory for the secondary instability, we solve (4.3)-(4.7) numerically. We consider the time-dependent solution for a single Fourier component:

$$(u', v', w', p', p') = [\tilde{u}(y, t), \tilde{v}(y, t), \tilde{w}(y, t), \tilde{p}(y, t), \tilde{p}(y, t)] e^{ik(x+\mu z)} + \text{c.c.} \quad (4.51)$$

Similar to (4.13), \tilde{p} can be expressed by other disturbances by

$$\begin{aligned} \tilde{p} = & -\frac{e^{-k\sqrt{1+\mu^2}y}}{2\sqrt{1+\mu^2}} \int_{-\infty}^y e^{k\sqrt{1+\mu^2}y'} \left[(f - 2U_y - 2\mu W_y) i\tilde{v} + f\sqrt{1+\mu^2}\tilde{u} - i\mu\tilde{p} \right] dy' \\ & -\frac{e^{k\sqrt{1+\mu^2}y}}{2\sqrt{1+\mu^2}} \int_y^{\infty} e^{-k\sqrt{1+\mu^2}y'} \left[(f - 2U_y - 2\mu W_y) i\tilde{v} - f\sqrt{1+\mu^2}\tilde{u} - i\mu\tilde{p} \right] dy'. \end{aligned} \quad (4.52)$$

To kick off the system, we impose an initial disturbance at $t = 0$:

$$\tilde{v} = \varepsilon^2 \exp \left[-\left(\frac{y - \mathcal{N}}{\delta_s} \right)^2 \right], \quad \tilde{u} = \frac{i}{k} \tilde{v}_y, \quad \tilde{w} = \tilde{p} = 0. \quad (4.53)$$

where $\delta_s = \varepsilon^{\frac{1}{2}}$ as obtained in Section 4.2. Since we are concerned with the exponential growth of the homogeneous solution, it is not very important how we start the system, though a disturbance localized in the critical layer is preferable, but the continuity condition (4.7) should be satisfied by the initial velocity field.

We solve the temporal evolution of (4.3)-(4.6) with (4.51) using a 4-th order Runge-Kutta method with a time step of 0.084. \tilde{p} is represented by (4.52) at each time step. For spatial discretization, we use a numerical scheme similar to that in the eigenvalue problem described in (4.14) and (4.15).

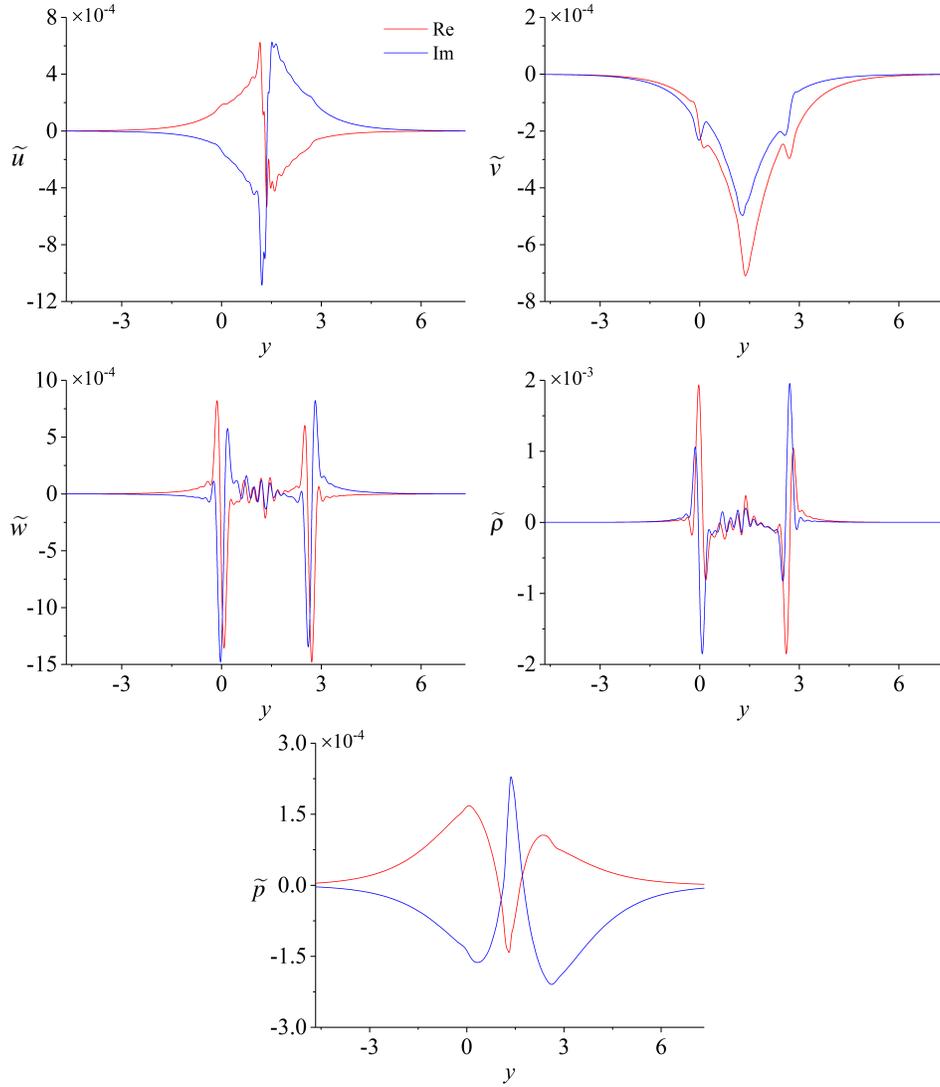


Figure 4.9: Numerical solution of the secondary instability at $t = 30$; $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$. Results are obtained by solving (4.3)-(4.6) with (4.51) numerically.

