Equatorial Symmetric Stability

by

Mark David Fruman

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

Equatorial Symmetric Stability

Doctor of Philosophy, 2005

Mark David Fruman, Department of Physics, University of Toronto

Abstract

Equatorial symmetric stability in the Earth's middle atmosphere, modelled by the inviscid adiabatic compressible Euler equations on a β-plane, is investigated, taking into account the previously neglected Coriolis force terms due to the component of the planetary rotation vector tangent to the surface. Using an energy-Casimir method based on the underlying Hamiltonian structure of the governing equations, two conditions that, taken together, are sufficient for linear stability are derived: the well known condition that potential vorticity be positive in the northern hemisphere and negative in the southern; and the condition that entropy must increase (decrease) in the direction of the local planetary rotation vector in the northern (southern) hemisphere. Far from the equator, the latter reduces to the familiar condition for static stability. Explicit steady solutions are found that are stable when the horizontal Coriolis force terms are neglected but unstable when they are included.

The same problem is considered using an anelastic equations model, and conditions for stability under finite amplitude perturbations are derived. It is argued that only steady solutions that are even functions of latitude can be stable in the sense of Lyapunov. States that are demonstrably Lyapunov stable are used to estimate the growth of disturbances to unstable equilibria. The short time evolution of the system away from an unstable state with linear meridional shear in the zonal velocity is explicitly calculated, and the normal mode solution exhibits features commonly associated with symmetric instability.

Finally, the Rayleigh criterion for linear stability of inviscid Taylor-Couette flow between rotating cylinders, that the magnitude of angular momentum increase with distance from the axis of rotation, is shown to be valid for finite amplitude disturbances provided the radial gradient of angular momentum is not too high.

Acknowledgements

I thank my supervisor Professor Theodore G. Shepherd for the idea for this work as well as for his guidance during its process. Thanks are also due to the faculty, staff and students in the Department of Physics at the University of Toronto for their integrity and hospitality. I am grateful in particular to my friends and colleagues Simal Saujani, Sorin Codoban, David Sankey and Kirill Semeniuk for their always ready help with matters scientific, technical, and other.

Above all, I thank my parents, my brother, and my grandmother for their constant love and support.

Financial support for this work was provided by my supervisor and scholarships from the National Science and Engineering Research Council of Canada, the Government of Ontario, and the Walter C. Sumner Foundation.

Contents

List of Figures

Chapter 1

Introduction

1.1 Symmetric stability

This study concerns the stability of a steadily rotating stratified three-dimensional fluid flow that is symmetric about the axis of rotation to disturbances that preserve the symmetry. The problem is known as "symmetric stability". Our main result is the demonstration of sufficient conditions on the velocity and thermodynamic fields for the symmetric stability of steady zonal flows in the Earth's equatorial atmosphere.

There is a good analogy between symmetric instability in an axially symmetric rotating unstratified fluid and adiabatic convection in a nonrotating fluid in terms of materially conserved variables and forces on displaced fluid parcels.

Fluid parcels undergoing adiabatic motion conserve *entropy*. An infinitesimal parcel of fluid displaced upwards in such a way that it conserves its entropy and does not disturb the local pressure field feels a downward buoyancy force proportional to its density, which is a function of its entropy and the ambient pressure, and an upward force due to the vertical pressure gradient, which balances the buoyancy force on the fluid *surrounding* it. The displaced parcel will therefore feel a net force towards its equilibrium position if its density at its new position is greater than the density of the surrounding fluid or, since density varies inversely with entropy, if its entropy is *less* than that of the surrounding fluid at its new position. Hence, a fluid at rest is convectively stable *if its entropy increases with height*.

Angular momentum plays the role in symmetric instability that entropy does in convection. Parcels of fluid in an axially symmetric flow, which due to the symmetry actually represent circular rings of fluid, conserve the component of angular momentum in the direction parallel to the axis of symmetry. A parcel displaced radially outward such that it conserves its angular momentum (to do which it slows down) and does not disturb the local pressure field feels an inward force due to the radial pressure gradient at its new position, which is such that the surrounding fluid is maintained in circular motion. If the angular momentum of the displaced parcel is greater than that of the surrounding fluid, then it will have a greater tangential velocity, the pressure gradient force will be too weak to keep it in circular motion, and it will drift further away from its equilibrium position. Alternatively, if its angular momentum is less than that of the surrounding fluid, the pressure gradient force will be greater than that needed to keep it in circular motion, and it will be pushed back towards its undisturbed position, leading to "inertial" oscillation about the equilibrium. The condition for symmetric stability is therefore that *the angular momentum in the unperturbed flow increases with distance from the axis of rotation*. This result is due to Lord Rayleigh (1916).

For an unstratified fluid, symmetric instability is more commonly referred to as *centrifugal* or *inertial* instability. For a stratified and rotating fluid, the term *symmetric stability* usually refers to the more general stability problem, considering both vertical and radial displacements and both buoyancy and inertial forces.

The two types of parcel instabilities have been studied extensively in the famous *Rayleigh-Bénard convection* and *Taylor-Couette* experiments. The former involves a layer of fluid confined between two horizontal plates being made unstable by heating the lower plate. When the temperature gradient exceeds a critical value, depending on the thermal diffusivity coefficient, the viscosity, and the thickness of the fluid layer, instability sets in, and cells of overturning motion develop. There is a rich literature on the scale, organization, and patterns of convection. The Taylor-Couette experiment involves an annular column of fluid confined between two rotating cylinders being made unstable by increasing the rotation rate of the inner cylinder. Instability initially takes the form of vertical rolls superposed on the initial azimuthal flow, known as *Taylor vortices*. The experiment is well known for the controllable bifurcations of the vortex pattern with progressive increase of the rotation rate of the inner cylinder, eventually leading to *chaos*. For discussions of the two experiments, see, for example, Chandrasekhar (1961) and Koschmieder (1993). Chapter 5 of this thesis deals with the stability of flow in the Taylor-Couette experiment.

1.2 Geophysical setting

In the case of a stratified and rotating fluid, stability depends on the gradients of both entropy and angular momentum. Chapters 3 and 4 concern the stability of steady rotating stratified flows in the equatorial middle atmosphere of the Earth (the region 15 – 100 km above the surface). It will be shown that there are two conditions that, taken together, are sufficient for stability: the entropy must increase in the direction of the local planetary rotation vector in the northern hemisphere and decrease in the southern; and the angular momentum must increase towards the equator on surfaces of constant entropy. The second condition is usually phrased in terms of *potential vorticity* (PV), a quantity that is related to the gradients of both angular momentum and entropy. For stability, PV must have the sign of latitude: positive in the northern hemisphere, negative in the southern. The equator features in the stability conditions because the orientation of the gravitational force relative to the centrifugal force changes at the equator.

1.2.1 Observational evidence of inertial instability

The importance of convection to atmospheric dynamics is well known. Heating of the surface by the sun warms the lowest layer of air in the troposphere. The resulting convection carries water vapour upwards, leading to the formation of clouds and the phenomena of thunderstorms and precipitation. Convection also plays a less obvious role as a mechanism for forcing waves in the upper troposphere and lower stratosphere. Such waves affect the large scale circulation in the middle atmosphere, driving it away from what would be expected on the grounds of radiative and dynamical balance.

The importance of inertial instability in atmospheric dynamics is much less well understood. One reason for this is that real atmospheric circulations are never exactly axially symmetric, so conditions are never such that the predictions of symmetric stability can be directly observed. However, the fact that the instability mechanism is inherently local suggests that (globally) axially symmetric results should be approximately accurate and at least qualitatively relevant when there is partial axial symmetry over an interval of longitude. Another reason for the difficulty of identifying inertial instability is that its effect on the velocity and temperature fields in the atmosphere is obscured by waves, which are always propagating through the atmosphere in all directions.

It is generally accepted now that inertial instability is directly observed occasionally by satellites measuring the temperature distribution in the middle atmosphere. The patterns that are identified as due to inertial instability are based on the important study by Dunkerton (1981), in which he addressed the simplest interesting case of symmetric instability, that of uniform stratification and constant meridional shear in the relative zonal (east-west) velocity. It is easily seen that this configuration is symmetrically unstable in an interval between the equator and a latitude proportional to the value of the velocity shear. Dunkerton solved the linearized primitive equations on an equatorial

β-plane[1], and found that the fastest growing mode exhibits a Taylor vortex-like column of vertically (radially) stacked cells of overturning motion centred over the unstable region. The corresponding signature in the zonal velocity is a column of oppositely signed anomalies on the equatorward side of the unstable region, and in the temperature field a checkerboard pattern of warm and cold anomalies on either side of the unstable region. See Figure 1.1 for a schematic picture of the features of Dunkerton's solution.

Hitchman et al. (1987) identified the characteristic inertial instability pattern in northern hemisphere winter temperature data from satellite observations of the lower mesosphere (approximately 50 km altitude or the level at which the pressure is about 1 hPa). Oppositely signed columns of temperature anomalies with amplitudes of about 5 K and vertical wavelengths of about 15 km were observed over the equator, persisting for periods of one to two weeks. These so-called *pancake structures* are distinguished from Rossby and Kelvin waves which are also observed in the same data by their large amplitude, stationarity, and confinement to low latitudes. Similar features were identified by Hayashi et al. (1998), with further evidence that they represent inertial instability provided by corresponding out of phase patterns in winter midlatitude, consistent with the Dunkerton solution (see Figure 1.2). Also, a similar signal was identified in southern hemisphere winter data. In both studies, the pancake structure events were seen to be preceded by breaking Rossby waves at midlatitudes which pull negative (positive) PV air into the northern (southern) hemisphere over a wide interval of longitude (about 60°), thus violating the symmetric stability condition.

Numerical simulations of the middle atmosphere show inertial instability activity. See Hunt (1981) (Figure 1.3), O'Sullivan and Hitchman (1992), and Semeniuk and Shepherd (2001).

[1] The (hydrostatic) *primitive equations* and the *equatorial β-plane* are defined precisely in Chapter 2. It will suffice here to say that they are widely used approximations that simplify the mathematics without losing most of the essential physics of the problem. We will not make the hydrostatic approximation for reasons that are discussed in Chapter 2.

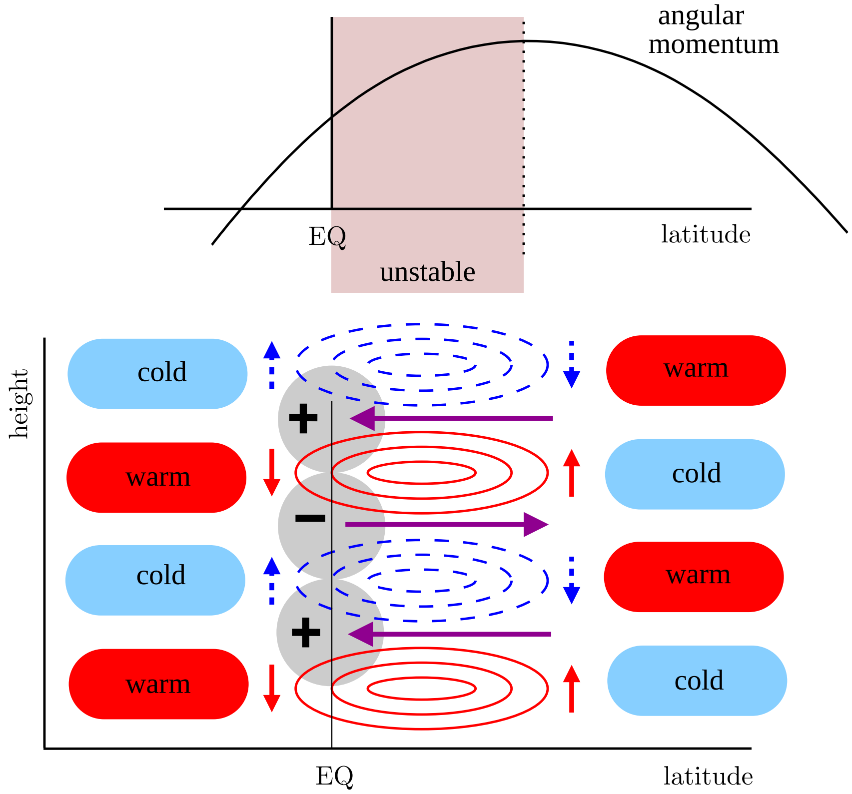


Figure 1.1: Schematic diagram showing features of the Dunkerton (1981) solution to the primitive equations linearized about state with linear meridional shear in the zonal velocity and stable stratification. The basic state is inertially unstable in the shaded interval. The fastest growing mode features a single column of Taylor vortices over the unstable region. Solid (dashed) contours are vortices with anticlockwise (clockwise) circulation. The shaded circles with + and − represent eastward and westward anomalies in the zonal wind on the equatorward side of the unstable region, associated with, respectively, equatorward and poleward transport of angular momentum (horizontal arrows). The **cold** and **warm** cells represent the "pancake structures" in the temperature field, associated with, respectively, upward and downward motion (vertical arrows).

The actual vertical scale of inertial instability cells in the middle atmosphere is an open question because both satellite data and numerical simulations are restricted to coarse vertical resolution, and the cells are always observed near the smallest resolvable vertical scale, causing numerical simulations to exhibit probably exaggerated vertical mixing. Indeed, Dunkerton's solution predicts that the fastest growing modes have vanishing

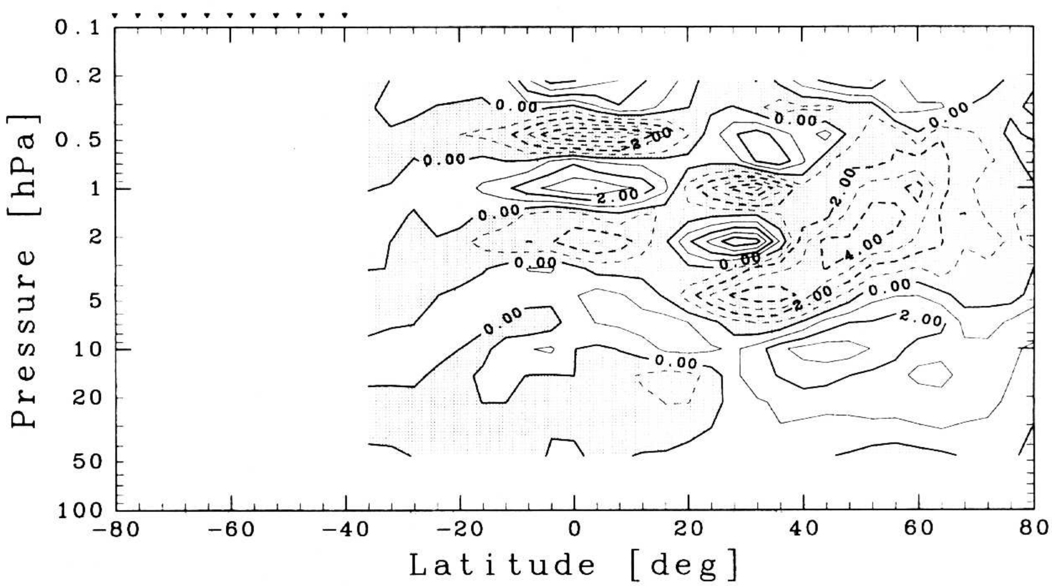


Figure 1.2: Latitude-height section of temperature anomaly (in K) from Cryogenic Limb Array Etalon Spectrometer (CLAES) data. Values are averaged over 7 days (December 14-20, 1992) and 60° longitudinal width centred on 225°E and filtered to reduce the signal of midlatitude planetary waves. The checkerboard pattern in the temperature anomaly is characteristic of inertial instability. From Hayashi et al. (1998).

vertical scale. Theories for the true preferred vertical scale are many and are based on, for example, effects of diffusion (Dunkerton, 1981), zonal asymmetry in the background state (Clark and Haynes, 1996), and secondary shear instability and nonlinear effects smoothing out small scale cells (Griffiths, 2003).

Role in solstice dynamics

Aside from direct observational evidence for inertial instability, there are features of the observed temperature and momentum fields in the middle atmosphere that suggest that inertial *adjustment* is an important part of the general circulation. In other words, the system is constantly being forced towards an inertially unstable state and *adjusting* itself towards a stable or neutrally stable state.

During solstice seasons, the estimated *radiative equilibrium* temperature profile (i.e.

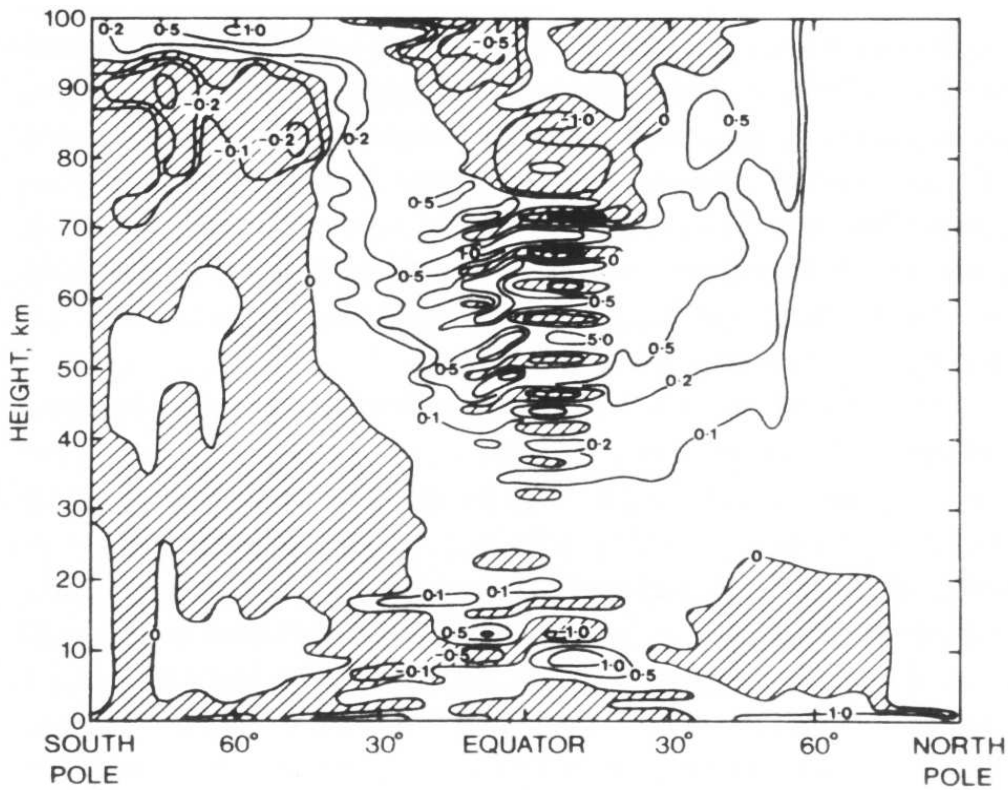


Figure 1.3: Time-averaged mean meridional wind (in ms^{-1}) distribution for January computed by a general circulation model. The column of cells of alternating sign on the winter side of the equator is believed to be due to inertial instability. From Hunt (1981).

the temperature profile that represents a balance between solar forcing, outgoing radiation and radiatively active chemical processes) has a maximum in the summer hemisphere (see, for example, Andrews et al., 1987). If the radiative equilibrium were realized, there would be a pressure gradient from the summer hemisphere to the winter hemisphere across the equator. However, a pressure gradient at the equator cannot be balanced by a purely zonal flow and there would therefore be a cross-equatorial meridional flow from summer to winter. If the angular momentum maximum is over the equator, as required by the Rayleigh criterion for stability, then a cross-equatorial flow tries to advect the angular momentum maximum into the winter hemisphere, leading to inertial instability. The solstice season temperature profiles from satellite data, and the angular momentum

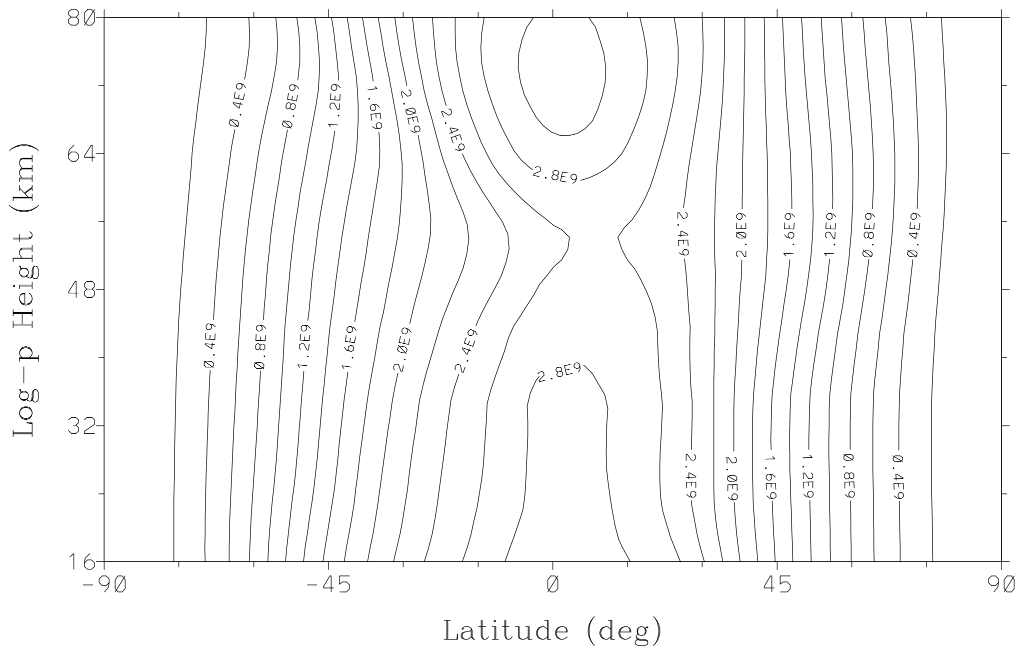


Figure 1.4: January mean absolute angular momentum contours (units m^2s^{-1}) from the Canadian Middle Atmosphere Model (CMAM). The smoothing of the horizontal absolute angular momentum gradient over the equator, especially near the stratopause ($\approx$ 50 km) is due to cross equatorial advection and inertial adjustment. Figure courtesy of K. Semeniuk (from Semeniuk and Shepherd, 2001).

profiles seen in numerical simulations (we do not have reliable measurements of horizontal winds in the tropical middle atmosphere) actually show approximately zero gradient across the equatorial region, and this is conjectured to be due to constant inertial adjustment mixing the air and smoothing the gradients (see Semeniuk and Shepherd, 2001, and Figure 1.4).

Inertial adjustment is thus part of the dynamics underlying the semi-annual oscillation in the zonal wind direction, eastward at equinox, westward at solstice (see, e.g., Shepherd, 2000). The summer-to-winter advection at solstice, maintained by inertial adjustment, carries relatively westward angular momentum from the summer hemisphere over the equator. (The eastward phase is caused by breaking equatorial Kelvin waves propagating upward from the troposphere.)

Inertial instability events also contribute to the excitation of the so-called two-day wave in the equatorial *summer* stratosphere by transporting relatively eastward momentum from the winter hemisphere towards the equator (see Figure 1.1), thereby increasing the curvature of the eastward jet in the summer hemisphere causing it to become barotropically unstable (Limpasuvan et al., 2000).

Inertial instability in the equatorial ocean

There is also evidence of inertial instability in the equatorial ocean. Inertial instability has been nominated by Hua et al. (1997) as an explanation for the pattern of alternately signed vertically stacked zonal jets below the equatorial thermocline (between 200 and 2000 m depth; see Firing, 1987). This feature is consistent with the Dunkerton solution (adapted for ocean fluid dynamics).

The same data shows a zonal velocity distribution corresponding to almost uniform angular momentum in the equatorial region (Hua et al., 1997). It is worth noting that the observed *vertical* gradient of zonal velocity approximately cancels out the vertical variation of planetary angular momentum, an effect which is neglected more often than not.

1.3 Structure of thesis

The thesis is organized as follows. In Chapter 2, some preliminary mathematical issues are covered. The equations governing inviscid, adiabatic flow in a rotating spherical shell are described, the validity of the widely used "traditional" hydrostatic approximation is discussed, the equatorial β-plane approximation used in the following chapters is defined and applied, and some useful ideas related to Hamiltonian fluid mechanics are introduced.

In Chapter 3, sufficient conditions for *linear* symmetric stability under the Euler equations are derived using a version of the energy-Casimir stability analysis method.

The result is novel in that it accounts for the effects of both components of the Earth's rotation vector. The same conditions are derived using a method based on fluid parcel displacements. Several examples of steady solutions to the symmetric Euler equations (velocity, pressure and temperature) are worked out and their stability assessed.

In Chapter 4, a Hamiltonian formulation of the symmetric anelastic equations on an equatorial β-plane is presented. Results of Chapter 3 are reproduced for the anelastic system and are extended to nonlinear stability (finite amplitude perturbations). The nonlinear stability result is used to calculate bounds on the available potential energy that can be released in adjustment from an unstable steady state ("saturation bounds"). The linearized equations for the anelastic version of the "Dunkerton problem" with constant meridional shear in the zonal velocity are solved and the solution compared with the hydrostatic case.

Finally, in Chapter 5, the stability of inviscid flow between coaxial rotating cylinders is considered. The Rayleigh condition for stability, that the magnitude of angular momentum increase with distance from the axis of rotation, is shown to be valid for finite amplitude disturbances provided the ratio of cylinder rotation rates is not too high. The finite amplitude result is used to show that the energy available to be released from an unstable flow into growing Taylor vortices approaches zero as the stability threshold is approached, demonstrating that the bifurcation from laminar Couette flow to Taylor vortex flow is supercritical. When the ratio of cylinder rotation rates is high, the finite amplitude stability result does not apply, allowing for the possibility of a subcritical instability suggested by experiments.

1.3.1 Some remarks on notation

The chapters in this thesis are for the most part self contained units. As such, some symbols and variables are redefined more than once. With a few exceptions, the meaning of a given symbol is the same in each chapter. We note the exceptions here, and again when they come up in the body of the thesis.

In Chapter 2, capital letters are used for "characteristic scales" of length, velocity, etc., with the time scale denoted by the lower case Greek τ to avoid confusion with temperature T. In Chapters 3, 4, and 5, capital letters represent *steady state* values of the variables denoted by the corresponding lower case letters. Hence, for example, U in Chapter 2 is a representative value of horizontal velocity, but in Chapter 3, $U \equiv U(y, z)$ is a steady state zonal velocity distribution.

In Chapter 3, rather than introduce a non-standard arbitrary symbol for temperature, and in order to adhere to the rule of lower case letters for time varying dependent variables and upper case for their steady state counterparts, $\tau \equiv \tau(\rho, \eta) \equiv \tau(y, z, t)$ is used for time varying temperature so that $T \equiv T(y, z)$ can represent the steady state temperature. Also, since upper case counterparts to ρ (density) and p (pressure) look the same, the steady state density is symbolized by D. The steady state entropy is symbolized by N rather than H (upper case η) for aesthetic reasons.

As a further confusion (hopefully not), in the nondimensionalization of the equations as part of the derivation of the anelastic equations in Chapter 4, Θ is first used as the characteristic temperature and potential temperature scale, and then later as $\Theta(y, z)$, the steady state value of the nondimensional potential temperature θ.

In Chapter 5, cylindrical polar coordinates are the natural choice for the geometry of the Taylor-Couette experiment. Velocity components in cylindrical coordinates are conventionally u, v, and w for, respectively, the radial, azimuthal and vertical components. In spherical polar coordinates (and on the β-plane), u is the azimuthal (zonal; x) component, v is the meridional (y) component and w the radial (vertical; z). These are standard, so confusion is unlikely, but because the component of absolute angular momentum parallel to the rotation axis, the most important quantity in this thesis, involves u in Chapters 2, 3, and 4, and v in Chapter 5, it bears mentioning.

Occasionally, similar looking symbols are used in close proximity and may lead to confusion. The symbol $\boldsymbol{x}$ represents position in space, but the similar looking $\mathbf{x}$ is the generalized independent variable (e.g. $\mathbf{x} \equiv (u, v, w, \rho, \eta)$) for the functional calculus of Hamiltonian mechanics. The small parameter in the derivation of the anelastic equations in Chapter 4 is ϵ (epsilon), and so we use ε ("varepsilon") whenever an arbitrary small real number is required in a definition or operation.

Chapter 2

The governing equations and approximate dynamical models

2.1 The unapproximated equations

For the purposes of this work, we will consider the inviscid, adiabatic Euler equations for a compressible fluid to exactly describe flow in a planetary atmosphere. We are thereby neglecting dissipative (frictional) forces and the radiation and irreversible diffusion of heat.

The first element in the statement of the Euler equations is that a fluid, which consists of haphazardly moving discrete particles, is considered to be a *continuum*, and that bulk descriptions such as temperature, pressure, and density can be applied equally well to volumes of fluid of any size. Infinitesimal volumes of fluid will be called *parcels*. The properties of the fluid within a parcel, including density, temperature, position and velocity, may be considered to be uniform.

The Euler equations are then simply the statements of conservation of mass, Newton's second law of motion applied to a fluid, the first law of thermodynamics, and an equation of state. Let $\rho(\boldsymbol{x}, t)$ be the density of the fluid parcel at position $\boldsymbol{x}$ at time t, and let

$\boldsymbol{v}(\boldsymbol{x}, t)$ be its instantaneous velocity vector. Conservation of mass means that the rate of change of mass inside any closed volume $\mathcal{V}$ (fixed with time) equals the inward flux of mass through the boundary of $\mathcal{V}$. In the limit as $\mathcal{V}$ becomes infinitesimally small, this implies the *continuity equation*

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0. \tag{2.1.1}$$

Newton's second law of motion for a fluid states that the rate of change of the linear momentum of the fluid parcels inside an arbitrary volume is equal to the inward flux of linear momentum through the boundary of the volume plus the sum of all external forces acting on the fluid parcels. In the absence of frictional forces, and viewed from an inertial reference frame, the only external forces we will consider are gravity $\rho\boldsymbol{g}$, where $\boldsymbol{g}(\boldsymbol{x})$ is the (constant in time) acceleration due to gravity vector, and the net pressure exerted by the fluid surrounding the volume. In differential form, the *momentum equation* is

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v} \cdot \nabla)\boldsymbol{v} - \boldsymbol{g} + \frac{1}{\rho}\nabla p = 0, \tag{2.1.2}$$

where $p(\boldsymbol{x}, t)$ is the fluid pressure.

We introduce the notation

$$\frac{D}{Dt} \equiv \frac{\partial}{\partial t} + (\boldsymbol{v} \cdot \nabla), \tag{2.1.3}$$

where D/Dt is called the *material* or *Lagrangian* time derivative. Df/Dt represents the time rate of change of the property f of a fluid parcel moving with velocity $\boldsymbol{v}$. The momentum equation (2.1.2) can then be written

$$\frac{D\boldsymbol{v}}{Dt} = \boldsymbol{g} - \frac{1}{\rho}\nabla p, \tag{2.1.4}$$

showing that the acceleration of fluid parcels is due to the influence of gravity and the *pressure gradient* force.

The final prognostic equation required is for the evolution of the thermodynamic fields. Since we are concerned with *adiabatic* motion, the thermodynamic variable that is most convenient is *entropy* η, which satisfies

$$\frac{D\eta}{Dt} = 0. \tag{2.1.5}$$

A property f satisfying $Df/Dt = 0$ is said to be *materially conserved*, or conserved *following the flow*.

The system is closed by an equation of state relating pressure, density and entropy:

$$F(p, \rho, \eta) = 0. \tag{2.1.6}$$

We will leave the equation of state unspecified as much as possible in this work so that the results may be applied to a broad class of fluids. An example of an equation of state is the ideal gas equation, which will be used in the examples in Chapter 3 and throughout Chapter 4. See Appendix 3.A for some useful thermodynamic identities.

2.1.1 Dynamics in a rotating reference frame

For obvious reasons, we prefer to describe the dynamics of planetary atmospheres in terms of a reference frame fixed with respect to the surface of the planet. Since the planet is rotating about an approximately fixed axis[1], the preferred reference frame is *noninertial*, and motion described with respect to it will appear to be influenced by inertial forces.

Let the rotation of the noninertial frame be described by the constant vector $\boldsymbol{\Omega}$, and consider an arbitrary time dependent vector $\boldsymbol{q}(t)$. It is readily shown (e.g., Batchelor, 1967) that the time derivatives of $\boldsymbol{q}$ in the two reference frames are related by

$$\frac{\mathrm{d}_r\boldsymbol{q}}{\mathrm{d}t} = \frac{\mathrm{d}_f\boldsymbol{q}}{\mathrm{d}t} - \boldsymbol{\Omega} \times \boldsymbol{q}, \tag{2.1.7}$$

where the reference frames are denoted by the subscripts r (rotating) and f (fixed). Notice that a vector fixed in the rotating frame precesses about $\boldsymbol{\Omega}$ at constant angular

[1] The axis of rotation is precessing slowly with time, and the planet is always accelerating in its motion about the sun, but these are relatively small effects.

speed $|\boldsymbol{\Omega}|$. Applying (2.1.7) twice, to the positions and then to the velocities of fluid parcels, we obtain the momentum equation in a rotating frame,

$$\frac{D\boldsymbol{v}}{Dt} = -2\boldsymbol{\Omega} \times \boldsymbol{v} + \boldsymbol{\Omega} \times \boldsymbol{\Omega} \times \boldsymbol{x} + \boldsymbol{g} - \frac{1}{\rho}\nabla p, \tag{2.1.8}$$

where $\boldsymbol{v}$ and $\boldsymbol{x}$ are measured relative to the rotating frame. The two inertial terms on the right hand side of (2.1.4) are referred to, respectively, as the *Coriolis* and *centrifugal* forces.

2.1.2 Spherical coordinates

For describing motion in the atmosphere around a planet, it is natural to work in spherical polar coordinates. In fact, since planets are rotating, they are not truly spherical but bulge at the equator. The bulging is such that the gravitational field is altered from that of a sphere in such a way that the net force on objects at rest on the surface of the planet is normal to the surface (i.e. the surface is an equipotential). Let $\boldsymbol{g}'$ denote the net force in the rotating frame on objects at rest (in the rotating frame). In the case of the Earth, the equatorial bulge is a small effect (approximately 21 km), and we may consider the surface to be a sphere to a reasonable level of approximation. To the same level of approximation, we may also assume that $\boldsymbol{g}'$ is in the radial direction and neglect the component of the centrifugal force in the meridional direction. This will be made explicit below. See Stommel and Moore (1989), Chapters 6-9, and Phillips (1973) for detailed discussions of this issue.

We proceed, then, by making the simplifying approximation that the Earth is a sphere with radius a (we also neglect topography on the surface of the earth). Let r measure distance from the centre of the Earth. We then define $z = r - a$ to be altitude above the Earth's surface. Let λ be longitude measured from an arbitrary angle, and let ϕ be latitude measured from the equatorial plane. Let $\hat{\mathbf{e}}_\lambda$, $\hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_z$ be the unit basis vectors in spherical coordinates, and let the components of velocity be

$$u \equiv r \cos\phi \,\frac{\mathrm{d}\lambda^{(p)}}{\mathrm{d}t}, \;\; v \equiv r \,\frac{\mathrm{d}\phi^{(p)}}{\mathrm{d}t}, \;\; w \equiv \frac{\mathrm{d}z^{(p)}}{\mathrm{d}t}, \tag{2.1.9}$$

where the superscript (p) indicates position of a fluid parcel,[1] so that

$$\boldsymbol{v} = u\,\hat{\mathbf{e}}_\lambda + v\,\hat{\mathbf{e}}_\phi + w\,\hat{\mathbf{e}}_z, \tag{2.1.10}$$

and the material derivative operator (2.1.3) becomes

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{u}{r\cos\phi}\frac{\partial}{\partial\lambda} + \frac{v}{r}\frac{\partial}{\partial\phi} + w\frac{\partial}{\partial z}. \tag{2.1.11}$$

The components of the momentum equation (2.1.8) in spherical coordinates are

$$\frac{Du}{Dt} = \frac{uv}{r}\tan\phi - \frac{uw}{r} + 2\Omega v\sin\phi - 2\Omega w\cos\phi - \frac{1}{\rho r\cos\phi}\frac{\partial p}{\partial\lambda}, \tag{2.1.12a}$$

$$\frac{Dv}{Dt} = -\frac{u^2}{r}\tan\phi - \frac{vw}{r} - 2\Omega u\sin\phi - \frac{1}{\rho r}\frac{\partial p}{\partial\phi}, \tag{2.1.12b}$$

$$\frac{Dw}{Dt} = \frac{u^2}{r} + \frac{v^2}{r} + 2\Omega u\cos\phi - \frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{2.1.12c}$$

where

$$g \equiv -|\boldsymbol{g}| - \Omega^2 r\cos^2\phi, \tag{2.1.13}$$

is the radial component of the net gravitational and centrifugal force, and we have neglected the meridional component of the centrifugal force. The magnitude of the gravitational force $|\boldsymbol{g}|$ varies like the inverse square of the distance to the centre of the Earth, but for most atmospheric and oceanic purposes, it may be taken to be constant. The centrifugal force term, which depends on position, is commonly neglected altogether. Actually, the centrifugal force term in (2.1.13) substantially cancels out the reduction in gravity due to the bulging of the Earth (about 0.3% at the equator), making the constant g approximation even more accurate.

[1] This prevents a circular definition of velocity and the material derivative D/Dt. Alternatively, we could have introduced a field of *parcel labels* $\boldsymbol{a}(\boldsymbol{x}, t)$, and velocity is then the partial time derivative of $\boldsymbol{x}(\boldsymbol{a}, t)$ with $\boldsymbol{a}$ held fixed. The Jacobian of the transformation from $\boldsymbol{x}$ coordinates to $\boldsymbol{a}$ coordinates is usually chosen to be the density. This formalism is convenient for the canonical Hamiltonian formulation of fluid mechanics in which the independent variable is $\boldsymbol{a}$ (Salmon, 1998).

The continuity equation can be expanded using

$$\nabla \cdot (\rho \boldsymbol{v}) = \frac{1}{r^2 \cos\phi} \left[\frac{\partial}{\partial \lambda}(\rho u r) + \frac{\partial}{\partial \phi}(\rho v r \cos\phi) + \frac{\partial}{\partial r}(\rho w r^2 \cos\phi) \right]. \tag{2.1.14}$$

2.1.3 Energy and angular momentum conservation

Assuming suitable boundary conditions, the system of equations (2.1.8), (2.1.1) and (2.1.5) conserves total energy and the component of absolute angular momentum parallel to $\boldsymbol{\Omega}$.

The total energy E is the sum of kinetic energy $\frac{1}{2}\rho \boldsymbol{v} \cdot \boldsymbol{v}$, potential energy $\rho\Phi(z)$ ($\Phi = gz$ when g is assumed to be constant[1]), and *internal* energy $\mathcal{E}(\rho, \eta)$, integrated over the entire domain $\mathcal{D}$ (or, equivalently, over all fluid parcels), i.e.

$$E(\boldsymbol{v}, \rho, \eta) = \int_{\mathcal{D}} \rho \left[\tfrac{1}{2} \boldsymbol{v} \cdot \boldsymbol{v} + \Phi(z) + \mathcal{E}(\rho, \eta) \right] \mathrm{d}\boldsymbol{x}. \tag{2.1.15}$$

$\mathcal{E}(\rho, \eta)$ satisfies the first law of thermodynamics for adiabatic processes,

$$\frac{D\mathcal{E}}{Dt} = \frac{p}{\rho^2} \frac{D\rho}{Dt}. \tag{2.1.16}$$

Using (2.1.8), (2.1.1), (2.1.5) and (2.1.16), and neglecting the centrifugal force and the radial dependence of gravity, it can be shown that

$$\frac{\mathrm{d}E}{\mathrm{d}t} = \int_{\partial\mathcal{D}} -\rho^2 \left[\frac{1}{2} \boldsymbol{v} \cdot \boldsymbol{v} + gz + \mathcal{E}(\rho, \eta) + \frac{p}{\rho^2} \right] \boldsymbol{v} \cdot \hat{\boldsymbol{\nu}} \, \mathrm{d}S(\boldsymbol{x}), \tag{2.1.17}$$

where $\hat{\boldsymbol{\nu}}(\boldsymbol{x})$ is the outward pointing unit normal vector to the domain $\mathcal{D}$, and $\mathrm{d}S(\boldsymbol{x})$ is an element of area on the surface $\partial\mathcal{D}$. For the case of a spherical shell, typical boundary conditions that imply conservation of energy are that the velocity is perpendicular to the inner and outer spherical boundaries, or if the outer boundary is infinite, that the density and pressure vanish sufficiently rapidly as z tends to infinity.