In figure 4.9, we show the numerical solution of (4.51) at $t = 30$ for the Fourier mode of the forcing. Parameters are the same as those in the normal mode analysis in Section 4.2. We see that the profiles are very similar to the normal mode instability shown in figure 4.4, hence for the bulk of the flow, the normal mode is

a good representation of the solution. This also justifies the reduced model using the steady wave solution for the outer region. The baroclinic critical layers at $y = 0$ and $y = 2\mathcal{N}$ of the numerical solution are less sharp than the normal mode. This is because in the setting of an initial value problem, the critical layer is associated with the continuous spectrum, which can smooth the solution when $c_i \rightarrow 0$ (*cf.* the linear forced critical layer (3.28)).

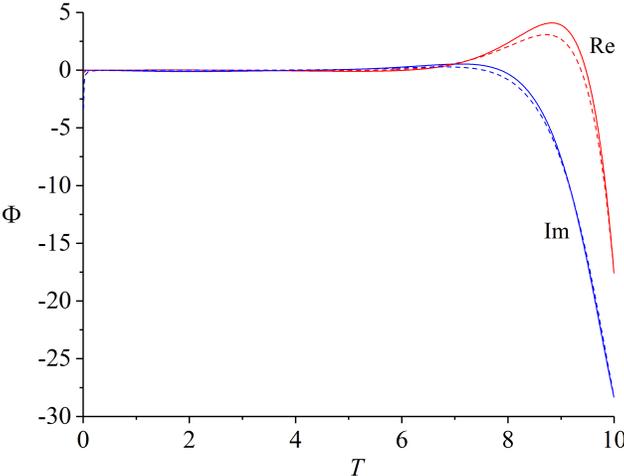


Figure 4.10: Evolution of $\Phi(T)$, $k = 1$, $m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$. Solid line is the numerical result of (4.51), and the dashed line is the WKB solution (4.48).

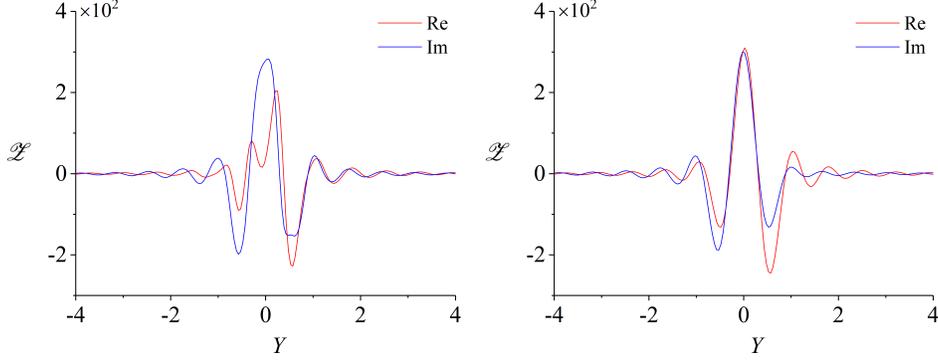


Figure 4.11: Vorticity of numerical solution (left) and the reduced model (right), $T = 9$, $k = 1$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 0.59$.

In figure 4.10 and 4.11, we show the results of Φ and \mathcal{Z} compared to the reduced model. The numerical predictions of Φ and \mathcal{Z} are computed by

$$\Phi = \frac{e^{ik\mathcal{N}t}}{ik\varepsilon} \tilde{v}|_{y=\mathcal{N}}, \quad \mathcal{Z} = -\varepsilon^{-\frac{1}{2}} e^{ik\mathcal{N}t} \tilde{u}_y. \quad (4.54)$$

For the reduced model, Φ is given by the WKB solution (4.48) with the constant Φ_0 adjusted to compare with the numerical results. We then plug the WKB solution of Φ into (4.40) to find \mathcal{Z} . We see that $\Phi(T)$ has very good agreement between the two results once the amplitude becomes large, so the WKB solution (4.48) indeed represents the exponential growth of the secondary instability. \mathcal{Z} of the two results have somewhat larger discrepancy, which we found was mainly caused by the vertical velocity and density disturbances, but the qualitative agreement is still good.

Finally, we briefly consider the numerical solutions of harmonics. In figure 4.12, we show the numerical solution of $|\Phi|$ for the harmonic $k = 2$, together with the fundamental mode $k = 1$. After an initial transient growth, the amplitude of the harmonic drops drastically, in a way much faster than the fundamental mode. This is an analogue of the Landau damping. In somewhat similar problems, *e.g.* Cheng and Knorr [79] and Balmforth et al. [26], it was shown that the Landau damping features a decay of $\log |\Phi| \sim -k^2 T^2$, which seems to well describe what we observe in figure 4.12. After about $T = 5$, the exponential growth of the harmonic emerges,

which is twice as fast as the fundamental mode as the WKB solution predicts. The results of vorticity are plotted in figures 4.13, which again demonstrate qualitative agreements between the two solutions.

If we pursue even higher harmonics, the length scale of the critical layer becomes much smaller, and our current desktop computer no longer has the capacity to resolve the full initial value problem. Further numerical investments are needed. However, we think the good agreements in figure 4.12 are solid evidence to confirm the WKB solution (4.48) and the effect of wavenumber it predicts.

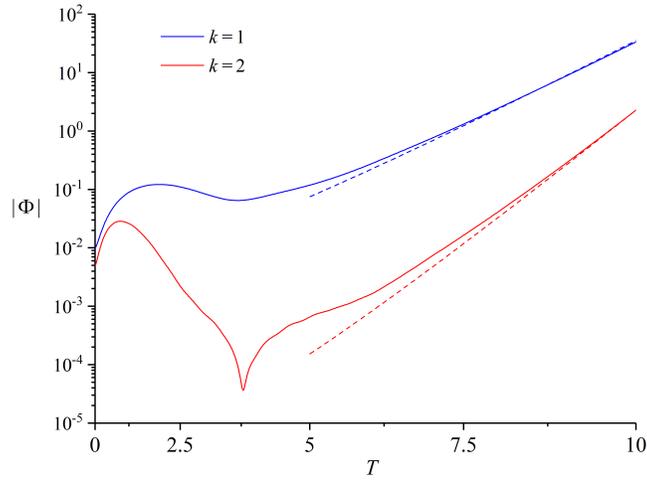


Figure 4.12: Evolution of Φ for $k = 1, 2$, $\mu = m = 1/2$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, Solid lines represent the numerical solution of (4.51)-(4.54), and dashed lines represent the WKB solution (4.48).

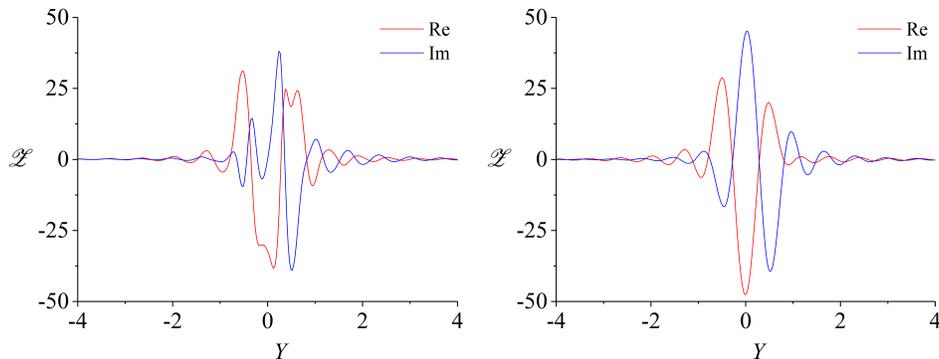


Figure 4.13: Vorticity of numerical solution (left) and the reduced model (right) for the harmonic $k = 2$, $\mu = m = 1/2$ at $T = 9.5$, $f = 4/3$, $\mathcal{N} = 4/3$, $\varepsilon = 0.01$, $\sigma = 2.36$.

4.5 Conclusion

In this chapter, we have shown that a secondary instability exists in forced baroclinic critical layers. The instability is caused by the mean-flow response of the baroclinic critical layer, and through a normal mode analysis where we parameterized time, we found that the instability features a horizontal shear instability, similar to the instability of defects in previous studies. Since the phase velocity of the secondary instability is close to the local mean-flow velocity, the baroclinic critical layer of the fundamental mode becomes the classical critical layer of the secondary instability. This shift is potentially important in the zombie vortex problem, because in this way the secondary instability can excite new baroclinic critical layers, and thus accomplish the replication process.

We adopted the matched asymptotic analysis of defect instability of Balmforth et al. [26] and derived a similar reduced model, but taking into account the unsteadiness of the mean flow. From the reduced model, we obtained a WKB solution for the secondary instability, which is confirmed by the numerical solution of the full linear equations. The WKB solution indicates that the unsteadiness of the basic flow renders an instability that grows faster than a common normal mode. This is an interesting conclusion in view that in general, the instability of unsteady flows is not well understood.

Finally, we studied the instability of various wavenumbers. We found that subharmonics are relatively weak. Harmonics remain weak in the early stage, but becomes very strong when they start to grow exponentially. At a later time, the exponential growth rate diverges as the order of harmonic increases. This behavior causes ill-posedness to the problem, and suggests that dissipative effects will be important. We finally inferred that the growing mean-flow defect, which causes ill-posedness in our current analysis, may be the reason that the secondary instability of baroclinic critical layers is much more violent than the Rossby wave critical layers as observed in previous numerical simulations.

Chapter 5

Summary and future directions

In this thesis, we have studied baroclinic critical layers in rotating stratified shear flows. Baroclinic critical levels are the locations where the Doppler shifted phase velocity matches the characteristic phase velocity of the internal gravity wave, and they are singularities of linear steady waves in inviscid fluid. The singularity is smoothed by weak dissipation, unsteadiness or nonlinearity, but there are strong wave amplitudes in their vicinity, which we refer to as the baroclinic critical layers. We have studied baroclinic critical layers in the strato-rotational instability, the forced baroclinic critical layers and the secondary instability of forced baroclinic critical layers.