[1] If the centrifugal force and the radial dependence of gravity are retained but the bulging of the planet is still neglected, then $\Phi(\boldsymbol{x}) \equiv GM/r - \frac{1}{2}\Omega^2 r^2 \cos^2\phi$, where G is the universal gravitational constant and M is the mass of the planet.

In spherical coordinates, the integral in (2.1.15) may be written as a triple integral over (λ, ϕ, r)

$$E(u,v,w,\rho,\eta) = \int_0^{2\pi}\int_{-\frac{\pi}{2}}^{\frac{\pi}{2}}\int_0^{\infty} \rho\left[\tfrac{1}{2}(u^2+v^2+w^2)+gz+\mathcal{E}(\rho,\eta)\right] r^2\cos\phi\,\mathrm{d}r\,\mathrm{d}\phi\,\mathrm{d}\lambda, \tag{2.1.18}$$

and (2.1.12) may be used to verify that E is conserved. Note that all of the dependent variables are periodic in λ with period 2π.

The component of absolute angular momentum parallel to $\mathbf{\Omega}$, due to motion in the zonal (east-west) direction, is given by

$$M \equiv (u+\Omega r\cos\phi) r\cos\phi. \tag{2.1.19}$$

From (2.1.11) and (2.1.12a),

$$\frac{DM}{Dt} = -\frac{1}{\rho}\frac{\partial p}{\partial\lambda} \tag{2.1.20}$$

and, since the lower boundary is independent of λ and pressure is periodic in λ,

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathcal{D}} \rho M \mathrm{d}\boldsymbol{x} = -\int_{\partial\mathcal{D}} \rho M \boldsymbol{v}\cdot\hat{\boldsymbol{\nu}}\mathrm{d}S(\boldsymbol{x}). \tag{2.1.21}$$

Hence, with energy conserving boundary conditions, the total absolute zonal angular momentum is also conserved.

2.2 Traditional hydrostatic approximation

The equations (2.1.12) are valid for flow in a general spherical shell rotating at a constant rate, with the only approximation being the neglect of the centrifugal force. It has long been recognized that planetary atmospheres do not exhibit all of the possible motions allowed by the exact equations. The most striking property of observed atmospheric behaviour is that horizontal accelerations are much greater than vertical accelerations. The forces on air parcels in the vertical (radial) direction are very nearly balanced. The balance between the largest terms in (2.1.12c),

$$\frac{\partial p}{\partial z} = -\rho g, \tag{2.2.1}$$

is commonly referred to as *hydrostatic balance.* It is the condition for equilibrium in a system at rest ($\boldsymbol{v} = 0$). The accuracy of (2.2.1) for a system in motion depends on the length, time, and velocity scales of the particular problem. It is appropriate for modelling extratropical cyclones which evolve slowly and over horizontal length scales of thousands of kilometres, but not, for example, for modelling thunderstorms, which have short horizontal length scales, evolve over periods of hours, and involve relatively strong vertical motion.

As we will see, the condition of near hydrostatic balance is related to the shallow atmosphere approximation, which we have already invoked in assuming that the acceleration due to gravity is approximately constant. Together with the (not just aesthetic) requirement of retaining principles of conservation of energy and angular momentum, either approximation — hydrostatic balance or the shallow atmosphere — may be used as the basis for simplifying the momentum equations. The two approaches end with similar but not identical systems of equations.

First, consider the legitimacy of assuming hydrostatic balance. Equation (2.2.1) agrees very well with observation — vertical accelerations are typically not more than 10^{-2} ms^{-2}, even in an intense thunderstorm, compared to $g \approx 10$ ms^{-2}. But the criterion for invoking the hydrostatic approximation should rather be that the vertical acceleration Dw/Dt be small compared to the *departure* of the terms in (2.2.1) from equilibrium. White (2002) argues for the case of an ideal gas as follows. Suppose the pressure and density vary slightly from a reference state according to

$$p = p_0(z) + p'(\lambda, \phi, z, t) \tag{2.2.2a}$$

$$\rho = \rho_0(z) + \rho'(\lambda, \phi, z, t), \tag{2.2.2b}$$

where $p_0(z)$ and $\rho_0(z)$ satisfy (2.2.1). Then, to first order in ρ' and p',

$$\frac{Dw}{Dt} \approx \frac{u^2}{r} + \frac{v^2}{r} + 2\Omega u \cos\phi - \frac{1}{\rho_0}\frac{\partial p'}{\partial z} - \frac{\rho'}{\rho_0}g. \tag{2.2.3}$$

The *potential temperature*, defined by

$$\theta(p, \rho) \equiv \left(\frac{p_{00}}{p}\right)^{\kappa} T = \frac{p_{00}^{\kappa}}{R}\frac{p^{1-\kappa}}{\rho}, \tag{2.2.4}$$

where T is temperature, and R and $\kappa \equiv R/c_p = 2/7$ are constants[1], is materially conserved by an adiabatic flow. It therefore satisfies

$$\frac{D\theta'}{Dt} + w\frac{\mathrm{d}\theta_0}{\mathrm{d}z} = 0, \tag{2.2.5}$$

where $\theta_0(z) = \theta(p_0, \rho_0)$ and $\theta' = \theta - \theta_0$. If the time scale of changes in θ is τ and is comparable to the time scale of changes in w, then we have, from (2.2.5),

$$\left|\frac{Dw}{Dt}\right| \sim \frac{\theta_0}{\tau^2(\mathrm{d}\theta_0/\mathrm{d}z)}\left|\frac{\theta'}{\theta_0}\right|. \tag{2.2.6}$$

From (2.2.4),

$$\frac{\theta'}{\theta_0} \approx (1-\kappa)\frac{p'}{p_0} - \frac{\rho'}{\rho}, \tag{2.2.7}$$

and if temperature $T = p/R\rho$ is approximately constant, then $|p'/p_0| \approx |\rho'/\rho_0| \sim |\theta'/\theta_0|$. Therefore,

$$\left|\frac{Dw}{Dt}\right| \sim \frac{1}{N^2\tau^2}\left|g\frac{\rho'}{\rho}\right|, \tag{2.2.8}$$

where $N^2 \equiv (g/\theta_0)\mathrm{d}\theta_0/\mathrm{d}z$ is the square of the *buoyancy* or *Brunt-Väisälä frequency*. N^2 is a measure of the stratification of the background hydrostatic state of the atmosphere.[2] If N^2 is positive, then it is the square of the frequency of oscillation of a vertically displaced fluid parcel. The criterion for discarding Dw/Dt is therefore that the dominant timescale in the problem τ must be much greater than N^{-1}.

[1] $R \approx 287$ JK^{-1}kg^{-1} is the ideal gas constant for air, and c_p is the heat capacity at constant pressure. For a diatomic gas, $c_p = \frac{7}{2}R$.

[2] p_0 and ρ_0 were arbitrary, but in practice, they might be chosen based on, say, the time mean pressure and density.

Alternatively, one could compare Dw/Dt to the pressure gradient term in (2.2.3), $(1/\rho_0)\partial p'/\partial z$. The two approaches must be equivalent since if Dw/Dt is small compared to each of the two large terms[1], they must be of the same order of magnitude. We present the argument from Phillips (1973).

From the thermodynamic equation (2.1.5) and the continuity equation (2.1.1),

$$\frac{Dp}{Dt} = \frac{Dp'}{Dt} + w\frac{\mathrm{d}p_0}{\mathrm{d}z} = c_s^2 \rho \nabla \cdot \boldsymbol{v}, \tag{2.2.9}$$

where $c_s^2 \equiv (\partial p/\partial \rho)|_\eta$ is the speed of sound. Therefore, if W and U are characteristic scales for vertical and horizontal velocity, and H and L for vertical and horizontal length, then

$$\left|\frac{Dw}{Dt}\right| \sim \frac{W}{\tau} = \left(\frac{1}{H} + \frac{g}{c_s^2}\right)^{-1} \left(\frac{U}{\tau L} + \frac{1}{c_s^2 \tau^2}\left|\frac{p'}{\rho}\right|\right). \tag{2.2.10}$$

If the pressure perturbation term on the right hand side of (2.2.10) is larger than $U/\tau L$, then the condition for the neglect of Dw/Dt, the condition that $|Dw/Dt| \ll (1/\rho_0)|\partial p'/\partial z|$, is that $\tau \gg \sqrt{HH'}/c_s$, where H' is the smaller of H and c_s^2/g. For typical Earth atmosphere values, $c_s \approx 320$ ms^{-1}, so $H \sim H' \sim 10^4$ m, and the condition is satisfied if the timescale of motion is much greater than 30 seconds. This condition is satisfied for all processes of interest.

If, on the other hand, the pressure perturbation term in (2.2.10) is smaller than $U/\tau L$, then the continuity equation implies $W \lesssim UH/L$ (the inequality allows for partial cancellation between $r\cos\phi\,\partial u/\partial\lambda$ and $r\,\partial v/\partial\phi$). From (2.1.12a) and (2.1.12b), the pressure perturbation satisfies

$$\frac{|p'|}{\rho} \sim \begin{cases} fUL, & f\tau > 1 \\ \dfrac{UL}{\tau}, & f\tau < 1 \end{cases}, \tag{2.2.11}$$

where $f \equiv 2\Omega\sin\phi$ and the metric terms due to the curvature of the coordinate system

[1] We are ignoring the other terms in (2.2.3): the metric terms are much smaller than g — $(u^2+v^2)/r \lesssim 10^{-4}$ ms^{-2} for Earth values; the Coriolis term $2\Omega u\cos\phi$ will be discussed in the next subsection.

have again been ignored. Therefore,

$$\left|\frac{Dw}{Dt}\right| \sim \frac{W}{\tau} \lesssim \begin{cases} \dfrac{H'H}{f\tau L^2}\dfrac{|p'|}{\rho H}, & f\tau > 1 \\ \dfrac{H'H}{L^2}\dfrac{|p'|}{\rho H}, & f\tau < 1 \end{cases} . \tag{2.2.12}$$

Hence, the other condition for neglecting Dw/Dt is that

$$L^2 \gg H'H. \tag{2.2.13}$$

This is a condition on the *aspect ratio* of the problem. If the vertical scale H is on the order of the depth of the atmospheric layer, then, since the horizontal scale L cannot be more than $2\pi a$, (2.2.13) implies $H \ll a$, a shallow atmosphere.

We have already mentioned that for most atmospheric applications, the acceleration due to gravity g may be taken to be a constant. This is dependent on the thinness of the atmospheric layer compared to the radius of the planet. The exact value of g outside a spherical planet of mass M and radius a is

$$g(r) = \frac{GM}{r^2} \approx \frac{GM}{a^2}\left(1 - 2\frac{z}{a}\right), \tag{2.2.14}$$

where G is the universal gravitational constant. Therefore, the constant g approximation depends on z/a being small. For the case of Earth, the atmosphere is approximately 10^2 km thick (depending on what is meant by atmosphere), compared to a mean radius of 6×10^3 km.

The shallow atmosphere approximation can be used to simplify the equations further (in addition to setting $g =$ constant) by replacing r^{-1} by a^{-1} wherever it appears in (2.1.1), (2.1.5) and (2.1.12), including in the material derivative (2.1.11) and divergence (2.1.14) operators. However, naively making the replacements and leaving the equations otherwise unchanged upsets the principle of angular momentum conservation.

Phillips (1966) (see also Müller, 1989) explains that the conservation principle is upset because replacing r^{-1} by a^{-1} in the component equations creates a system of equations that are not the components of a vector equation. Instead, he writes (2.1.8) in the form

$$\frac{\partial \boldsymbol{v}}{\partial t} = -\nabla(\tfrac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v}) + \boldsymbol{v}\times[\nabla\times(\boldsymbol{v}+\boldsymbol{v}_{\boldsymbol{\Omega}})] + \boldsymbol{g}' - \frac{1}{\rho}\nabla p, \tag{2.2.15}$$

where $\boldsymbol{v}_{\boldsymbol{\Omega}} \equiv \Omega r\cos\phi\,\hat{\mathbf{e}}_\lambda$ is the zonal component of the velocity of the rotating reference frame. The Coriolis force comes from the term

$$\begin{aligned}\boldsymbol{v}\times(\nabla\times\boldsymbol{v}_{\boldsymbol{\Omega}}) \;&=\; (2\Omega\sin\phi v - 2\Omega\cos\phi w)\,\hat{\mathbf{e}}_\lambda \\ &\quad - 2\Omega\sin\phi u\,\hat{\mathbf{e}}_\phi + 2\Omega\cos\phi u\,\hat{\mathbf{e}}_r. \end{aligned}\tag{2.2.16}$$

The shallow atmosphere approximation is then made in the definition of the curvilinear coordinates (r,ϕ,λ) by defining the scale factors

$$h_\lambda \equiv a\cos\phi, \qquad h_\phi \equiv a, \qquad h_r \equiv 1 \tag{2.2.17}$$

(see, e.g. Arfken, 1985, Chapter 2, for discussion of scale factors in curvilinear coordinates) and by setting $\boldsymbol{v}_{\boldsymbol{\Omega}} \equiv \Omega a\cos\phi\,\hat{\mathbf{e}}_\lambda$[1]. Then the velocity components become (c.f. (2.1.9))

$$u \equiv a\cos\phi\frac{\mathrm{d}\lambda^{(p)}}{\mathrm{d}t},\;\; v \equiv a\frac{\mathrm{d}\phi^{(p)}}{\mathrm{d}t},\;\; w \equiv \frac{\mathrm{d}r^{(p)}}{\mathrm{d}t} = \frac{\mathrm{d}z^{(p)}}{\mathrm{d}t}, \tag{2.2.18}$$

and the gradient and divergence operators have r^{-1} replaced by a^{-1}. The Coriolis term is then simply

$$\boldsymbol{v}\times(\nabla\times\boldsymbol{v}_{\boldsymbol{\Omega}}) = 2\Omega\sin\phi v\,\hat{\mathbf{e}}_\lambda - 2\Omega\sin\phi u\,\hat{\mathbf{e}}_\phi. \tag{2.2.19}$$

Evidently, the shallow atmosphere approximation has forced the exclusion of the Coriolis force terms proportional to $\cos\phi$, terms due to the component of $\boldsymbol{\Omega}$ parallel to the surface of the planet. Also, most of the metric terms disappear since the r dependence is gone from the curl operator. The approximated momentum equations are

$$\frac{D_a u}{Dt} = \frac{uv}{a}\tan\phi + 2\Omega v\sin\phi - \frac{1}{\rho a\cos\phi}\frac{\partial p}{\partial\lambda}, \tag{2.2.20a}$$

$$\frac{D_a v}{Dt} = -\frac{u^2}{a}\tan\phi - 2\Omega u\sin\phi - \frac{1}{\rho a}\frac{\partial p}{\partial\phi}, \tag{2.2.20b}$$

$$\frac{D_a w}{Dt} = -\frac{1}{\rho}\frac{\partial p}{\partial z} - g, \tag{2.2.20c}$$

[1]The shallow atmosphere system is effectively a family of spheres, all of radius a, all rotating with the same angular speed, with the family parameterized by the coordinate z.

where D_a/Dt is the material derivative with r replaced by a in the coefficients, and the approximate angular momentum

$$M_a \equiv (u + \Omega a \cos\phi) a \cos\phi \tag{2.2.21}$$

satisfies $D_a M_a/Dt = -\rho^{-1}\partial p/\partial\lambda$. The shallow atmosphere angular momentum is the angular momentum a parcel would have if it were at the surface of the planet. Note that the shallow atmosphere approximation does not force us to assume hydrostatic balance. However, when the shallow atmosphere approximation is appropriate, the vertical acceleration Dw/Dt is, as we have seen, much smaller than the other terms, and is often neglected. In that case, the angular momentum principle is not affected, but to retain conservation of energy, $\frac{1}{2}w^2$ is dropped from the kinetic energy.

Had we decided to assume hydrostatic balance without making the shallow atmosphere approximation, and simply replaced (2.1.12c) with (2.2.1), then the energy principle would have been upset. Dropping $\frac{1}{2}w^2$ from the kinetic energy and the terms proportional to w from the right hand sides of (2.1.12a,b) restores the energy principle, but in turn upsets the angular momentum principle, which is restored finally by making the shallow atmosphere approximation.

The resulting equations (2.2.20a,b) and (2.2.1) are known as the *hydrostatic primitive equations* (or simply the *primitive equations*). Note that in the new system, there is no longer a prognostic equation for the vertical velocity. It is obtained by requiring that hydrostatic balance is maintained, i.e. that

$$\frac{\partial}{\partial t}\left(\frac{\partial p}{\partial z} + \rho g\right) = 0. \tag{2.2.22}$$

An advantage of the primitive equations is that if pressure is used as the vertical coordinate, the continuity equation (2.1.1) becomes simply a statement that the divergence of velocity vanishes. Note that pressure is necessarily decreasing with height by virture of (2.2.1), so this change of coordinates is always well defined. A drawback of using pressure coordinates is that the lower boundary becomes a function of time. See Lorenz (1967) for a clear discussion of the hydrostatic approximation, the primitive equations and pressure coordinates.

2.2.1 Neglect of $\cos\phi$ Coriolis force terms

In deducing the primitive equations from the hydrostatic approximation or Phillips' system from the shallow atmosphere approximation, we were forced to drop the Coriolis force terms due to the component of $\mathbf{\Omega}$ parallel to the surface of the planet, terms proportional to $\cos\phi$, in order to retain the conservation of angular momentum principle. Unlike the neglect of the spherical "metric" terms proportional to $1/r$, the neglect of the $\cos\phi$ Coriolis terms is not obviously justified on scaling grounds.

Eckart (1960) calls attention to the precariousness of the approximation by referring to the neglect of the $\cos\phi$ Coriolis force terms in the zonal and vertical momentum equations as the "traditional" approximation. According to Eckart, the approximation is made in order to remove "mathematical difficulties" from his *field equations* (essentially the Euler equations linearized about a rest state). The field equations may be solved for motion bounded by spherical level surfaces by the method of separation of variables only if the traditional approximation is made.

Certainly at midlatitudes, where $\sin\phi \sim \cos\phi$, and under circumstances in which the predominant vertical velocity scale W is much less than the horizontal velocity scale U, the neglected term in the zonal momentum equation $2\Omega w \cos\phi$ is relatively small compared to $2\Omega v \sin\phi$. However, near the equator, where the other ($\sin\phi$) term becomes small, and the $\cos\phi$ term is the largest inertial force term (indeed, at the equator, $\mathbf{\Omega}$ is tangent to the surface), it is not clear that the approximation is justified.

Veronis (1963) considers the expression for potential vorticity,

$$q \equiv \frac{(2\mathbf{\Omega} + \nabla \times \boldsymbol{v}) \cdot \nabla \eta}{\rho}, \tag{2.2.23}$$

in spherical coordinates, and argues that the contribution from the horizontal component of $\mathbf{\Omega}$ can be neglected wherever

$$\frac{\dfrac{1}{r}\dfrac{\partial\eta}{\partial\phi}}{\dfrac{\partial\eta}{\partial r}} \sim \frac{H}{L} \ll \tan\phi, \tag{2.2.24}$$

where H and L are the vertical and meridional length scales of motion. For large scale motion in the Earth's atmosphere, $H/L \sim 10^{-3}$, in which case (2.2.24) recommends the traditional approximation outside of a band of about 1^{o} latitude about the equator.

Colin de Verdière and Schopp (1994) find a similar criterion by looking at the vorticity version of the Euler equations. For motion near the equator, $\tan\phi \lesssim L/a$, so the condition (2.2.24) may be written $\sqrt{Ha} \ll L$. For the Earth's atmosphere, $\sqrt{Ha} \approx 240$ km, and for the ocean, $\sqrt{Ha} \approx 80$ km, a scale on which "considerable energy is found at low latitudes in the form of zonal jets" (Colin de Verdière and Schopp, 1994, p244).

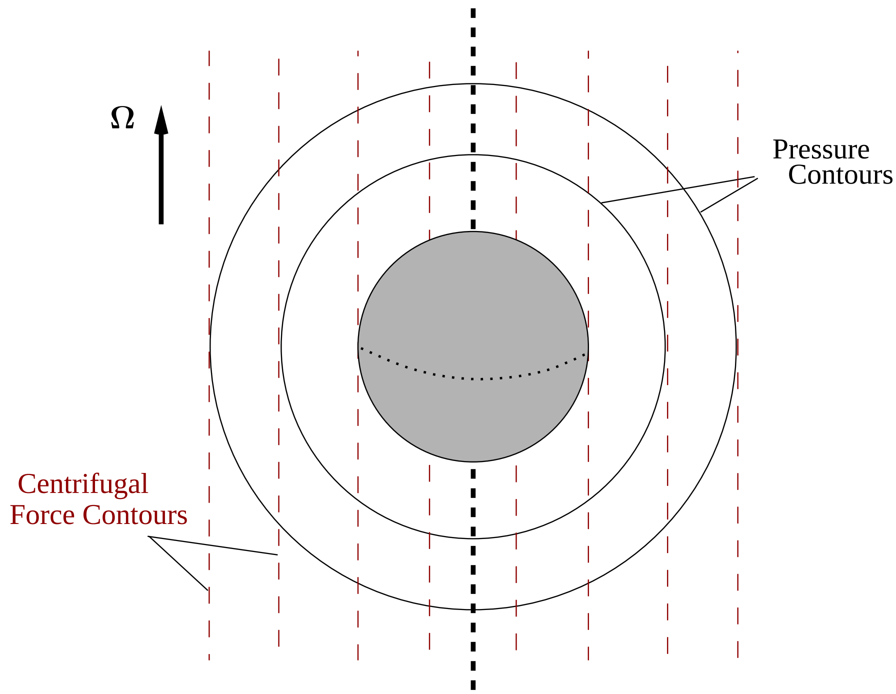
In a comment on Phillips (1966), Veronis (1968) objects to the claim that conservation of angular momentum is justification for the traditional approximation when assuming a shallow atmosphere because (2.1.20) is valid to the same degree of approximation as the approximate equations themselves. The implication is that the need to have an exact conservation principle corresponding to angular momentum is an aesthetic consideration. In his response, Phillips (1968) concedes that the approximation is not *justified* on the basis of retaining the angular momentum principle, but argues that it is (perhaps) justified based on the following physical grounds (see also Gill, 1982). In the dispersion relation for plane-wave solutions to the shallow Euler equations (i.e. with r replaced by a but without the traditional approximation, the system suggested by Veronis) linearized about a stratified atmosphere at rest, the $\cos\phi$ terms are insignificant if $4\Omega^2 \ll N^2$, where N is again the buoyancy frequency (a measure of stratification). Therefore, the traditional approximation is justified when the stratification is such that $N^2 \gg 10^{-8}\ \text{s}^{-2}$. Typical in the middle atmosphere is $N^2 \approx 10^{-4}\ \text{s}^{-2}$.

The above argument is based on the effects of the Coriolis terms on oscillations about a state of rest, but the $\cos\phi$ terms can have a significant effect on the mean flow that obtains in the atmosphere. Assuming motion on a time scale L/U and the approximate nondivergence condition $W \lesssim HU/L$, White and Bromley (1995) estimate that the term $2\Omega w \cos\phi$ in the zonal momentum equation is much smaller than $|Du/Dt|$ and (equivalently) the term $2\Omega u \cos\phi$ in the vertical equation is much smaller than $|\rho'/\rho_0 g|$ when

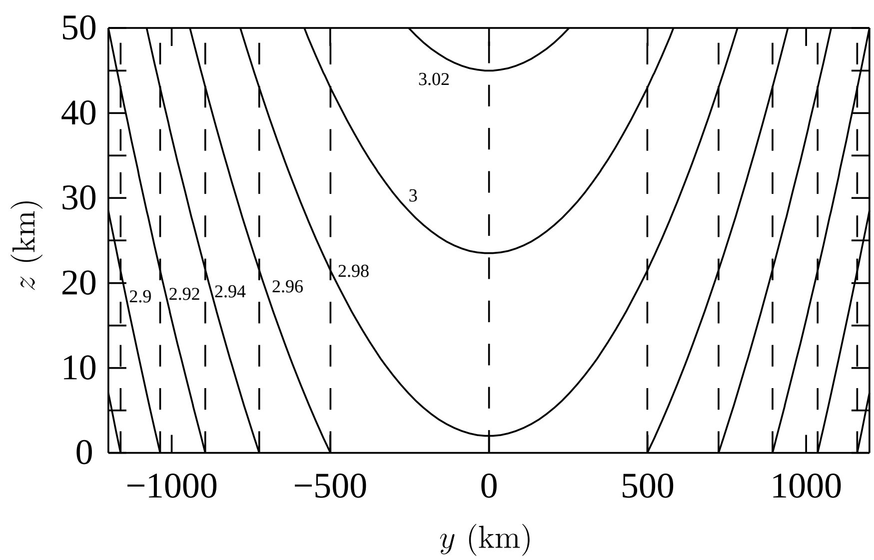
$$\frac{2\Omega H \cos\phi}{U} \ll 1. \tag{2.2.25}$$

For $H \approx 10$ km and $U \approx 10$ ms^{-1}, $2\Omega H \cos\phi/U \approx 0.14 \cos\phi$. Close to the equator, they argue that this is certainly too large to ignore in simulations of atmospheric motion. They offer the *quasihydrostatic* equations, which are the Euler equations without Dw/Dt in the vertical momentum equation, and claim that they are not much more computationally expensive than the primitive equations. A pressure-coordinate based form of the quasihydrostatic equations are used in the United Kingdom Meteorological Office forecast model (White, 1999).

Advocates of neglecting the $\cos\phi$ Coriolis terms appeal to particular situations in which the terms are insignificant as reasons to neglect them in general. The prudent course would seem to be to retain all of the terms unless they cause insurmountable "mathematical difficulties". In the following chapters, in which we will consider the stability of steady zonal solutions around the equator, we retain a representation of both components of the rotation vector $\mathbf{\Omega}$. We will find that stability depends on the orientation of gradients of angular momentum and entropy relative to $\mathbf{\Omega}$ and surfaces of constant pressure. In Figure 2.1a, approximate pressure contours (spheres) and contours of centrifugal force (cylinders parallel to $\mathbf{\Omega}$) are shown. Without the $\cos\phi$ terms, constant pressure surfaces would be everywhere perpendicular to $\mathbf{\Omega}$. Notice that at the equator, the two sets of surfaces are actually tangent. This has consequences both for which equilibrium states are possible and for which of them are stable.



(a)



(b)

Figure 2.1: (a) Approximate contours of centrifugal force (cylinders) and pressure (spheres). (b) Contours of planetary angular momentum $a(\Omega a - \frac{1}{2}\beta y^2 + \gamma z)$ with γz term (solid lines) and without (dashed). Contours are in 10^9 m^{2}s^{-1}. Contours which meet at the surface have the same value.

2.3 The equatorial β-plane

In analysing fluid motion in the equatorial region, particularly in studying equatorial wave phenomena, it is very common to consider the problem on an equatorial β-plane instead of, or as a precursor to, the corresponding problem in a spherical shell.

Recall that in spherical coordinates, the components of the planetary rotation vector $\mathbf{\Omega}$ are functions of position:

$$\mathbf{\Omega} = \Omega(\cos\phi\,\hat{\mathbf{e}}_\phi + \sin\phi\,\hat{\mathbf{e}}_r). \tag{2.3.1}$$

The β-plane approximation replaces the spherical coordinates used in a spherical shell geometry with rectangular Cartesian coordinates but retains a linear approximation to the variation of the rotation vector with position. The applicability of the approximation depends on the smallness of the meridional length scale compared to the radius of the planet. The following derivation is based on Grimshaw (1975).

The first step is to write the equations (2.1.12) in terms of (r, λ, μ), where $\mu(\phi)$ is the *Mercator* latitude (Eckart, 1960), defined by

$$\mathrm{d}\phi = \cos\phi\,\mathrm{d}\mu, \qquad \mu(0) = 0. \tag{2.3.2}$$

Integrating (2.3.2) gives the relations

$$\cosh\mu = \sec\phi, \qquad \sinh\mu = \tan\phi. \tag{2.3.3}$$

Equal (infinitesimal) intervals of longitude λ and Mercator latitude μ represent equal arclengths on the surface of a sphere. The material derivative can be written

$$\frac{D}{Dt} = \frac{\partial}{\partial t} + \frac{1}{r\cos\phi}\left(u\frac{\partial}{\partial\lambda} + v\frac{\partial}{\partial\phi}\right) + w\frac{\partial}{\partial z}. \tag{2.3.4}$$

Next, we introduce the dimensionless coordinates which will be shown to behave like rectangular coordinates in the β-plane limit

$$x^* \equiv \left(\frac{a}{L}\right)\lambda, \qquad y^* \equiv \left(\frac{a}{L}\right)\mu, \qquad z^* \equiv \left(\frac{a}{H}\right)\ln\left(\frac{r}{a}\right), \tag{2.3.5}$$

where a is the radius of the planet, and L and H are appropriate horizontal and vertical length scales for the particular problem being studied. Note that the vertical coordinate z^* is, to first order in (H/a), equal to $(r-a)/H$. Also, let U be a characteristic horizontal velocity scale, and let

$$t^* \equiv \left(\frac{U}{L}\right) t, \qquad \frac{D}{Dt^*} \equiv \left(\frac{L}{U}\right) \frac{D}{Dt}, \qquad \Omega^* \equiv \left(\frac{L}{U}\right) \Omega. \tag{2.3.6}$$

The components of velocity are defined by

$$u^* \equiv \frac{Dx^*}{Dt^*}, \qquad v^* \equiv \frac{Dy^*}{Dt^*}, \qquad w^* \equiv \frac{Dz^*}{Dt^*}, \tag{2.3.7}$$

so that

$$u = \left[U\left(\frac{r}{a}\right)\cos\phi\right] u^*, \qquad v = \left[U\left(\frac{r}{a}\right)\cos\phi\right] v^*, \qquad w = \left[U\left(\frac{r}{a}\right)\left(\frac{H}{L}\right)\right] w^*, \tag{2.3.8}$$

and the material derivative is

$$\frac{D}{Dt^*} \equiv \left(\frac{L}{U}\right)\frac{D}{Dt} = \frac{\partial}{\partial t^*} + u^*\frac{\partial}{\partial x^*} + v^*\frac{\partial}{\partial y^*} + w^*\frac{\partial}{\partial z^*}. \tag{2.3.9}$$

The continuity equation becomes

$$\begin{aligned} \frac{D\rho}{Dt^*} &= -\rho\left\{\frac{\partial u^*}{\partial x^*} + \frac{\partial v^*}{\partial y^*} + \frac{\partial w^*}{\partial z^*}\right. \\ &\qquad \left.\left(\frac{L}{a}\right)\left[-2v^*\sin\phi + \left(\frac{H}{L}\right)(3r^3 w^*)\right]\right\}. \end{aligned} \tag{2.3.10}$$

The zonal momentum equation (2.1.12a) becomes

$$\begin{aligned} &\frac{Du^*}{Dt^*} + \left(\frac{L}{a}\right)\left[\left(\frac{H}{L}\right)u^*w^* - u^*v^*\sin\phi\right] \\ &\qquad = 2\Omega^*\left[v^*\sin\phi - \left(\frac{H}{L}\right)w^*\right] - \left(\frac{a^2}{U^2}\right)\left(\frac{1}{\rho r^2\cos^2\phi}\right)\frac{\partial p}{\partial x^*} \\ &\qquad\qquad + \left(\frac{L}{a}\right)\left[u^*v^*\tan\phi - \left(\frac{H}{L}\right)u^*w^*\right], \end{aligned} \tag{2.3.11}$$

with similar equations for the meridional and vertical momentum equations.

Expanding $\sin\phi$, $\cos\phi$, and r/a about $L/a = 0$,

$$\sin\phi = \left(\frac{L}{a}\right) y^* + \mathcal{O}\left[\left(\frac{L}{a}\right)^3\right], \tag{2.3.12a}$$

$$\cos\phi = 1 - \frac{1}{2}\left(\frac{L}{a}\right)^2 y^{*2} + \mathcal{O}\left[\left(\frac{L}{a}\right)^4\right], \tag{2.3.12b}$$

$$\frac{r}{a} = 1 + \left(\frac{H}{L}\right)\left(\frac{L}{a}\right) z^* + \frac{1}{2}\left(\frac{H}{L}\right)^2\left(\frac{L}{a}\right)^2 z^{*2} + \mathcal{O}\left[\left(\frac{L}{a}\right)^3\right], \tag{2.3.12c}$$

so that to $\mathcal{O}[(L/a)^2]$,

$$\frac{r}{a}\cos\phi = 1 + \left(\frac{L}{a}\right)\left(\frac{H}{L}\right) z^* + \left(\frac{L}{a}\right)^2\left[\frac{1}{2}\left(\frac{H}{L}\right)^2 z^{*2} - \frac{1}{2}y^{*2}\right], \tag{2.3.13a}$$

$$\left(\frac{r}{a}\cos\phi\right)^2 = 1 + \left(\frac{L}{a}\right)\left[2\left(\frac{H}{L}\right) z^*\right] + \left(\frac{L}{a}\right)^2\left[2\left(\frac{H}{L}\right)^2 z^{*2} - y^{*2}\right]. \tag{2.3.13b}$$

Substituting (2.3.12) and (2.3.13) into (2.3.11) gives

$$\begin{aligned} &\frac{Du^*}{Dt^*} + \left(\frac{L}{a}\right)\left(\frac{H}{L}\right) u^* w^* \\ &\quad = 2\Omega^*\left[\left(\frac{L}{a}\right) y^* v^* - \left(\frac{H}{L}\right) w^*\right] - \left(\frac{1}{U^2}\right)\frac{1}{\rho}\frac{\partial p}{\partial x^*} \\ &\qquad 2\Omega^*\sin\phi + \left(\frac{L}{a}\right)\left[-\left(\frac{H}{L}\right) u^* w^* - 2\left(\frac{H}{L}\right) z^*\left(\frac{1}{U^2}\right)\frac{1}{\rho}\frac{\partial p}{\partial x^*}\right] \\ &\qquad\qquad + \mathcal{O}\left[\left(\frac{L}{a}\right)^2\right]. \end{aligned} \tag{2.3.14}$$

The β-plane equations are obtained by taking the limit as L/a approaches zero, *while* $\Omega^* L/a = \Omega U/a$ *remains finite*. In that limit,

$$\frac{D\rho}{Dt^*} = -\rho\left(\frac{\partial u^*}{\partial x^*} + \frac{\partial v^*}{\partial y^*} + \frac{\partial w^*}{\partial z^*}\right), \tag{2.3.15}$$

and

$$\frac{Du^*}{Dt^*} = \left(\frac{2\Omega^* L}{a}\right) y^* v^* - \left(\frac{2\Omega^* H}{L}\right) w^* - \left(\frac{1}{U^2}\right)\frac{1}{\rho}\frac{\partial p}{\partial x^*}. \tag{2.3.16}$$

Dimensions can be restored to the equations by defining

$$(x, y) \equiv L(x^*, y^*), \qquad z \equiv H z^*, \tag{2.3.17a}$$

$$(\tilde{u}, \tilde{v}) \equiv U(u^*, v^*), \qquad \tilde{w} \equiv U\left(\frac{H}{L}\right) w^*, \tag{2.3.17b}$$

$$\beta \equiv \frac{2\Omega}{a}, \qquad \gamma \equiv 2\Omega. \tag{2.3.17c}$$

Assuming the scalings introduced earlier, $(\tilde{u}, \tilde{v}, \tilde{w})$ is the leading order β-plane approximation to the exact dimensional velocity $\boldsymbol{v} = (u, v, w)$. The full set of approximate equations is

$$\frac{D\tilde{u}}{Dt} = \beta y \tilde{v} - \gamma \tilde{w} - \frac{1}{\rho}\frac{\partial p}{\partial x}, \tag{2.3.18a}$$

$$\frac{D\tilde{v}}{Dt} = -\beta y \tilde{u} - \frac{1}{\rho}\frac{\partial p}{\partial y}, \tag{2.3.18b}$$

$$\frac{D\tilde{w}}{Dt} = \gamma \tilde{u} - g - \frac{1}{\rho}\frac{\partial p}{\partial z}, \tag{2.3.18c}$$

$$\frac{D\rho}{Dt} = -\rho\left(\frac{\partial \tilde{u}}{\partial x} + \frac{\partial \tilde{v}}{\partial y} + \frac{\partial \tilde{w}}{\partial z}\right). \tag{2.3.18d}$$

Notice that the Coriolis force terms associated with the horizontal component of $\boldsymbol{\Omega}$ is retained in these equations. Based on (2.3.16), it should be kept or may be dropped depending on the largeness of the *aspect ratio* L/H. The effect of neglecting these terms on the stability problems considered in the following chapters may be observed by setting γ equal to zero.

2.3.1 Hamiltonian representation

The analysis in the following chapters is based on the fact that the adiabatic, inviscid equations are a Hamiltonian system. In this subsection, we present the β-plane equations derived above in *Hamiltonian form* without belaboring the details of the mathematics of Hamiltonian fluid mechanics. Elements of the theory will be introduced as needed throughout the thesis. For detailed discussions of Hamiltonian systems of equations, see, e.g., Arnold (1989), and for fluid systems, see Morrison (1998), Salmon (1998), and Shepherd (1990).

We now show that the β-plane equations (2.3.18) can be written in the Hamiltonian form

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{J}\frac{\delta \mathcal{H}}{\delta \mathbf{x}}, \tag{2.3.19}$$

where $\mathbf{x} \equiv (u, v, w, \rho, \eta)^T$ is the independent variable, $\mathcal{J}$ is the Poisson tensor, and $\mathcal{H}$ is the Hamiltonian. The Hamiltonian for the β-plane equations is just the total energy (dropping tildes from the velocity components)

$$\mathcal{H}(u, v, w, \rho, \eta) = \iiint \rho \left[\tfrac{1}{2}(u^2 + v^2 + w^2) + gz + \mathcal{E}(\rho, \eta)\right] \mathrm{d}x\, \mathrm{d}y\, \mathrm{d}z. \tag{2.3.20}$$

$\delta\mathcal{H}/\delta\mathbf{x}$ is the functional gradient of $\mathcal{H}$ and has components

$$\frac{\delta\mathcal{H}}{\delta\boldsymbol{v}} = \rho\boldsymbol{v}, \tag{2.3.21a}$$
$$\frac{\delta\mathcal{H}}{\delta\rho} = \frac{1}{2}\boldsymbol{v}\cdot\boldsymbol{v} + gz + \mathcal{E} + \rho\frac{\partial\mathcal{E}}{\partial\rho}, \tag{2.3.21b}$$
$$\frac{\delta\mathcal{H}}{\delta\eta} = \rho\frac{\partial\mathcal{E}}{\partial\eta}. \tag{2.3.21c}$$

It may be verified that (2.3.18) have the form (2.3.19) if the Poisson tensor is (see Shepherd, 1990, ,§4.5)

$$\mathcal{J} = \begin{bmatrix} 0 & \rho^{-1}\omega_3 & -\rho^{-1}\omega_2 & -\partial_x & \rho^{-1}\eta_x \\ -\rho^{-1}\omega_3 & 0 & \rho^{-1}\omega_1 & -\partial_y & \rho^{-1}\eta_y \\ \rho^{-1}\omega_2 & -\rho^{-1}\omega_1 & 0 & -\partial_z & \rho^{-1}\eta_z \\ -\partial_x & -\partial_y & -\partial_z & 0 & 0 \\ -\rho^{-1}\eta_x & -\rho^{-1}\eta_y & -\rho^{-1}\eta_z & 0 & 0 \end{bmatrix}, \tag{2.3.22}$$

where

$$(\omega_1, \omega_2, \omega_3) = (w_y - v_z, u_z - w_x + \gamma, v_x - u_y + \beta y). \tag{2.3.23}$$

2.3.2 Angular momentum conservation

Recall that in the exact Euler equations, the absolute angular momentum (2.1.19) obeyed a conservation principle. The conservation of the (integrated) angular momentum is connected to the fact that the equations are Hamiltonian, and do not explicitly depend on λ.[1] The equations are thus invariant under a continuous *point transformation* of the

[1] The conservation of angular momentum also requires that the *boundary conditions* are independent of λ.

dependent variables, namely $[\boldsymbol{v},\rho,\eta](\lambda,\phi,r,t) \to [\boldsymbol{v},\rho,\eta](\lambda+\delta\lambda,\phi,r,t)$. According to *Noether's Theorem* (see, e.g., Arnold, 1989), such symmetries may be associated with conserved functionals. More precisely, if the equations are Hamiltonian and are invariant under a point transformation $\mathbf{x} \to \mathbf{x}+\delta\mathbf{x}$, and the functional $\mathcal{F}(\mathbf{x})$ satisfies

$$-\epsilon\mathcal{J}\frac{\delta\mathcal{F}}{\delta\mathbf{x}} = \delta\mathbf{x}, \tag{2.3.24}$$

for some ϵ, then $\mathcal{F}(\mathbf{x})$ is conserved in time.