5.1 Strato-rotational instability associated with baroclinic critical layers

In Chapter 2, we have studied baroclinic critical layers in strato-rotational instability (SRI). Previous studies of SRI showed that the instability is caused by the resonance between neutral modes localized near the two boundaries [45, 46]. As a result, unstable modes are characterized by strong amplitudes near the boundaries and narrow unstable wavebands where the phase velocities of the two neutral modes match. Baroclinic critical levels were outside the domain in previous studies. In our current study, we considered the range of parameters for which the baroclinic levels are located in the domain and investigate how they will influence

the instability and resonance mechanism.

We found that the baroclinic critical level destroys one of the neutral modes and thus precludes any possibility of resonance. The other mode, however, survives and becomes unstable. We therefore have a new type of strato-rotational instability induced by baroclinic critical levels. We considered the instability in relatively strong and weak stratification compared to rotation, and found different roles played by the baroclinic critical level. The new instability does not require the resonance condition, so it occupies an extensive broad band of unstable wavenumbers. We also used the conservation of pseudomomentum to rationalize the new instability: the baroclinic critical layer provides a strong localized source of pseudomomentum that balances the pseudomomentum of the basic neutral mode, and thus the zero-pseudomomentum condition of unstable modes is satisfied.

Since the strato-rotational instability induced by a baroclinic critical level is a new instability, it brings many new interesting questions. First, our current analysis is for inviscid fluid, so it is natural to ask if it can survive small viscosity in realistic fluids, especially given that viscosity may play crucial roles in the critical layer. Secondly, it will also be interesting to study the instability in other configurations, for example, in Taylor-Couette flow which is most common in experiments. A question of physical relevance is whether the new instability can be more unstable than the traditional SRI, so it may exhibit as the dominant feature. We may also explore whether it can exist in the flow of protoplanetary disks, and therefore become a possible mechanism for the accretion process. Thirdly, it will be interesting to explore the subsequent nonlinear evolution of the new instability. Will the flow evolve to a new steady state or transit into turbulence? Finally, it will be exciting if such a new instability can be observed in experiments.

5.2 Nonlinear dynamics of forced baroclinic critical layers and their secondary instability

In Chapter 3, we have studied the nonlinear dynamics of forced baroclinic critical layers. We drive the flow by a weak periodic forcing, and it establishes steady waves through the bulk of the flow. The baroclinic critical layer, however, remains unsteady, characterized by secular growth of wave amplitude and a decreasing

thickness. Such linear dynamics is similar to previous studies on classical critical layers of Rossby waves [1, 2] and internal gravity waves [16, 17]. When the critical layer becomes nonlinear, the streamwise mean-flow modification drives the nonlinearity but harmonics remain small. The vorticity field features a pattern of a dipolar stripe, representing a jet-like defect in the streamwise mean-flow velocity. Such behaviors divorce from classical critical-layer theories where all harmonics are excited, and this is because all harmonics have different baroclinic critical levels. The secular growth of the linear disturbance is arrested by the nonlinearity, but at later times, an exponential focussing takes place and finally makes the model break down. If we add small thermal diffusion to the system, then the focussing is again arrested. A drifting coherent structure emerges instead, leaving a mean-flow defect that gradually deepens and widens.

Motivated by the strong mean-flow vorticity gradient in the baroclinic critical layer, we further considered its secondary instability in Chapter 4. Through a normal mode analysis where we parameterized time in the unsteady basic flow, we found the essential properties of the secondary instability. Similar to the defect instabilities in previous studies [25, 26], the secondary instability is a two-dimensional horizontal shear instability. The phase velocity of the secondary instability is close to the local basic flow velocity, so interestingly, the baroclinic critical layer of the original forced waves becomes the classical critical layer of the secondary instability. This also makes the secondary instability have new baroclinic critical levels. We further did a matched asymptotic analysis to explore the actual temporal evolution of the secondary instability upon the unsteady mean-flow defect. We found it can be described by a WKB solution, which demonstrates a growth that is faster than a normal mode. Finally, we studied the instability of various wavenumbers. We found that short-wave harmonics have very strong exponential growth, which can render the linear problem ill-posed.

We believe that the forced baroclinic critical layers we study represent the fundamental dynamics of the zombie vortices of Marcus et al. [6]. In the linear and weakly nonlinear stage, we have obtained a pattern of vorticity field similar to their numerical simulation, the most distinctive one being the dipolar stripe. The secondary instability may be the reason for the rollup of the dipolar stripe into new vortices, and the ill-posedness may explain why the rollup is so violent. More-

over, the new baroclinic critical levels of the secondary instability may represent the replication of zombie vortices. Our analysis may therefore provide a theoretical foundation for the zombie vortex instability.

In the future work, we should consider the effects of dissipation, *i.e.*, viscosity and thermal diffusion on the secondary instability. Dissipation will effectively remove the ill-posedness we found in non-dissipative flows. It is also of great importance regarding the threshold of dissipation that suppresses the secondary instability completely. Further more, we may build a reduced model for the nonlinear evolution of the secondary instability, similar to the one of Balmforth and Korycansky [74]. Hopefully, we might realize the roll-up of the critical layer as in the full numerical simulation of Marcus et al. [6]. If this can be accomplished, we may gain significant advance in the nonlinear baroclinic critical layers and zombie vortices.

Bibliography

- [1] K. Stewartson. The evolution of the critical layer of a Rossby wave. *Geophys. Astrophys. Fluid Dyn.*, 9:185–200, 1978. → pages x, 3, 4, 13, 39, 40, 42, 44, 70, 73, 74, 75, 95, 106
- [2] T. Warn and H. Warn. The evolution of a nonlinear critical level. *Stud. Appl. Math*, 59:37–71, 1978. → pages x, 3, 4, 13, 39, 40, 70, 73, 74, 75, 95, 106
- [3] P. H. Haynes. The effect of barotropic instability on the nonlinear evolution of a Rossby-wave critical layer. *J. Fluid Mech.*, 207:231–266, 1989. → pages x, 4, 71, 73, 74, 96
- [4] N. Boulanger, P. Meunier, and S. Le Dizès. Structure of a stratified tilted vortex. *J. Fluid Mech.*, 583:443–458, 2007. → pages x, 6, 40, 65
- [5] M. Le Bars and P. Le Gal. Experimental analysis of the stratorotational instability in a cylindrical Couette flow. *Phys. Rev. Lett.*, 99:064502, 2007. → pages x, 9, 10, 14
- [6] P. S. Marcus, S. Pei, C-H. Jiang, and Hassanzadeh P. Three-dimensional vortices generated by self-replication in stably stratified rotating shear flows. *Phys. Rev. Lett.*, 111:084501, 2013. → pages x, 5, 10, 11, 12, 13, 42, 72, 73, 81, 96, 106, 107
- [7] S. A. Maslowe. Critical layers in shear flows. *Ann. Rev. Fluid Mech.*, 18: 405–432, 1986. → page 1
- [8] Lord Rayleigh. On the stability, or instability, of certain fluid motions. *Proc. London Math. Soc.*, 11:57–70, 1880. → page 2
- [9] J. W. Miles. On the stability of heterogeneous shear flows. *J. Fluid Mech.*, 10:496–508, 1961. → pages 2, 7
- [10] W. Tollmien. The production of turbulence. *Tech. Memor. Nat. Adv. Comm. Aero., Wash.*, 609, 1931. → pages 2, 3

- [11] C. C. Lin. On the stability of two-dimensional parallel flows: Part I. General theory. *Quart. Appl. Math.*, 3:117–142, 1945. → pages 2, 3
- [12] P. G. Drazin and W. H. Reid. *Hydrodynamic stability*. Cambridge University Press, 1982. ISBN 9781316604922. → pages 2, 74
- [13] F. P. Bretherton. The propagation of groups of internal gravity waves in a shear flow. *Quart. J. Roy. Met. Soc.*, 92:466–480, 1966. → pages 2, 5, 40
- [14] D. J. Benney and R. F. Bergeron. A new class of nonlinear waves in parallel flows. *Stud. Appl. Math*, 48:181–204, 1969. → page 3
- [15] R. Harberman. Critical layers in parallel flows. *Stud. Appl. Math*, 51: 139–161, 1972. → page 3
- [16] J. R. Booker and F. P. Bretherton. The critical layer for internal gravity waves in a shear flow. *J. Fluid Mech.*, 27:513–539, 1967. → pages 3, 13, 39, 40, 42, 44, 70, 106
- [17] S. N. Brown and K. Stewartson. On the nonlinear reflexion of a gravity wave at a critical level. Part 1. *J. Fluid Mech.*, 100:577–595, 1980. → pages 3, 13, 40, 70, 106
- [18] S. N. Brown and K. Stewartson. On the nonlinear reflection of a gravity wave at a critical level. Part 2. *J. Fluid Mech.*, 115:217–230, 1982. → pages 3, 40
- [19] S. N. Brown and K. Stewartson. On the nonlinear reflection of a gravity wave at a critical level. Part 3. *J. Fluid Mech.*, 115:231–250, 1982. → pages 3, 39, 40
- [20] F. J. Hickernell. Time-dependent critical layers in shear flows on the beta-plane. *J. Fluid Mech.*, 142:431–449, 1984. → page 3
- [21] M. E. Goldstein and S. J. Leib. Nonlinear evolution of oblique waves on compressible shear layers. *J. Fluid Mech.*, 207:73–96, 1989. → page 3
- [22] S. M. Churilov and I. G. Shukhman. Nonlinear spatial evolution of helical disturbances to an axial jet. *J. Fluid Mech.*, 281:371–402, 1994. → page 3
- [23] X. Wu. Nonlinear theories for shear flow instabilities: physical insights and practical implications. *Ann. Rev. Fluid Mech.*, 51:451–485, 2019. → page 3