Consider the transformation corresponding to a translation in the x direction, i.e. $[\tilde{\boldsymbol{v}},\rho,\eta](x,y,z,t) \to [\tilde{\boldsymbol{v}},\rho,\eta](x+\epsilon,y,z,t)$. The corresponding conserved functional $\mathcal{M}(\mathbf{x})$ satisfies

$$-\epsilon\mathcal{J}\frac{\delta\mathcal{M}}{\delta\mathbf{x}} = \epsilon\frac{\partial\mathbf{x}}{\partial x}. \tag{2.3.25}$$

An $\mathcal{M}(\mathbf{x})$ satisfying (2.3.25) has functional gradient

$$\frac{\delta\mathcal{M}}{\delta\mathbf{x}} = [\rho, 0, 0, \tilde{u} - \tfrac{1}{2}\beta y^2 + \gamma z, 0]. \tag{2.3.26}$$

Inverting the functional gradient operation gives the conserved functional

$$\mathcal{M} = \iiint \rho(u - \tfrac{1}{2}\beta y^2 + \gamma z)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{2.3.27}$$

By analogy with (2.1.21), the angular momentum associated with the β-plane equations may therefore be defined to be[1]

$$m \equiv u - \tfrac{1}{2}\beta y^2 + \gamma z. \tag{2.3.28}$$

We might ask how m is related to the β-plane limit of the exact angular momentum M given by (2.1.19). It may be verified that

$$M = \frac{U}{L}a^2\Omega^* + Ua\left\{u^* + \left(\frac{H}{L}\right)(2\Omega^* z^*) + \left(\frac{\Omega^* L}{a}\right)\left[2\left(\frac{H}{L}\right)^2 z^{*2} - y^{*2}\right] + \mathcal{O}\left[\left(\frac{L}{a}\right)\right]\right\}, \tag{2.3.29}$$

[1] Alternative choices, differing from (2.3.28) by the addition of a *Casimir invariant*, are possible but exactly equivalent for the purpose of the present discussion.

which would suggest the dimensional version

$$m' = u + \gamma z + \beta(z^2 - \tfrac{1}{2}y^2), \tag{2.3.30}$$

where the additive and multiplicative constants have been dropped. Evidently, the β-plane approximation is missing the βz^2 term, associated with variation of $\mathbf{\Omega}$ with r. Assuming $H/L \ll 1$ (shallow atmosphere), this term is small compared to γz, so the difference between the conserved quantity m and the approximate angular momentum m' is small.

Except near the equator, the term γz is small compared to $\frac{1}{2}\beta y^2$, but we feel that it should be retained for the problem of equatorial inertial stability for which the state of the atmosphere in the vicinity of the equator is critical. As we will show in Chapter 3, the criterion for stability will depend on the orientation of the gradients of entropy and angular momentum relative to contours of constant pressure and "planetary angular momentum" $M^{(\mathrm{p})} \equiv \gamma z - \frac{1}{2}\beta y^2$. The significance of retaining the γz term may be felt by referring to Figure 2.1b. The parabolae are contours of constant $M^{(\mathrm{p})}$, and the vertical lines are the same contours with $\gamma \equiv 0$ (curves that meet at the surface have the same value of $M^{(\mathrm{p})}$). Notice also that Figure 2.1b resembles Figure 2.1a: the lines of constant height (pressure) are now (approximately) flat, and the lines of constant $M^{(\mathrm{p})}$ (centrifugal force) curve upward away from the equator.

The inclusion of the γz term also places restrictions on what velocity, density and entropy fields satisfy conditions for equilibrium, and a stability criterion is only a meaningful test if equilibrium solutions exist to which it can be applied.

Chapter 3

Symmetric stability of steady zonal solutions to the compressible Euler equations

In this chapter, we derive sufficient conditions for the linear stability of axisymmetric steady state solutions to the fully compressible Euler equations on an equatorial β-plane with otherwise arbitrary velocity and temperature fields to disturbances that preserve the axial symmetry. What is novel in the calculation is that the Coriolis force terms due to the horizontal component of the planetary rotation vector, the terms proportional to $\cos\phi$, which are neglected in most applications, are included here. The well-known condition for symmetric stability, that the potential vorticity Q have the same sign as latitude (Stevens, 1983), is shown to apply to this more general case, and together with a generalized static stability condition, is sufficient for linear stability.

The main effects of retaining the $\cos\phi$ terms are that the structure of the steady solutions in the vicinity of the equator is significantly more restricted, and the widths of regions of anomalous Q depend (weakly) on the reintroduced terms.

The approach used in this chapter, known as the energy-Casimir method, due originally to Arnold (see Shepherd, 1990), exploits the Hamiltonian nature of the Euler equations as well as the material conservation of absolute zonal angular momentum and entropy associated with, respectively, the zonal symmetry and the adiabatic nature of the flow. For a given steady solution $\mathbf{X}$ of the governing equations, we construct a functional $\mathcal{A}$ of the dynamical fields (velocity, density, and entropy) which has a critical point, i.e. either a local extremum or a saddle point, at $\mathbf{X}$. The required functional is a combination of the Hamiltonian (total energy) and a Casimir invariant. The condition that $\mathbf{X}$ be a critical point corresponds to conditions on the gradient of the Casimir evaluated at $\mathbf{X}$. The further condition that $\mathbf{X}$ be a global minimum of $\mathcal{A}$ (more precisely, that $\mathcal{A}(\mathbf{x}) - \mathcal{A}(\mathbf{X})$ be bounded from above and below by the square of a norm $||\mathbf{x} - \mathbf{X}||^2$ for all $\mathbf{x} \neq \mathbf{X}$) corresponds to sufficient conditions on $\mathbf{X}$ for stability.

Normally, the method is used to derive conditions on $\mathbf{X}$ for stability to finite amplitude perturbations ("nonlinear stability"), with the small amplitude result ("linear stability") emerging as a by-product. However, due to technical issues which we will outline, the finite amplitude problem is not easily tractable in this case. However, we apply a version of the energy-Casimir method directly to the problem linearized about $\mathbf{X}$ and derive rigorous conditions for small amplitude stability.

Further details of the method may be found in Shepherd (1990) and Holm et al. (1985). It has been successfully applied to the symmetric stability problem in the Boussinesq equations on the f-plane (Cho et al., 1993; Mu et al., 1996) and, with restrictions, in the hydrostatic equations in a shallow spherical shell (Bowman and Shepherd, 1995).

The structure of the chapter is as follows. In Section 3.1, the conditions for linear stability are calculated using both the energy-Casimir method and a heuristic approach based on forces on displaced fluid parcels, and a sequence of instructive examples are presented. In Section 3.2, aspects of the problem of nonlinear stability are discussed.

3.1 Linear stability of zonal equilibrium

3.1.1 Symmetric equations

We seek conditions for the stability of zonally symmetric steady solutions to the β-plane equations (derived in the previous chapter) under zonally symmetric perturbations. The governing equations are (2.3.18) but with $\partial_x \equiv 0$. Consider then the following equations governing adiabatic, compressible x-symmetric flow in a domain with rectangular cross-section $\mathcal{D} = \{(y,z) | -L \le y \le L, 0 \le z \le H\}$:

$$\begin{aligned} u_t &= -vu_y - wu_z + \beta yv - \gamma w, && (3.1.1a)\\ v_t &= -vv_y - wv_z - \beta yu - \frac{1}{\rho}p_y, && (3.1.1b)\\ w_t &= -vw_y - ww_z + \gamma u - g - \frac{1}{\rho}p_z, && (3.1.1c)\\ \rho_t &= -(\rho v)_y - (\rho w)_z, && (3.1.1d)\\ \eta_t &= -v\eta_y - w\eta_z, && (3.1.1e) \end{aligned}$$

where u, v and w are the components of velocity in the x, y and z directions; ρ, p and η are density, pressure and entropy; and subscripts indicate partial differentiation. The flow is subject to the no-normal-flow boundary condition $v = 0$ on $y = \pm L$ and $w = 0$ on $z = 0, H$. The system is closed by an equation of state

$$F(\rho, p, \eta) = 0 \qquad (3.1.2)$$

such as the ideal gas law. Temperature τ may be obtained from any two of ρ, p and η using the laws of thermodynamics (see Appendix 3.A).

The equations conserve energy

$$\mathcal{H}(\mathbf{x}) = \iint\limits_{\mathcal{D}} \rho \left[\tfrac{1}{2}(u^2+v^2+w^2) + gz + \mathcal{E}(\rho,\eta)\right] \mathrm{d}y\,\mathrm{d}z, \qquad (3.1.3)$$

where $\mathbf{x} \equiv (u, v, w, \rho, \eta)$, and $\mathcal{E}(\rho,\eta)$ is internal energy, satisfying the thermodynamic identity

$$\mathrm{d}\mathcal{E} = \frac{p}{\rho^2}\mathrm{d}\rho + \tau \mathrm{d}\eta, \qquad (3.1.4)$$

and functionals of the form

$$\mathcal{C}(\mathbf{x}) = \iint_{\mathcal{D}} \rho C(m, \eta)\, \mathrm{d}y\, \mathrm{d}z, \tag{3.1.5}$$

where $C(m, \eta)$ is any function, and m is defined by

$$m \equiv u - \tfrac{1}{2}\beta y^2 + \gamma z. \tag{3.1.6}$$

Note that m is proportional to the β-plane approximation to the component of absolute angular momentum parallel to the Earth's rotation axis (2.3.30), and will be referred to simply as *angular momentum* henceforth.

Functionals of the form (3.1.5) are related to the Casimir invariants of the noncanonical Hamiltonian representation of the system (the same as in Section 2.3.1, but with $\partial_x \equiv 0$), and their conservation is a consequence of the conservation of m and η by fluid parcels. We note also that fluid parcels conserve potential vorticity

$$q = \frac{1}{\rho}\partial(\eta, m). \tag{3.1.7}$$

Here $\partial(\cdot, \cdot)$ is the Jacobian operator, defined by

$$\partial(f, g) = f_y g_z - f_z g_y = (\nabla f \times \nabla g) \cdot \hat{\imath}, \tag{3.1.8}$$

where $\hat{\imath}$ is the unit vector in the x direction. The sign of $\partial(f, g)$ is given by the "right-hand-rule" applied to ∇f and ∇g: positive if ∇f points in the semicircle clockwise of ∇g.

3.1.2 Steady solution and linearized equations

We seek to decide the stability of an equilibrium solution $\mathbf{X}$ of (3.1.1) with $u = U(y, z)$, $v = w = 0$, $\eta = N(y, z)$ and $\rho = D(y, z)$, with associated $m = M(U; y, z)$ and $\tau = T(D, N)$. The pressure field $P(D, N)$ balances the velocity and mass fields: from (3.1.1b) and (3.1.1c),

$$-\beta y U - \frac{1}{D}P_y = 0, \tag{3.1.9a}$$
$$\gamma U - g - \frac{1}{D}P_z = 0, \tag{3.1.9b}$$

which may be combined by "cross-differentiation" to get the thermal wind balance relation,

$$\partial(M, M^{(\mathrm{p})}) = \frac{1}{D^2}\partial(P, D), \tag{3.1.10}$$

which relates the *baroclinic vector* $\nabla P \times \nabla D$ to the basic state velocity field.

In this section, we derive conditions on $\mathbf{X}$ such that the zero solution to the equations (3.1.1) linearized about $\mathbf{X}$ is stable with respect to arbitrary disturbances. These may be interpreted as conditions under which $\mathbf{X}$ is stable to infinitesimal disturbances. The linearized equations are

$$u'_t = -U_y v' - U_z w' + \beta y v' - \gamma w', \tag{3.1.11a}$$
$$v'_t = -\beta y u' - \frac{1}{D}p'_y + \frac{1}{D^2}P_y\rho', \tag{3.1.11b}$$
$$w'_t = \gamma u' - \frac{1}{D}p'_z + \frac{1}{D^2}P_z\rho', \tag{3.1.11c}$$
$$\rho'_t = -(Dv')_y - (Dw')_z, \tag{3.1.11d}$$
$$\eta'_t = -N_y v' - N_z w', \tag{3.1.11e}$$

where primed quantities represent departure from the basic state $\mathbf{X}$. Note that the perturbation pressure p' is related to the perturbation density ρ' and entropy η' by[1]

$$p' = \left(\frac{\partial P}{\partial D}\right)_N \rho' + \left(\frac{\partial P}{\partial N}\right)_D \eta'. \tag{3.1.12}$$

Assume that the mapping from (y, z) to (N, M) has nonzero Jacobian everywhere in $\mathcal{D}$ except perhaps on a finite number of curves. Partition $\mathcal{D}$ into a finite number of subregions $\mathcal{D}^{(i)}$, $i = 1, \ldots, n$, such that

$$Q \equiv \frac{1}{D}\partial(N, M) \neq 0 \tag{3.1.13}$$

[1]We employ the notation of using capital letters in derivatives such as $(\partial P/\partial N)_D$ when the expression is to be evaluated at the basic state, rather than writing, say, $(\partial p/\partial \eta)_\rho|_{(D,N)}$.

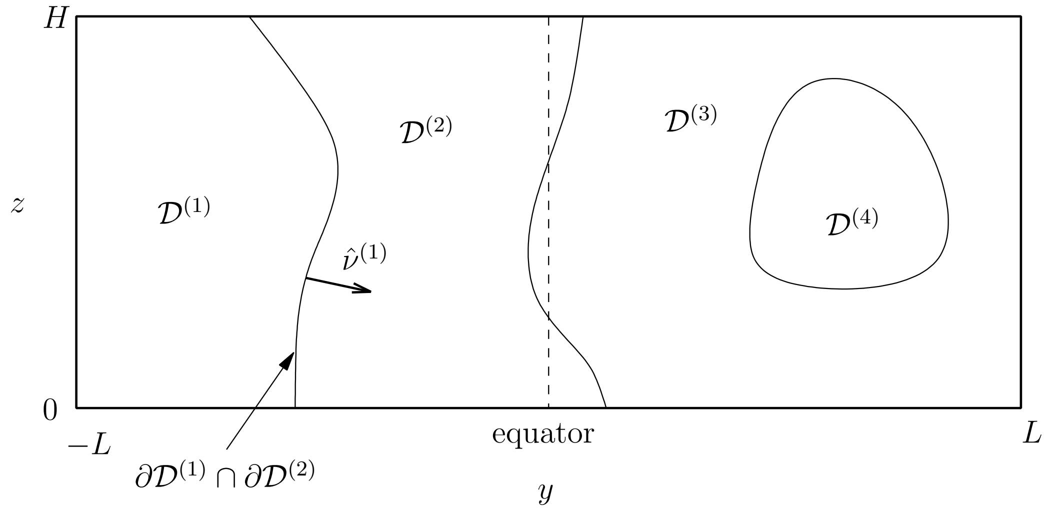


Figure 3.1: Sample partition of $\mathcal{D}$ into regions with nonzero Q.

inside each of the $\mathcal{D}^{(i)}$ (see Figure 3.1). $Q(y, z)$ is the potential vorticity associated with the basic state. By construction, the mapping from (y, z) to (N, M) has a unique inverse inside each $\mathcal{D}^{(i)}$, which we denote by $[Y^{(i)}(N, M), Z^{(i)}(N, M)]$.

3.1.3 A conservation law for the linearized dynamics

We employ the energy-Casimir approach to derive linear stability conditions. Usually (Holm et al., 1985; Cho et al., 1993), this involves constructing a functional which is exactly conserved by the nonlinear equations and finding conditions under which the functional is convex at the basic state, implying linear stability of the basic state. In our case, finding such a functional is problematic (see Section 3.2 below), but we can still apply the "nonlinear method" to the linear problem using a functional which is conserved by the *linearized* equations. Consider

$$\mathcal{C}_L = \sum_{i=1}^{n} \iint_{\mathcal{D}^{(i)}} \rho C^{(i)}(m, \eta) \, \mathrm{d}y \, \mathrm{d}z \tag{3.1.14}$$

where each of the $C^{(i)}$ are arbitrary twice-differentiable functions of m and η. We observe that $\mathcal{C}_L$ is not conserved by the nonlinear system (3.1.1). Differentiating $\mathcal{C}_L$ with respect to time and substituting from (3.1.1) gives

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}_L &= \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}}\left[\rho\left(C_m^{(i)}m_t + C_\eta^{(i)}\eta_t\right) + C^{(i)}\rho_t\right]\,\mathrm{d}y\,\mathrm{d}z \\ &= -\sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}}\left[C_m^{(i)}\rho\boldsymbol{v}\cdot\nabla m + C_\eta^{(i)}\rho\boldsymbol{v}\cdot\nabla\eta + C^{(i)}\nabla\cdot(\rho\boldsymbol{v})\right]\,\mathrm{d}y\,\mathrm{d}z \quad (3.1.15)\end{aligned}$$

where $\boldsymbol{v} \equiv (v, w)$ and $\nabla \equiv (\partial_y, \partial_z)$. Applying the divergence theorem to the last term in (3.1.15) gives

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{C}_L = \frac{1}{2}\sum_{i,j=1}^{n}\int\limits_{\partial\mathcal{D}^{(i)}\cap\partial\mathcal{D}^{(j)}}\rho\left(C^{(j)} - C^{(i)}\right)\boldsymbol{v}\cdot\hat{\boldsymbol{\nu}}^{(i)}\mathrm{d}l^{(i)}(y, z), \quad (3.1.16)$$

where $\partial\mathcal{D}^{(i)}$ is the boundary of the region $\mathcal{D}^{(i)}$, $\hat{\boldsymbol{\nu}}^{(i)}$ is the outward pointing unit vector normal to $\partial\mathcal{D}^{(i)}$, and $\mathrm{d}l^{(i)}$ is the element of arclength along $\partial\mathcal{D}^{(i)}$. The outer boundary terms vanish because velocity $\boldsymbol{v}$ is orthogonal to the boundary of $\mathcal{D}$ by assumption. $\mathrm{d}\mathcal{C}_L/\mathrm{d}t$ vanishes if the functions $C^{(i)}$ in neighbouring regions always match along the boundaries. Since the $C^{(i)}$ are in general different, this is not generally the case. We can, however, choose the $C^{(i)}$ so that they match along the boundaries when evaluated *at the basic state*. It is reasonable to expect that a functional $\mathcal{C}_L$ so constructed will be relevant for small amplitude perturbations to the basic state.

We choose the functions $C^{(i)}$ so that $\mathcal{C}_L$ is tangent to $\mathcal{H}$ (in the sense of functionals) at the basic state. This ensures that the combined functional $\mathcal{H} + \mathcal{C}_L$ has a critical point at the basic state.[1] That is

$$\delta(\mathcal{H} + \mathcal{C}_L)|_{\mathbf{X}} = 0, \quad (3.1.17)$$

where $\delta(\mathcal{H} + \mathcal{C}_L)$ is the first variation of $(\mathcal{H} + \mathcal{C}_L)$. For arbitrary $\mathbf{x} = (u, v, w, \rho, \eta)$,

$$\begin{aligned}\delta(\mathcal{H} + \mathcal{C}_L) &= \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}}\Big\{\delta\rho\left[\tfrac{1}{2}(u^2 + v^2 + w^2) + gz + \mathcal{E} + \rho\mathcal{E}_\rho + C^{(i)}\right] \\ &\quad + \rho\left[(u + C_m^{(i)})\delta u + v\delta v + w\delta w + (\mathcal{E}_\eta + C_\eta^{(i)})\delta\eta\right]\Big\}\,\mathrm{d}y\,\mathrm{d}z, \quad (3.1.18)\end{aligned}$$

[1]The rationale for this choice of $\mathcal{C}_L$ will made more clear in Chapter 4, where the Hamiltonian structure underlying the energy-Casimir method will be discussed in more detail.

so at $\mathbf{X}$,

$$
\begin{aligned}
&\delta(\mathcal{H}+\mathcal{C}_L)|_{\mathbf{X}} \\
&\quad = \sum_{i=1}^{n} \iint_{\mathcal{D}^{(i)}} \left\{ \delta\rho \left[\tfrac{1}{2}(U^2 + gZ^{(i)}(N,M) + \mathcal{E}(D,N) + D\mathcal{E}_\rho(D,N) + C^{(i)}(M,N) \right] \right. \\
&\qquad\qquad \left. + D\left[(U + C^{(i)}_m(M,N))\delta u + (\mathcal{E}_\eta(D,N) + C^{(i)}_\eta(M,N))\delta\eta \right] \right\} \mathrm{d}y\,\mathrm{d}z. \quad (3.1.19)
\end{aligned}
$$

Hence $\delta(\mathcal{H}+\mathcal{C}_L)$ vanishes at $\mathbf{X}$ if

$$
\begin{aligned}
C^{(i)}(M,N) &= -\left[\tfrac{1}{2}U^2 + gZ^{(i)}(N,M) + \mathcal{E}(D,N) + D\mathcal{E}_\rho(D,N) \right], && (3.1.20a)\\
C^{(i)}_m(M,N) &= -U, && (3.1.20b)\\
C^{(i)}_\eta(M,N) &= -\mathcal{E}_\eta(D,N) && (3.1.20c)
\end{aligned}
$$

for each i. Note that $U(y,z)$ and $D(y,z)$ are implicit functions of (N,M) in each $\mathcal{D}^{(i)}$ through the inverse mappings $[y,z] = [Y^{(i)}(N,M), Z^{(i)}(N,M)]$.

Before proceeding, we should check that the three conditions (3.1.20) are mutually consistent. Differentiating (3.1.20a) with respect to M keeping N fixed gives

$$
\begin{aligned}
C^{(i)}_m(M,N) &= -\frac{\partial}{\partial M}\left(\tfrac{1}{2}U^2 + gZ^{(i)} + \mathcal{E} + D\mathcal{E}_\rho \right)\Big|_N \\
&= -U\left[1 + \beta Y^{(i)}\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N - \gamma\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \right] - g\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \\
&\quad - (2\mathcal{E}_\rho + D\mathcal{E}_{\rho\rho})\left[D_y\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N + D_z\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \right] \\
&= -U + \frac{1}{D}\left[P_y\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N + P_z\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \right] \\
&\quad - \frac{1}{D}\left(\frac{\partial P}{\partial D}\right)_N \left[D_y\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N + D_z\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \right] \qquad (3.1.21)
\end{aligned}
$$

where all functions are evaluated at $\mathbf{X}$. In the last step we have used (3.1.9) and (3.A.5). Writing P derivatives in terms of D and N derivatives,

$$
C^{(i)}_m(M,N) = -U + \frac{1}{D}\left(\frac{\partial P}{\partial N}\right)_D \left[N_y\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N + N_z\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N \right]. \qquad (3.1.22)
$$

Finally, using (3.B.4b),

$$
\begin{aligned}
C_m^{(i)}(M,N) &= -U + \frac{1}{D^2 Q}\left(\frac{\partial P}{\partial N}\right)_D (-N_y N_z + N_z N_y) \\
&= -U,
\end{aligned}
\tag{3.1.23}
$$

which agrees with (3.1.20b). Similarly, differentiating (3.1.20a) with respect to N keeping M fixed gives

$$
\begin{aligned}
C_\eta^{(i)}(M,N) &= -\frac{\partial}{\partial N}\left(\tfrac{1}{2}U^2 + gZ^{(i)} + \mathcal{E} + D\mathcal{E}_\rho\right)\Big|_M \\
&= -U\left[\beta Y^{(i)}\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M - \gamma\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] - g\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M \\
&\quad - (2\mathcal{E}_\rho + D\mathcal{E}_{\rho\rho})\left[D_y\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M + D_z\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] \\
&\quad - (\mathcal{E}_\eta + D\mathcal{E}_{\rho\eta}) \\
&= -\mathcal{E}_\eta + \frac{1}{D}\left[P_y\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M + P_z\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] \\
&\quad - \frac{1}{D}\left(\frac{\partial P}{\partial D}\right)_N\left[D_y\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M + D_z\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] - \frac{1}{D}\left(\frac{\partial P}{\partial N}\right)_D \\
&= -\mathcal{E}_\eta - \frac{1}{D}\left(\frac{\partial P}{\partial N}\right)_D + \frac{1}{D}\left(\frac{\partial P}{\partial N}\right)_D\left[N_y\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M + N_z\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] \\
&= -\mathcal{E}_\eta - \frac{1}{D}\left(\frac{\partial P}{\partial N}\right)_D\left[1 - \frac{1}{DQ}(N_y M_z - N_z M_y)\right] \\
&= -\mathcal{E}_\eta,
\end{aligned}
\tag{3.1.24}
$$

by virtue of (3.1.13). Therefore, conditions (3.1.20) are mutually consistent.

We now construct a quadratic invariant for the linearized equations based on the second variation of $\mathcal{H} + \mathcal{C}_L$. The second variation evaluated at arbitrary $\mathbf{x}$ is

$$
\begin{aligned}
\delta^2(\mathcal{H} + \mathcal{C}_L) &= \sum_{i=1}^{n} \iint_{\mathcal{D}^{(i)}} \Big\{\rho(1 + C_{mm}^{(i)})(\delta u)^2 + \rho(\delta v)^2 + \rho(\delta w)^2 \\
&\quad + (2\mathcal{E}_\rho + \rho\mathcal{E}_{\rho\rho})(\delta\rho)^2 + \rho(\mathcal{E}_{\eta\eta} + C_{\eta\eta}^{(i)})(\delta\eta)^2 \\
&\quad + 2(u + C_m^{(i)})\delta u\delta\rho + 2v\delta v\delta\rho + 2w\delta w\delta\rho \\
&\quad + 2(\mathcal{E}_\eta + \rho\mathcal{E}_{\rho\eta} + C_\eta^{(i)})\delta\rho\delta\eta + 2\rho C_{m\eta}^{(i)}\delta u\delta\eta\Big\}\, \mathrm{d}y\, \mathrm{d}z,
\end{aligned}
\tag{3.1.25}
$$

which we evaluate at $\mathbf{X}$, finding

$$\begin{aligned}\delta^2(\mathcal{H}+\mathcal{C}_L)\big|_{\mathbf{X}} \;=\; \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}} \big\{ & D(1+C^{(i)}_{mm}(M,N))(\delta u)^2 + D(\delta v)^2 + D(\delta w)^2 \\ & + (2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))(\delta\rho)^2 \\ & + D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))(\delta\eta)^2 \\ & + 2D\mathcal{E}_{\rho\eta}(D,N)\delta\rho\delta\eta + 2DC^{(i)}_{m\eta}(M,N)\delta u\delta\eta\big\}\,\mathrm{d}y\,\mathrm{d}z. \qquad (3.1.26)\end{aligned}$$

Identifying the primed variables in the linear system (3.1.11) with the variations in (3.1.26), we define

$$\begin{aligned}\mathcal{H}_L(\mathbf{x}';\mathbf{X}) \;\equiv\; \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}} \big\{ & D(1+C^{(i)}_{mm}(M,N))u'^2 + Dv'^2 + Dw'^2 \\ & + (2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))\rho'^2 \\ & + D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))\eta'^2 \\ & + 2D\mathcal{E}_{\rho\eta}(D,N)\rho'\eta' + 2DC^{(i)}_{m\eta}(M,N)u'\eta'\big\}\,\mathrm{d}y\,\mathrm{d}z, \qquad (3.1.27)\end{aligned}$$

where $\mathbf{x}' \equiv (u', v', w', \rho', \eta')$. In Appendix 3.C, we show that $\mathcal{H}_L$ is conserved by the linear equations. In service of that proof and much of what follows, we compute the second partial derivatives of the functions $C^{(i)}(m,\eta)$:

$$\begin{aligned}C^{(i)}_{mm}(M,N) &= -\left[\frac{\partial}{\partial M}\left(M + \tfrac{1}{2}\beta Y^{(i)2} - \gamma Z^{(i)}\right)\right]_N \\ &= -1 - \beta Y^{(i)}\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N + \gamma\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N. \qquad (3.1.28)\end{aligned}$$

Using (3.B.4b), we can write the last expression in terms of derivatives of $N(y,z)$:

$$C^{(i)}_{mm}(M,N) = -1 + \frac{1}{DQ}\left(\beta y N_z + \gamma N_y\right). \qquad (3.1.29)$$

Similarly, using (3.B.4),

$$
\begin{aligned}
C^{(i)}_{m\eta}(M,N) &= -\left[\frac{\partial}{\partial N}\left(M+\tfrac{1}{2}\beta Y^{(i)2}-\gamma Z^{(i)}\right)\right]_M \\
&= -\beta Y^{(i)}\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M+\gamma\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M \qquad (3.1.30)\\
&= \frac{1}{DQ}\left(-\beta y M_z-\gamma M_y\right), \qquad (3.1.31)
\end{aligned}
$$

$$
\begin{aligned}
C^{(i)}_{\eta m}(M,N) &= -\left[\frac{\partial}{\partial M}\mathcal{E}_\eta(D,N)\right]_N \\
&= -\mathcal{E}_{\eta\rho}(D,N)\left[D_y\left(\frac{\partial Y^{(i)}}{\partial M}\right)_N+D_z\left(\frac{\partial Z^{(i)}}{\partial M}\right)_N\right] \qquad (3.1.32)\\
&= \frac{1}{DQ}\mathcal{E}_{\eta\rho}(D,N)\left(D_yN_z-D_zN_y\right), \qquad (3.1.33)
\end{aligned}
$$

$$
\begin{aligned}
C^{(i)}_{\eta\eta}(M,N) &= -\left[\frac{\partial}{\partial N}\mathcal{E}_\eta(D,N)\right]_M \\
&= -\mathcal{E}_{\eta\eta}(D,N) \\
&\quad -\mathcal{E}_{\eta\rho}(D,N)\left[D_y\left(\frac{\partial Y^{(i)}}{\partial N}\right)_M+D_z\left(\frac{\partial Z^{(i)}}{\partial N}\right)_M\right] \qquad (3.1.34)\\
&= -\mathcal{E}_{\eta\eta}(D,N)-\mathcal{E}_{\eta\rho}(D,N)\frac{1}{DQ}\left(D_yM_z-D_zM_y\right). \qquad (3.1.35)
\end{aligned}
$$

3.1.4 Conditions for linear stability

Notice that the integrands of the terms in $\mathcal{H}_L$ are quadratic with respect to the components of $\mathbf{x}'$. Rearranging $\mathcal{H}_L$, we write

$$
\begin{aligned}
\mathcal{H}_L(\mathbf{x}') = \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}}\Bigg\{ & DU_0^2\left[1+C^{(i)}_{mm}(M,N)\right]\left(\frac{u'}{U_0}\right)^2+Dv'^2+Dw'^2 \\
&+D_0^2\left[2\mathcal{E}_\rho(D,N)+D\mathcal{E}_{\rho\rho}(D,N)\right]\left(\frac{\rho'}{D_0}\right)^2 \\
&+DN_0^2\left[\mathcal{E}_{\eta\eta}(D,N)+C^{(i)}_{\eta\eta}(M,N)\right]\left(\frac{\eta'}{N_0}\right)^2 \\
&+2DD_0N_0\mathcal{E}_{\rho\eta}(D,N)\frac{\rho'\eta'}{D_0N_0}+2DU_0N_0C^{(i)}_{m\eta}(M,N)\frac{u'\eta'}{U_0N_0}\Bigg\}\,\mathrm{d}y\,\mathrm{d}z, \qquad (3.1.36)
\end{aligned}
$$

where U_0, D_0 and N_0 are arbitrary positive constants with the dimensions of velocity, density and entropy respectively. $\mathcal{H}_L$ itself will be strictly positive if the integrands are strictly positive, or equivalently, if (U_0, D_0, N_0) can be found such that the matrices

$$\Lambda^{(i)}(D,N,M) = \begin{bmatrix} \frac{D_0^2}{D^2}c_s^2(D,N) & D_0N_0\mathcal{E}_{\rho\eta}(D,N) & 0 \\ D_0N_0\mathcal{E}_{\rho\eta}(D,N) & N_0^2(\mathcal{E}_{\eta\eta}(D,N)+C_{\eta\eta}^{(i)}(M,N)) & U_0N_0C_{m\eta}^{(i)}(M,N) \\ 0 & U_0N_0C_{m\eta}^{(i)}(M,N) & U_0^2(1+C_{mm}^{(i)}(M,N)) \end{bmatrix} \tag{3.1.37}$$

are positive definite for all triples (D,N,M) which occur in the corresponding region $\mathcal{D}^{(i)}$.

The condition for all of the matrices to be positive definite is equivalent to the condition that all subdeterminants which include the top-left (or bottom-right) element be positive (see, e.g., Perlis, 1952, p. 103). Hence,

$$\frac{D_0^2}{D^2}c_s^2(D,N) > 0, \tag{3.1.38a}$$

$$\det\begin{bmatrix} \frac{D_0^2}{D^2}c_s^2(D,N) & D_0N_0\mathcal{E}_{\rho\eta}(D,N) \\ D_0N_0\mathcal{E}_{\rho\eta}(D,N) & N_0^2[\mathcal{E}_{\eta\eta}(D,N)+C_{\eta\eta}^{(i)}(M,N)] \end{bmatrix} > 0, \tag{3.1.38b}$$

$$\det\begin{bmatrix} \frac{D_0^2}{D^2}c_s^2(D,N) & D_0N_0\mathcal{E}_{\rho\eta}(D,N) & 0 \\ D_0N_0\mathcal{E}_{\rho\eta}(D,N) & N_0^2[\mathcal{E}_{\eta\eta}(D,N)+C_{\eta\eta}^{(i)}(M,N)] & U_0N_0C_{m\eta}^{(i)}(M,N) \\ 0 & U_0N_0C_{m\eta}^{(i)}(M,N) & U_0^2[1+C_{mm}^{(i)}(M,N)] \end{bmatrix} > 0. \tag{3.1.38c}$$

Before interpreting these conditions, we show that positive definiteness and boundedness of $\mathcal{H}_L$ implies the stability of $\mathbf{x}' = 0$ with respect to a suitably defined norm.

Under the hypothesis that the $\Lambda^{(i)}$ are symmetric, positive definite matrices, their eigenvalues $\lambda^{(i)}$ are all positive, and there exists a complete set of mutually orthogonal eigenvectors $\boldsymbol{\xi}_j$ $[j = 1, 2, 3]$, for each $\Lambda^{(i)}$. Any vector $\mathbf{x}'_1 \equiv (\rho'/D_0, \eta'/N_0, u'/U_0)$ can be written as a linear combination of the $\boldsymbol{\xi}_j$ (for any particular i). Thus

$$\mathbf{x}'_1 = \sum_{j=1}^{3} \alpha_j \boldsymbol{\xi}_j, \tag{3.1.39}$$

and

$$\mathbf{x}'^T_1 \Lambda^{(i)} \mathbf{x}'_1 = \sum_{j=1}^{3} \lambda_j^{(i)} {\alpha_j}^2 |\boldsymbol{\xi}_j|^2 \tag{3.1.40}$$

because the $\boldsymbol{\xi}_j$ are mutually orthogonal, and

$$\lambda_{\text{min}}^{(i)} |\mathbf{x}'_1|^2 \le \mathbf{x}'^T_1 \Lambda^{(i)} \mathbf{x}'_1 \le \lambda_{\text{max}}^{(i)} |\mathbf{x}'_1|^2, \tag{3.1.41}$$

where $\lambda_{\text{min}}^{(i)}$ and $\lambda_{\text{max}}^{(i)}$ are the smallest and largest of the eigenvalues of $\Lambda^{(i)}$. Note that the matrices vary with D, N, and M, and so do the eigenvalues. Let λ_- and λ_+ be the smallest and largest eigenvalues of the $\Lambda^{(i)}$ considering all triples (D, N, M) and all i. Define the norm $||\mathbf{x}'||_\lambda$ by

$$||\mathbf{x}'||_\lambda^2 = \iint_{\mathcal{D}} D \left\{ \lambda \left[\left(\frac{\rho'}{D_0} \right)^2 + \left(\frac{\eta'}{N_0} \right)^2 + \left(\frac{u'}{U_0} \right)^2 \right] + v'^2 + w'^2 \right\} \mathrm{d}y \, \mathrm{d}z, \tag{3.1.42}$$

and we have that

$$||\mathbf{x}'||_{\lambda_-}^2 \le \mathcal{H}_L(\mathbf{x}') \le ||\mathbf{x}'||_{\lambda_+}^2. \tag{3.1.43}$$

But this is true for all time, and since $\mathcal{H}_L$ is conserved in time, we have

$$||\mathbf{x}'(t)||_{\lambda_-}^2 \le \mathcal{H}_L[\mathbf{x}'(t)] = \mathcal{H}_L[\mathbf{x}'(0)] \le ||\mathbf{x}'(0)||_{\lambda_+}^2 \le \frac{\lambda_+}{\lambda_-} ||\mathbf{x}'(0)||_{\lambda_-}^2, \tag{3.1.44}$$

so that for any ε, if $||\mathbf{x}'(0)||_{\lambda_-} < \sqrt{\frac{\lambda_-}{\lambda_+}} \varepsilon$, then $||\mathbf{x}'(t)||_{\lambda_-} < \varepsilon$ for all times t.

Therefore, if the conditions (3.1.38) are satisfied, and all coefficients are bounded, then $\mathbf{x}' = 0$ is stable, and we say that the solution to the full equations $\mathbf{X}$ is *linearly stable*.

We now turn to the physical interpretation of the stability conditions. (3.1.38a) is satisfied by any gas. For example, for an ideal gas with equation of state

$$p = R\rho\tau, \tag{3.1.45}$$

where R is a constant, it can be readily shown that $c_s^2 = (c_p/c_v)R\tau > 0$, where c_p and c_v are the specific heat capacities of the gas at constant pressure and volume respectively (see Appendix 3.A).

Using (3.A.3), (3.1.35) and (3.1.9), condition (3.1.38b) can be written

$$\frac{1}{D^3 Q}\left(\frac{\partial P}{\partial N}\right)_D \partial(M,P) > 0, \tag{3.1.46}$$

and condition (3.1.38c) can be written

$$\frac{g\beta}{D}\left(\frac{\partial P}{\partial N}\right)_D \frac{y}{Q} > 0. \tag{3.1.47}$$

To interpret these conditions, we need to know the sign of $(\partial P/\partial N)_D$. Using (3.A.4), (3.A.7b), and (3.A.10),

$$\left(\frac{\partial P}{\partial N}\right)_D = \left(\frac{\partial T}{\partial N}\right)_D \left(\frac{\partial P}{\partial T}\right)_D = \frac{T}{c_v}\left(\frac{\partial P}{\partial T}\right)_D . \tag{3.1.48}$$

where the heat capacity c_v need not be constant, but is surely positive. The derivative $(\partial P/\partial T)_D$ is obviously positive — pressure increases with increasing temperature for fixed density. Hence, $(\partial P/\partial N)_D > 0$.

(3.1.47) is therefore the well known symmetric stability condition that potential vorticity be positive in the northern hemisphere and negative in the southern (Stevens, 1983). Assuming that Q is continuous across the equator, this implies that $Q = 0$ on the equator. Equation (3.1.46) may then be interpreted as a generalization of the Rayleigh criterion for *inertial stability*: the angular momentum gradient must be clockwise of the pressure gradient in the northern hemisphere, and anticlockwise of the pressure gradient in the southern, when viewed with the northern hemisphere on the right (Figure 3.2). For example, if the pressure gradient is directly downwards, then ∇M must be towards the equator for stability. The shape of the pressure contours depends on the velocity field U. From (3.1.9),

$$\frac{P_z}{P_y} = \frac{g - \gamma U}{\beta y U}, \tag{3.1.49}$$

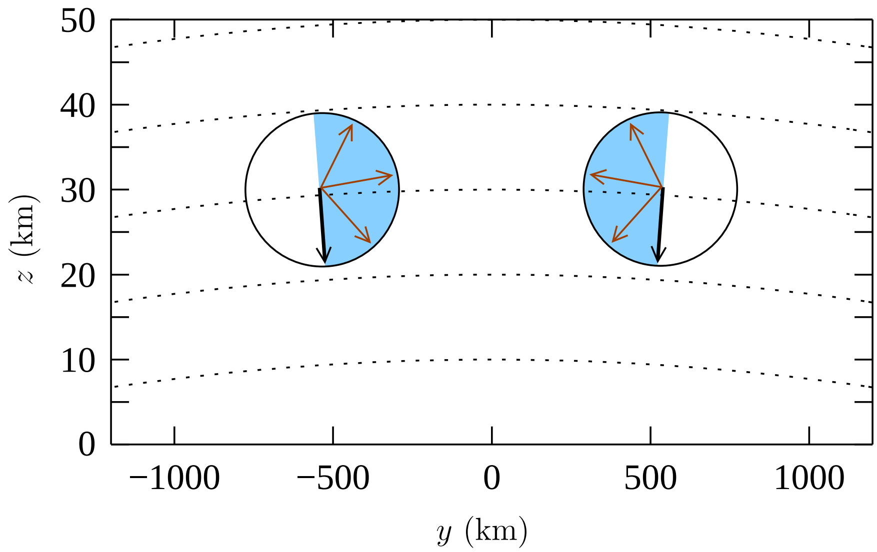


Figure 3.2: The dotted curves are contours of constant pressure, with the curvature exaggerated by a factor of 10 compared to a typical atmospheric state. The pressure contours are concave down, implying $U > 0$. The thick black arrow is the gradient of pressure. ∇M must be in the shaded semicircles for stability.

so the pressure contours are concave down for eastward flow ($U > 0$) and concave up for westward flow ($U < 0$). The effect of γ is to steepen the pressure contours for eastward flow and flatten them for westward flow (we have assumed that $g > \gamma|U|$). See Figure 3.2. Note, however, that the effect of the curved pressure contours on the stability conditions is small for scales relevant to the equatorial stratosphere. Choosing constant $U = 40$ ms^{-1}, and $y = 2000$ km, $P_y/P_z \approx 10^{-4}$, and pressure contours rise or sink by only about 0.1 km over the whole equatorial region. For extreme values of U, the pressure contours become severely curved, and as we will see in an example below, when γ is included in the equations, this affects the sizes of regions of instability.

(3.1.46) and (3.1.47) are formally the same stability conditions as in Bowman and Shepherd (1995) (where the hydrostatic approximation is used), namely $yQ > 0$ and $y(\partial M/\partial y)|_p < 0$, but of course, the definitions of M and Q are different because of the difference in the underlying physics.

If the $\Lambda^{(i)}$ are positive definite, then it also follows that

$$1 + C^{(i)}_{mm}(M, N) > 0, \tag{3.1.50}$$

which is equivalent to

$$\frac{1}{DQ}\left(\beta y N_z + \gamma N_y\right) = \frac{1}{DQ}\partial(N, M^{(\mathrm{p})}) > 0, \tag{3.1.51}$$

which may be interpreted as follows: the component of the entropy gradient in the direction of the local planetary rotation vector $\mathbf{\Omega} = \frac{1}{2}(\gamma\hat{\mathbf{e}}_y + \beta y\hat{\mathbf{e}}_z)$ must be positive in the northern hemisphere, and negative in the southern. Equivalently, the entropy gradient must be clockwise of the planetary angular momentum gradient in the northern hemisphere, and anticlockwise in the southern; see Figure 3.3. The difference between the hemispheres is due to the coordinate system (exactly the same result would obtain if the rotation of the planet were reversed). This is a generalization of static stability since the "static" state in the rotating frame is determined by the planetary rotation parameters β and γ. For example, if the γ term is neglected (as in the traditional hydrostatic primitive equations), then $\nabla M^{(\mathrm{p})}$ is towards the equator, and (3.1.51) reduces to $N_z > 0$. The effect of γ is to create the slope in the planetary angular momentum contours (see Figure 2.1). This effect of including the $\cos\phi$ terms was discussed by Sun (1994), who explained that the vertical gradient of entropy competes with the centrifugal effect of vertical shear in velocity which is linked with N_y through thermal wind balance. Note that the potential vorticity condition (3.1.47) and either of (3.1.46) and (3.1.51) implies the remaining condition. However, it is possible for a state to be inertially and statically stable but to violate the potential vorticity condition. Such a situation is depicted in Figure 3.4. See also Emanuel (1983) for further dicussion.

Recall that the coefficients in $\mathcal{H}_L$ have to be bounded to ensure that the matrices $\Lambda^{(i)}$ have a maximum eigenvalue. This restricts the functional form of $M(y, z)$ and $N(y, z)$ in the limit as $y \to 0$. This is discussed briefly in Example 4 in Section 3.1.6 below.

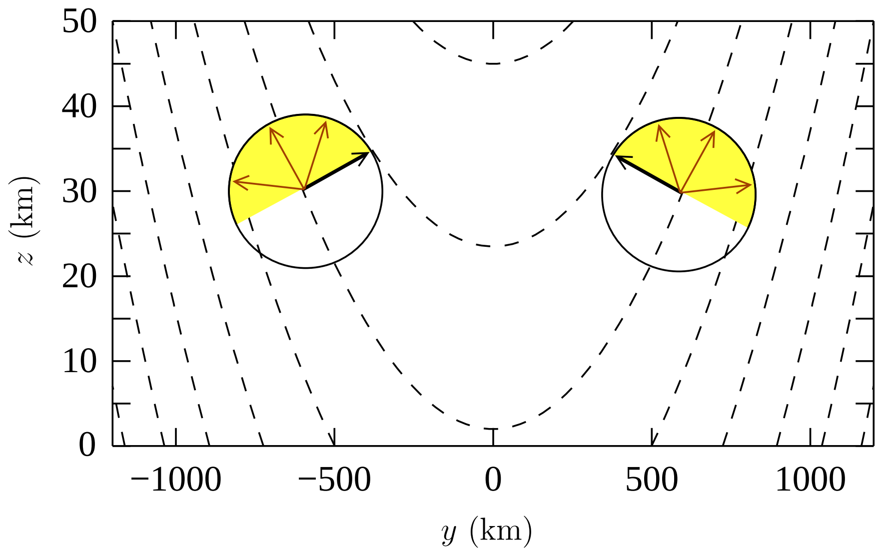


Figure 3.3: The dashed curves are contours of constant planetary angular momentum. The thick black arrow is the gradient of planetary angular momentum. ∇N must be in the shaded regions for stability.

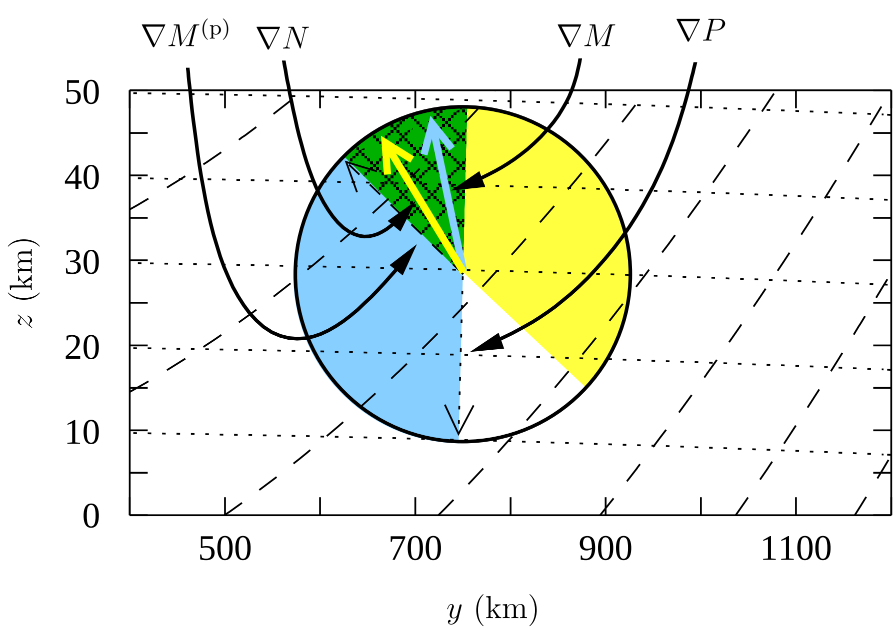


Figure 3.4: An example of a basic state which satisfies the inertial and static stability conditions but fails the "symmetric" stability condition. The shaded regions are as in Figures 3.2 and 3.3. If they fall in the hatched region, it is possible to have both ∇N and ∇M satisfy their respective stability criteria, but be such that Q has the wrong sign.

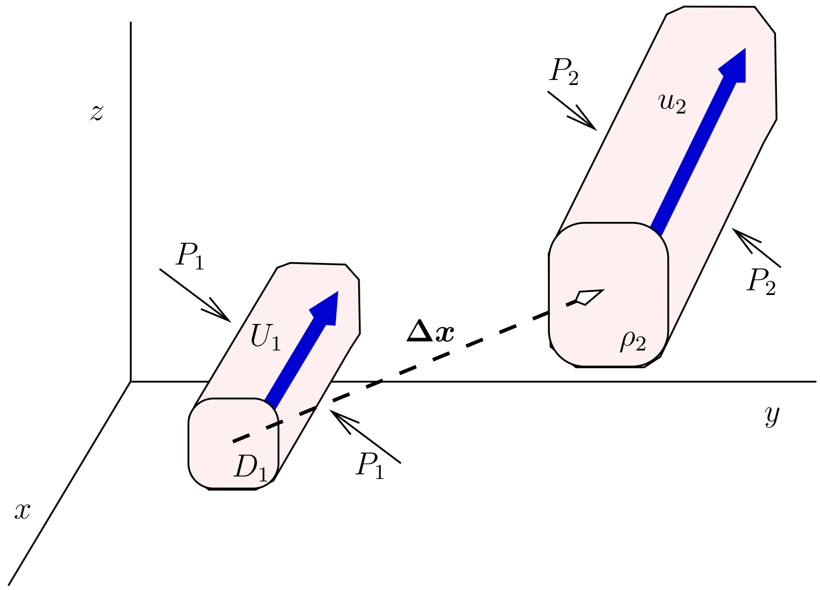


Figure 3.5: A parcel of fluid (a ring of fluid on the sphere, a tube on the β-plane) is displaced from its equilibrium position. It is assumed that the disturbance does not affect the background pressure field. The density and zonal velocity of the ring change in such a way that entropy and angular momentum are conserved.

3.1.5 A parcel displacement derivation of symmetric stability conditions

Condition (3.1.47) may also be derived by calculating the acceleration of a thin ring of fluid which is displaced from its equilibrium position. Let us call an equilibrium stable if the component of acceleration in the direction of the displacement is in the opposite direction to the displacement for all possible infinitesimal displacements. This is a generalization of the arguments applied separately to inertial and static stability in Section 1.1.

Consider again an equilibrium with $u = U$ $(m = M)$, $\rho = D$, and $\eta = N$ $(p = P)$. Suppose a ring of fluid at (y_1, z_1) is displaced to position (y_2, z_2), conserving its angular momentum and entropy, and that the ring is thin enough so that the pressure field is not disturbed by the displacement (Figure 3.5).

The acceleration of the parcel at its new position is

$$\left.\frac{Dv}{Dt}\right|_2 = -\beta y_2 u_2 - \frac{1}{\rho_2}\left.\frac{\partial P}{\partial y}\right|_2, \tag{3.1.52a}$$

$$\left.\frac{Dw}{Dt}\right|_2 = \gamma u_2 - g - \frac{1}{\rho_2}\left.\frac{\partial P}{\partial z}\right|_2, \tag{3.1.52b}$$

where the subscripts indicate at which point the function is evaluated. From (3.1.9),

$$-\left.\frac{\partial P}{\partial y}\right|_2 = \beta y_2 D_2 U_2, \tag{3.1.53a}$$

$$-\left.\frac{\partial P}{\partial z}\right|_2 = -D_2(\gamma U_2 - g). \tag{3.1.53b}$$

Therefore,

$$\rho_2 \left.\frac{Dv}{Dt}\right|_2 = -\beta y_2(\rho_2 u_2 - D_2 U_2), \tag{3.1.54a}$$

$$\rho_2 \left.\frac{Dw}{Dt}\right|_2 = \rho_2(\gamma u_2 - g) - D_2(\gamma U_2 - g). \tag{3.1.54b}$$

Since the parcel conserves its entropy, its density at (y_2, z_2) is, to first order

$$\begin{aligned}
\rho_2 &\approx D_1 + \left.\left(\frac{\partial \rho}{\partial p}\right)_\eta\right|_{(P_2,\eta_2)} (P_2 - P_1) + \left.\left(\frac{\partial \rho}{\partial \eta}\right)_p\right|_{(P_2,\eta_2)} (\eta_2 - N_1) \\
&\approx D_1 + \left[\left.\left(\frac{\partial \rho}{\partial p}\right)_\eta\right|_{(P_2,N_2)} + \left.\frac{\partial^2 \rho}{\partial p \partial \eta}\right|_{(P_2,N_2)} (N_1 - N_2)\right](P_2 - P_1) \\
&\approx D_1 + \left.\left(\frac{\partial \rho}{\partial p}\right)_\eta\right|_{(P_2,N_2)} \left[\left.\left(\frac{\partial p}{\partial \rho}\right)_\eta\right|_{(P_2,N_2)} (D_2 - D_1) + \left.\left(\frac{\partial p}{\partial \eta}\right)_\rho\right|_{(P_2,N_2)} (N_2 - N_1)\right] \\
&= D_2 - \left.\left(\frac{\partial \rho}{\partial \eta}\right)_p\right|_{(P_2,N_2)} (N_2 - N_1).
\end{aligned} \tag{3.1.55}$$

In going from the first line to the second, we have used that $\eta_2 = N_1$. To conform to the notation of the earlier sections, we define

$$\left.\left(\frac{\partial \rho}{\partial \eta}\right)_p\right|_{(P_2,N_2)} \equiv \left(\frac{\partial D}{\partial N}\right)_P. \tag{3.1.56}$$

Similarly, since it conserves angular momentum,

$$u_2 = U_2 - (M_2 - M_1). \tag{3.1.57}$$

Therefore, to first order in displacement quantities,

$$\begin{aligned} \rho_2 \left.\frac{Dv}{Dt}\right|_2 &= \beta y_2 \left[U_2 \left(\frac{\partial D}{\partial N}\right)_P \nabla N|_2 + D_2 \, \nabla M|_2 \right] \cdot \boldsymbol{\Delta x} && (3.1.58a) \\ \rho_2 \left.\frac{Dw}{Dt}\right|_2 &= \left[(g - \gamma U_2) \left(\frac{\partial D}{\partial N}\right)_P \nabla N|_2 - \gamma D_2 \, \nabla M|_2 \right] \cdot \boldsymbol{\Delta x}, && (3.1.58b) \end{aligned}$$

where $\boldsymbol{\Delta x} \equiv (y_2 - y_1, z_2 - z_1)$. We may write (3.1.58) more suggestively as

$$\rho_2 \begin{bmatrix} \left.\dfrac{Dv}{Dt}\right|_2 \\ \left.\dfrac{Dw}{Dt}\right|_2 \end{bmatrix} = S\boldsymbol{\Delta x} \equiv \begin{bmatrix} \beta y_2 U_2 \left(\dfrac{\partial D}{\partial N}\right)_P \nabla N|_2 + \beta y_2 D_2 \, \nabla M|_2 \\ (g - \gamma U_2) \left(\dfrac{\partial D}{\partial N}\right)_P \nabla N|_2 - \gamma D_2 \, \nabla M|_2 \end{bmatrix} \boldsymbol{\Delta x}, \tag{3.1.59}$$

where the rows of the matrix S have been written as vectors. The projection of the acceleration onto the displacement is then

$$\rho_2 \boldsymbol{\Delta x}^T \begin{bmatrix} \left.\dfrac{Dv}{Dt}\right|_2 \\ \left.\dfrac{Dw}{Dt}\right|_2 \end{bmatrix} = \boldsymbol{\Delta x}^T S \boldsymbol{\Delta x}. \tag{3.1.60}$$

For the component of the acceleration in the direction of the displacement to be negative (i.e. for the parcel to feel a restoring force) for all possible displacements, the matrix S must be *negative* definite. This is equivalent to the conditions

$$\begin{aligned} \det S &> 0, && (3.1.61a) \\ \operatorname{trace} S &< 0. && (3.1.61b) \end{aligned}$$

Now,

$$\det S = -g\beta y_2 D_2 \left(\frac{\partial D}{\partial N}\right)_P \partial(N, M)|_2 \, , \tag{3.1.62}$$

and

$$\operatorname{trace} S = -\frac{1}{D_2} \left(\frac{\partial D}{\partial N}\right)_P \nabla P|_2 \cdot \nabla N|_2 - D_2 \, \nabla M^{(\mathrm{p})}\big|_2 \cdot \nabla M|_2 \tag{3.1.63}$$

We now show that the conditions (3.1.62) and (3.1.63) are in fact identical to the conditions derived using the energy-Casimir method.

Firstly, since

$$\left(\frac{\partial D}{\partial N}\right)_P = -\left(\frac{\partial D}{\partial P}\right)_N \left(\frac{\partial P}{\partial N}\right)_D = -\frac{1}{c_s^2}\left(\frac{\partial P}{\partial N}\right)_D < 0 \tag{3.1.64}$$

(see (3.1.48)), $\det S > 0$ is equivalent to (3.1.47). Next, we rewrite the conditions in terms of the angles that the various gradient vectors make with each other. Let θ_M be the angle from $\nabla M^{(\mathrm{p})}$ to ∇M and let θ_N be the angle from ∇P to ∇N. The thermal wind balance condition (3.1.10) can then be written

$$|\nabla M^{(\mathrm{p})}||\nabla M|\sin\theta_M = -\frac{1}{D^2}\left(\frac{\partial D}{\partial N}\right)_P |\nabla P||\nabla N|\sin\theta_N. \tag{3.1.65}$$

This implies that θ_M and θ_N have the same sign at equilibrium. (3.1.63) can be rewritten (dropping the "2" subscripts):

$$\operatorname{trace} S = -\frac{1}{D}\left(\frac{\partial D}{\partial N}\right)_P |\nabla P||\nabla N|\cos\theta_N - D|\nabla M^{(\mathrm{p})}||\nabla M|\cos\theta_M < 0. \tag{3.1.66}$$

Combining (3.1.65) and (3.1.66) gives

$$\operatorname{trace} S = D|\nabla M^{(\mathrm{p})}||\nabla M|\frac{\sin(\theta_M - \theta_N)}{\sin(\theta_N)} < 0. \tag{3.1.67}$$

We now determine the possible pairs of angles for which (3.1.62), (3.1.65), and (3.1.67) are satisfied. The three conditions can be summarized:

$$\sin\theta_M \sin\theta_N > 0, \tag{3.1.68a}$$

$$(\theta_N + \theta_0) - \theta_M > 180^\circ, \tag{3.1.68b}$$

$$|\theta_M| < |\theta_N|, \tag{3.1.68c}$$

where θ_0 is the angle from $\nabla M^{(\mathrm{p})}$ to ∇P, and θ_N, θ_M and θ_0 are taken to be in $[-180^\circ, 180^\circ]$. Assuming $y > 0$ (the corresponding result for $y < 0$ is obvious from symmetry), (3.1.68b) is the statement that ∇N must be "clockwise" of ∇M (equivalent to $\det S > 0$).

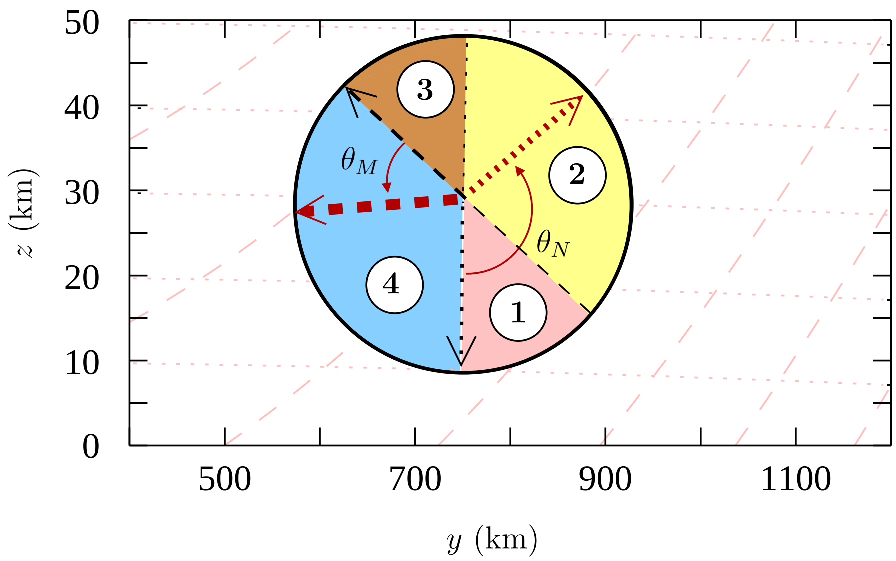


Figure 3.6: Division of circle $[-180^o, 180^o]$ into four intervals determined by ∇P and $\nabla M^{(\mathrm{p})}$. θ_M measures from $\nabla M^{(\mathrm{p})}$ to ∇M, and θ_N from ∇P to ∇N. It is shown that for stability, θ_N must be in interval **2** or **3**, with θ_M in **4** or **3**, respectively.

In Figure 3.6, the circle $[-180°, 180°]$ has been divided into four intervals divided by ∇P and $\nabla M^{(\mathrm{p})}$.

If θ_N is in interval **1**, (3.1.68a) and (3.1.68b) cannot both be satisfied.

If ∇N is parallel to $-\nabla M^{(\mathrm{p})}$, then the only direction of (marginal) stability for ∇M is parallel to $\nabla M^{(\mathrm{p})}$. As θ_N increases into interval **2**, θ_M can be in interval **4**, provided (3.1.68b) is satisfied.

If θ_N is in interval **3**, then θ_M must also be in interval **3**, provided it satisfies (3.1.68b).

If θ_N is in interval **4**, then no value of θ_M can satisfy all three of the conditions — as θ_N decreases from 0, the maximum θ_M satisfying (3.1.68b) decreases from $\theta_0 - 180°$, but then $|\theta_M| > |\theta_N|$, violating (3.1.68c).

Therefore, the conditions derived by the parcel method are identical with the conditions derived above with the energy-Casimir method. The two derivations are complementary in that the energy-Casimir method is based on a rigorous definition of stability, while the parcel method implies a physical mechanism for stability (and instability when the conditions are violated).

3.1.6 Examples

In each of the following examples, we assume that the atmosphere is an ideal gas. A useful identity that will be used frequently in calculating Jacobians is

$$\nabla N = -\frac{R}{P}\nabla P + \frac{c_p}{T}\nabla T. \tag{3.1.69}$$

The identity may be derived using (3.A.3b), (3.A.12), (3.A.5) and (3.A.13).

Example 1: Isothermal atmosphere in solid-body rotation

First, we show that an isothermal atmosphere in solid body rotation is linearly stable. The nonlinear stability of the solid-body rotation state in the (viscous) Couette-Taylor problem is a well known result (see Joseph, 1976). Consider an atmosphere with temperature $T = T_{00}$ = constant, in solid-body rotation. In the equatorial β-plane system, we mean by this that the relative velocity is uniform, i.e. $U = U_{00}$.

By the ideal gas equation of state (3.A.8) and the conditions for balance (3.1.9), we have

$$\frac{\partial}{\partial y}\ln P = -\left(\frac{\beta U_{00}}{RT_{00}}\right)y, \tag{3.1.70a}$$

$$\frac{\partial}{\partial z}\ln P = -\left(\frac{g-\gamma U_{00}}{RT_{00}}\right), \tag{3.1.70b}$$

from which we can solve for the basic state pressure field:

$$P(y,z) = P_{00}\exp\left[-\frac{1}{2}\left(\frac{\beta U_{00}}{RT_{00}}\right)y^2 - \left(\frac{g-\gamma U_{00}}{RT_{00}}\right)z\right], \tag{3.1.71}$$

where P_{00} is the pressure at the origin. Since $\nabla T = 0$, we have from (3.1.69) that

$$\nabla N = -\frac{1}{T_{00}D}\nabla P. \tag{3.1.72}$$

The potential vorticity satsifies

$$DQ = \partial(N, M) = -\frac{1}{T_{00}D}\partial(P, M^{(\mathrm{p})}) = \left(\frac{\beta g}{T_{00}}\right) y. \tag{3.1.73}$$

To compute the coefficient matrix Λ (note that the basic state is an even function of y, so there is only one function $C(m, \eta)$ required), we need (3.A.13), (3.A.12), and, from (3.1.29)-(3.1.35),

$$\begin{aligned} \mathcal{E}_{\eta\eta}(D, N) + C_{\eta\eta}(M, N) &= -\frac{1}{DQ}\mathcal{E}_{\rho\eta}\partial(D, M^{(\mathrm{p})}) = \frac{T_{00}}{c_v}, && (3.1.74\mathrm{a}) \\ C_{m\eta}(M, N) &= \frac{1}{DQ}\mathcal{E}_{\rho\eta}\partial(D, N) = 0, && (3.1.74\mathrm{b}) \\ 1 + C_{mm}(M, N) &= \frac{1}{DQ}\partial(N, M^{(\mathrm{p})}) = 1. && (3.1.74\mathrm{c}) \end{aligned}$$

Hence,

$$\Lambda = \begin{bmatrix} D_0^2 \dfrac{c_p}{c_v}\dfrac{RT_{00}}{D^2} & D_0 N_0 \dfrac{1}{c_v}\dfrac{RT_{00}}{D} & 0 \\ D_0 N_0 \dfrac{1}{c_v}\dfrac{RT_{00}}{D} & N_0^2 \dfrac{T_{00}}{c_v} & 0 \\ 0 & 0 & U_0^2 \end{bmatrix}. \tag{3.1.75}$$

Choosing[1]

$$N_0 = \sqrt{Rc_v}, \quad U_0 = \sqrt{RT_{00}}, \quad D_0 = \sqrt{\frac{c_v}{R}} D_{00}, \tag{3.1.76}$$

where $D_{00} = P_{00}/RT_{00}$, we get

$$\Lambda = RT_{00} \begin{bmatrix} \dfrac{c_p}{R}\left(\dfrac{D_{00}}{D}\right)^2 & \dfrac{D_{00}}{D} & 0 \\ \dfrac{D_{00}}{D} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{3.1.77}$$

The eigenvalues of Λ are $\lambda_3 = RT_{00}$ and

$$\lambda_{\frac{1}{2}} = RT_{00} \times \frac{1}{2}\left[(ax^2 + 1) \pm \sqrt{(ax^2 - 1)^2 + 4x^2}\right], \tag{3.1.78}$$

[1] Recall that the constants D_0, N_0 and U_0 are arbitrary. They are there to make the dimensions of the elements in Λ uniform.

where $a = c_p/R = 7/2$ and $x = D_{00}/D$. By the triangle inequality,

$$\lambda_1 \leq RT_{00}(ax^2 + x), \tag{3.1.79a}$$

$$\lambda_2 \geq RT_{00}(1 + x). \tag{3.1.79b}$$

The minimum value of x in the domain occurs at the origin and it is unity. A smaller upper bound on λ_1 is the trace of the upper-left principal subdeterminant of Λ,

$$\lambda_1 < RT_{00} \text{ trace} \begin{bmatrix} ax^2 & x \\ x & 1 \end{bmatrix} = RT_{00}(1 + ax^2). \tag{3.1.80}$$

We therefore have that the minimum and maximum eigenvalues are bounded by

$$\lambda_- = RT_{00}, \tag{3.1.81a}$$

$$\lambda_+ = RT_{00} \left\{ 1 + \frac{c_p}{R} \exp\left[\left(\frac{\beta U_{00}}{RT_{00}} \right) L^2 + 2 \left(\frac{g - \gamma U_{00}}{RT_{00}} \right) H \right] \right\}. \tag{3.1.81b}$$

We have shown that this basic state is linearly stable by explicitly calculating upper and lower positive bounds on the conserved functional $\mathcal{H}_L$. Depending on the height of the domain, the potential linear amplification factor $\sqrt{\lambda_+/\lambda_-}$ may be very large. For example, if the domain height H is one density scale-height, i.e. $H = RT_{00}/g$, then $\sqrt{\lambda_+/\lambda_-} \approx 5$, but if the basic state is colder, such that, say $H = 5RT_{00}/g$, then $\sqrt{\lambda_+/\lambda_-} \approx 300$. The other terms in (3.1.81b) are much smaller for realistic values of L, U_{00} and T_{00}. In particular, the γ correction is significant only if $U_{00} \sim g/\gamma \approx 10^5$ ms^{-1} (for Earth), which is orders of magnitude larger than realistic values.

Example 2: Linear velocity shear at equator: $U = \lambda y$ with $T(y, 0) = T_{00}$

In the next example we show that constant meridional velocity shear at the equator is inertially unstable. This is the case considered by Dunkerton (1981).

Consider an equilibrium with velocity $U(y, z) = \lambda y$, where $\lambda < g/\gamma L$ is a constant ($\lambda > g/\gamma L$ would imply negative temperatures in the domain; for Earth values of g and γ, and $L \approx 10^3$ km, $g/\gamma L \sim 100$ (ms^{-1})km^{-1}).

To proceed, we must specify pressure and density fields that are in balance with U, i.e. which satisfy (3.1.9). Eliminating density D between (3.1.9a) and (3.1.9b), we find

$$\left(1-\frac{\gamma\lambda}{g}y\right)P_y+\left(-\frac{\beta\lambda}{g}y^2\right)P_z=0. \tag{3.1.82}$$

We solve for $P(y,z)$ using the method of characteristics (see, e.g., Zauderer, 1989). Families of curves of constant P have parametric equations

$$y(s) = \frac{g}{\gamma\lambda}\left[1-c_y\exp\left(-\frac{\gamma\lambda}{g}s\right)\right], \tag{3.1.83a}$$

$$z(s) = c_z-\left(\frac{\beta\lambda}{g}\right)\left(\frac{g}{\gamma\lambda}\right)^2\left[s+2\frac{g}{\gamma\lambda}c_y\exp\left(-\frac{\gamma\lambda}{g}s\right)-\frac{1}{2}\frac{g}{\gamma\lambda}c_y^2\exp\left(-2\frac{\gamma\lambda}{g}s\right)\right], \tag{3.1.83b}$$

where s is a parameter, and c_y and c_z are integration constants. We fix the constants by specifying an initial curve (i.e. a curve in (y,z,P) through which all of the isobars pass at $s=0$). We choose (somewhat arbitrarily) the pressure over the equator from Example 1 as the initial curve. It has parametric equations

$$y(0,r) = 0, \tag{3.1.84a}$$

$$z(0,r) = r, \tag{3.1.84b}$$

$$P(0,r) = P_{00}\exp\left(-\frac{g}{RT_{00}}r\right), \tag{3.1.84c}$$

where r parameterizes the initial curve. Substituting $s=0$ into (3.1.83) and using (3.1.84), we find

$$c_y=1,\quad c_z=r+\frac{3}{2}\frac{\beta g^2}{\gamma^3\lambda^2}. \tag{3.1.85}$$

Solving for r in terms of y and z, and substituting into (3.1.84c) gives

$$P(y,z) = P_{00}\left(1-\frac{\gamma\lambda}{g}y\right)^{\left(\frac{g}{RT_{00}}\right)\left(\frac{\beta}{\gamma}\right)\left(\frac{g}{\gamma\lambda}\right)^2} \times\exp\left\{-\frac{g}{RT_{00}}\left[z-\frac{\beta}{\gamma}\left(\frac{g}{\gamma\lambda}y+\frac{1}{2}y^2\right)\right]\right\}. \tag{3.1.86}$$

The corresponding density and pressure fields are

$$D(y,z) = \frac{P}{RT_{00}} \frac{1}{\left(1 - \frac{\gamma\lambda}{g} y\right)}, \tag{3.1.87a}$$

$$T(y,z) = T_{00}\left(1 - \frac{\gamma\lambda}{g} y\right). \tag{3.1.87b}$$

Notice that there is a very slight meridional temperature gradient to balance the velocity shear (a physically reasonable value of λ would be $\lesssim 10^{-4}$ s^{-1}, making $\gamma\lambda/g \lesssim 10^{-9}$, a temperature gradient of $0.001\,T_{00}$ per 1000 km). This simplest of solutions could also have been derived by requiring that $T(y,z) = T(y)$, and rearranging (3.1.10) to get

$$\frac{\mathrm{d}}{\mathrm{d}y}[\ln T(y)] = -\frac{\mathrm{d}U/\mathrm{d}y}{g/\gamma - U(y)}, \tag{3.1.88}$$

but the method of characteristics is necessary if $T_z \neq 0$ or $U_z \neq 0$ (see Examples 5 and 6).

We confirm that $P(y,z)$ and $D(y,z)$ are in thermal wind balance by computing

$$\partial(M, M^{(\mathrm{p})}) = \frac{1}{D^2}\partial(P,D) = \gamma\lambda, \tag{3.1.89}$$

which agrees with (3.1.10). In Dunkerton (1981), it is shown that this basic state velocity field is linearly unstable (in a hydrostatic model), with perturbations leading to the formation of Taylor vortices in the latitude interval $0 < y < \lambda/\beta$ and the associated changes to the m and η fields. The potential vorticity is

$$Q = \frac{\beta g}{DT}\left[\left(1 + \frac{\gamma\lambda^2}{\beta g}\right) y - \frac{\lambda}{\beta}\left(1 + \frac{\gamma^2}{g^2} c_p T_{00}\right)\right], \tag{3.1.90}$$

where (3.1.9) and several relations from Appendix 3.A have been used. For the γ correction to the interval of anomalous potential vorticity to be significant, we would need $\lambda \sim \sqrt{g/a}$. For Earth, this is approximately 1 (ms^{-1})km^{-1}, which is possible, although not over a wide latitude interval (it would imply unrealistically large velocities). The

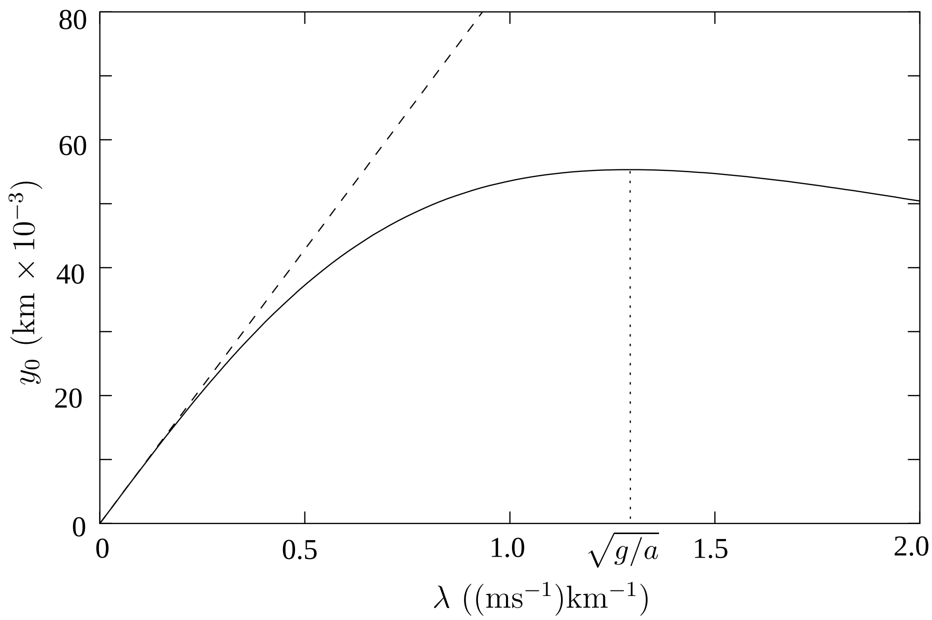


Figure 3.7: Interval of instability y_0 as a function of meridional velocity shear λ for the Dunkerton inertial instability problem. The dashed line is the prediction of the hydrostatic calculation. The inclusion of the γ term causes the interval of instability to grow more slowly with increased shear, eventually tending to zero as $\lambda \to \infty$.

latitude y_0 at which $Q = 0$ (the width of the unstable latitude interval) is

$$y_0 = \frac{\lambda}{\beta}\left(\frac{1 + \dfrac{\gamma^2}{g^2} c_p T_{00}}{1 + \dfrac{\gamma\lambda^2}{\beta g}}\right), \tag{3.1.91}$$

which has its maximum value ($\approx \lambda/2\beta$) for $\lambda = \sqrt{g/a} \approx \lambda/2\beta$ (see Figure 3.7). For $\lambda \ll \sqrt{g/a}$, $y_0 \approx \lambda/\beta$, and for $\lambda \gg \sqrt{g/a}$, $y_0 \to 0$ like λ^{-1}.

This effect is due to the increasing curvature of the pressure contours with increasing velocity. At the equator, ∇P points vertically down, so the clockwise edge of the "semicircle of stability" (see Figure 3.2) is vertically up. ∇M points clockwise of vertically up at the equator (for $\lambda > 0$), and turns towards the vertical linearly with y. Meanwhile, the semicircle of stability rotates clockwise (towards ∇M) quadratically with y. ∇M enters the semicircle of stability at y_0. Note that the effect of the curved pressure surfaces is not due to the *shear* in the basic state. An effect of comparable order is achieved by adding a constant U_{00} to the basic state velocity, where $U_{00} \sim \lambda L$.

Taking a realizable Earth value of $\lambda \approx 10^{-2}$ (ms^{-1})km^{-1}, we find that $DQ < 0$ in $0 < y \lesssim \lambda/\beta$, with a very small correction due to γ.

Example 3: Velocity profile $U = \frac{1}{2}\beta' y^2$ with $T(y,z) = T(y)$

Now consider an equilibrium with $U(y,z) = \frac{1}{2}\beta' y^2$, where $\beta' < \frac{1}{2}(g/\gamma L^2)$ is a constant, and temperature depending only on y. Without the effect of the γ terms, this state would be stable if $\beta' < \beta$. Using (3.1.88), we calculate fields of temperature,

$$T(y) = T_{00}\left(1 - \frac{\beta'\gamma}{2g}y^2\right), \tag{3.1.92}$$

and pressure,

$$P(y,z) = P_{00}\left(1 - \frac{\beta'\gamma}{2g}y^2\right)^{\left(\frac{g}{RT_{00}}\right)\left(\frac{\beta g}{\beta'\gamma^2}\right)} \exp\left[-\frac{g}{RT_{00}}\left(z - \frac{\beta}{2\gamma}y^2\right)\right], \tag{3.1.93}$$

which balance the velocity. The potential vorticity is

$$Q = \frac{\beta g}{DT}\left\{\left[1 - \frac{\beta'}{\beta}\left(1 - \frac{\gamma^2}{g^2}c_p T_{00}\right)\right]y + \frac{1}{2}\left(\frac{\beta'}{\beta}\right)\left(\frac{\gamma}{g}\right)\beta' y^3\right\}. \tag{3.1.94}$$

The coefficients of y and y^3 are both positive provided $\beta' \leq (1 + \gamma^2 c_p T_{00}/g^2)^{-1}\beta$. If not, then Q will have the wrong sign near the equator in both hemispheres, and the equilibrium is not linearly stable. The γ correction $\gamma^2 c_p T_{00}/g^2$ is very small for typical Earth values of the parameters. For example, for $c_p \approx 10^3$ JK^{-1}, $T_{00} \approx 300$ K, $\gamma \approx 1.4 \times 10^{-4}$ s^{-1} and $g \approx 10$ ms^{-2}, the correction is $6 \times 10^{-5} \ll 1$. A notable consequence of the inclusion of γ is that a flat Q profile ($\beta' \equiv 0$) is unstable at the equator.