- [24] P. D. Killworth and M. E. McIntyre. Do Rossby-wave critical layers absorb, reflect, or over-reflect? *J. Fluid Mech.*, 161:449–492, 1985. → pages 4, 70, 71, 73, 74, 75, 76
- [25] A. E. Gill. A mechanism for instability of plane Couette flow and of Poiseuille flow in a pipe. *J. Fluid Mech.*, 21:503–511, 1965. → pages 4, 5, 74, 76, 77, 86, 87, 90, 106
- [26] N. J. Balmforth, D. del Castillo-Negrete, and W. R. Young. Dynamics of vorticity defects in shear. *J. Fluid Mech.*, 33:197–230, 1997. → pages 4, 5, 74, 76, 77, 81, 86, 87, 90, 92, 100, 102, 106
- [27] P. H. Haynes. Nonlinear instability of a Rossby-wave critical layer. *J. Fluid Mech.*, 161:493–551, 1985. → pages 4, 73, 96
- [28] S. A. Thorpe. An experimental study of critical layers. *J. Fluid Mech.*, 103:321–344, 1981. → page 4
- [29] C. G. Koop. A preliminary investigation of the interaction of internal gravity waves with a steady shearing motion. *J. Fluid Mech.*, 113:347–386, 1981. → page 4
- [30] C. G. Koop and B. GcGee. Measurements of internal gravity waves in a continuously stratified shear flow. *J. Fluid Mech.*, 172:453–480, 1986. → page 4
- [31] C. L. Lin, J. H. Ferziger, J. R. Koseff, and S. G. Monismith. Simulation and stability of two-dimensional internal gravity waves in a stratified shear flow. *Dynam. Atmos. Oceans*, 19:1–4, 1993. → page 4
- [32] O. M. Umurhan, K. Shariff, and J. N. Cuzzi. Critical layers and protoplanetary disk turbulence. *Astrophys. J.*, 830:95, 2016. → pages 5, 13, 74, 77, 84
- [33] D. J. Olbers. The propagation of internal waves in a geostrophic current. *J. Phys. Oceanogr.*, 11:1224–1233, 1981. → pages 5, 40
- [34] A. Ya. Basovich and L. Sh. Tsimring. Internal waves in a horizontally inhomogeneous flow. *J. Fluid Mech.*, 142:233–249, 1984. → pages 5, 40
- [35] S. I. Badulin, V. I. Shrira, and L. Sh. Tsimring. The trapping and vertical focusing of internal waves in a pycnocline due to the horizontal inhomogeneities of density and currents. *J. Fluid Mech.*, 158:199–218, 1985. → pages 5, 40

- [36] C. Staquet and G. Huerre. On transport across a barotropic shear flow by breaking inertia-gravity waves. *Phys. Fluids*, 14:1993–2006, 2002. → pages 5, 40
- [37] N. R. Edwards and C. Staquet. Focusing of an inertia-gravity wave packet by a baroclinic shear flow. *Dyn. Atmos. Oceans*, 40:91–113, 2005. → pages 5, 40
- [38] N. Boulanger, P. Meunier, and S. Le Dizès. Tilt-induced instability of a stratified vortex. *J. Fluid Mech.*, 596:1–20, 2008. → pages 6, 75, 76
- [39] L. N. Howard. Note on a paper of John W. Miles. *J. Fluid Mech.*, 10: 509–512, 1961. → page 7
- [40] J. P. Williams and L. A. Cieza. Protoplanetary disks and their evolution. *Annu. Rev. Astron. Astrophys.*, 49:67–117, 2011. → page 7
- [41] W. Lyra and O. M. Umurhan. The initial conditions for planet formation: turbulence driven by hydrodynamical instabilities in disks around young stars. *PASP*, 131(1001), 2019. → page 8
- [42] S. A. Balbus and J. F. Hawley. A powerful local shear instability in weakly magnetized disks. I. linear analysis. *Astrophys. J.*, 376:214, 1991. → page 8
- [43] C. F. Gammie. Layered accretion in T tauri disks. *Astrophys. J.*, 457:355, 1996. → page 8
- [44] M. J. Molemaker, J. C. McWilliams, and I. Yavneh. Instability and equilibration of centrifugally stable stratified Taylor-Couette flow. *Phys. Rev. Lett.*, 86:5270–5273, 2001. → pages 9, 14
- [45] I. Yavneh, J. C. McWilliams, and M. J. Molemaker. Non-axisymmetric instability of centrifugally stable stratified Taylor-Couette flow. *J. Fluid Mech.*, 448:1–21, 2001. → pages 9, 42, 104
- [46] B. Dubrulle, L. Marié, Ch. Normand, D. Richard, F. Hersant, and J.-P. Zahn. A hydrodynamic shear instability in stratified disks. *Astronomy & Astrophysics*, 429:1–13, 2005. → pages 9, 15, 104
- [47] D. Shalybkov and G. Rüdiger. Stability of density-stratified viscous Taylor-Couette flows. *Astronomy & Astrophysics*, 438:411–417, 2005. → page 9

- [48] J. Park and P. Billant. The stably stratified Taylor–Couette flow is always unstable except for solid-body rotation. *J. Fluid Mech.*, 725:262–280, 2013. → pages 28, 37
- [49] C. Leclercq, F. Nguyen, and R. R. Kerswell. Connections between centrifugal, stratorotational, and radiative instabilities in viscous Taylor–Couette flow. *Phys. Rev. E*, 94(4):043103, 2016. → page 14
- [50] J. Park, P. Billant, and J-J. Baik. Instabilities and transient growth of the stratified Taylor–Couette flow in a Rayleigh-unstable regime. *J. Fluid Mech.*, 822:80–108, 2017. → page 9
- [51] R. Ibanez, H. L. Swinney, and B. Rodenborn. Observations of the stratorotational instability in rotating concentric cylinders. *Phys. Rev. Fluids*, 1:053601, 2016. → pages 9, 14
- [52] J. Park, P. Billant, J-J. Baik, and J. M. Seo. Competition between the centrifugal and strato-rotational instabilities in the stratified Taylor–Couette flow. *J. Fluid Mech.*, 840:5–24, 2018. → pages 9, 10
- [53] J. Vanneste and I. Yavneh. Unbalanced instabilities of rapidly rotating stratified shear flows. *J. Fluid. Mech.*, 584:573–596, 2007. → pages 10, 15, 17, 18, 24, 34, 48
- [54] G. Facchini, B. Favier, P. Le Gal, M. Wang, and M. Le Bars. The linear instability of the stratified plane Couette flow. *J. Fluid Mech.*, 853:205–234, 2018. → page 10
- [55] S. Le Dizès and P. Billant. Radiative instability in stratified vortices. *Phys. Fluids*, 21:096602, 2009. → pages 10, 14, 37
- [56] P. Billant and S. Le Dizès. Waves on a columnar vortex in a strongly stratified fluid. *Phys. Fluids*, 21:106602, 2009. → pages 11, 14
- [57] S. Le Dizès and X. Riedinger. The strato-rotational instability of Taylor–Couette and keplerian flows. *J. Fluid Mech.*, 660:147–161, 2010. → pages 11, 14
- [58] P. S. Marcus, S. Pei, C-H. Jiang, and Barranco J. A. Zombie vortex instability. I. a purely hydrodynamic instability to resurrect the dead zones of protoplanetary disks. *Astrophys. J.*, 808:87, 2015. → page 12
- [59] P. S. Marcus, S. Pei, C-H. Jiang, and Barranco J. A. Zombie vortex instability. II. thresholds to trigger instability and the properties of zombie