Now consider the other two subdeterminants of the coefficient matrix. The inertial stability discriminant (3.1.46) is

$$\frac{1}{D^3 Q}\left(\frac{\partial P}{\partial N}\right)_D \partial(M,P) = \frac{PT}{c_v D}\left\{1 + \frac{(\beta'\gamma^2/g)c_p T_{00} y}{\left[g(\beta - \beta') - (\beta'\gamma^2/g)c_p T_{00}\right]y + \frac{1}{2}\gamma\beta'^2 y^3}\right\}, \tag{3.1.95}$$

and the static stability discriminant (3.1.51) is

$$\frac{1}{DQ}\partial(N, M^{(\mathrm{p})}) = 1 + \frac{g\beta' y - \frac{1}{2}\gamma\beta'^2 y^3}{\left[g(\beta-\beta') - (\beta'\gamma^2/g)c_p T_{00}\right] y + \frac{1}{2}\gamma\beta'^2 y^3}, \tag{3.1.96}$$

which are both positive if $\beta' < (1+\gamma^2 c_p T_{00}/g^2)^{-1}\beta$, and finite as y approaches zero.[1]

Example 4: Velocity profile $U = \frac{1}{2}\beta y^2 - \alpha|y|^k$ with $T(y,z) = T(y)$

The following example has an angular momentum profile with a power law dependency on y different from y^2. This might arise if an adjustment process has flattened the angular momentum across the equator such that, for example, $U = \frac{1}{2}\beta y^2$, and then the angular momentum is further perturbed towards stability, so that M decreases away from the equator but not in a quadratic way.

We now show (albeit only in the case of temperature being independent of z) that if $M \propto -|y|^k$, where $k > 2$, the potential vorticity stability condition (3.1.47) is violated in the neighbourhood of the equator, while the other two conditions are not violated anywhere. This is an example of the situation described in Figure 3.4. If $1 < k < 2$, the stability conditions are all satisfied.

Consider

$$U(y,z) = \tfrac{1}{2}\beta y^2 - \alpha|y|^k, \tag{3.1.97}$$

where $\alpha > 0$ and $k > 0$ are constants. We again assume that temperature is a function of y only. From (3.1.88), we calculate

$$T(y) = T_{00}\left[1 - \frac{\gamma}{g}U(y)\right] \tag{3.1.98}$$

and pressure

$$P(y,z) = P_{00}\exp\left[-\frac{g}{RT_{00}}\left(z + \frac{1}{2}\frac{\beta}{\gamma}y^2 - \frac{\beta}{\gamma}\int_0^y \frac{y'}{1 - \frac{1}{2}\frac{\beta\gamma}{g}y'^2 + \frac{\alpha\gamma}{g}|y'|^k}\,\mathrm{d}y'\right)\right]. \tag{3.1.99}$$

[1]The concern is that as $\det\Lambda$ tends to 0 at the equator, one of the other two subdeterminants may tend to infinity, but that is not so in this case, nor in any of the other examples.

The potential vorticity is

$$
\begin{aligned}
Q \;=\; \frac{1}{DT}\Bigg\{ & -\beta\gamma^2\left(\frac{c_pT_{00}}{g}\right)y + \frac{1}{2}\beta^2\gamma y^3 + k\alpha\left[g + \gamma^2\left(\frac{c_pT_{00}}{g}\right)\right]\frac{|y|^k}{y} \\
& - \beta\gamma\alpha\left(1+\frac{k}{2}\right)y|y|^k + \gamma k\alpha^2\frac{|y|^{2k}}{y}\Bigg\}.
\end{aligned}
\tag{3.1.100}
$$

Substituting $\alpha = \frac{1}{2}(\beta - \beta')$ and $k = 2$ recovers (3.1.94) from Example 3.

Case $k > 2$

It appears from (3.1.100) that if $k > 2$, the leading behaviour of Q near the equator is like $-y$ as y approaches zero, and sufficiently close to the equator, the symmetric stability condition is violated. Is this instability of the "inertial" type or the "static" type? Checking the other criteria, we find that

$$
y\partial(M, P) = D\left[k\alpha g|y|^k + \tfrac{1}{2}\beta^2\gamma y^4 - \alpha\beta\gamma(1+\tfrac{k}{2})|y|^{k+2} + k\alpha^2\gamma|y|^{2k}\right] \tag{3.1.101}
$$

is positive for all y if $L^2 < [2k/(k+2)](g/\beta\gamma)$, which is true for all k if $L < 10^5$ km. Similarly,

$$
y\partial(N, M^{(\mathrm{p})}) = \frac{1}{T}\left\{k\alpha\gamma^2\left(\frac{c_pT_{00}}{g}\right)|y|^k + \beta g\left[1 - \frac{\gamma^2}{g}\left(\frac{c_pT_{00}}{g}\right)y^2\right]\right\} \tag{3.1.102}
$$

is positive for all y if $T_{00} < g^2/\gamma^2 c_p \approx 2\times 10^7$ K.

Therefore, both the static stability and Rayleigh criteria are satisfied. The case must be as in Figure 3.4. The gradients of M and N satisfy their individual conditions for stability, but for small y, they must be unstably oriented relative to each other. The gradients, normalized to have a z component of unity, to leading order in y are

$$
\nabla_n M \;\sim\; -\left(k\frac{\alpha}{\gamma}\frac{|y|^k}{y}\right)\hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z, \tag{3.1.103a}
$$

$$
\nabla_n N \;\sim\; -\left(\frac{\beta\gamma}{g^2}c_pT_{00}y\right)\hat{\mathbf{e}}_y + \hat{\mathbf{e}}_z, \tag{3.1.103b}
$$

where ∇_n is the normalized gradient. ∇N tips towards the equator increasingly like $|y|$, while ∇M tips like $|y|^{k-1}$, which is slower for small y. The interval over which this is a factor is small for Earth values. For $k = 4$, the leading order gradients are colinear at $y = \sqrt{\beta\gamma^2 c_p T_{00}/4\alpha g^2}$. A value of α that would produce a difference in velocity of 10 ms^{-1} between the equator and $y = 1000$ km is 10^{-23} m^{-3}s^{-1}, making the interval of anomalous Q only about 100 m in width.

Case $1 < k < 2$

If $1 < k < 2$, then the leading behaviour of the potential vorticity is $Q \sim \operatorname{sgn}(y)\,|y|^{k-1}$, and the symmetric stability condition is satisfied in the neighbourhood of the equator. We next evaluate the subdeterminants of the coefficient matrix to check that they are also positive. First, we reorder the terms in Q in ascending order of powers of y for this case,

$$\begin{aligned} Q &= \frac{1}{DT}\left\{k\alpha\left[g+\gamma^2\left(\frac{c_pT_{00}}{g}\right)\right]\frac{|y|^k}{y} - \beta\gamma^2\left(\frac{c_pT_{00}}{g}\right)y\right. \\ &\qquad \left. - \beta\gamma\alpha\left(1+\frac{k}{2}\right)y|y|^k + \gamma k\alpha^2\frac{|y|^{2k}}{y} + \frac{1}{2}\beta^2\gamma y^3\right\}. \end{aligned} \tag{3.1.104}$$

Then (3.1.46) is

$$\frac{1}{D^3Q}\left(\frac{\partial P}{\partial N}\right)_D \partial(M,P) = \frac{PT}{c_vD}\left\{1 - \frac{\gamma^2\left(\frac{c_pT_{00}}{g}\right)\left(k\alpha\frac{|y|^k}{y} - \beta y\right)}{k\alpha\left[g+\gamma^2\left(\frac{c_pT_{00}}{g}\right)\right]\frac{|y|^k}{y} - \beta\gamma^2\left(\frac{c_pT_{00}}{g}\right)y + \dots}\right\}, \tag{3.1.105}$$

where the $\dots$ represents the higher order terms in TDQ. The right hand side of (3.1.105) is positive for all geophysically relevant values of the parameters. It is positive for all α if $L < \sqrt{g/2\beta\gamma} \approx 10^5$ km. Taking the limit as $y \to 0$, we find

$$\lim_{y\to 0}\frac{1}{D^3Q}\left(\frac{\partial P}{\partial N}\right)_D \partial(M,P) = \frac{PT}{c_vD}\left[1 - \frac{1}{1+\frac{g}{\gamma^2}\left(\frac{g}{c_pT_{00}}\right)}\right], \tag{3.1.106}$$

which is positive (in particular, it is not 0). (3.1.51) is

$$\frac{1}{DQ}\partial(N, M^{(\mathrm{p})}) = 1 - \frac{-g\beta y + k\alpha g\dfrac{|y|^k}{y} + \dots}{k\alpha\left[g + \gamma^2\left(\dfrac{c_p T_{00}}{g}\right)\right]\dfrac{|y|^k}{y} - \beta\gamma^2\left(\dfrac{c_p T_{00}}{g}\right)y + \dots}, \tag{3.1.107}$$

which is certainly positive provided $T_{00} < g^2/\gamma^2 c_p \approx 2\times 10^7$ K. Taking the limit as $y \to 0$, we find

$$\lim_{y\to 0}\frac{1}{DQ}\partial(N, M^{(\mathrm{p})}) = 1 - \frac{1}{1 + \dfrac{\gamma^2}{g}\left(\dfrac{RT_{00}}{g}\right)}, \tag{3.1.108}$$

which is non-zero. Hence, the stability criteria are satisfied by this basic state over the whole domain.

Example 5: An example with temperature increasing with z

Next, consider an equilibrium with temperature increasing linearly with height. This is perhaps more representative of the stratosphere than the previous examples. For purposes of comparison, we use the same velocity as in Example 3, $U(y) = \frac{1}{2}\beta' y^2$. We use the method of characteristics to solve

$$\left(1 - \frac{\beta'\gamma}{2g}y^2\right)P_y + \left(-\frac{\beta\beta'}{g}y^3\right)P_z = 0 \tag{3.1.109}$$

and find that curves of constant P have parametric equations

$$y(s) = \sqrt{\frac{2g}{\beta'\gamma}}\left[\frac{c_y\exp\left(\sqrt{2\beta'\gamma g}\,s\right) - 1}{c_y\exp\left(\sqrt{2\beta'\gamma g}\,s\right) + 1}\right], \tag{3.1.110a}$$

$$z(s) = c_z + \frac{\beta g}{\beta'\gamma^2}\left\{\sqrt{2\beta'\gamma g}s - 2\ln\left[\exp\left(\sqrt{2\beta'\gamma g}\,s\right) + 1\right] - \frac{4\exp\left(\sqrt{2\beta'\gamma g}\,s\right)}{\left[\exp\left(\sqrt{2\beta'\gamma g}\,s\right) + 1\right]^2}\right\}, \tag{3.1.110b}$$

where c_y and c_z are constants. This time, we choose as the initial curve the pressure at the equator corresponding to $U(0,z) = 0$ and $T(0,z) = T_{00}(1 + \sigma z)$, which is

$$P(0,z) = P_{00}\left(1 + \sigma z\right)^{-\dfrac{g}{RT_{00}\sigma}}. \tag{3.1.111}$$

Hence, the initial curve is

$$
\begin{aligned}
y(0,r) &= 0, && (3.1.112a)\\
z(0,r) &= r, && (3.1.112b)\\
P(0,r) &= P_{00}\,(1+\sigma r)^{-\dfrac{g}{RT_{00}\sigma}}, && (3.1.112c)
\end{aligned}
$$

from which

$$
c_y = 1, \quad c_z = r + \frac{\beta g}{\beta'\gamma^2}(1 + 2\ln 2). \qquad (3.1.113)
$$

Solving for r in terms of y and z, and substituting into (3.1.111) gives

$$
P(y,z) = P_{00}\left\{1 + \sigma\left[z - \frac{\beta g}{\beta'\gamma^2}\left(\ln\left(1 - \frac{\beta'\gamma}{2g}y^2\right) + \frac{\beta'\gamma}{2g}y^2\right)\right]\right\}^{-\dfrac{g}{RT_{00}\sigma}}. \qquad (3.1.114)
$$

The corresponding density and pressure fields are

$$
\begin{aligned}
D(y,z) &= \frac{P}{RT_{00}}\left(\frac{1}{1 - \frac{\beta'\gamma}{2g}y^2}\right)\left\{\frac{1}{1 + \sigma\left[z - \frac{\beta g}{\beta'\gamma^2}\left(\ln\left(1 - \frac{\beta'\gamma}{2g}y^2\right) + \frac{\beta'\gamma}{2g}y^2\right)\right]}\right\}, && (3.1.115a)\\
T(y,z) &= T_{00}\left(1 - \frac{\beta'\gamma}{2g}y^2\right)\left\{1 + \sigma\left[z - \frac{\beta g}{\beta'\gamma^2}\left(\ln\left(1 - \frac{\beta'\gamma}{2g}y^2\right) + \frac{\beta'\gamma}{2g}y^2\right)\right]\right\}. && (3.1.115b)
\end{aligned}
$$

The potential vorticity is

$$
\begin{aligned}
Q = \frac{1}{DT}\Bigg\{&\left[(g + \sigma c_p T_{00})(\beta - \beta') - \beta'\gamma^2\left(\frac{c_p T_{00}}{g}\right)(1 + \sigma z)\right] y\\
&+ \frac{1}{2}\left[\gamma\beta'^2 + \sigma\gamma\beta'\left(\frac{c_p T_{00}}{g}\right)(\beta + \beta')\right] y^3\\
&+ (\sigma\beta c_p T_{00})\, y \ln\left(1 - \frac{\beta'\gamma}{2g}y^2\right)\Bigg\}. \qquad (3.1.116)
\end{aligned}
$$

Now, yQ is positive if

$$
\beta' < \left\{\frac{g + \sigma c_p T_{00}}{g + c_p T_{00}\left[\sigma + \frac{\gamma^2}{g}(1 + \sigma H)\right]}\right\}\beta. \qquad (3.1.117)
$$

A reasonable value of σ might be 10^{-5} (temperature changing by 10% over 10 km). This condition is not significantly different from the $\sigma = 0$ case in Example 3. We now check the other two conditions. The quantity

$$y\partial(M,P) = D\left[g(\beta-\beta')y^2 + \tfrac{1}{2}\gamma\beta'^2y^4\right] \tag{3.1.118}$$

is positive for all y if $\beta' < \beta$, while

$$\begin{aligned} y\partial(N,M^{(\mathrm{p})}) &= \frac{1}{T}\left\{\left[g\beta + c_pT_{00}\left(\beta\sigma - \frac{\beta'\gamma^2}{g}(1+\sigma z)\right)\right]y^2\right. \\ &\quad \left. + \left[\sigma c_pT_{00}\left(\frac{\beta'\gamma}{2g}\right)\right]y^4 + \left[\beta\sigma c_pT_{00}\right]y^2\ln\left(1-\frac{\beta'\gamma}{2g}y^2\right)\right\} \end{aligned} \tag{3.1.119}$$

is positive if

$$\beta' < \left[\frac{g+\sigma c_pT_{00}}{g + c_pT_{00}\dfrac{\gamma^2}{g}(1+\sigma H)}\right]\beta. \tag{3.1.120}$$

Both conditions are satisfied if β' satisfies (3.1.117), and again there is an interval of β' values for which only the potential vorticity condition is violated.

Example 6: An example with velocity changing with z

The final example is one in which the basic state velocity is a linear function of z, $U(z) = U_{00} + \gamma' z$, where $\gamma' > -|U_{00}|/H$. Using the method of characteristics with the surface pressure from Example 1 as initial curve, we calculate

$$\begin{aligned} P(y,z) &= P_{00}(1+\gamma' z)^{-\left(\frac{g}{RT_{00}}\right)\left(\frac{U_{00}}{\gamma'}\right)} \\ &\qquad \times\exp\left[-\frac{1}{2}\left(\frac{\beta U_{00}}{RT_{00}}\right)y^2 + \left(\frac{\gamma U_{00}}{RT_{00}}\right)z\right], \end{aligned} \tag{3.1.121a}$$

$$D(y,z) = \left(\frac{1}{RT_{00}}\right)\frac{P}{1+\dfrac{\gamma'}{U_{00}}z}, \tag{3.1.121b}$$

$$T(y,z) = T_{00}\left(1+\frac{\gamma'}{U_{00}}z\right). \tag{3.1.121c}$$

The potential vorticity is

$$Q = \frac{\beta y}{DT}\left[g + \gamma'(U_{00}+\gamma' z) + \frac{\gamma'}{U_{00}}c_pT_{00}\right]. \tag{3.1.122}$$

$yQ > 0$ for all y and z if γ' and U_{00} satisfy

$$\gamma' U_{00} > -\frac{g}{1 + \frac{c_p T_{00}}{U_{00}^2}}, \tag{3.1.123}$$

which can fail only if γ' and U_{00} are of opposite signs (i.e. if the magnitude of U is decreasing with z). Note that it is possible to violate (3.1.123) while respecting $\gamma' > -|U_{00}|/H$, since

$$-\frac{g}{1 + \frac{c_p T_{00}}{U_{00}^2}} = -\frac{U_{00}^2}{H}\left(\frac{gH}{U_{00}^2 + c_p T_{00}}\right), \tag{3.1.124}$$

which is greater than $-U_{00}^2/H$ for realistic values of U_{00}, T_{00} and H. However, a vertical shear that violates (3.1.123) is unrealistically large (> 100 (ms^{-1})km^{-1}).

The inertial stability condition $y\partial(M, P) > 0$ is satisfied if $\gamma' U_{00} > -g$, and the static stability condition $y\partial(N, M^{(\mathrm{p})}) > 0$ is satisfied if $\gamma' U_{00} > -gU_{00}^2/c_p T_{00}$. Both are satisfied if (3.1.123) is.

Note that there is no difference between the stability conditions for basic states with eastward or westward velocities and that the conditions are independent of γ. This is perhaps surprising in light of the Rayleigh criterion which refers to the radial derivative of absolute angular momentum, which depends on the sign of γ' ($M_z = \gamma + \gamma'$). The explanation is that the planetary part of the angular momentum gradient is balanced by the pressure field, i.e. $\partial(M^{(\mathrm{p})}, P) = \beta y g D$, which is independent of γ.

3.2 Remarks on nonlinear stability

In similar problems (such as Mu et al., 1996), the conditions for linear stability can be extended to finite amplitude, "nonlinear" stability, meaning that if a basic state $\mathbf{X}$ satisfies appropriate conditions, then arbitrary, even large, perturbations to $\mathbf{X}$ will remain bounded for all time (as governed by the full nonlinear equations). The usual approach is to define an exact invariant $\mathcal{A}(\mathbf{x}; \mathbf{X})$, called the *pseudoenergy*, which is zero if $\mathbf{x} = \mathbf{X}$

and to find conditions on $\mathbf{X}$ such that $\mathcal{A}$ is strictly positive for all other choices of $\mathbf{x}$ (Shepherd, 1990).

Since linear stability is a necessary condition for nonlinear stability (provided the norm for the linear problem is the small amplitude limit of the finite amplitude norm in terms of which nonlinear stability is defined), in this case we would only consider $\mathbf{X}$ having $Q = 0$ on $y = 0$ and nowhere else. Recalling that q is materially conserved, one can show that the functional

$$\mathcal{C}_{NL} = \iint_{\mathcal{D}} \rho \left\{ C^-(m,\eta) + H(q) \left[C^+(m,\eta) - C^-(m,\eta) \right] \right\} \, \mathrm{d}y \, \mathrm{d}z, \tag{3.2.1}$$

where

$$H(q) = \begin{cases} 0, q < 0 \\ 1, q \geq 0 \end{cases}, \tag{3.2.2}$$

is conserved by (3.1.1). We would then choose C^- and C^+ to satisfy (3.1.20) and define

$$\mathcal{A}(\mathbf{x};\mathbf{X}) = (\mathcal{H} + \mathcal{C}_{NL})(\mathbf{x}) - (\mathcal{H} + \mathcal{C}_{NL})(\mathbf{X}). \tag{3.2.3}$$

By construction, $\mathcal{A}(\mathbf{X};\mathbf{X}) = 0$ and $\mathcal{A}$ has a critical point at $\mathbf{x} = \mathbf{X}$. However, it is not a simple matter to find norms to bound $\mathcal{A}$ from above and below for all times like we did with $\mathcal{H}_L$ in the linear case. There are two separate difficulties. The first is connected to the asymmetry of the basic state and hence to the difference between C^- and C^+. If the flow evolves in such a way that regions of $q > 0$ develop in the southern hemisphere, then the corresponding contribution to $\mathcal{A}$ from those regions will depend on C^+ and not on C^- as it did at the basic state. In this way, it is possible for $\mathcal{A}$ to be negative, which would prevent any rigorous Lyapunov stability result by this method. This problem is addressed in more detail in the next chapter.

The second difficulty is related to the fact that we are free to choose the $C^{(i)}(m,\eta)$ provided we satisfy (3.1.20) and the matching condition at boundaries between the $\mathcal{D}^{(i)}$ at equilibrium (lines of $Q = 0$), but we cannot choose the dependence of the internal energy

$\mathcal{E}$ on ρ and η and hence cannot bound certain terms in $\mathcal{A}$ for all possible perturbations of ρ and η outside of the ranges of $N(y,z)$ and $D(y,z)$.

One approach to generalizing the linear conditions to finite amplitude is to Taylor expand the integrand of $\mathcal{A}$ about $\mathbf{x} = \mathbf{X}$. To avoid the first difficulty, suppose that $\mathbf{X}(y,z)$ is an *even* function of y, i.e. $\mathbf{X}(-y,z) = \mathbf{X}(y,z)$. In that case, $C^-(M,N) = C^+(M,N) = C(M,N)$. Expanding the integrand (minus the v and w terms, which are clearly positive) to second order, and using Taylor's Remainder Theorem,

$$\begin{aligned}
&\rho\left[\tfrac{1}{2}u^2 + gz + \mathcal{E}(\rho,\eta) + C(m,\eta)\right] \\
&\quad - D\left[\tfrac{1}{2}U^2 + gZ(M,N) + \mathcal{E}(D,N) + C(M,N)\right] \\
&\qquad = \frac{1}{2}\Big\{ \left[\tilde{\rho}(1 + C_{mm}(\tilde{m},\tilde{\eta}))\right](u-U)^2 \\
&\qquad\quad + \left[2\mathcal{E}_\rho(\tilde{\rho},\tilde{\eta}) + \tilde{\rho}\mathcal{E}_{\rho\rho}(\tilde{\rho},\tilde{\eta})\right](\rho - D)^2 + \left[\tilde{\rho}(\mathcal{E}_{\eta\eta}(\tilde{\rho},\tilde{\eta}) + C_{\eta\eta}(\tilde{\rho},\tilde{\eta}))\right](\eta - N)^2 \\
&\qquad\quad + 2\left[C_m(\tilde{m},\tilde{\eta}) + \tilde{u}\right](u-U)(\rho-D) + 2\left[\tilde{\rho}C_{m\eta}(\tilde{m},\tilde{\eta})\right](u-U)(\eta-N) \\
&\qquad\quad + 2\left[\tilde{\rho}\mathcal{E}_{\rho\eta} + C_\eta(\tilde{m},\tilde{\eta}) + \mathcal{E}_\eta(\tilde{\rho},\tilde{\eta})\right](\rho-D)(\eta-N)\Big\},
\end{aligned} \tag{3.2.4}$$

where $(\tilde{m},\tilde{\rho},\tilde{\eta})$ is a point on the line joining (M,D,N) and (m,ρ,η). Note that $(\tilde{m},\tilde{\rho},\tilde{\eta})$ is a function of (y,z,t). The linear terms vanish in (3.2.4) because C was chosen to satisfy (3.1.20).

Observe that the quadratic form in the integrand of $\mathcal{A}$ has coefficient matrix

$$\tilde{\Lambda}(\tilde{m},\tilde{\rho},\tilde{\eta};M,D,N) =$$
$$\begin{bmatrix}
\frac{D_0^2}{\tilde{\rho}}c_s^2(\tilde{\rho},\tilde{\eta}) & D_0N_0(\tilde{\rho}\mathcal{E}_{\rho\eta}(\tilde{\rho},\tilde{\eta}) + C_\eta(\tilde{m},\tilde{\eta}) + \mathcal{E}_\eta(\tilde{\rho},\tilde{\eta})) & D_0U_0(C_m(\tilde{m},\tilde{\eta}) + \tilde{u}) \\
D_0N_0(\tilde{\rho}\mathcal{E}_{\rho\eta}(\tilde{\rho},\tilde{\eta}) + C_\eta(\tilde{m},\tilde{\eta}) + \mathcal{E}_\eta(\tilde{\rho},\tilde{\eta})) & N_0^2\tilde{\rho}(\mathcal{E}_{\eta\eta}(\tilde{\rho},\tilde{\eta}) + C_{\eta\eta}(\tilde{m},\tilde{\eta})) & N_0U_0\tilde{\rho}C_{m\eta}(\tilde{m},\tilde{\eta}) \\
D_0U_0(C_m(\tilde{m},\tilde{\eta}) + \tilde{u}) & N_0U_0\tilde{\rho}C_{m\eta}(\tilde{m},\tilde{\eta}) & U_0^2\tilde{\rho}(1 + C_{mm}(\tilde{m},\tilde{\eta}))
\end{bmatrix}, \tag{3.2.5}$$

where $\tilde{u} = \tilde{m} + \frac{1}{2}\beta y^2 - \gamma z$, and $C_m(\tilde{m},\tilde{\eta}) = -U(Y(\tilde{m},\tilde{\eta}), Z(\tilde{m},\tilde{\eta}))$, which do not generally cancel except at the basic state. Similarly, $C_\eta(\tilde{m},\tilde{\eta}) = -\mathcal{E}_\eta[\rho(Y(\tilde{m},\tilde{\eta}), Z(\tilde{m},\tilde{\eta})),\tilde{\eta}]$ and $\mathcal{E}_\eta(\tilde{\rho},\tilde{\eta})$ do not in general cancel except at the basic state.

The difficulty arises because in principle, (u, ρ, η) can be anything (as long as ρ and η are positive), so we have almost no information about the likely values of $(\tilde{m}, \tilde{\rho}, \tilde{\eta})$. We can control the behaviour of $C(m, \eta)$ outside of the ranges of $M(y, z)$ and $N(y, z)$, but we cannot control $\mathcal{E}(\rho, \eta)$. In addition, while the possible values that m and η can take are determined by the initial conditions (since they are materially conserved), we do not know the set of values that ρ will take. The crucial point is that conditions for the positive definiteness of $\tilde{\Lambda}$ depend not only on the basic state, but on the states that the system might pass through over time, the details of which we do not know.

Example: Nonlinear stability of isothermal atmosphere in solid-body rotation

Consider again the case of an isothermal atmosphere in solid-body rotation (Example 1 in Section 3.1). The matrix of coefficients in the integrand of $\mathcal{A}$ is

$$\tilde{\Lambda} = \begin{bmatrix} D_0^2 \dfrac{Rc_p}{c_v} \dfrac{\tilde{\tau}}{\tilde{\rho}} & D_0 N_0 \left(\dfrac{c_p}{c_v} \tilde{\tau} - T_{00} \right) & D_0 U_0 \left(\tilde{u} - U_{00} \right) \\ D_0 N_0 \left(\dfrac{c_p}{c_v} \tilde{\tau} - T_{00} \right) & N_0^2 \dfrac{1}{c_v} \tilde{\rho} \tilde{\tau} & 0 \\ D_0 U_0 \left(\tilde{u} - U_{00} \right) & 0 & U_0^2 \tilde{\rho} \end{bmatrix}, \tag{3.2.6}$$

where $\tilde{\tau} = \tau(\tilde{\rho}, \tilde{\eta})$ is the temperature corresponding to the state which makes the Taylor expansion (3.2.4) exact. The bottom right element of $\tilde{\Lambda}$ is positive, as is the bottom right 2×2 subdeterminant

$$\det \begin{bmatrix} N_0^2 \dfrac{1}{c_v} \tilde{\rho} \tilde{\tau} & 0 \\ 0 & U_0^2 \tilde{\rho} \end{bmatrix} = N_0^2 U_0^2 \frac{\tilde{\rho}^2 \tilde{\tau}}{c_v}. \tag{3.2.7}$$

The matrix is therefore positive definite if and only if its determinant is positive,

$$\det \tilde{\Lambda} = D_0^2 N_0^2 U_0^2 \tilde{\rho} \left[\frac{R}{c_v} T_{00}^2 - \frac{c_p}{c_v} (\tilde{\tau} - T_{00})^2 - \frac{1}{c_v} \tilde{\tau} (\tilde{u} - U_{00})^2 \right] > 0. \tag{3.2.8}$$

Clearly, $\det \tilde{\Lambda}$ is not positive for all possible values of $\tilde{\tau}$ and $\tilde{u}$. We note that as long as the system evolves such that the variables stay within realistic limits, $\det \tilde{\Lambda}$ is always positive. For example, suppose $T_{00} = 300$ K and assume that the temperature does not change by more than about 100 K anywhere and that the velocity does not change by more than about 100 ms^{-1}. Then

$$\begin{aligned}\frac{R}{c_v}T_{00}^2 &= 3.6\times10^4\ \mathrm{K}^2, &&(3.2.9a)\\ \frac{c_p}{c_v}(\tilde\tau-T_{00})^2 &\lesssim 2\times10^4\ \mathrm{K}^2, &&(3.2.9b)\\ \frac{1}{c_v}\tilde\tau(\tilde u-U_{00})^2 &\lesssim 6\times10^3\ \mathrm{K}^2, &&(3.2.9c)\end{aligned}$$

so the positive term is greater than the upper limit of the negative terms. Neglecting the velocity term, we can call the equilibrium nonlinearly stable as long as the temperature deviation satisfies

$$\frac{|\Delta\tau|}{T_{00}}\lesssim\sqrt{\frac{R}{c_p}}\approx 0.5 \tag{3.2.10}$$

everywhere for all time, i.e. as long as the temperature does not change by half its value. Note, however, that this result does not represent a true estimate of the growth of a disturbance (not even an upper bound) without knowing how the system evolves. If the evolution of the system satisfies (3.2.10), then that is itself a statement of stability, and the pseudoenergy result adds nothing to it.

3.3 Summary

Conditions for linear stability of a purely zonal equilibrium to the x-symmetric adiabatic compressible Euler equations on a β-plane, including a leading order representation of the $\cos\phi$ Coriolis force terms (controlled by the parameter $\gamma\equiv2\Omega$), have been calculated. Formally, there is no change to the inertial stability condition from the traditional hydrostatic result. The condition for stability is still the Rayleigh criterion that angular momentum in the basic state must increase with latitude towards the equator on surfaces of constant pressure, although the definition of angular momentum is slightly different with the inclusion of the γ effect. The "static" stability conditions derived are rather different from those in the traditional system due to the modification in the surfaces of constant planetary angular momentum due to the $\cos\phi$ terms. We find that for static stability, the gradient of entropy must be clockwise of the gradient of planetary angular momentum in the northern hemisphere and anticlockwise in the southern hemisphere. The symmetric stability condition that potential vorticity be positive in the northern hemisphere and negative in the southern generalizes to the nonhydrostatic system if the definition of potential vorticity is modified to account for the nonhydrostatic terms.

Several examples were presented. In each case, pressure and temperature fields are found which satisfy thermal wind balance. There are two notable examples in which the γ effect is decisive. In the case of a basic state with angular momentum profile higher than quadratic in y, conditions for both inertial and static stability are satisfied, but the potential vorticity condition fails. We show that this is due to the tipping of the angular momentum gradient more quickly than the entropy gradient as y increases away from the equator, an effect dependent on the inclusion of the γ terms in the dynamical equations. In the case of a basic state with linear velocity shear across the equator, the famous example of Dunkerton (1981), the width of the latitude interval of instability varies with the velocity shear if γ is included in the equations, approaching zero as λ gets large.

Steps for extending the result to finite amplitude disturbances were outlined, and the example of an isothermal atmosphere in solid-body rotation discussed. However, technical details associated with asymmetric basic states and the evolution of the density field prevent a general result. In the following chapter, we repeat the same problem, but we make the anelastic approximation, in which the density field is constrained, and half of the obstacle to the nonlinear result is removed.

3.A Thermodynamics relations

The specific internal energy of the fluid $\mathcal{E}(\rho, \eta)$ is assumed to be a twice differentiable function of density ρ and entropy η, with continuous second partial derivatives. The first law of thermodynamics may be written

$$\mathrm{d}\mathcal{E} = \frac{p}{\rho^2}\mathrm{d}\rho + \tau \mathrm{d}\eta, \tag{3.A.1}$$

where p is pressure and τ is temperature. The first term in (3.A.1) represents work done on the fluid by compression, and the second represents heating of the fluid. Implicit in (3.A.1) are

$$\mathcal{E}_\rho \equiv \left(\frac{\partial \mathcal{E}}{\partial \rho}\right)_\eta = \frac{p}{\rho^2}, \qquad \mathcal{E}_\eta \equiv \left(\frac{\partial \mathcal{E}}{\partial \eta}\right)_\rho = \tau. \tag{3.A.2}$$

The second partial derivatives of $\mathcal{E}$ are

$$\mathcal{E}_{\rho\rho} = \frac{1}{\rho^2}\left(\frac{\partial p}{\partial \rho}\right)_\eta - \frac{2p}{\rho^3}, \qquad \mathcal{E}_{\eta\eta} = \left(\frac{\partial \tau}{\partial \eta}\right)_\rho, \tag{3.A.3a}$$

$$\mathcal{E}_{\rho\eta} = \frac{1}{\rho^2}\left(\frac{\partial p}{\partial \eta}\right)_\rho, \qquad \mathcal{E}_{\eta\rho} = \left(\frac{\partial \tau}{\partial \rho}\right)_\eta. \tag{3.A.3b}$$

The requirement that the second partial derivatives of $\mathcal{E}$ be continuous implies the *Maxwell relation*

$$\frac{1}{\rho^2}\left(\frac{\partial p}{\partial \eta}\right)_\rho = \left(\frac{\partial \tau}{\partial \rho}\right)_\eta. \tag{3.A.4}$$

A combination of terms that appears frequently is

$$2\rho\mathcal{E}_\rho + \rho^2\mathcal{E}_{\rho\rho} = \left(\frac{\partial p}{\partial \rho}\right)_\eta = c_s^2, \tag{3.A.5}$$

where c_s^2 is the speed of sound in the fluid.

Further Maxwell relations are obtained by Legendre transformation of $\mathcal{E}$ with respect to ρ and η (see any thermodynamics text):

$$\mathrm{d}\left(\mathcal{E} + \frac{p}{\rho}\right) = \tau\mathrm{d}\eta + \frac{1}{\rho}\mathrm{d}p, \tag{3.A.6a}$$

$$\mathrm{d}(\mathcal{E} - \tau\eta) = \frac{p}{\rho^2}\mathrm{d}\rho - \eta\mathrm{d}\tau, \tag{3.A.6b}$$

$$\mathrm{d}\left(\mathcal{E} + \frac{p}{\rho} - \tau\eta\right) = \frac{1}{\rho}\mathrm{d}p - \eta\mathrm{d}\tau, \tag{3.A.6c}$$

whence, respectively,

$$\left(\frac{\partial\tau}{\partial p}\right)_\eta = -\frac{1}{\rho^2}\left(\frac{\partial\rho}{\partial\eta}\right)_p, \quad (3.A.7a)$$

$$\frac{1}{\rho^2}\left(\frac{\partial p}{\partial\tau}\right)_\rho = -\left(\frac{\partial\eta}{\partial\rho}\right)_\tau, \quad (3.A.7b)$$

$$\frac{1}{\rho^2}\left(\frac{\partial\rho}{\partial\tau}\right)_p = \left(\frac{\partial\eta}{\partial p}\right)_\tau. \quad (3.A.7c)$$

Relations for an ideal gas

An ideal gas is defined by the equation of state

$$p = R\rho\tau, \quad (3.A.8)$$

where $R \approx 287$ JK^{-1}kg^{-1} is the universal gas constant 8.314 JK^{-1} mol^{-1} divided by the molar mass of dry air (approximately 29 g mol^{-1}), and by the property that the internal energy is a linear function of temperature, with

$$c_v \equiv \left(\frac{\partial\mathcal{E}}{\partial\tau}\right)_\rho = \text{constant}. \quad (3.A.9)$$

c_v is called the specific heat capacity at constant volume. By the chain rule,

$$c_v = \left(\frac{\partial}{\partial\tau}\mathcal{E}[\rho,\eta(\rho,\tau)]\right)_\rho = \left(\frac{\partial\mathcal{E}}{\partial\eta}\right)_\rho\left(\frac{\partial\eta}{\partial\tau}\right)_\rho \quad (3.A.10)$$

whence

$$\left(\frac{\partial\tau}{\partial\eta}\right)_\rho = \frac{\tau}{c_v}. \quad (3.A.11)$$

From (3.A.8) and (3.A.11),

$$\mathcal{E}_{\rho\rho} = \left(\frac{R}{c_v} - 1\right)\frac{p}{\rho^3}, \qquad \mathcal{E}_{\eta\eta} = \frac{\tau}{c_v}, \qquad \mathcal{E}_{\eta\rho} = \mathcal{E}_{\eta\rho} = \frac{1}{c_v}\frac{p}{\rho^2}, \quad (3.A.12)$$

and the speed of sound squared is

$$c_s^2 = \frac{c_p}{c_v}R\tau, \quad (3.A.13)$$

where $c_p = R + c_v$. c_p is called the specific heat at constant pressure.

3.B Derivatives with respect to M and N

Given distributions of entropy $N(y,z)$ and angular momentum $M(y,z)$ in a domain in which $DQ = \partial(N,M) \neq 0$, there exist unique inverse functions $Y(M,N)$ and $Z(M,N)$. In particular, partial derivatives with respect to y and z can be rewritten in terms of partial derivatives with respect to N and M using a two-dimensional version of the chain rule:

$$\begin{bmatrix} \partial_N \\ \partial_M \end{bmatrix} = \begin{bmatrix} Y_N & Z_N \\ Y_M & Z_M \end{bmatrix} \begin{bmatrix} \partial_y \\ \partial_z \end{bmatrix}, \tag{3.B.1}$$

where ∂_y means partial derivative with respect to y keeping z fixed, ∂_N means partial derivative with respect to N keeping M fixed, etc. Similarly,

$$\begin{bmatrix} \partial_y \\ \partial_z \end{bmatrix} = \begin{bmatrix} N_y & M_y \\ N_z & M_z \end{bmatrix} \begin{bmatrix} \partial_N \\ \partial_M \end{bmatrix}, \tag{3.B.2}$$

so that

$$\begin{bmatrix} \partial_N \\ \partial_M \end{bmatrix} = \begin{bmatrix} N_y & M_y \\ N_z & M_z \end{bmatrix}^{-1} \begin{bmatrix} \partial_y \\ \partial_z \end{bmatrix} = \frac{1}{DQ} \begin{bmatrix} M_z & -M_y \\ -N_z & N_y \end{bmatrix} \begin{bmatrix} \partial_y \\ \partial_z \end{bmatrix}. \tag{3.B.3}$$

In (3.1.29), (3.1.31), (3.1.33) and (3.1.35), we convert expressions involving Y_N to expressions involving M_z etc. Comparing (3.B.2) with (3.B.3), we find

$$\left(\frac{\partial Y}{\partial N}\right)_M = \frac{1}{DQ}\left(\frac{\partial M}{\partial z}\right)_y, \qquad \left(\frac{\partial Z}{\partial N}\right)_M = -\frac{1}{DQ}\left(\frac{\partial M}{\partial y}\right)_z \tag{3.B.4a}$$

$$\left(\frac{\partial Y}{\partial M}\right)_N = -\frac{1}{DQ}\left(\frac{\partial N}{\partial z}\right)_y, \qquad \left(\frac{\partial Z}{\partial M}\right)_N = \frac{1}{DQ}\left(\frac{\partial N}{\partial y}\right)_z. \tag{3.B.4b}$$

3.C Proof of conservation of $\mathcal{H}_L$ by linearized dynamics

We show that $\mathcal{H}_L$ is conserved by the linear equations. From (3.1.27),

$$\begin{aligned}\mathcal{H}_L(\mathbf{x}';\mathbf{X}) \equiv \sum_{i=1}^{n}\iint_{\mathcal{D}^{(i)}} & \big\{D(1+C^{(i)}_{mm}(M,N))u'^2 + Dv'^2 + Dw'^2 \\ &+ (2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))\rho'^2 \\ &+ D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))\eta'^2 \\ &+ 2D\mathcal{E}_{\rho\eta}(D,N)\rho'\eta' + 2DC^{(i)}_{m\eta}(M,N)u'\eta'\big\}\,\mathrm{d}y\,\mathrm{d}z. \qquad (3.C.1)\end{aligned}$$

Since the regions of integration are fixed, we may differentiate under the integral:

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_L = \sum_{i=1}^{n}\iint_{\mathcal{D}^{(i)}} & \big\{2D(1+C^{(i)}_{mm}(M,N))u'u'_t + 2Dv'v'_t + 2Dw'w'_t \\ &+ 2(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))\rho'\rho'_t \\ &+ 2D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))\eta'\eta'_t \\ &+ 2D\mathcal{E}_{\rho\eta}(D,N)(\rho'_t\eta' + \rho'\eta'_t) + 2DC^{(i)}_{m\eta}(M,N)(u'_t\eta' + u\eta'_t)\big\}\,\mathrm{d}y\,\mathrm{d}z. \qquad (3.C.2)\end{aligned}$$

Substituting from (3.1.11) and collecting terms,

$$\begin{aligned}\frac{d}{dt}\mathcal{H}_L = \sum_{i=1}^{n}\iint_{\mathcal{D}^{(i)}} & \big\{u'v'\big[2D(1+C^{(i)}_{mm}(M,N))(-U_y+\beta y) \\ &\qquad - 2D\beta y - 2DC^{(i)}_{m\eta}(M,N)N_y\big] \qquad (3.C.3a) \\ &+ u'w'\big[2D(1+C^{(i)}_{mm}(M,N))(-U_z-\gamma) \\ &\qquad - 2D\gamma - 2DC^{(i)}_{m\eta}(M,N)N_z\big] \qquad (3.C.3b) \\ &+ \eta'v'\big[-2D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))N_y \\ &\qquad - 2DC^{(i)}_{m\eta}(M,N)(U_y-\beta y) \\ &\qquad - 2D\mathcal{E}_{\rho\eta}(D,N)D_y\big] \qquad (3.C.3c)\end{aligned}$$

$$
\begin{aligned}
&+ \eta' w' \left[-2D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))N_z \right.\\
&\qquad\qquad - 2DC^{(i)}_{m\eta}(M,N)(U_y+\gamma)\\
&\qquad\qquad\qquad \left. - 2D\mathcal{E}_{\rho\eta}(D,N)D_z\right]\Big\}\, \mathrm{d}y\,\mathrm{d}z \qquad (3.C.3d)
\end{aligned}
$$

$$
\begin{aligned}
+ \iint_{\mathcal{D}} \Big\{ &\rho' v' \left[\frac{2}{D}P_y - 2D\mathcal{E}_{\rho\eta}(D,N)N_y \right.\\
&\qquad \left. - 2(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))D_y \right] \qquad (3.C.3e)\\
&+ \rho' w' \left[\frac{2}{D}P_z - 2D\mathcal{E}_{\rho\eta}(D,N)N_z \right.\\
&\qquad \left. - 2(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))D_z \right] \qquad (3.C.3f)\\
&+ v'_y \left[-2D(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N)\rho' \right.\\
&\qquad \left. - 2D^2\mathcal{E}_{\rho\eta}\eta'\right] - 2p'_y v' \qquad (3.C.3g)\\
&+ w'_z \left[-2D(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N)\rho' \right.\\
&\qquad \left. - 2D^2\mathcal{E}_{\rho\eta}\eta'\right] - 2p'_z w' \Big\}\, \mathrm{d}y\,\mathrm{d}z. \qquad (3.C.3h)
\end{aligned}
$$

Each of the expressions in square brackets (3.C.3a)-(3.C.3f) vanishes identically, as we now show. From (3.1.29) and (3.1.31), (3.C.3a) is

$$
\begin{aligned}
&2D(1 + C^{(i)}_{mm}(M,N))(-U_y + \beta y) - 2D\beta y - 2DC^{(i)}_{m\eta}(M,N)N_y\\
&\quad = 2D\left(\frac{1}{DQ}\right)(-\beta y N_z M_y - \gamma N_y M_y - \beta y DQ + \beta y M_z N_y + \gamma M_y N_y)\\
&\quad = 0, \qquad (3.C.4a)
\end{aligned}
$$

and (3.C.3b) is

$$
\begin{aligned}
&2D(1 + C^{(i)}_{mm}(M,N))(-U_z - \gamma) + 2D\gamma - 2DC^{(i)}_{m\eta}(M,N)N_z\\
&\quad = 2D\left(\frac{1}{DQ}\right)(-\beta y N_z M_z - \gamma N_y M_z + \gamma DQ + \beta y M_z N_z + \gamma M_y N_z)\\
&\quad = 0. \qquad (3.C.4b)
\end{aligned}
$$

From (3.A.3a), (3.1.35) and (3.A.3b), (3.C.3c) and (3.C.3d) are

$$\begin{aligned}
-2D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))N_y - 2DC^{(i)}_{m\eta}(M,N)(U_y - \beta y) - 2D\mathcal{E}_{\rho\eta}(D,N)D_y \\
= \frac{2}{D}\left(\frac{1}{DQ}\right)\left(\frac{\partial P}{\partial N}\right)_D (D_yM_zN_y - D_zM_yN_y - D_yN_zM_y + D_zN_yM_y - D_yDQ) \\
= 0, \qquad \text{(3.C.4c)}
\end{aligned}$$

$$\begin{aligned}
-2D(\mathcal{E}_{\eta\eta}(D,N) + C^{(i)}_{\eta\eta}(M,N))N_z - 2DC^{(i)}_{m\eta}(M,N)(U_z + \gamma) - 2D\mathcal{E}_{\rho\eta}(D,N)D_z \\
= \frac{2}{D}\left(\frac{1}{DQ}\right)\left(\frac{\partial P}{\partial N}\right)_D (D_yM_zN_z - D_zM_yN_z - D_yN_zM_z + D_zN_yM_z - D_zDQ) \\
= 0. \qquad \text{(3.C.4d)}
\end{aligned}$$

Using (3.A.3b) and (3.A.5), (3.C.3e) and (3.C.3f) are

$$\begin{aligned}
\frac{2}{D}P_y - 2D\mathcal{E}_{\rho\eta}(D,N)N_y - 2(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))D_y \\
= \frac{2}{D}\left(\frac{\partial P}{\partial D}\right)_N (D_y + N_y - N_y - D_y) \\
= 0, \qquad \text{(3.C.4e)}
\end{aligned}$$

$$\begin{aligned}
\frac{2}{D}P_z - 2D\mathcal{E}_{\rho\eta}(D,N)N_z - 2(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N))D_z \\
= \frac{2}{D}\left(\frac{\partial P}{\partial D}\right)_N (D_z + N_z - N_z - D_z) \\
= 0. \qquad \text{(3.C.4f)}
\end{aligned}$$

Therefore, we have that

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_L = \iint_{\mathcal{D}} \{v'_y\left[-2D(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N)\rho' - 2D^2\mathcal{E}_{\rho\eta}\eta'\right] - 2p'_yv' \\
+ w'_z\left[-2D(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N)\rho' - 2D^2\mathcal{E}_{\rho\eta}\eta'\right] - 2p'_zw'\}\, \mathrm{d}y\, \mathrm{d}z.
\end{aligned} \qquad \text{(3.C.5)}$$

Integrating the p'_y and p'_z terms by parts and applying the no normal flow boundary condition to eliminate the surface terms, we find

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{H}_L = \iint_{\mathcal{D}} (v'_y + w'_z)\left[-2D(2\mathcal{E}_\rho(D,N) + D\mathcal{E}_{\rho\rho}(D,N)\rho'\right. \\
\left. - 2D^2\mathcal{E}_{\rho\eta}\eta' + 2p'\right]\, \mathrm{d}y\, \mathrm{d}z, \qquad \text{(3.C.6)}
\end{aligned}$$

which vanishes because of (3.A.5), (3.A.3b) and (3.1.12). Therefore, $\mathcal{H}_L$ is conserved by the linearized dynamics.

Chapter 4

Symmetric stability in the anelastic system

In the previous chapter, we derived sufficient conditions for the linear stability of a steady zonal solution to the compressible Euler equations on an equatorial β-plane. We could not extend the result to finite amplitude because of two difficulties: "uncontrollability" of density perturbations, and asymmetric steady states. In this chapter, we avoid the first difficulty by making the *anelastic* approximation, and we address the second difficulty in greater detail.

The anelastic approximation, so called because the energy conserved by the approximate equations does not contain the "elastic energy" term responsible for pressure fluctuations during sound wave propagation, is based on two assumptions. The first is that the relative departure of potential temperature from a constant value is never large, and the second is that the time scale on which the velocity and thermodynamic fields vary is that of gravity wave propagation and is *slow*. The two assumptions are related — the first implies a weak stratification which in turn implies a slow gravity wave propagation speed, much as weak tension in a guitar string implies slow propagation of transverse waves.

By using the anelastic equations for our symmetric stability problem, we avoid the density perturbation difficulty that we met using the Euler equations. The reason is that density is not a *prognostic* variable in the anelastic system, but is given by a *diagnostic* (time-independent) equation. Variations in density are determined by variations in velocity and potential temperature.

The chapter is organized as follows. In Section 4.1, we introduce the modified anelastic equations (Wilhelmson and Ogura, 1972; Lipps and Hemler, 1982) for an ideal gas on the equatorial β-plane, of which the classical "deep" equations of Ogura and Phillips (1962) are a special case. We develop a Hamiltonian form of the longitudinally symmetric version of the modified anelastic equations in terms of angular momentum, vorticity, and potential temperature. In Section 4.2, conditions for symmetric stability under the linearized symmetric equations are derived using the same energy-Casimir approach as was used in Chapter 3, and then conditions for nonlinear stability of basic states which are symmetric about the equator are derived. In section 4.3, an exact solution to the anelastic version of the linear meridional shear problem (the "Dunkerton problem") is calculated. Finally, in section 4.4, an important application of the nonlinear stability result, the calculation of *saturation bounds* on instabilities (based on Shepherd, 1988), is applied to this problem. A derivation of the classical anelastic equations is presented in Appendix 4.A.

4.1 The anelastic equations on the equatorial β-plane

4.1.1 The anelastic approximation and scaling

The motivation for the classical anelastic approximation of Ogura and Phillips (1962) was to construct a system which filters sound waves without imposing the hydrostatic approximation. The resulting equations allow for relatively strong vertical accelerations and are suitable for modelling deep convection. They may be compared wih the Boussinesq equations (Spiegel and Veronis, 1960), used in modelling ocean dynamics, but the anelastic equations allow for greater variation in density.

The central assumption leading to the anelastic equations is that the potential temperature, defined by

$$\theta \equiv \left(\frac{p_{00}}{p}\right)^{\kappa} T, \tag{4.1.1}$$

where p_{00} is a constant reference pressure and $\kappa \equiv R/c_p$, where R is the gas constant and c_p the heat capacity at constant pressure, does not vary strongly from a constant value, the weak stratification allowing for deep vertical motion. We define the parameter

$$\epsilon \equiv \frac{\Delta\theta}{\Theta}, \tag{4.1.2}$$

where $\Delta\theta$ is the width of the range of values to which θ is presumed to be limited. The anelastic equations obtain in the limit of ϵ being small.

Consider again the Euler equations for an ideal gas on an equatorial β-plane (see (2.3.18)). We nondimensionalize the equations by introducing the characteristic potential temperature scale Θ and scaling (x, y) and z by the meridional domain half-length L and domain height H respectively. The parameter ϵ enters the equations through the time scale τ, defined by

$$\tau = \frac{L}{H} N^{-1} = \sqrt{\frac{L^2}{gH\epsilon}}, \tag{4.1.3}$$

where N is the Brunt-Väisällä frequency, given by

$$N^2 \equiv \frac{g}{\Theta}\frac{\Delta\theta}{H} = \frac{g}{H}\epsilon. \tag{4.1.4}$$

N is the frequency of vertically propagating internal gravity waves in a linearly stratified fluid. When ϵ is small, the gravity wave time scale is long. In particular, it is long compared to acoustic oscillations.

Pressure is replaced by the nondimensional Exner pressure,

$$\pi \equiv \left(\frac{p}{p_{00}}\right)^{\kappa}. \tag{4.1.5}$$

Expanding all variables in power series in ϵ:

$$\boldsymbol{v} = \boldsymbol{v}_0 + \epsilon \boldsymbol{v}_1 + \epsilon^2 \boldsymbol{v}_2 + \ldots, \tag{4.1.6a}$$

$$\pi = \pi_0 + \epsilon \pi_1 + \epsilon^2 \pi_2 + \ldots, \tag{4.1.6b}$$

$$\theta = 1 + \epsilon \theta_1 + \epsilon^2 \theta_2 + \ldots, \tag{4.1.6c}$$

where $\boldsymbol{v}$ and θ are dimensionless, and substituting into the nondimensional Euler equations yields, to $\mathcal{O}(\epsilon)$, the anelastic equations of Ogura and Phillips. We will use instead a modified form of the equations (Lipps and Hemler, 1982), in which the $\mathcal{O}(1)$ potential temperature has a prescribed $\mathcal{O}(\epsilon)$ variation in z. That is,

$$\begin{aligned} \theta &= 1 + \epsilon(\bar{\theta_1}(z) + \theta_1) + \epsilon^2 \theta_2 + \ldots \\ &\equiv \theta_0(z) + \epsilon \theta_1 + \epsilon^2 \theta_2 + \ldots, \end{aligned} \tag{4.1.7}$$

where θ_1 has been redefined from (4.1.6c) implicitly to be the departure of the potential temperature perturbation from the specified profile $\bar{\theta_1}(z)$. The classical anelastic equations are recovered by setting $\theta_0(z) \equiv 1$.

The *modified* anelastic equations are

$$\frac{D_0 u_0}{Dt} - \beta_\delta y v_0 + \gamma_\alpha w_0 + \frac{1}{B}\theta_0 \frac{\partial \pi_1}{\partial x} = 0, \tag{4.1.8a}$$

$$\frac{D_0 v_0}{Dt} + \beta_\delta y u_0 + \frac{1}{B}\theta_0 \frac{\partial \pi_1}{\partial y} = 0, \tag{4.1.8b}$$

$$\frac{D_0 w_0}{Dt} + \alpha^2 \left\{ -\gamma_\alpha u_0 + \frac{1}{B}\left[\frac{\partial}{\partial z}(\theta_0 \pi_1) + \frac{\mathrm{d}\pi_0}{\mathrm{d}z}\theta_1\right]\right\} = 0, \tag{4.1.8c}$$

$$\frac{D_0 \theta_1}{Dt} + \frac{w_0}{\epsilon}\frac{\mathrm{d}\theta_0}{\mathrm{d}z} = 0, \tag{4.1.8d}$$

$$\frac{\partial}{\partial x}(\rho_0 u_0) + \frac{\partial}{\partial y}(\rho_0 v_0) + \frac{\partial}{\partial z}(\rho_0 w_0) = 0, \tag{4.1.8e}$$

where all variables are dimensionless, and

$$B \equiv \frac{Hg}{c_p \Theta}. \tag{4.1.9}$$

B may be called the *domain thickness* parameter. It is the ratio of the domain height to the maximum height of an atmosphere with flat bottom topography at $z = 0$ and uniform potential temperature Θ in hydrostatic balance. The nondimensional Coriolis parameters in (4.1.8) are defined by

$$\beta_\delta \equiv \frac{1}{S\delta}, \qquad \gamma_\alpha \equiv \frac{1}{S\alpha}, \tag{4.1.10}$$

where

$$S \equiv \frac{H}{L}\left(\frac{N}{2\Omega}\right), \qquad \alpha \equiv \frac{L}{H}, \qquad \delta \equiv \frac{a}{L}. \tag{4.1.11}$$

α is the aspect ratio of the flow, and δ is the planetary aspect ratio. S is a form of the Burger number and is a measure of the assumed maximum stratification.

The $\mathcal{O}(1)$ pressure and density fields π_0 and ρ_0 are functions of z only and are determined by the hydrostatic balance relation

$$B + \theta_0 \frac{\mathrm{d}\pi_0}{\mathrm{d}z} = 0, \tag{4.1.12}$$

and the $\mathcal{O}(\epsilon)$ pressure and density fields π_1 and ρ_1 are obtained by requiring that the $\mathcal{O}(1)$ continuity equation (4.1.8e) is preserved with time, i.e. by solving the elliptic equation

$$\frac{\partial}{\partial t}\left[\frac{\partial}{\partial x}(\rho_0 u_0) + \frac{\partial}{\partial y}(\rho_0 v_0) + \frac{\partial}{\partial z}(\rho_0 w_0)\right] = 0. \tag{4.1.13}$$

The system (4.1.8) conserves the energy integral

$$E = \iiint \rho_0 \left\{\frac{1}{2}\left[u_0^2 + v_0^2 + \frac{1}{\alpha^2} w_0^2\right] + \frac{1}{B}\pi_0\theta_1)\right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.1.14}$$

Note that E is not the $\mathcal{O}(\epsilon)$ approximation to the energy conserved by the Euler equations. It is missing the *elastic energy* term proportional to the integral of pressure fluctuations. Hence the name *anelastic* equations. See Appendix 4.A for a derivation of the anelastic equations and their energetics.

4.1.2 Symmetric equations

We will soon return to the problem of the stability of a zonally symmetric zonal flow to zonally symmetric perturbations but in the context of the anelastic system. In preparation for that analysis, we present a Hamiltonian representation of the symmetric anelastic equations. Recall that the underlying symmetries and conservation laws on which the energy-Casimir method is based are connected with the Hamiltonian structure of the equations.

The symmetric version of the equations is obtained by setting $\partial_x \equiv 0$ in (4.1.8). In terms of Eulerian time derivatives, the symmetric equations are

$$(u_0)_t = -v_0(u_0)_y - w_0(u_0)_z + \beta_\delta y v_0 - \gamma_\alpha w_0, \tag{4.1.15a}$$

$$(v_0)_t = -v_0(v_0)_y - w_0(v_0)_z - \beta_\delta y u_0 - \frac{1}{B}\theta_0(\pi_1)_y, \tag{4.1.15b}$$

$$\begin{aligned}(w_0)_t &= -v_0(w_0)_y - w_0(w_0)_z \\ &\quad + \alpha^2 \left\{ -\gamma_\alpha u_0 + \frac{1}{B}\left[(\theta_0\pi_1)_z + \frac{\mathrm{d}\pi_0}{\mathrm{d}z}\theta_1 \right] \right\},\end{aligned} \tag{4.1.15c}$$

$$(\theta_1)_t = -v_0(\theta_1)_y - w_0(\theta_1)_z - \frac{w_0}{\epsilon}\frac{\mathrm{d}\theta_0}{\mathrm{d}z}, \tag{4.1.15d}$$

$$(\rho_0 v_0)_y + (\rho_0 w_0)_z = 0, \tag{4.1.15e}$$

where Latin subscripts denote partial differentiation.

Hamiltonian representation

Observe that in the symmetric version of the equations, two fields are materially conserved — the first order potential temperature,

$$\theta \equiv \theta_0(z) + \epsilon\theta_1, \tag{4.1.16}$$

and the component of absolute angular momentum associated with motion in the x-direction

$$m \equiv u_0 - \tfrac{1}{2}\beta_\delta y^2 + \gamma_\alpha z. \tag{4.1.17}$$

The 2-D continuity equation (4.1.15e) invites the definition of a vertical-meridional stream function $\psi(y, z, t)$ defined implicitly by

$$v_0 = \frac{1}{\rho_0}\frac{\partial \psi}{\partial z}, \qquad w_0 = -\frac{1}{\rho_0}\frac{\partial \psi}{\partial y}. \tag{4.1.18}$$

Lastly, we introduce the x-component of the relative vorticity

$$\zeta \equiv \frac{1}{\alpha^2}\frac{\partial w_0}{\partial y} - \frac{\partial v_0}{\partial z}. \tag{4.1.19}$$

Using (4.1.18),

$$\zeta = -\left[\frac{1}{\alpha^2}\frac{\partial}{\partial y}\left(\frac{1}{\rho_0}\frac{\partial \psi}{\partial y}\right) + \frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial \psi}{\partial z}\right)\right]. \tag{4.1.20}$$

The equations (4.1.15) can be written in terms of the new variables as

$$m_t = \frac{1}{\rho_0}\partial(\psi, m), \tag{4.1.21a}$$

$$\begin{aligned} \zeta_t = {} & \partial\left(\psi, \frac{1}{\rho_0}\zeta\right) + \partial\left[\frac{1}{\rho_0}\left(\frac{1}{\epsilon B}\rho_0\pi_0\right), \theta\right] \\ & + \partial\left[\frac{1}{\rho_0}\left[\rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)\right], m\right], \end{aligned} \tag{4.1.21b}$$

$$\theta_t = \frac{1}{\rho_0}\partial(\psi, \theta). \tag{4.1.21c}$$

We now show that the system (4.1.21) can be written in the Hamiltonian form

$$\frac{\partial \mathbf{x}}{\partial t} = \mathcal{J}\frac{\delta \mathcal{H}}{\delta \mathbf{x}}, \tag{4.1.22}$$

where $\mathbf{x} \equiv (m, \zeta, \theta)^T$ is the generalized independent variable, $\mathcal{J}$ is the Poisson tensor, and $\mathcal{H}$ is the Hamiltonian. In this case, the Hamiltonian is

$$\begin{aligned} \mathcal{H} &= E' + \text{constant} \\ &= \iint \left\{\rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)m + \frac{1}{2\rho_0}\left[\left(\frac{\partial \psi}{\partial z}\right)^2 + \frac{1}{\alpha^2}\left(\frac{\partial \psi}{\partial y}\right)^2\right] + \frac{1}{\epsilon B}\rho_0\pi_0\theta\right\} \mathrm{d}y\,\mathrm{d}z \end{aligned} \tag{4.1.23}$$

where integration over the horizontal coordinate x has been suppressed because of the symmetry in that direction. $\mathcal{H}$ is simply the total energy (4.1.14) plus a constant (a Casimir invariant, to be defined below). $\delta\mathcal{H}/\delta\mathbf{x}$ is the functional gradient of $\mathcal{H}$, defined in terms of the inner product

$$\left(\frac{\delta \mathcal{H}}{\delta \mathbf{x}}, \mathbf{w}\right) \equiv \iint \left(\frac{\delta \mathcal{H}}{\delta \mathbf{x}} \cdot \mathbf{w}\right) \mathrm{d}y\,\mathrm{d}z \equiv \left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0} \mathcal{H}(\mathbf{x} + \varepsilon\mathbf{w}), \tag{4.1.24}$$

where the "direction" $\mathbf{w}$ is a vector in the same space as $\mathbf{x}$. Note that $\delta\mathcal{H}/\delta\mathbf{x}$ is independent of $\mathbf{w}$. The functional gradient of $\mathcal{H}$ is calculated in Appendix 4.B. It has components

$$\frac{\delta\mathcal{H}}{\delta m} = \rho_0(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z), \qquad \frac{\delta\mathcal{H}}{\delta\zeta} = \psi, \qquad \frac{\delta\mathcal{H}}{\delta\theta} = \frac{1}{\epsilon B}\rho_0\pi_0. \tag{4.1.25}$$

Comparison of (4.1.25) with (4.1.21) shows that

$$\begin{aligned} m_t &= \frac{1}{\rho_0}\partial\left(\frac{\delta\mathcal{H}}{\delta\zeta}, m\right), &(4.1.26a)\\ \zeta_t &= \partial\left(\frac{1}{\rho_0}\frac{\delta\mathcal{H}}{\delta m}, m\right) + \partial\left(\frac{\delta\mathcal{H}}{\delta\zeta}, \frac{1}{\rho_0}\zeta\right) + \partial\left(\frac{1}{\rho_0}\frac{\delta\mathcal{H}}{\delta\theta}, \theta\right), &(4.1.26b)\\ \theta_t &= \frac{1}{\rho_0}\partial\left(\frac{\delta\mathcal{H}}{\delta\zeta}, \theta\right), &(4.1.26c)\end{aligned}$$

which is in the form (4.1.22) if we define

$$\mathcal{J} = \begin{bmatrix} 0 & \rho^{-1}\partial(\cdot, m) & 0 \\ \partial(\rho^{-1}\cdot, m) & \partial(\cdot, \rho^{-1}\zeta) & \partial(\rho^{-1}\cdot, \theta) \\ 0 & \rho^{-1}\partial(\cdot, \theta) & 0 \end{bmatrix}. \tag{4.1.27}$$

Casimirs

Hamiltonian representations of fluid systems are typically noncanonical. That is to say, the state of the system is not specified by the canonical positions and momenta characteristic of Hamiltonian particle mechanics. In the present case, the dynamics of the velocity related variables m and ζ and the potential temperature θ are governed by the equations (4.1.26) which do not refer to positions of fluid parcels. If positions of fluid parcels are assigned at a given time, then they can be recovered, in principle, by integrating the solution of (4.1.26) forward and backward in time. This separability is a consequence of the so called parcel relabelling symmetry. The parcel relabelling symmetry is associated with the material conservation of potential vorticity q (Salmon, 1982), defined by

$$q = \frac{(\boldsymbol{\Omega} + \nabla \times \boldsymbol{v}_0)\cdot\nabla\theta}{\rho_0}, \tag{4.1.28}$$

where $\boldsymbol{\Omega} \equiv (0, \gamma_\alpha, \beta_\delta y)$ is the local planetary rotation vector. Using $\partial_x \equiv 0$, we find the symmetric, nondimensional version of q to be

$$\begin{aligned} q &= \frac{1}{\rho_0}\left[\left(\frac{\partial u_0}{\partial z} + \gamma_\alpha\right)\frac{\partial \theta}{\partial y} - \left(\frac{\partial u_0}{\partial y} - \beta_\delta y\right)\frac{\partial \theta}{\partial z}\right] \\ &= \frac{1}{\rho_0}\partial(\theta, m). \end{aligned} \tag{4.1.29}$$

It may be verified that

$$q_t = \frac{1}{\rho_0}\partial(\psi, q), \tag{4.1.30}$$

i.e. that q is indeed a material invariant. In the symmetric equations, there are relabelling type symmetries associated also with m and with θ. Noether's theorem associates conserved functionals with continuous symmetries. The conserved functional associated wih the relabelling symmetries takes the form

$$\mathcal{C}_1(m, \theta) = \iint \rho_0 C_1(m, \theta, q)\, \mathrm{d}y\, \mathrm{d}z, \tag{4.1.31}$$

where $C_1(m, \theta, q)$ is an arbitrary function. Functionals of the form

$$\mathcal{C}(m, \theta) = \iint \rho_0 C(m, \theta)\, \mathrm{d}y\, \mathrm{d}z, \tag{4.1.32}$$

where $C(m, \theta)$ is an arbitrary differentiable function, are *Casimirs* of the system[1], meaning that they satisfy

$$\mathcal{J}\frac{\delta \mathcal{C}}{\delta \mathbf{x}} = \mathbf{0}, \tag{4.1.33}$$

i.e. that the gradients of $\mathcal{C}$ are in the nullspace of the operator $\mathcal{J}$. The components of the functional gradient of $\mathcal{C}$ are

$$\frac{\delta \mathcal{C}}{\delta m} = \rho_0 \frac{\partial C}{\partial m}, \qquad \frac{\delta \mathcal{C}}{\delta \zeta} = 0, \qquad \frac{\delta \mathcal{C}}{\delta \theta} = \rho_0 \frac{\partial C}{\partial \theta}, \tag{4.1.34}$$

from which it may be verified that $\mathcal{C}$ satisfies (4.1.33).

[1] In a canonical formulation of the equations, $\mathcal{C}$ would be a conserved functional associated with the relabelling symmetries through Noether's theorem. In the noncanonical representation, the symmetry associated with $\mathcal{C}$ is a transformation of a variable that has been reduced from the system, namely parcel labels. Conserved functionals associated with symmetries in variables which have been reduced from the system become Casimirs.

The condition (4.1.33) is not satisfied if the Casimir density function depends on q (as in (4.1.31)) unless the space of variations is restricted such that $\delta m = \delta\theta = 0$ on the boundary, or the boundary conditions chosen such that $\nabla m = \nabla\theta = 0$ on the boundary. Neither choice is natural for our purposes. In spite of that fact, for the nonlinear stability calculation in Subsection 4.2.2, we use a functional that depends on q. The only property that is essential for our purpose is that $\mathrm{d}\mathcal{C}_1/\mathrm{d}t = 0$.

4.2 Symmetric stability of zonal equilibrium

In this section, we repeat the calculation of the previous chapter, but because of the simplification of the anelastic approximation, we make somewhat more progress. Once again, we investigate the stability of a zonally symmetric zonal flow at the equator to zonally symmetric perturbations in the velocity and potential temperature fields.

Consider a steady solution of the system (4.1.21) in the domain $\mathcal{D} = \{(y,z)| -1 \leq y \leq 1, 0 \leq z \leq 1\}$, with $\mathbf{x} = \mathbf{X} = (M(y,z), 0, \Theta(y,z))^T$ and associated potential vorticity $Q = (1/\rho_0)\partial(\Theta, M)$. The velocity is strictly in the x-direction, so $\zeta = 0$, and $\psi =$ constant, which we may take to be zero.[1]

$\zeta_t = 0$ and (4.1.21b) imply the condition

$$\left(\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\right)\frac{\partial\Theta}{\partial y} - \partial\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z, M\right) = 0, \tag{4.2.1}$$

known as thermal wind balance because the gradient of the velocity field balances the meridional gradient of the temperature field.

Since $\mathbf{X}$ is a steady solution to a system of the form (4.1.22), it satisfies

$$\left.\left[\mathcal{J}\,\frac{\delta\mathcal{H}}{\delta\mathbf{x}}\right]\right|_{\mathbf{x}=\mathbf{X}} = 0, \tag{4.2.2}$$

which is to say that $\delta\mathcal{H}/\delta\mathbf{x}|_{\mathbf{x}=\mathbf{X}}$ is in the nullspace of $\mathcal{J}|_{\mathbf{x}=\mathbf{X}}$. Therefore, by definition,

[1] Note that the steady state potential temperature $\Theta(y,z)$ is not related to the potential temperature scaling factor Θ introduced in the derivation of the anelastic system.

there is a Casimir $\mathcal{C}$ such that

$$\left.\frac{\delta\mathcal{H}}{\delta\mathbf{x}}\right|_{\mathbf{x}=\mathbf{X}} = -\left.\frac{\delta\mathcal{C}}{\delta\mathbf{x}}\right|_{\mathbf{x}=\mathbf{X}}. \tag{4.2.3}$$

Comparison of (4.1.25) with (4.1.34) at $\mathbf{X}$ gives the following conditions on the partial derivatives of the function $C(m,\theta)$ so that $\mathcal{C}$ satisfies (4.2.3):

$$\rho_0 C_m|_X = -\rho_0(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z), \tag{4.2.4a}$$

$$\rho_0 C_\theta|_X = -\frac{1}{\epsilon B}\rho_0\pi_0. \tag{4.2.4b}$$

The functional $(\mathcal{H}+\mathcal{C})(m,\zeta,\theta)$ then has a critical "point" at $\mathbf{x}=\mathbf{X}$.[1]

Note that the right hand sides of (4.2.4) are in terms of y and z, not M and Θ. We might be tempted to assume that there is a one-to-one inverse of the mapping from (y,z) to $[M(y,z),\Theta(y,z)]$, which we might write as $[Y(M,\Theta),Z(M,\Theta)]$. However, for a realistic basic state, there is not a one-to-one inverse. Recall that the Rayleigh criterion requires that $M(y,z)$ have a maximum at $y=0$. Evidently, a single function $C(m,\theta)$ cannot typically satisfy (4.2.4) at all points (y,z).

We address this problem in slightly different ways for proving *linear* and *nonlinear* stability of $\mathbf{X}$. For linear stability, we partition the domain into regions in which $Q(y,z)\neq 0$ and define a functional $\mathcal{C}_L$ which is the sum of functionals satisfying (4.2.3) in each region, as we did in the Chapter 3. Using $\mathcal{H}$ and $\mathcal{C}_L$, we construct a functional $\mathcal{H}_L$ which is conserved by the linearized equations. For nonlinear stability, we derive criteria for Lyapunov stability (to be defined later) only for those basic states which satisfy the linear stability criteria and are even functions of y by defining a functional of the form (4.1.31) whose integrand depends on q as well as on m and θ. For basic states which are not even symmetric functions of y but satisfy the linear stability criteria, we show that the growth of disturbances is bounded but not arbitrarily close to the basic state. The latter

[1] The Hamiltonian of a canonical system will itself have a critical point at a steady solution. This is a consequence of the operator $\mathcal{J}$ being invertible for canonical systems.

is an unexpected result; in most similar problems, the linear stability conditions are also sufficient for nonlinear stability.

4.2.1 Linear stability

To find conditions for linear stability, we construct a functional that is conserved by the linearized equations and vanishes at the basic state. We identify conditions for the functional to be positive definite with conditions for the stability of the basic state solution.

The linearized equations are

$$m'_t = \frac{1}{\rho_0}\partial(\psi', M), \tag{4.2.5a}$$

$$\zeta'_t = \partial\left[\frac{1}{\rho_0}\left(\frac{1}{\epsilon B}\rho_0\pi_0\right), \theta'\right] + \partial\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), m'\right], \tag{4.2.5b}$$

$$\theta'_t = \frac{1}{\rho_0}\partial(\psi', \Theta), \tag{4.2.5c}$$

where primed quantities represent departures from the basic state $\mathbf{X}$.

As in Chapter 3, we assume that the mapping from (y, z) to (M, Θ) has nonzero Jacobian everywhere in $\mathcal{D}$ except perhaps on a finite number of curves, and partition $\mathcal{D}$ into a finite number of subregions $\mathcal{D}^{(i)}$, $i = 1, \ldots, n$, such that

$$Q \equiv \frac{1}{\rho_0}\partial(\Theta, M) \neq 0 \tag{4.2.6}$$

inside each of the $\mathcal{D}^{(i)}$ (see Figure 3.1). Repeating the procedure from Chapter 3, we define the functional

$$\mathcal{C}_L = \sum_{i=1}^{n} \iint_{\mathcal{D}^{(i)}} \rho_0 C^{(i)}(m, \theta)\,\mathrm{d}y\,\mathrm{d}z, \tag{4.2.7}$$

where each of the $C^{(i)}$ are arbitrary twice-differentiable functions of m and θ. We observe that $\mathcal{C}_L$ is *not* conserved by the equations (4.1.21) but that $\mathrm{d}\mathcal{C}_L/\mathrm{d}t|_{\mathbf{x}=\mathbf{X}} = 0$, so we expect $\mathcal{C}_L$ to be relevant for small amplitude perturbations to the steady state $\mathbf{X}$.

Again, we choose the functions $C^{(i)}$ so that $\mathcal{C}_L$ is tangent to $\mathcal{H}$ at the basic state. That is

$$\delta(\mathcal{H}+\mathcal{C}_L)|_{\mathbf{X}} = 0, \tag{4.2.8}$$

where $\delta(\mathcal{H}+\mathcal{C}_L)$ is the first variation of $(\mathcal{H}+\mathcal{C}_L)$. For arbitrary (m,ζ,θ),

$$\begin{aligned}\delta(\mathcal{H}+\mathcal{C}_L) &= \iint\limits_{\mathcal{D}} \left\{\rho_0(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z)\delta m + \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}\psi_y\delta(\psi_y) + \psi_z\delta(\psi_z)\right] + \frac{1}{\epsilon B}\rho_0\pi_0\delta\theta\right\}\mathrm{d}y\,\mathrm{d}z \\ &+ \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}} \rho_0\left(C_m^{(i)}\delta m + C_\theta^{(i)}\delta\theta\right)\mathrm{d}y\,\mathrm{d}z,\end{aligned} \tag{4.2.9}$$

so at $\mathbf{X}$,

$$\begin{aligned}\delta(\mathcal{H}+\mathcal{C}_L)|_{\mathbf{X}} &= \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}} \rho_0\left\{\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right) + C_m^{(i)}(M,\Theta)\right]\delta m\right. \\ &\left. + \left[\frac{1}{\epsilon B}\pi_0 + C_\theta^{(i)}(M,\Theta)\right]\delta\theta\right\}\mathrm{d}y\,\mathrm{d}z.\end{aligned} \tag{4.2.10}$$

Hence $\delta(\mathcal{H}+\mathcal{C}_L)$ vanishes at $\mathbf{X}$ if

$$C_m^{(i)}(M,\Theta) = -\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), \quad \text{and} \quad C_\theta^{(i)}(M,\Theta) = -\frac{1}{\epsilon B}\pi_0 \tag{4.2.11}$$

for each i. This is well defined because $Q \neq 0$ within each $\mathcal{D}^{(i)}$, so y and z can be expressed as functions of M and Θ. Note that (4.2.11) implies that at $\mathbf{X}$, the first partial derivatives of the functions $C^{(i)}$ in adjacent regions match on their common boundaries.

We now construct a quadratic invariant for the linearized equations based on the second variation of $(\mathcal{H}+\mathcal{C}_L)$. The second variation evaluated at $\mathbf{X}$ is

$$\begin{aligned}\delta^2(\mathcal{H}+\mathcal{C}_L)\big|_{\mathbf{X}} &= \iint\limits_{\mathcal{D}} \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}(\delta\psi_y)^2 + (\delta\psi_z)^2\right]\mathrm{d}y\,\mathrm{d}z \\ &+ \sum_{i=1}^{n}\iint\limits_{\mathcal{D}^{(i)}} \rho_0\left[C_{mm}^{(i)}(M,\Theta)(\delta m)^2 + 2C_{\theta m}^{(i)}(M,\Theta)\delta\theta\delta m + C_{\theta\theta}^{(i)}(M,\Theta)(\delta\theta)^2\right]\mathrm{d}y\,\mathrm{d}z.\end{aligned} \tag{4.2.12}$$

The first integral is obviously strictly positive for any perturbation $\delta\psi$, and the second integral is positive if for all i, the matrix

$$\Lambda^{(i)} = \begin{bmatrix} C_{mm}^{(i)} & C_{m\theta}^{(i)} \\ C_{\theta m}^{(i)} & C_{\theta\theta}^{(i)} \end{bmatrix} \tag{4.2.13}$$

is positive definite at all points $(y, z) = [Y(M, \Theta), Z(M, \Theta)]$.