- turbulence in the dead zones of protoplanetary disks. *Astrophys. J.*, 883:2, 2016. → pages 12, 71
- [60] J. A. Barranco, S. Pei, and P. S. Marcus. Zombie vortex instability. III. persistence with nonuniform stratification and radiative damping. *Astrophysical J.*, 869(2):127, 2018. → page 12
- [61] C. Wang and N. J. Balmforth. Strato-rotational instability without resonance. *J. Fluid Mech.*, 846:815–833, 2018. → pages 14, 42, 43, 48
- [62] O. Bühler. *Waves and Mean Flows*. Cambridge University Press, 2 edition, 2014. ISBN 9781107669666. → pages 17, 52
- [63] C. M. Bender and S. A. Orszag. *Advanced mathematical methods for scientists and engineers I: Asymptotic methods and perturbation theory*. Springer Science & Business Media, 2013. → page 25
- [64] N. J. Balmforth. Shear instability in shallow water. *J. Fluid Mech.*, 387: 97–127, 1999. → page 37
- [65] A. Kubokawa. Instability of a geostrophic front and its energetics. *Geophys. Astrophys. Fluid Dyn.*, 33:223–257, 1984. → page 37
- [66] C. Wang and N. J. Balmforth. Nonlinear dynamics of forced baroclinic critical layers. *J. Fluid Mech.*, 883, 2020. → page 39
- [67] T. Warn and H. Warn. On the development of a Rossby wave critical level. *J. Atmos. Sci.*, 33:2021–2024, 1976. → pages 39, 40, 44, 70
- [68] K. A. Emanuel. *Atmospheric convection*. Oxford University Press (New York), 1994. ISBN 0195066308. → page 42
- [69] M. Béland. Numerical study of the nonlinear Rossby wave critical level development in a barotropic zonal flow. *J. Atmos. Sci.*, 33:2066–2078, 1976. → page 43
- [70] R. Haberman and E. K. Ho. Higher-order change of the Hamiltonian between saddle approaches by phase plane matching. *Int. J. Nonlin. Mech.*, 31(3):239–254, 1996. → page 63
- [71] D. C. Diminnie and R. Haberman. Slow passage through the nonhyperbolic homoclinic orbit of the saddle-center Hamiltonian bifurcation. *Stud. Appl. Math.*, 108(1):65–75, 2002. → page 63

- [72] S. N. Brown and K. Stewartson. The evolution of the critical layer of a Rossby wave. Part II. *Geophys. Astrophys. Fluid Dyn.*, 10:1–24, 1978. → pages 69, 95, 96, 117
- [73] M. Wang. *Baroclinic Critical Layers and Zombie Vortex Instability in Stratified Rotational Shear Flow*. PhD thesis, University of California, Berkeley, 2016. → pages 71, 81
- [74] N. J. Balmforth and D. G. Korycansky. Non-linear dynamics of the corotation torque. *Mon. Not. Roy. Astron. Soc.*, 326:833–851, 2001. → pages 71, 107, 117
- [75] T. B. Benjamin and F. Ursell. The stability of the plane free surface of a liquid in vertical periodic motion. *Proc. Roy. Soc. Lond. A*, 225:505–515, 1954. → page 74
- [76] S. F. Shen. Some considerations of laminar stability of incompressible time-dependent basic flows. *J. Aerospace Sci.*, 28:397–404, 1961. → pages 74, 96
- [77] T. B. Benjamin. Stability of solitary waves. *Proc. Roy. Soc. Lond. A*, 328:153–183, 1972. → page 74
- [78] S. N. Brown and K. Stewartson. On the algebraic decay of disturbances in a stratified shear flow. *J. Fluid Mech.*, 100:811–816, 1980. → page 91
- [79] C. Z. Cheng and G. Knorr. The integration of the Vlasov equation in configuration space. *J. Comput. Phys.*, 22(3):330–351, 1976. → pages 100, 117