That $\delta^2(\mathcal{H} + \mathcal{C}_L)$ is positive definite does not obviously imply that $\mathbf{X}$ is (even linearly) stable because $\mathcal{C}_L$ is not conserved by the nonlinear dynamics (4.1.21). However, identifying the primed variables of (4.2.5) with the variations in (4.2.12), we find that

$$\begin{aligned} \mathcal{H}_L \;=\;& \iint_{\mathcal{D}} \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}(\psi_y')^2 + (\psi_z')^2\right] \mathrm{d}y\,\mathrm{d}z \\ &+ \sum_{i=1}^{n} \iint_{\mathcal{D}^{(i)}} \rho_0\left[C_{mm}^{(i)}(M,\Theta)m'^2 + 2C_{\theta m}^{(i)}(M,\Theta)\theta' m' + C_{\theta\theta}^{(i)}(M,\Theta)\theta'^2\right] \mathrm{d}y\,\mathrm{d}z \end{aligned} \tag{4.2.14}$$

is conserved by the linearized system (4.2.5). To verify that $\mathcal{H}_L$ is conserved, we need to calculate the second partial derivatives of the $C^{(i)}$:

$$C_{mm}^{(i)}(M,\Theta) = -\frac{\partial}{\partial M}\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)\right] = -\frac{1}{\rho_0 Q}\partial\left[\Theta, \left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)\right], \tag{4.2.15a}$$

$$C_{m\theta}^{(i)}(M,\Theta) = -\frac{\partial}{\partial \Theta}\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right) = -\frac{1}{\rho_0 Q}\partial\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), M\right], \tag{4.2.15b}$$

$$C_{\theta m}^{(i)}(M,\Theta) = \frac{\partial}{\partial M}\left(-\frac{\pi_0}{\epsilon B}\right) = -\frac{1}{\rho_0 Q}\partial\left(\Theta, \frac{\pi_0}{\epsilon B}\right), \tag{4.2.15c}$$

$$C_{\theta\theta}^{(i)}(M,\Theta) = \frac{\partial}{\partial \Theta}\left(-\frac{\pi_0}{\epsilon B}\right) = -\frac{1}{\rho_0 Q}\partial\left(\frac{\pi_0}{\epsilon B}, M\right), \tag{4.2.15d}$$

where partial derivatives with respect to M are taken with Θ fixed, and vice versa. See Appendix 4.C for the calculation showing $\mathrm{d}\mathcal{H}_L/\mathrm{d}t = 0$.

We notice that the contribution to $\mathcal{H}_L$ from each $\mathcal{D}^{(i)}$ is quadratic in m' and θ' and hence that $\mathcal{H}_L$ is positive definite if all of the matrices $\Lambda^{(i)}$ are positive definite. It is an easy matter to define a norm on the space of disturbances (m', ζ', θ') and identify Lyapunov stability of the linearized equations with positive definiteness and boundedness of $\mathcal{H}_L$. Let

$$||\mathbf{x}'||^2_{(L)} = \iint\limits_{\mathcal{D}} \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}(\psi'_y)^2 + (\psi'_z)^2\right] \mathrm{d}y\,\mathrm{d}z + \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \rho_0\lambda^{(i)}\left[m'^2 + \theta'^2\right] \mathrm{d}y\,\mathrm{d}z \quad (4.2.16)$$

where $\lambda^{(i)}$ is the minimum of the eigenvalues of $\Lambda^{(i)}$ for all of the values of M and Θ inside $\mathcal{D}^{(i)}$. Since $\mathcal{H}_L$ is conserved, it follows that

$$||\mathbf{x}'(t)||^2_{(L)} \le \mathcal{H}_L[\mathbf{x}'(t)] = \mathcal{H}_L[\mathbf{x}'(0)] \le \frac{\lambda_+}{\lambda_-}||\mathbf{x}'(\mathbf{0})||^2_{(L)}, \quad (4.2.17)$$

where λ_+ is the maximum and λ_- the minimum among all of the eigenvalues of the $\Lambda^{(i)}$. To derive (4.2.17), we have used the fact that the $\Lambda^{(i)}$ are symmetric matrices and hence that any pair (m', θ') can be expressed as a linear combination of orthogonal eigenvectors of any of the $\Lambda^{(i)}$. The condition that the $\Lambda^{(i)}$ be positive definite is equivalent to

$$C^{(i)}_{mm}(M,\Theta) = -\frac{1}{\rho_0 Q}\partial\left(\Theta, \tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right) \quad > \quad 0, \quad (4.2.18a)$$

$$C^{(i)}_{\theta\theta}(M,\Theta) = -\frac{1}{\rho_0 Q}\partial\left(\frac{\pi_0}{\epsilon B}, M\right) = \frac{1}{\rho_0 Q}\left(\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\right) M_y \quad > \quad 0, \quad (4.2.18b)$$

$$C^{(i)}_{mm}(M,\Theta)C^{(i)}_{\theta\theta}(M,\Theta) - {C^{(i)}_{m\theta}}^2(M,\Theta) = \left(-\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\right)\frac{\beta_\delta y}{\rho_0 Q} \quad > \quad 0. \quad (4.2.18c)$$

In general, from (4.1.12), $\mathrm{d}\pi_0/\mathrm{d}z < 0$ (pressure decreases with height). The condition (4.2.18c) is the familiar symmetric stability condition that potential vorticity be positive in the northern hemisphere and negative in the southern for stability. (4.2.18b) is the Rayleigh condition applied to flow in a spherical shell, namely that angular momentum decrease in magnitude with the absolute value of latitude. (4.2.18a) is a generalization of static stability, stating that for stability, the potential temperature gradient must be clockwise of the gradient of planetary angular momentum $M^{(\mathrm{p})} \equiv -\frac{1}{2}\beta_\delta y^2 + \gamma_\alpha z$ in the northern hemisphere and anticlockwise in the southern. Far from the equator, where $\frac{1}{2}\beta_\delta y^2 \gg \gamma_\alpha z$, this reduces to the usual static stability condition $\Theta_z > 0$.

Conditions (4.2.18a) and (4.2.18c) are formally identical with (3.1.51) and (3.1.47), the corresponding conditions derived for the Euler equations in Chapter 3, with the caveat that the definition of potential vorticity is slightly different in the anelastic system.

The inertial stability condition (4.2.18b), however, is fundamentally different from (3.1.46) in that it does not refer to pressure surfaces. Note that since static stability (4.2.18a) and "symmetric" stability (4.2.18c) together imply (4.2.18b) because of thermal wind balance (4.2.1), an essential difference must be contained in all three conditions.

The counterintuitive examples from Chapter 3 with surprising stability properties due to the curvature of pressure surfaces behave "normally" here. For example, a steady state of the form of $M = -y^4$ (see Example 4 in Chapter 3) is inertially stable everywhere with respect to the anelastic system, and the interval of instability of the steady state with $U = \lambda y$ (Example 2 in Chapter 3) grows linearly with λ even as λ gets very large (cf. Figure 3.7).

4.2.2 Nonlinear stability

To establish nonlinear stability using the energy-Casimir method, we must find a functional conserved by the nonlinear equations (4.1.21) which is bounded from above and below by disturbance norms (as mentioned above, $\mathcal{H}_L$ is not conserved by the full equations).

For Lyapunov stability under the nonlinear equations, a basic state must be stable under the linearized equations. Therefore, in this section we need only consider basic states satisfying the linear stability criteria.[1] In particular, assume that $Q = 0$ on the equator and $yQ > 0$ everywhere else in the domain. We now introduce the functional

$$\mathcal{C}_{NL}(m,\theta) \equiv \iint_{\mathcal{D}} \rho_0 \left\{ C^-(m,\theta) + H(q) \left[C^+(m,\theta) - C^-(m,\theta) \right] \right\} \, \mathrm{d}y \, \mathrm{d}z, \tag{4.2.19}$$

where

$$H(q) = \begin{cases} 0, & q < 0 \\ 1, & q \geq 0 \end{cases}, \tag{4.2.20}$$

[1] To exclude a solution from consideration, we would have to show linear *instability*, which is to say that there exists a solution to the linear equations that grows monotonically with time (with respect to some norm). We have not done that (see Ooyama, 1966, for an example of this type of argument).

and C^- and C^+ satisfy

$$\begin{aligned}
C_m^-(M,\Theta) &= -\left[\tfrac{1}{2}\beta_\delta\left[Y^-(M,\Theta)\right]^2 - \gamma_\alpha Z^-(M,\Theta)\right], && (4.2.21a)\\
C_\theta^-(M,\Theta) &= -\frac{1}{\epsilon B}\pi_0\left[Z^-(M,\Theta)\right], && (4.2.21b)\\
C_m^+(M,\Theta) &= -\left[\tfrac{1}{2}\beta_\delta\left[Y^+(M,\Theta)\right]^2 - \gamma_\alpha Z^+(M,\Theta)\right], && (4.2.21c)\\
C_\theta^+(M,\Theta) &= -\frac{1}{\epsilon B}\pi_0\left[Z^+(M,\Theta)\right], && (4.2.21d)
\end{aligned}$$

where (Y^-, Z^-) and (Y^+, Z^+) are the inverse functions defined by $[M(y,z), \Theta(y,z)]$ in the regions with $Q < 0$ and $Q \geq 0$ respectively. Note that we do not explicitly require that C^- and C^+ and their derivatives match along the curve $Q = 0$. Unlike the linear case, $\mathcal{C}_{NL}$ is conserved without the matching condition. As we will see, though, the matching condition is trivially satisfied by Lyapunov stable states, for which $C^- \equiv C^+$.

Now consider the *pseudoenergy* $\mathcal{A}(\mathbf{x};\mathbf{X})$, defined by

$$\begin{aligned}
\mathcal{A}(\mathbf{x};\mathbf{X}) &= (\mathcal{H}+\mathcal{C}_{NL})(m,\zeta,\theta) - (\mathcal{H}+\mathcal{C}_{NL})(M,0,\Theta)\\
&= \iint_{\mathcal{D}} \left\{ \frac{1}{2\rho_0}\left[\frac{1}{\alpha^2}(\psi_y)^2 + (\psi_z)^2\right]\right.\\
&\qquad + \rho_0\left[(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z)(m-M) + \frac{1}{\epsilon B}\pi_0(\theta-\Theta)\right.\\
&\qquad\quad + C^-(m,\theta) - C^-(M,\Theta)\\
&\qquad\quad + H(q)\left(C^+(m,\theta) - C^-(m,\theta)\right)\\
&\qquad\quad \left.\left. - H(Q)\left(C^+(M,\Theta) - C^-(M,\Theta)\right)\right]\right\} \mathrm{d}y\,\mathrm{d}z. \qquad (4.2.22)
\end{aligned}$$

By construction, $\mathcal{A}(\mathbf{X};\mathbf{X}) = 0$. We rewrite $\mathcal{A}$, using Taylor's Remainder Theorem, as

$$\begin{aligned}
C^-(m,\theta) - C^-(M,\Theta) &= \\
& C_m^-(M,\Theta)(m-M) + C_\theta^-(M,\Theta)(\theta-\Theta)\\
&+ \frac{1}{2}\left[C_{mm}^-(\tilde{m}^{(-)},\tilde{\theta}^{(-)})(m-M)^2 + 2C_{m\theta}^-(\tilde{m}^{(-)},\tilde{\theta}^{(-)})(m-M)(\theta-\Theta)\right.\\
&\qquad \left. + C_{\theta\theta}^-(\tilde{m}^{(-)},\tilde{\theta}^{(-)})(\theta-\Theta)^2\right], \qquad (4.2.23)
\end{aligned}$$

where $\tilde{m}^{(-)}(y,z,t) \in [M,m]$ and $\tilde{\theta}^{(-)}(y,z,t) \in [\Theta,\theta]$ (with a similar expansion for C^+). Inserting (4.2.23) into (4.2.22), and using (4.2.21),

$$
\begin{aligned}
\mathcal{A} \;=\; & \iint\limits_{\mathcal{D}} \left\{ \frac{1}{2\rho_0} \left[\frac{1}{\alpha^2}(\psi_y)^2 + (\psi_z)^2 \right] \right\} \mathrm{d}y\,\mathrm{d}z \\
& + \iint\limits_{\mathcal{D}^{(-)}} \left\{ \frac{1}{2}\rho_0 \left[C^-_{mm}(\tilde{m}^{(-)},\tilde{\theta})(m-M)^2 + 2C^-_{m\theta}(\tilde{m}^{(-)},\tilde{\theta}^{(-)})(m-M)(\theta-\Theta) \right.\right. \\
& \qquad \left.\left. + \, C^-_{\theta\theta}(\tilde{m}^{(-)},\tilde{\theta}^{(-)})(\theta-\Theta)^2 \right] + \rho_0 H(q) \left[C^+(m,\theta) - C^-(m,\theta) \right] \right\} \mathrm{d}y\,\mathrm{d}z \\
& + \iint\limits_{\mathcal{D}^{(+)}} \left\{ \frac{1}{2}\rho_0 \left[C^+_{mm}(\tilde{m}^{(+)},\tilde{\theta}^{(+)})(m-M)^2 + 2C^+_{m\theta}(\tilde{m}^{(+)},\tilde{\theta}^{(+)})(m-M)(\theta-\Theta) \right.\right. \\
& \qquad \left.\left. + \, C^+_{\theta\theta}(\tilde{m}^{(+)},\tilde{\theta}^{(+)})(\theta-\Theta)^2 \right] + \rho_0 [1-H(q)] \left[C^-(m,\theta) - C^+(m,\theta) \right] \right\} \mathrm{d}y\,\mathrm{d}z,
\end{aligned}
\tag{4.2.24}
$$

where $\mathcal{D}^{(-)}$ and $\mathcal{D}^{(+)}$ are the regions of $y<0$ ($Q<0$) and $y>0$ ($Q>0$) repectively. Note that outside of the ranges of $M(y,z)$ and $\Theta(y,z)$, C^- and C^+ can be defined arbitrarily without compromising (4.2.21), and because m and θ are materially conserved variables, we need only extend the definitions of C^- and C^+ to include the ranges of $m(y,z,t=0)$ and $\theta(y,z,t=0)$. If C^- can be extended such that (4.2.18) are satisfied for all values of m and θ accessible to the flow, then all terms in $\mathcal{A}$ are positive except those which depend on the extent to which C^- and C^+ have been mixed, and hence on the asymmetry of the basic state.[1]

The terms in (4.2.24) of indefinite sign are

$$
\begin{aligned}
\mathcal{A}_a \;=\; & \iint\limits_{\mathcal{D}^{(-)}} \rho_0 H(q) \left[C^+(m,\theta) - C^-(m,\theta) \right] \mathrm{d}y\,\mathrm{d}z \\
& + \iint\limits_{\mathcal{D}^{(+)}} \rho_0 [1-H(q)] \left[C^-(m,\theta) - C^+(m,\theta) \right] \mathrm{d}y\,\mathrm{d}z.
\end{aligned}
\tag{4.2.25}
$$

[1] More precisely, (4.2.18) must be satisfied by all $(\tilde{m},\tilde{\theta})$ in a convex region containing $\{\,\mathrm{ran}\,[M(y,z)] \cup \mathrm{ran}\,[m(y,z,t=0)]\,\} \times \{\,\mathrm{ran}\,[\Theta(y,z)] \cup \mathrm{ran}\,[\theta(y,z,t=0)]\,\}$, where the set $\{\mathrm{ran}\,F\}$ is the range of the function F, so that Taylor's theorem may be applied safely in (4.2.23).

While $\mathcal{A}_a$ is not sign definite, it is bounded from above and below. If C^- and C^+ can be extended outside of the ranges of $M(y,z)$ and $\Theta(y,z)$ in such a way that the difference $|C^- - C^+|$ is maximum within the basic state ranges, then we can write bounds on $\mathcal{A}_a$ which depend only on the basic state:

$$\mathcal{A}_a \geq -\max_{(M,\Theta)} \left|C^+(M,\Theta) - C^-(M,\Theta)\right| \left[2\int_0^H \rho_0 \mathrm{d}z\right], \tag{4.2.26a}$$

$$\mathcal{A}_a \leq \max_{(M,\Theta)} \left|C^+(M,\Theta) - C^-(M,\Theta)\right| \left[2\int_0^H \rho_0 \mathrm{d}z\right]. \tag{4.2.26b}$$

We will elaborate on the case of $\mathbf{X}$ not being an even function of y in Section 4.5. In the interim, the ideas will be made more concrete with Examples 3 and 4 below.

In the special case of $\mathbf{X}$ being even symmetric about $y = 0$, $C^- = C^+$ and $\mathcal{A}_a = 0$. We then define the norm on phase space displacements $\Delta\mathbf{x} = \mathbf{x} - \mathbf{X} = (m - M, \zeta, \theta - \Theta)$,

$$||\Delta\mathbf{x}||_\lambda^2 \equiv \iint_{\mathcal{D}} \left\{ \frac{1}{2\rho_0}\left[\frac{1}{\alpha^2}(\psi_y)^2 + (\psi_z)^2\right] + \lambda\frac{\rho_0}{2}\left[(m-M)^2 + (\theta-\Theta)^2\right]\right\} \mathrm{d}y\,\mathrm{d}z. \tag{4.2.27}$$

Let λ_- be the minimum of the eigenvalues of the matrix $\Lambda(m,\theta)$ (see (4.2.13)) over all possible values of m and θ (see the footnote on page 104). If $\Lambda(m,\theta)$ is positive definite over the extended domain, then λ_- is positive, and we have that

$$||\Delta\mathbf{x}(t)||_{\lambda_-}^2 \leq \mathcal{A}[\mathbf{x}(t)] = \mathcal{A}[\mathbf{x}(0)] \leq \frac{\lambda_+}{\lambda_-}||\Delta\mathbf{x}(0)||_{\lambda_-}^2, \tag{4.2.28}$$

where λ_+ is the maximum of the eigenvalues of $\Lambda(m,\theta)$. (4.2.28) implies that $||\Delta\mathbf{x}(t)||_\lambda$ is bounded for all t in terms of its initial value $||\Delta\mathbf{x}(0)||_\lambda$. In particular, $\mathbf{x}(t)$ can be bound as close to $\mathbf{X}$ as desired by choosing $\mathbf{x}(0)$ close enough to $\mathbf{X}$. Since the system is conservative, the initial time is arbitrary. If the solution *ever* passes within δ of equilibrium, then it will always be within ϵ. This is "stability in the sense of Lyapunov".

The statement (4.2.28) actually represents a stronger statement than just Lyapunov stability because it implies that the solution is bound near equilibrium no matter how *far* from equilibrium it starts. This fact will be used in Section 4.4 when we calculate saturation bounds on unstable equilibria.

4.2.3 Examples

In all of the following examples, we set $\theta_0(z) = 1$ (i.e. we use the classical anelastic approximation). All figures are plotted using $\beta_\delta = \gamma_\alpha = a = b = \epsilon = \Gamma = 1$, and at $\theta = 0$ (we focus on the m-dependence of the $C^{(i)}(m, \theta)$).

Example 1: Even basic state with quadratic $C(m, \theta)$

Consider the equilibrium state $\mathbf{X}_1$, with

$$M(y, z) = -\tfrac{1}{2} b y^2 \tag{4.2.29a}$$

$$\Theta(y, z) = (\epsilon\gamma)(\tfrac{1}{2} b y^2) + \epsilon \Gamma z, \tag{4.2.29b}$$

$$Q = \frac{1}{\rho_0}(\epsilon \Gamma b) y, \tag{4.2.29c}$$

where b and Γ are constants. Since $\mathbf{X}_1$ is an even function of y, we can define a single function $C(M, \Theta)$ for both $q > 0$ and $q < 0$. Inverting (4.2.29) gives

$$Y^2(M, \Theta) = -\frac{2M}{b}, \tag{4.2.30a}$$

$$Z(M, \Theta) = \frac{\gamma_\alpha}{\Gamma} M + \frac{1}{\epsilon\Gamma}\Theta, \tag{4.2.30b}$$

and applying (4.2.21), we find

$$C_m(M, \Theta) = \left(\frac{\beta_\delta}{b} + \frac{\gamma_\alpha^2}{\Gamma}\right) M + \frac{\gamma_\alpha}{\epsilon\Gamma}\Theta, \tag{4.2.31a}$$

$$C_\theta(M, \Theta) = -\frac{1}{\epsilon B} + \frac{\gamma_\alpha}{\epsilon\Gamma} M + \frac{1}{\epsilon^2 \Gamma}\Theta, \tag{4.2.31b}$$

and

$$C_{mm}(M, \Theta) = \frac{\beta_\delta}{b} + \frac{\gamma_\alpha^2}{\Gamma}, \tag{4.2.32a}$$

$$C_{\theta\theta}(M, \Theta) = \frac{1}{\epsilon^2 \Gamma}, \tag{4.2.32b}$$

$$C_{m\theta}(M, \Theta) = \frac{\gamma_\alpha}{\epsilon\Gamma}. \tag{4.2.32c}$$

Since all of the second derivatives are constants, we can safely extend the definition of $C(m, \theta)$ to apply to values outside of the ranges of $M(y, z)$ and $\Theta(y, z)$ without affecting the maximum and minimum values of the eigenvalues of Λ. Λ is positive definite if and only if $b > 0$ and $\Gamma > 0$, which agrees with our physical intuition: Θ must increase with z, and M must be maximum at the equator and decrease towards either pole.

Example 2: Even basic state with non-quadratic $C(m, \theta)$

Consider the equilibrium state $\mathbf{X}_2$, with

$$M(y,z) = -\tfrac{1}{2}by^2 - \tfrac{1}{4}ay^4, \tag{4.2.33a}$$

$$\Theta(y,z) = (\epsilon\gamma)(\tfrac{1}{2}by^2 + \tfrac{1}{4}ay^4) + \epsilon\Gamma z, \tag{4.2.33b}$$

$$Q = \frac{1}{\rho_0}(\epsilon\Gamma)(by + ay^3), \tag{4.2.33c}$$

$\mathbf{X}_2$ is also an even function of y, so we can define a single function $C(M, \Theta)$ for both $q > 0$ and $q < 0$. Inverting (4.2.33),

$$Y^2(M,\Theta) = \frac{1}{a}\sqrt{b^2 - 4aM} - \frac{b}{a}, \tag{4.2.34a}$$

$$Z(M,\Theta) = \frac{\gamma_\alpha}{\Gamma}M + \frac{1}{\epsilon\Gamma}\Theta, \tag{4.2.34b}$$

and we find

$$C_m(M,\Theta) = \frac{\beta_\delta b}{2a} - \frac{\beta_\delta}{2a}\sqrt{b^2 - 4aM} + \frac{\gamma_\alpha^2}{\Gamma}M + \frac{\gamma_\alpha}{\epsilon\Gamma}\Theta, \tag{4.2.35a}$$

$$C_\theta(M,\Theta) = -\frac{1}{\epsilon B} + \frac{\gamma_\alpha}{\epsilon\Gamma}M + \frac{1}{\epsilon^2\Gamma}\Theta, \tag{4.2.35b}$$

and

$$C_{mm}(M,\Theta) = \frac{\gamma_\alpha^2}{\Gamma} + \beta_\delta(b^2 - 4aM)^{-1/2}, \tag{4.2.36a}$$

$$C_{\theta\theta}(M,\Theta) = \frac{1}{\epsilon^2\Gamma}, \tag{4.2.36b}$$

$$C_{m\theta}(M,\Theta) = \frac{\gamma_\alpha}{\epsilon\Gamma}. \tag{4.2.36c}$$

Integrating (4.2.35) gives

$$C(M,\Theta) = \frac{\gamma_\alpha^2}{2\Gamma}M^2 + \frac{\beta_\delta b}{2a}M + \frac{\beta_\delta}{12a}(b^2-4aM)^{3/2} + \frac{\gamma_\alpha}{\epsilon\Gamma}\Theta M + \frac{\gamma_\alpha}{2\epsilon^2\Gamma}\Theta^2 - \frac{\beta_\delta b^3}{12a^2}. \quad (4.2.37)$$

Outside the ranges of $M(y,z)$ and $\Theta(y,z)$, we may define $C(m,\theta)$ in any way, provided it remains twice differentiable (this ensures that the Taylor expansion (4.2.23) remains valid). The simplest choice that does not change the bounds on the norm is to fix the second derivatives of $C(m,\theta)$ at their values on the perimeter of the domain. Hence, set

$$C_{mm}(m,\theta) = \begin{cases} \frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{a+b}, & m < -\frac{b}{2} - \frac{a}{4} \\ \frac{\gamma_\alpha^2}{\Gamma} + \beta_\delta(b^2-4am)^{-1/2}, & -\frac{b}{2} - \frac{a}{4} \le m < 0 \\ \frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{b}, & m \ge 0 \end{cases}, \quad (4.2.38a)$$

$$C_{\theta\theta}(m,\theta) = \frac{\gamma_\alpha}{\epsilon^2\Gamma}, \quad (4.2.38b)$$

$$C_{m\theta}(m,\theta) = \frac{\gamma_\alpha}{\epsilon\Gamma}, \quad (4.2.38c)$$

which corresponds to

$$C(m,\theta) = \begin{cases} \frac{1}{2}\left(\frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{a+b}\right)(m + \frac{b}{2} + \frac{a}{4})^2 \\ \quad + \left[-\frac{\beta_\delta}{2} - \frac{\gamma_\alpha^2}{\Gamma}(\frac{b}{2}+\frac{a}{4}) + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta\right](m + \frac{b}{2} + \frac{a}{4}) \\ \quad + \left[\frac{\gamma_\alpha^2}{2\Gamma}\left(\frac{b}{2}+\frac{a}{4}\right)^2 - \left(\frac{\beta_\delta b}{2a} + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta\right)\left(\frac{b}{2}+\frac{a}{4}\right) + \frac{\beta_\delta}{12a}(a+b)^3\right] \\ \quad + \frac{\gamma_\alpha}{2\epsilon^2\Gamma}\theta^2 - \frac{\beta_\delta b^3}{12a^2}, & m < -\frac{b}{2} - \frac{a}{4} \\ \frac{\gamma_\alpha^2}{2\Gamma}m^2 + \frac{\beta_\delta b}{2a}m + \frac{\beta_\delta}{12a}(b^2-4am)^{3/2} + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta m + \frac{\gamma_\alpha}{2\epsilon^2\Gamma}\theta^2 - \frac{\beta_\delta b^3}{12a^2}, & -\frac{b}{2} - \frac{a}{4} \le m < 0 \\ \left(\frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{b}\right)m^2 + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta m + \frac{\gamma_\alpha}{2\epsilon^2\Gamma}\theta^2, & m \ge 0 \end{cases}. \quad (4.2.39)$$

(see Figure 4.1). The eigenvalues of Λ in this case are

$$\lambda_\pm(C_{mm}) = \frac{1}{2}\left[\frac{1}{\epsilon^2\Gamma} + C_{mm} \pm \sqrt{\left(\frac{1}{\epsilon^2\Gamma}\right)^2 + 4\left(\frac{\gamma_\alpha}{\epsilon\Gamma}\right)^2 + C_{mm}^2 - 2\left(\frac{1}{\epsilon^2\Gamma}\right)C_{mm}}\right]. \quad (4.2.40)$$

C_{mm} is between $\frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{a+b}$ and $\frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{b}$ for all m. The smallest and largest eigenvalues can be found from (4.2.40), and an upper bound on the growth factor can be computed.

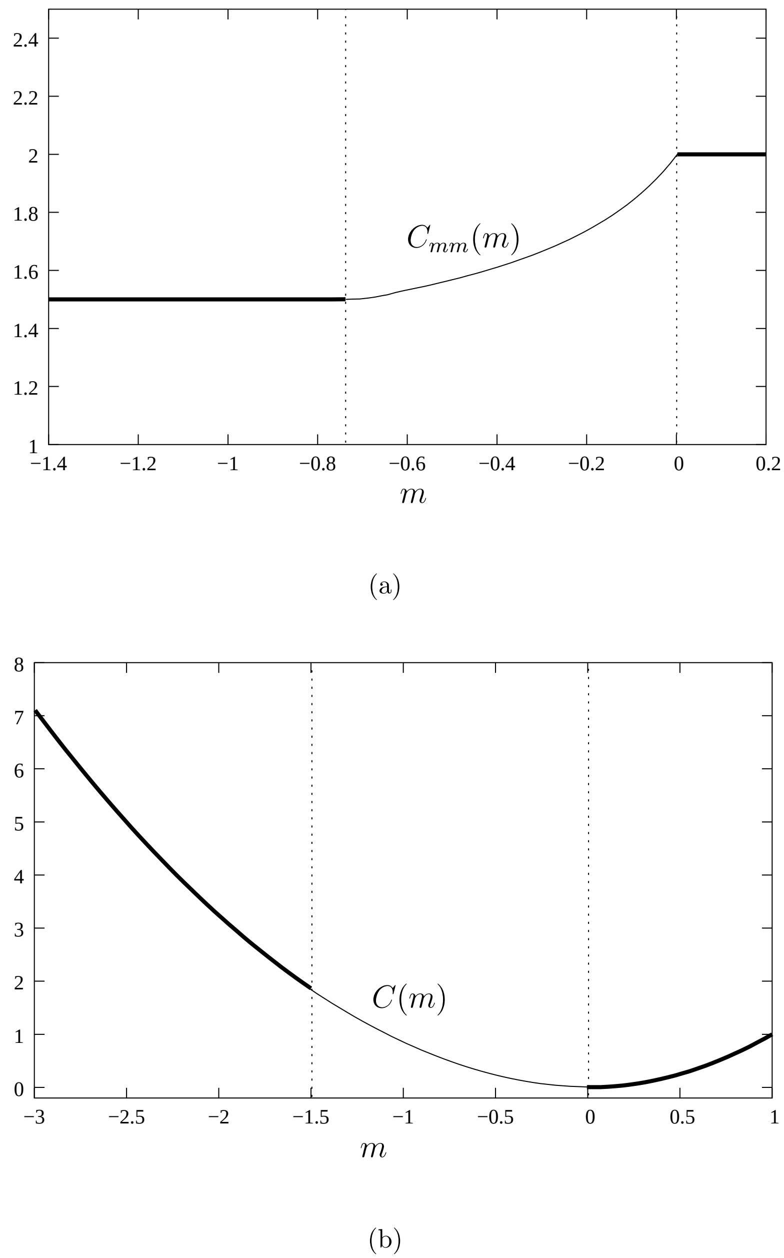


Figure 4.1: Casimir density function for Example 2: (a) C_{mm} over the full domain of all m. The heavy curves are the extensions beyond the range of $M(y, z)$. (b) $C(m, \theta)$ over the full domain of all m for $\theta = 0$. The heavy curves are the extensions beyond the range of $M(y, z)$.

Example 3: Asymmetric basic state

We illustrate the ideas related to asymmetric basic states with the following example, which is a combination of the first two. For this example we proceed using $\beta_\delta = \gamma_\alpha = \epsilon = 1$ so as to eliminate the complication of carrying all of the constants throughout the analysis.

Consider the basic state $\mathbf{X}_3$, with (Figure 4.2a)

$$M(y,z) = \begin{cases} -\frac{1}{2}y^2, & y < 0 \\ -\frac{1}{2}y^2 - \frac{1}{4}y^4, & y \geq 0 \end{cases}, \tag{4.2.41a}$$

$$\Theta(y,z) = \begin{cases} \frac{1}{2}y^2 + z, & y < 0 \\ \frac{1}{2}y^2 + \frac{1}{4}y^4 + z, & y \geq 0 \end{cases}, \tag{4.2.41b}$$

$$Q(y,z) = \begin{cases} \dfrac{1}{\rho_0}y, & y < 0 \\ \dfrac{1}{\rho_0}(y + y^3), & y \geq 0 \end{cases}. \tag{4.2.41c}$$

We present this example to illustrate the idea of defining multiple Casimir density functions for multiple regions of invertibility of the basic state. Note that this state satisfies the linear stability conditions but does not satisfy conditions for Lyapunov stability because of the asymmetry in y.

Inverting (4.2.41) in two parts

$$(Y^-)^2(M,\Theta) = -2M, \qquad Z^-(M,\Theta) = M + \Theta, \tag{4.2.42a}$$

$$(Y^+)^2(M,\Theta) = \sqrt{1-4M} - 1, \qquad Z^+(M,\Theta) = M + \Theta \tag{4.2.42b}$$

(Figure 4.2b), and applying (4.2.21), we find

$$C_m^-(M,\Theta) = 2M + \Theta, \qquad C_\theta^-(M,\Theta) = M + \Theta, \tag{4.2.43a}$$

$$C_m^+(M,\Theta) = \tfrac{1}{2} - \tfrac{1}{2}\sqrt{1-4M} + M + \Theta, \qquad C_\theta^+(M,\Theta) = M + \Theta, \tag{4.2.43b}$$

and hence

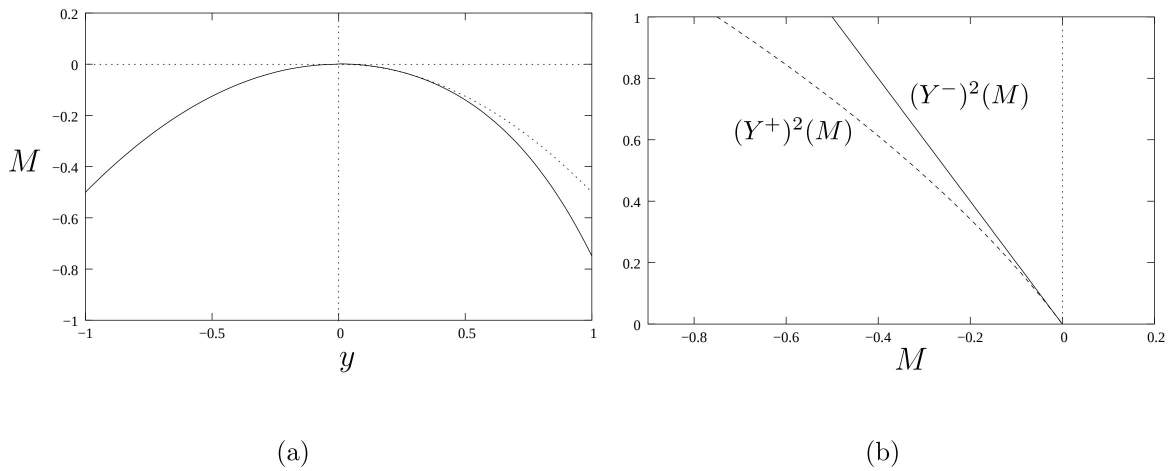


Figure 4.2: (a) Basic state $M(y)$, which is not symmetric about $y = 0$ (Example 3). Dotted line is reflection of $y < 0$ portion of $M(y)$. (b) Inverse functions of $M(y)$ for $y < 0$ (solid line) and for $y \geq 0$ (dashed line) (the inverse functions squared are plotted).

$$C^-_{mm}(M,\Theta) = 2, \qquad C^-_{\theta\theta}(M,\Theta) = 1, \qquad C^-_{m\theta}(M,\Theta) = 1, \quad (4.2.44a)$$

$$C^+_{mm}(M,\Theta) = 1 + (1-4M)^{-1/2}, \qquad C^+_{\theta\theta}(M,\Theta) = 1, \qquad C^+_{m\theta}(M,\Theta) = 1 \quad (4.2.44b)$$

(Figure 4.3a,b). Note that the derivatives and second derivatives of C^- and C^+ match where $Q = 0$ (which is at $M = 0$) and that $\det \Lambda > 0$ so the linear stability conditions are satisfied. We may integrate (4.2.43) to get

$$C^-(M,\Theta) = M^2 + \Theta M + \tfrac{1}{2}\Theta^2, \quad (4.2.45a)$$

$$C^+(M,\Theta) = \tfrac{1}{2}M^2 + \tfrac{1}{2}M + \tfrac{1}{12}(1-4M)^{3/2} + \Theta M + \tfrac{1}{2}\Theta^2 - \tfrac{1}{12}, \quad (4.2.45b)$$

where the additive constant $-\frac{1}{12}$ is included in C^+ so that it matches C^- at $M = 0$ (see Figure 4.3c).

In this case, it is convenient to extend the definition of $C^-(m,\theta)$ so that it takes the same functional form (4.2.45a) over its whole domain (namely, all real m and θ) and to extend $C^+(m,\theta)$ so that the maximum difference between C^- and C^+ is never greater than it is in the range of (M,Θ) (this ensures that $\mathcal{A}_a$ is bounded). For example, extending C^+_{mm} such that

$$C^+_{mm}(m,\theta) = \begin{cases} 2, & m < -\frac{3}{4} - 2\varepsilon \\ \left(\frac{1}{4}\varepsilon^{-2} - \frac{1}{4}\varepsilon^{-1}\right) m + \frac{3}{16}\varepsilon^{-2} - \frac{11}{16}\varepsilon^{-1} + \frac{5}{2}, & -\frac{3}{4} - 2\varepsilon \le m < -\frac{3}{4} - \varepsilon \\ \left(-\frac{1}{4}\varepsilon^{-2} + \frac{3}{4}\varepsilon^{-1}\right) m - \frac{3}{16}\varepsilon^{-2} + \frac{9}{16}\varepsilon^{-1} + \frac{3}{2}, & -\frac{3}{4} - \varepsilon \le m < -\frac{3}{4} \\ 1 + (1-4m)^{-1/2}, & -\frac{3}{4} \le m < 0 \\ 2, & m \ge 0 \end{cases}, \tag{4.2.46}$$

where ε is an arbitrary positive number (not to be confused with ϵ, the asymptotic parameter in the anelastic equations) so that

$$C^-(m,\theta) = m^2 + \theta m + \tfrac{1}{2}\theta^2, \tag{4.2.47a}$$

$$C^+(m,\theta) = \begin{cases} m^2 - \frac{7}{96} + \frac{1}{4}\varepsilon - \frac{1}{6}\varepsilon^2 \\ \qquad + \theta m + \frac{1}{2}\theta^2, & m < -\frac{3}{4} - 2\varepsilon \\ \mathrm{int1}(m + \frac{3}{4}) + \theta m + \frac{1}{2}\theta^2, & -\frac{3}{4} - 2\varepsilon \le m < -\frac{3}{4} - \varepsilon \\ \mathrm{int2}(m + \frac{3}{4}) + \theta m + \frac{1}{2}\theta^2, & -\frac{3}{4} - \varepsilon \le m < -\frac{3}{4} \\ \frac{1}{2}m^2 + \frac{1}{2}m + \frac{1}{12}(1-4m)^{3/2} \\ \qquad + \theta m + \frac{1}{2}\theta^2 - \frac{1}{12}, & -\frac{3}{4} \le m < 0 \\ m^2 + \theta m + \frac{1}{2}\theta^2, & m \ge 0 \end{cases}, \tag{4.2.47b}$$

where

$$\begin{aligned} \mathrm{int1}(m') &= \left(\frac{1}{24}\varepsilon^{-2} - \frac{1}{24}\epsilon^{-1}\right) m'^3 + \left(-\frac{1}{4}\varepsilon^{-1} + \frac{5}{4}\right) m'^2 \\ &\quad + \left(-1 - \tfrac{1}{2}\varepsilon\right) m' + \frac{47}{96} - \frac{1}{12}\varepsilon + \frac{1}{6}\varepsilon^2, \end{aligned} \tag{4.2.48a}$$

$$\mathrm{int2}(m') = \left(\frac{-1}{24}\varepsilon^{-2} + \frac{3}{24}\varepsilon^{-1}\right) m'^3 + \frac{3}{4}m'^2 - \frac{5}{4}m' - \frac{43}{96} \tag{4.2.48b}$$

are the cubic spline interpolations corresponding to the spike in C^+_{mm} (see Figure 4.3d,e). The interpolation has been designed so that C^+, C^+_m and C^+_{mm} are continuous over all values of m. ε may be taken as small as desired so that the transition from the forced functional form of $C^+(M,\Theta)$ to the simple quadratic form outside of the range of M can be virtually instantaneous in m so that the upper bound on $|\mathcal{A}_a|$ is as small as possible. Alternatively, ε may be taken as large as desired so that C^+_{mm} is kept small so that the upper bound on the symmetric part of $\mathcal{A}$ is as small as possible.