Appendix A

The critical-layer vorticity distribution

The reconstruction of the critical-layer vorticity from the matched asymptotics is:

$$\zeta = \zeta_0 + \left[\varepsilon^{\frac{1}{3}} \zeta_1 e^{ix+imz} + \varepsilon^{\frac{2}{3}} \zeta_2 e^{2ix+2imz} + \text{c.c.} \right], \quad (\text{A.1})$$

where ζ_0 is given by (3.68),

$$\zeta_1 = i \left(\frac{m^2}{2\mathcal{N}} \right)^{\frac{2}{3}} \left[(f-1)\gamma + \gamma \frac{\partial |\gamma|^2}{\partial \eta} + \frac{2}{m^2} v_1 \frac{\partial^2 |\gamma|^2}{\partial \eta^2} \right], \quad (\text{A.2})$$

$$\zeta_2 = (f-1) \left(\frac{m^2}{2\mathcal{N}^4} \right)^{\frac{1}{3}} \left(\frac{1}{2} m^2 \gamma^2 - v_1 \gamma \eta \right) + \frac{i}{(2m\mathcal{N})^{\frac{2}{3}}} \frac{\partial}{\partial \eta} \left[v_1 \zeta_1 + \left(\frac{m^4 \mathcal{N}}{4} \right)^{\frac{1}{3}} u_1 \gamma \right], \quad (\text{A.3})$$

and the leading-order fundamental components of the critical-layer horizontal velocity are

$$v_1 = \frac{im^2 A}{2} \left[\log \left| \left(\frac{\varepsilon^2 m^2}{2\mathcal{N}} \right)^{\frac{1}{3}} \eta \right| + 1 \right] - \frac{m^2}{2} \mathcal{P} \int_{\eta}^{\infty} \left(\gamma - \frac{iA}{\eta'} \right) d\eta' + \frac{iA(\alpha\mathcal{N} - f)}{\mathcal{N}^2 - f(f-1)}, \quad (\text{A.4})$$

$$u_1 = \frac{(f-1)v_1 - iA + v_1 \zeta_0}{i\mathcal{N}}, \quad (\text{A.5})$$

where the \mathcal{P} on the integral sign implies principal value.

Appendix B

A nonlinear model for the secondary instability

In view that the exponential growth of the secondary instability will make it nonlinear, we build a nonlinear model in this appendix. Due to the issue of numerical instability, the model did not reveal significant insights to the nonlinear evolution. But we document it here as it provides important information on the ill-posedness and provide the foundation for future work.

For nonlinear evolution, we assume the stream function has the form:

$$\psi = \varepsilon[\tilde{\Phi}(\theta, T) + \varepsilon^{\frac{1}{2}}\tilde{\Psi}(\theta, Y, T)], \quad \theta = x + \mu z - \mathcal{N}t. \quad (\text{B.1})$$

Now $\tilde{\Phi}$ and $\tilde{\Psi}$ include all the Fourier components, *e.g.* for $\tilde{\Phi}$

$$\tilde{\Phi} = \sum_{n=1}^{\infty} \Phi_n e^{ink\theta} + \text{c.c.}, \quad (\text{B.2})$$

where the n -th harmonic Φ_n corresponds to Φ in Chapter 4 with horizontal wavenumber k replaced by nk . In Section 4.3, we have shown that the effects of vertical velocity and density do not contribute to the leading order of the reduced model, and one can confirm that this is the case for nonlinear evolution as well. An elementary analysis tells us that the leading-order nonlinearity of (4.8) is the term $-v'u'_{yy}$ which should be added to the left hand side, and this gives us the nonlinear

vorticity equation:

$$\frac{\partial \widetilde{\mathcal{Z}}}{\partial T} + Y \frac{\partial \widetilde{\mathcal{Z}}}{\partial \theta} + \frac{\partial \widetilde{\Phi}}{\partial \theta} \frac{\partial \widetilde{\mathcal{Z}}}{\partial Y} - \frac{\partial \widetilde{\Phi}}{\partial \theta} \frac{\partial^2 \mathcal{U}}{\partial Y^2} = 0 \quad (\text{B.3})$$

where $\widetilde{\mathcal{Z}} = \widetilde{\Psi}_{YY}$. We then match each Fourier component of the ensemble integral of $\widetilde{\mathcal{Z}}$ to the jump of streamwise velocity of the outer flow, in the same manner as in Chapter 4, and this determines Φ_n :

$$\Phi_n = \frac{kf}{2\pi(f-1)(\alpha_{n+} - \alpha_{n-})} \int_0^{\frac{2\pi}{k}} \int_{-\infty}^{\infty} \widetilde{\mathcal{Z}} e^{-ikn\theta} dY d\theta, \quad (\text{B.4})$$

where α_{n+} and α_{n-} correspond to α_+ and α_- in (4.27) with horizontal wavenumber nk .

(B.2)-(B.4) constitute a reduced model for the nonlinear evolution of the secondary instability. They are similar to previous studies on the nonlinear evolution of classical critical layers, for example Brown and Stewartson [72] and Balmforth and Korycansky [74], and we solved the problem using the numerical method of Cheng and Knorr [79].

Our initial intention was that maybe the nonlinear term $\widetilde{\Phi}_\theta \widetilde{\mathcal{Z}}_Y$ could stop the exponential growth and cure the ill-posedness of the problem. However, this did not happen. A grid-size numerical instability invariably appears before we could study the nonlinear evolution. As we discussed in Section 4.3, to stop the ill-posedness, we may need to add diffusion or viscosity to the baroclinic critical layer as we did in Section 3.6. The ill-posedness suggests the secondary instability is a very strong one, and may partially rationalize why the baroclinic critical layers evolve much more violently than the Rossby wave critical layers.