Example 4: A case of negative $\mathcal{A}$

In this example, we show explicitly how a perturbation to a linearly stable state can have negative pseudoenergy. Consider the basic state with

$$M(y,z) = \begin{cases} -\frac{1}{2}b^- y^2, & y < 0 \\ M_1(y), & 0 \le y < y_1 \\ -\frac{1}{2}b^+ y^2, & y \ge y_1 \end{cases}, \tag{4.2.49a}$$

$$\Theta(y,z) = \begin{cases} \epsilon\gamma_\alpha(\frac{1}{2}b^- y^2) + \epsilon\Gamma z, & y < 0 \\ \Theta_1(y,z), & 0 \le y < y_1 \\ \epsilon\gamma_\alpha(\frac{1}{2}b^+ y^2) + \epsilon\Gamma z, & y < 0 \end{cases}, \tag{4.2.49b}$$

where $b^+ < b^-$, and $M_1(y)$ and $\Theta_1(y,z)$ are smooth functions, smoothly connecting the two pieces that are defined explicitly, and satisfying thermal wind balance. We will consider a perturbation which only differs from the basic state for $y > y_1$, so for this exercise, the exact form of the basic state is not required.

We evaluate $\mathcal{A}(\mathbf{x};\mathbf{X})$, where $\mathbf{x} = (m(y), 0, \Theta(y,z))$ and $m(y)$ is defined by

$$m(y) = \begin{cases} M(y), & y < y_1 \\ M(y_1) + \lambda_1(y - y_1), & y_1 \le y < y_2 \\ M(y_3) - \lambda_2(y - y_3), & y_2 \le y < y_3 \\ M(y), & y \ge y_3 \end{cases} \tag{4.2.50}$$

(Figure 4.4a). Clearly, the only nonzero contribution to $\mathcal{A}$ is from the interval in which

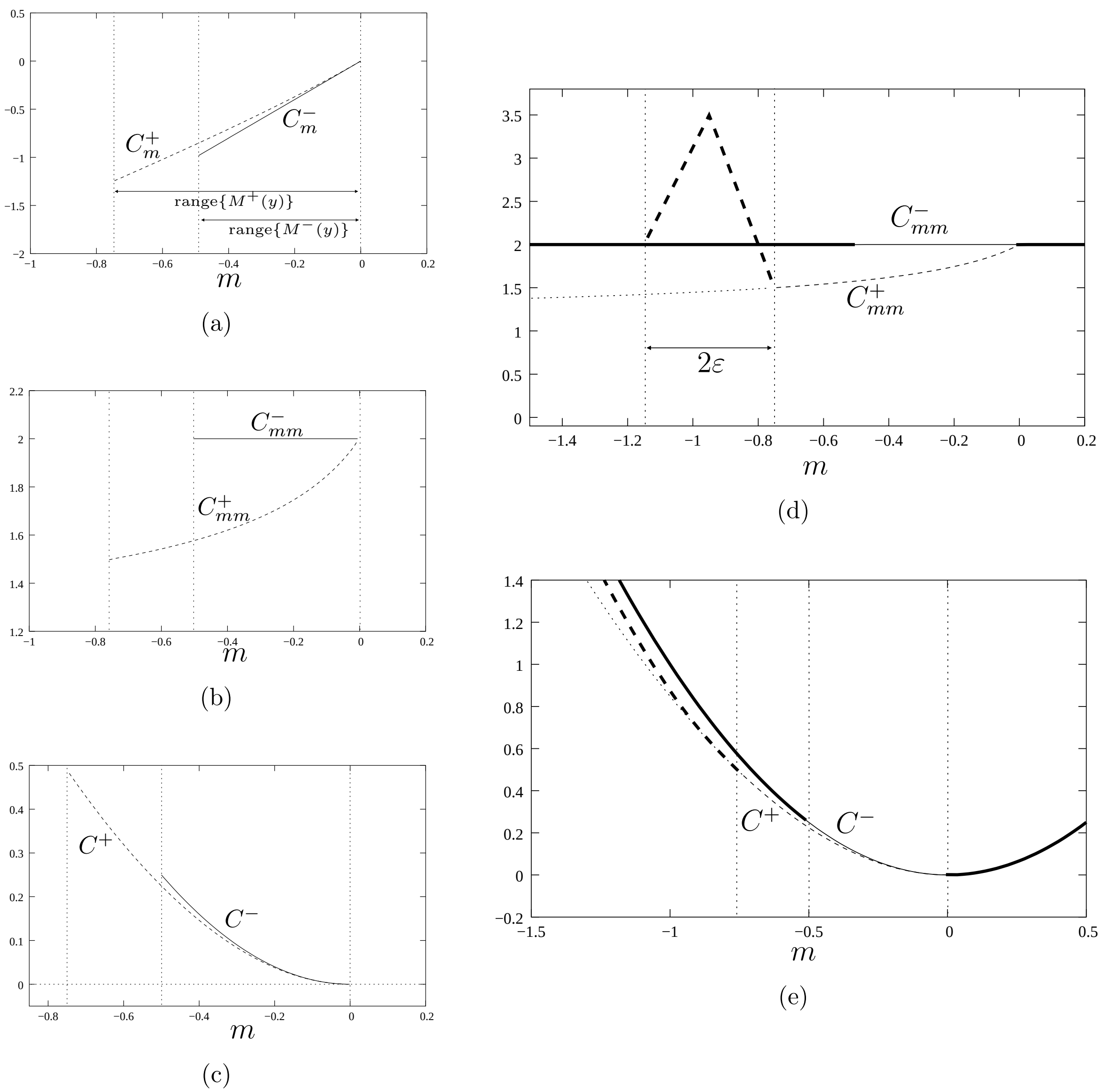


Figure 4.3: Plots of Casimir density functions for Example 3. $\theta = 0$ in all cases. (a) C_m^- and C_m^+ required to satisfy tangency of $\mathcal{C}$ with $\mathcal{H}$ at $\mathbf{X}_3$. Each function is only plotted for the values of m included in the range of $M(y)$ for $y < 0$ and $y \geq 0$ respectively. (b)-(c) Corresponding C_{mm}^-, C_{mm}^+, C^- and C^+. (d) C_{mm}^- and C_{mm}^+ over the full domain (all m). The heavy curves are the extensions beyond the range of $M(y,z)$, and the dotted curve is the continuation of $C_{mm}^+(M)$ outside of the range of $M(y)$. (e) C^- and C^+ over all m. The heavy curves are the extensions beyond the range of $M(y,z)$, and the dotted curve is the continuation of $C^+(M)$ outside of the range of $M(y)$.

$\mathbf{x}$ and $\mathbf{X}$ differ, namely for $y_1 < y < y_3$. Therefore,

$$\begin{aligned}
\mathcal{A} &= \int_0^1\int_{y_1}^{y_2} \rho_0 \left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(m-M) + C^-(m,\Theta) - C^+(M,\Theta)\right] \mathrm{d}y\,\mathrm{d}z \\
&\quad + \int_0^1\int_{y_2}^{y_3} \rho_0 \left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(m-M) + C^+(m,\Theta) - C^+(M,\Theta)\right] \mathrm{d}y\,\mathrm{d}z \\
&= \int_0^1\int_{y_1}^{y_2} \rho_0 \left[\tfrac{1}{2}C^+_{mm}(m-M)^2 + C^-(m,\Theta) - C^+(M,\Theta)\right] \mathrm{d}y\,\mathrm{d}z \\
&\quad + \int_0^1\int_{y_2}^{y_3} \rho_0 \left[\tfrac{1}{2}C^+_{mm}(m-M)^2\right] \mathrm{d}y\,\mathrm{d}z,
\end{aligned} \tag{4.2.51}$$

where the Taylor expansion of C^+ about (M,Θ) has been used, and the derivatives of C^+ have been taken from (4.2.21). Since we seek a perturbation for which $\mathcal{A}$ is negative, we let $\lambda_2 \to -\infty$, and henceforth neglect the integration from y_2 to y_3 (which is positive). Integrating (4.2.21) for this case, we find

$$C^-(m,\theta) = \frac{1}{2}\left(\frac{\beta_\delta}{b^-} + \frac{\gamma_\alpha^2}{\Gamma}\right)m^2 + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta m + \frac{\gamma_\alpha}{2\epsilon\Gamma}\theta^2, \tag{4.2.52a}$$

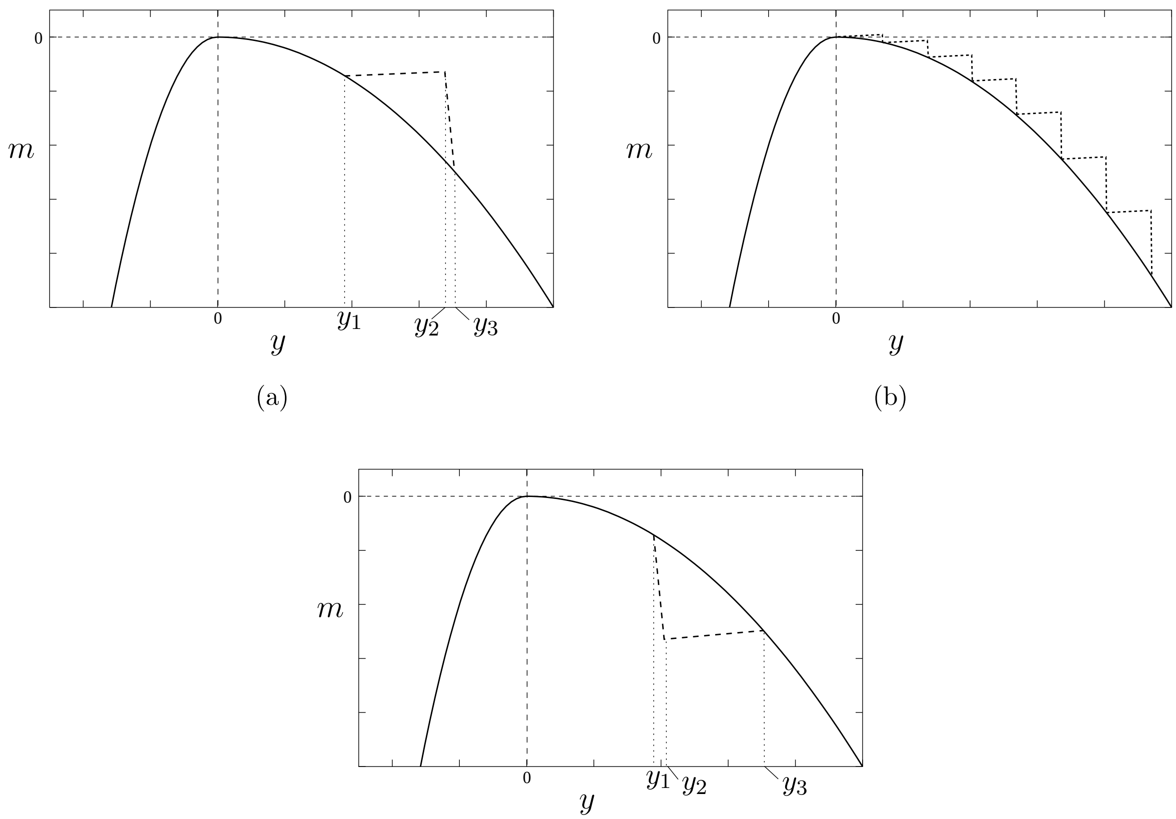
$$C^+(m,\theta) = \frac{1}{2}\left(\frac{\beta_\delta}{b^+} + \frac{\gamma_\alpha^2}{\Gamma}\right)m^2 + \frac{\gamma_\alpha}{\epsilon\Gamma}\theta m + \frac{\gamma_\alpha}{2\epsilon\Gamma}\theta^2. \tag{4.2.52b}$$

Substituting (4.2.52) into (4.2.51) gives

$$\mathcal{A} \approx \int_0^1\int_{y_1}^{y_2} \rho_0 \left[\frac{1}{2}\left(\frac{\beta_\delta}{b^-} + \frac{\gamma_\alpha^2}{\Gamma}\right)(m-M)^2 - \frac{\beta_\delta}{2}\left(\frac{1}{b^+} - \frac{1}{b^-}\right)m^2\right] \mathrm{d}y\,\mathrm{d}z. \tag{4.2.53}$$

The difference $C^+ - C^-$ is an increasing function of m^2, so the negative term in the integrand of (4.2.53) is largest if we let $\lambda_1 \to 0$ and hence $m(y) \to M(y_1) = -\frac{1}{2}b^+y_1^2$. As the width of the "step" becomes smaller, $m(y)$ and $M(y)$ become closer together, and eventually, the sum becomes negative (see Figure 4.5).

We might next consider a perturbation with $m(y)$ being composed of many steps (see Figure 4.4b). In the limit of infinitely many short steps, the first term in (4.2.53)



(a) (b) (c)

Figure 4.4: Perturbations for which $\mathcal{A}$ may be negative. The heavy dashed lines are the perturbed field $m(y)$. $m(y)$ and $M(y)$ coincide everywhere else. (a) A single "over" perturbation. (b) A staircase of "over" perturbations. In the limit of infinitely many steps, the most negative $\mathcal{A}$ obtains. (c) An "under" perturbation.

disappears, and $\mathcal{A} \to \mathcal{A}_{\text{lim}}$, where

$$\mathcal{A}_{\text{lim}} = -\int_0^1 \rho_0 \left[\frac{\beta_\delta b^+}{10} \left(1 - \frac{b^+}{b^-} \right) \right] \mathrm{d}z. \tag{4.2.54}$$

One might object to this, arguing that a state with infinitesimal steps is a small amplitude perturbation, and that a nonsymmetric state such as (4.2.49) has previously been shown to be stable under the linearized equations. However, the linearized equations assume that $\mathbf{x}$ is close to $\mathbf{X}$ and all *derivatives* of $\mathbf{x}$ are close to the corresponding derivatives of $\mathbf{X}$. Clearly, the derivatives of the step-like state are not close to those in the steady

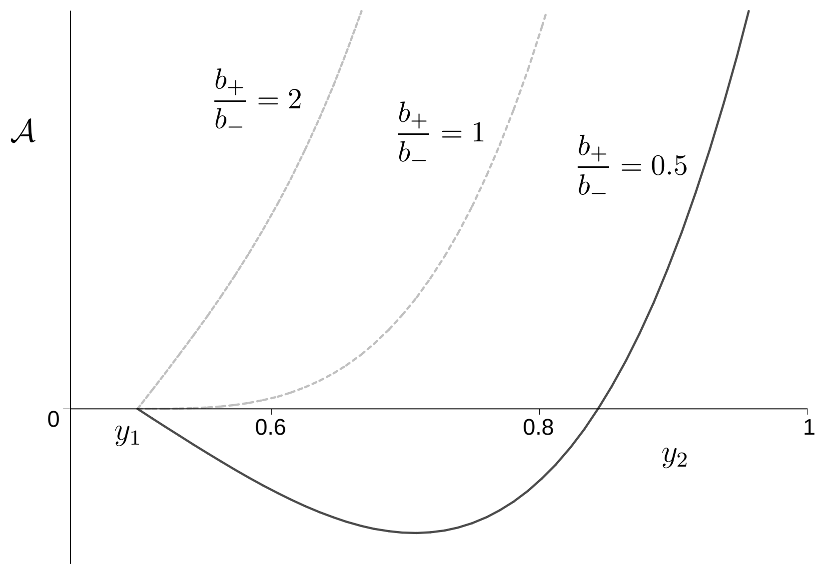


Figure 4.5: $\mathcal{A}$ versus y_2 for $y_1 = 0.5$ (see (4.2.53) and Figure 4.4a) for three values of b_+/b_-. Notice that $\mathcal{A}$ is negative for small enough "step" size for $b_+/b_- = 0.5$ (the heavy curve).

state.

$\mathcal{A}_{\mathrm{lim}}$ is the most negative value of $\mathcal{A}$ that can be achieved with this construction. However, depending on Γ and the ratio b^+/b^-, it might be possible to achieve a more negative $\mathcal{A}$ with a perturbation in the form of a "step" below the curve $M(y)$, namely

$$m(y) \approx \begin{cases} M(y), & y < y_1 \\ M(y_2), & y_1 < y < y_2 \\ M(y), & y \geq y_2 \end{cases} \tag{4.2.55}$$

(Figure 4.4c), where by the $\approx$ symbol, we are indicating that this is the limiting form of $m(y)$ when the slopes of the side and base of the step approach $-\infty$ and zero respectively. In this case,

$$\mathcal{A} = \int_0^1 \int_{y_1}^{y_2} \rho_0 \left[\frac{1}{2} \left(\frac{\beta_\delta}{b^+} + \frac{\gamma_\alpha^2}{\Gamma} \right) \left(-\tfrac{1}{2} b^+ y^2 + \tfrac{1}{2} b^+ y_2^2 \right)^2 - \frac{1}{2} \left(\frac{\beta_\delta}{b^+} - \frac{\beta_\delta}{b^-} \right) \left(\tfrac{1}{2} b^+ y_2^2 \right)^2 \right] \mathrm{d}y \, \mathrm{d}z. \tag{4.2.56}$$

If we assume that $\gamma_\alpha^2/\Gamma \ll \beta_\delta/b^+$, then for fixed y_2, $\mathcal{A}$ is minimized (most negative) for

$$y_1^{(\min)} = \left(\sqrt{1 - \sqrt{1 - \frac{b^+}{b^-}}}\right) y_2. \tag{4.2.57}$$

4.3 Anelastic version of Dunkerton (1981) problem

The linearized anelastic equations (4.2.5) can be solved analytically for the simplest choice of reference potential temperature $\theta_0(z) \equiv 1$ (recovering the classical anelastic equations) and the basic state angular momentum profile

$$M(y,z) = -\tfrac{1}{2}by^2 + \lambda y, \tag{4.3.1}$$

where b and λ are positive constants. $M_y = -by + \lambda$, so the state fails the stability criterion (4.2.18b) in the interval $0 < y < \lambda/b$, where M decreases away from the equator. This state has a vertical velocity shear $U_z = -\gamma$ to balance the planetary angular momentum shear.

The choice of $\theta_0(z) \equiv 1$ corresponds to the reference pressure and density profiles

$$\pi_0(z) = 1 - Bz, \tag{4.3.2a}$$

$$\rho_0(z) = (1 - Bz)^{c_v/R}. \tag{4.3.2b}$$

We choose the basic state potential temperature profile

$$\Theta(y,z) = (\epsilon\gamma_\alpha)(\tfrac{1}{2}by^2 - \lambda y) + \epsilon\Gamma z, \tag{4.3.3}$$

where Γ is another positive constant, which satisfies thermal wind balance (4.2.1) and is statically stable. The linearized equations (4.2.5) are then

$$m'_t = \frac{1}{\rho_0}(by - \lambda)\psi'_z, \tag{4.3.4a}$$

$$\zeta'_t = \frac{1}{\epsilon}\theta'_y + \beta_\delta y m'_z + \gamma_\alpha m'_y, \tag{4.3.4b}$$

$$\theta'_t = -\frac{1}{\rho_0}(\epsilon\gamma_\alpha)(by - \lambda)\psi'_z + \frac{1}{\rho_0}\epsilon\Gamma\psi'_y, \tag{4.3.4c}$$

which are combined to get

$$\rho_0\left(\frac{1}{\rho_0}\psi'_z\right)_{ztt}+\frac{1}{\alpha^2}\psi'_{yytt}=-\beta_\delta by\left(y-\frac{\lambda}{b}\right)\rho_0\left(\frac{1}{\rho_0}\psi'_z\right)_z-\Gamma\psi'_{yy}, \tag{4.3.5}$$

where (4.1.20) has been used to write ζ' in terms of ψ'.

We solve (4.3.5) over the domain $y\in(-\infty,+\infty)$ with the boundary condition that $\psi'\to 0$ as $y\to\pm\infty$.[1] We use the method of separation of variables, substituting

$$\psi'(y,z,t)=Y_\psi(y)Z_\psi(z)T_\psi(t) \tag{4.3.6}$$

into (4.3.5) and find the three ordinary differential equations

$$\begin{aligned}\frac{\mathrm{d}^2Y_\psi}{\mathrm{d}y^2}+\left(\frac{k^2}{\Gamma-\dfrac{\omega^2}{\alpha^2}}\right)\left[\omega^2-\beta_\delta by\left(y-\frac{\lambda}{b}\right)\right]Y_\psi &= 0,\\ \rho_0\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\rho_0}\frac{\mathrm{d}Z_\psi}{\mathrm{d}z}\right)+k^2Z_\psi &= 0,\\ \frac{\mathrm{d}^2T_\psi}{\mathrm{d}t^2}+\omega^2T_\psi &= 0,\end{aligned}$$

where k and ω are constants. We obtain the normal mode solution

$$\begin{aligned}\psi'(y,z,t) &= \exp\left[i\omega t-\frac{1}{2}\sqrt{\frac{\beta_\delta bk^2}{\Gamma-\dfrac{\omega^2}{\alpha^2}}}\left(y-\frac{\lambda}{2b}\right)^2\right]H_n\left[\left(\frac{\beta_\delta bk^2}{\Gamma-\dfrac{\omega^2}{\alpha^2}}\right)^{-\frac{1}{4}}\left(y-\frac{\lambda}{2b}\right)\right]\\ &\quad\times(1-Bz)^{\frac{1}{2}(1+\frac{c_v}{R})}\left\{c_{1kn}J_{\frac{1}{2}(1+\frac{c_v}{R})}\left[\frac{k}{B}(1-Bz)\right]\right.\\ &\quad\left.+\ c_{2kn}Y_{\frac{1}{2}(1+\frac{c_v}{R})}\left[\frac{k}{B}(1-Bz)\right]\right\},\end{aligned} \tag{4.3.7}$$

where J_ν and Y_ν are νth order Bessel functions of the first and second kinds, n is a non-negative integer satisfying the *characteristic equation*

$$2n+1=\sqrt{\frac{k^2}{\Gamma-\dfrac{\omega^2}{\alpha^2}}}\left[\omega^2+\beta_\delta b\left(\frac{\lambda}{2b}\right)^2\right], \tag{4.3.8}$$

[1] These boundary conditions are chosen rather than the fixed finite walls used in the stability analysis so that we can get an exact analytical solution of simple form. Neither choice is physically correct, and we hope that the behaviour near the equator is not substantially changed either way.

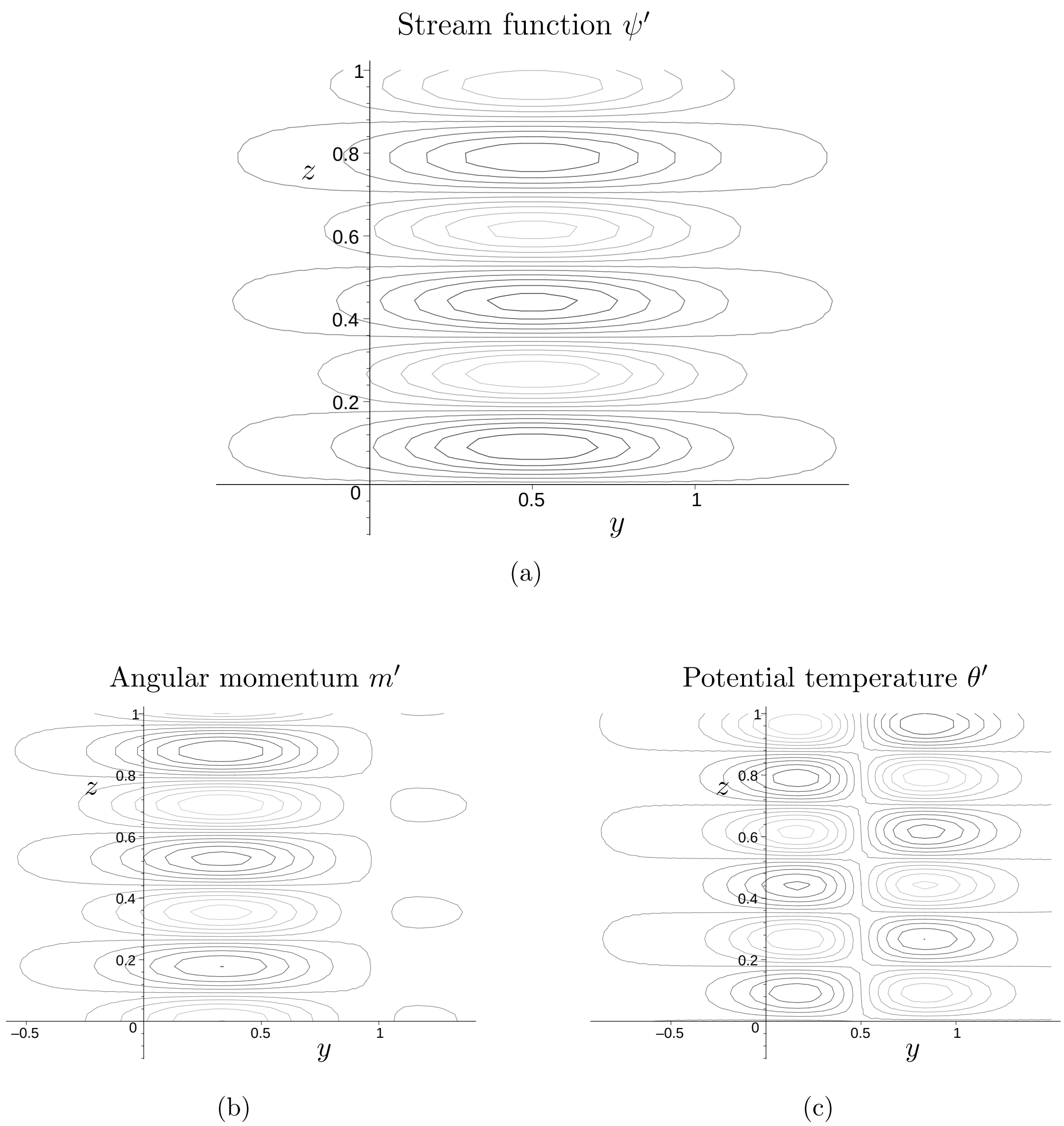


Figure 4.6: (a) Contours of stream function ψ' for the $n = 0$, $k = k_c$ (see Figure 4.7) mode of solution (4.3.7) and the parameters listed in Table 4.1. Dark contours represent negative values (clockwise circulation) and light contours represent positive values (anticlockwise circulation). The vertical scale k_c^{-1} is chosen to be that of the deepest unstable mode. (b) Angular momentum perturbation m' contours; note the zonal jets on equatorward side of unstable region. (c) Potential tempertaure perturbation θ' contours; note the checkerboard pattern of pancake structures (cf. Figure 1.2).

Scaling	Nondimensional (derived)	Steady State $\mathbf{X}$
$H = 2 \times 10^4$ m	$\alpha = 150$	$b = \beta_\delta$
$L = 3 \times 10^6$ m	$\delta = 2.1$	$\lambda = \beta_\delta$
$a = 6.4 \times 10^6$ m	$B = 0.2$	$\Gamma = 1$
$g = 10$ ms^{-2}	$\epsilon = 0.9$	
$\Omega = 7 \times 10^{-5}$ s^{-1}	$S = 0.95$	
$N = 2 \times 10^{-2}$ s^{-1}	$\beta_\delta = 0.5$	
$c_p = 10^3$ J K^{-1}	$\gamma_\alpha = 0.007$	
$\Theta = 10^3$ K	$H_s/H = 1.5$	

Table 4.1: Parameters used in plots of Figures 4.6 and 4.8.

H_n is the nth degree Hermite polynomial, and c_{1kn} and c_{2kn} are normalization constants. Figure 4.6 shows contour plots of stream function, angular momentum, and potential temperature for the $n = 0$ mode of the solution (4.3.7) given the choices of parameter values listed in Table 4.1. The discrete mode structure of the solution is the result of the condition that ψ' remain finite as $y \to \pm\infty$ (see any reference on Sturm-Liouville theory; e.g. Arfken, 1985). The solution is unstable, i.e. there exists a mode which is monotonically growing with time, if $\omega^2 < 0$. Rearranging (4.3.9) gives

$$\begin{aligned}\omega^2 \;&=\; -\left[\beta_\delta b\left(\frac{\lambda}{2b}\right)^2 + \frac{(2n+1)^2}{2\alpha^2 k^2}\right] \\ &\qquad + \frac{1}{2}\left(\frac{2n+1}{k}\right)\sqrt{\frac{1}{\alpha^2}\left[4\beta_\delta b\left(\frac{\lambda}{2b}\right)^2 + \frac{(2n+1)^2}{k^2\alpha^2}\right] + 4\Gamma}\;. \qquad (4.3.9)\end{aligned}$$

Inspection of (4.3.9) reveals that ω^2 will be negative for large enough vertical wavenumber k for any values of velocity shear $\lambda \neq 0$, potential temperature gradient $\Gamma > 0$, and meridional index n. For any given k, the most negative ω^2, i.e. the fastest growing mode, has $n = 0$. Figure (4.7) shows ω^2 as a function of vertical wavenumber k for $n = 0$. Modes with $k > k_c$ are unstable. For the parameter choices listed in Table 4.1, k_c corresponds

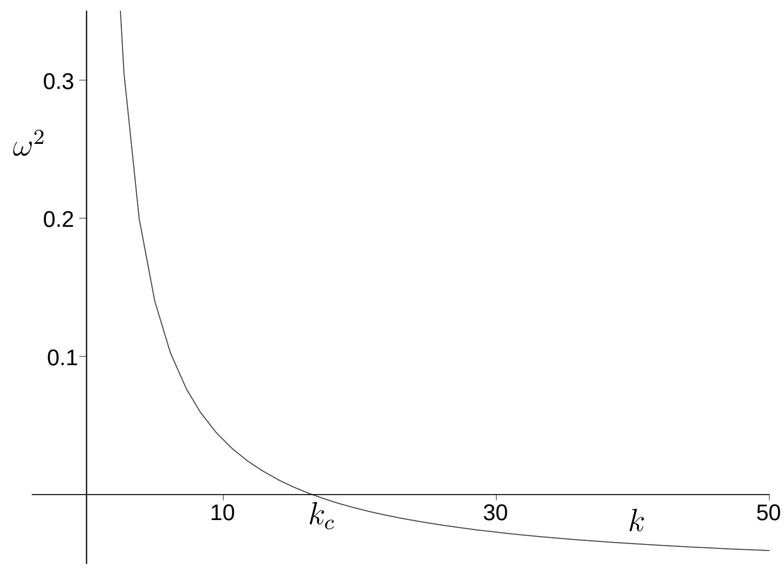


Figure 4.7: Frequency squared ω^2 of $n = 0$ mode as a function of vertical wavenumber k. Modes with $k > k_c$ are unstable (growing).

to a vertical wavelength of about 7 km.

In the limit of $k \gg \alpha^{-1}$ (the vertical wavelength being much smaller than the aspect ratio L/H), (4.3.9) becomes

$$\omega^2 \approx \frac{\sqrt{\Gamma}}{k}(2n+1) - \beta_\delta b \left(\frac{\lambda}{2b}\right)^2, \tag{4.3.10}$$

which is almost identical to the characteristic equation in the hydrostatic case (see Dunkerton, 1981). This is expected because the terms involving α come from the vertical velocity term $\rho_0^{-1}\alpha^{-2}\psi'_{yytt} = \alpha^{-2}w'_{ytt}$ in (4.3.5), which is negligible in the hydrostatic system. The effects of using the anelastic approximation rather than the hydrostatic, namely allowing deep vertical motion, are not important for small vertical wavelength modes.

The meridional velocity solution in the hydrostatic case, at fixed y and t, satisfies

$$v'_{\text{hydro}} \propto \exp\left(\frac{1}{2}\frac{H}{H_s}z + i\sqrt{k^2 - \frac{1}{4(H_s/H)^2}}z\right), \tag{4.3.11}$$

where $H_s \equiv R\Theta/g =$ constant is the scale height. For $H_s \ll H$, the structure is close to sinusoidal in z, and for large H_s, the amplitude grows with z. In the anelastic system,

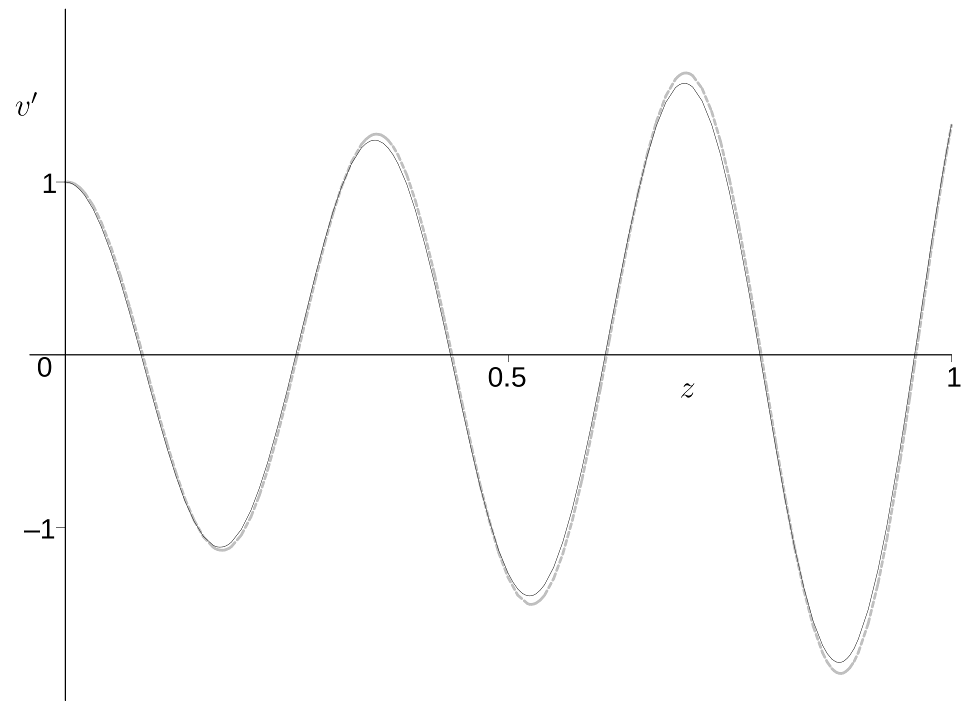


Figure 4.8: z-dependence of meridional velocity perturbation in anelastic solution (solid) and hydrostatic solution (dashed).

the meridional velocity at fixed y and t satisfies

$$v'_{\text{anelastic}} \propto (1 - Bz)^{-c_v/R} \frac{\mathrm{d}}{\mathrm{d}z} J_{\frac{1}{2}(1+\frac{c_v}{R})} \left[\frac{k}{B}(1 - Bz) \right]. \tag{4.3.12}$$

The domain depth parameter $B = gH/c_p\Theta$ plays the role of H_s in the anelastic solution: for $B \ll 1$ (a shallow domain), the solution is approximately sinusoidal, and for $B \lesssim 1$, it grows with z. In Figure 4.8, the two solutions are plotted for the values of the parameters in Table 4.1. Properly normalized, they overlap almost exactly, with the amplitude of v'_{hydro} growing slightly faster with z.

Qualitatively, the familiar signatures of inertial instability are evident in the solution

— compare Figures 4.6 with Figure 1.1, with Taylor vortices stacked over the centre of the unstable region (Figure 4.6a), oppositely signed zonal jets stacked on the equatorward side of the unstable region (Figure 4.6b), and the characteristic pattern of pancake structures in the temperature field (Figure 4.6c) which has been observed in satellite data of the stratopause region (see Section 1.2).

The parameters were chosen so that the unstable region covers about 30° of latitude in the northern hemisphere, and the vertical wavelength of the largest scale unstable modes for $n = 0$ is similar to what is observed ($\sim$ 10 km; see Figure 1.2), although it should be noted that (4.3.9) implies that the modes with *smallest* vertical scale have the highest growth rates. Why the observed vertical scale is as large as it is is still an open question.

4.4 Saturation bounds

The pseudoenergy $\mathcal{A}$ was introduced in order to demonstrate Lyapunov stability of steady states which are even functions of y and satisfy the linear stability conditions (4.2.18). It was argued that the statement of stability (4.2.28) implies that solutions are bound arbitrarily close to equilibrium (in terms of the norm (4.2.27)) provided they are initially close enough to equilibrium: for any ε, there is a δ such that $||\Delta\mathbf{x}(0)|| < \delta$ implies $||\Delta\mathbf{x}(t)|| < \varepsilon$ for all times t.

In Section 4.2.2, it was noted that (4.2.28) also implies that no matter how *far* from equilibrium the system starts, it is bound for all time within a definite radius: for any δ, there is an ε such that $||\Delta\mathbf{x}(0)|| < \delta$ implies $||\Delta\mathbf{x}(t)|| < \varepsilon$.

This suggests the following interesting application (following, e.g., Mu et al., 1996). Consider an initial state which is close to an *unstable* equilibrium. The pseudoenergy consists of a positive definite term due to the kinetic energy in the meridional and vertical velocity components,

$$\mathcal{K}_\perp(\mathbf{x}) \equiv \iint_{\mathcal{D}} \left\{ \frac{1}{2\rho_0} \left[\frac{1}{\alpha^2}(\psi_y)^2 + (\psi_z)^2 \right] \right\} \mathrm{d}y\,\mathrm{d}z, \tag{4.4.1}$$

and the term

$$\begin{aligned} \mathcal{APE}(\mathbf{x};\mathbf{X}) \;\; \equiv \;\; & \iint_{\mathcal{D}} \left\{ \rho_0 \left[(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z)(m - M) \right.\right. \\ & \left.\left. + \; \frac{1}{\epsilon B}\pi_0(\theta - \Theta) + C(m,\theta) - C(M,\Theta) \right] \right\} \mathrm{d}y\,\mathrm{d}z, \end{aligned} \tag{4.4.2}$$

which may be called *available potential energy* when the reference state $\mathbf{X}$ is nonlinearly stable. Since $\mathcal{A}$ is conserved and $\mathcal{APE}$ is strictly positive, $\mathcal{K}_\perp$ is bound from above by $\mathcal{A}$. In particular, for an unstable equilibrium $\mathbf{X}_U$, $\mathcal{K}_\perp$ is bound from above by $\mathcal{A}(\mathbf{X}_U;\mathbf{X})$, where $\mathbf{X}$ is *any* nonlinearly stable equilibrium.

Shepherd (1988) calls the lowest such value of $\mathcal{A}$ the *saturation bound* on the disturbance amplitude in the transverse plane. The short time evolution of the system from near the unstable equilibrium follows approximately the solution of the linearized equations (see Section 4.3 for a particular case) for which $\mathcal{K}_\perp$ grows monotonically in time. The long time evolution of $\mathcal{K}_\perp$ in the nonlinear system, while unknown in detail, is bound by this result.

In practice, all of $\mathcal{APE}$ is not converted into $\mathcal{K}_\perp$ because $\mathbf{X}$ had to be chosen such that $\mathcal{APE}(\mathbf{x};\mathbf{X})$ was positive for any perturbed state $\mathbf{x}$, and the true evolution does not sample every state. Also, in reality there is inevitably some diffusion of heat and momentum which limits the production of $\mathcal{K}_\perp$.

If the saturation amplitude calculated with this method is found to be close to what is observed (in high resolution simulations, for example), then it can potentially be part of a parameterization scheme in coarse resolution models which cannot satisfactorily simulate inertial adjustment. Specifically, the kinetic energy estimate can be used to estimate an effective diffusivity for vertical mixing due to inertial adjustment.

As an example, consider again the anelastic system with $\theta_0(z) \equiv 1$, and the unstable basic state $\mathbf{X}_U$ with

$$M_U(y,z) = -\tfrac{1}{2} b_U y^2 + \lambda_U y, \tag{4.4.3a}$$

$$\Theta_U(y,z) = (\epsilon\gamma_\alpha)(\tfrac{1}{2} b_U y^2 - \lambda_U y) + \epsilon\Gamma_U z. \tag{4.4.3b}$$

We seek to minimize $\mathcal{A}(\mathbf{X}_U;\mathbf{X})$ where $\mathbf{X}$ is chosen from the class of nonlinearly stable basic states of the form

$$M(y,z) = M_0 - \tfrac{1}{2} b y^2, \tag{4.4.4a}$$

$$\Theta(y,z) = \Theta_0 + (\epsilon\gamma_\alpha)(\tfrac{1}{2} b y^2) + \epsilon\Gamma z \tag{4.4.4b}$$

(see Example 1 above). From the earlier calculation, we have the second partial derivatives of $C(m,\theta)$ (the constants M_0 and Θ_0 do not change the Casimir density function), and since they are constants, the second order Taylor expansion equals $C(m,\theta)$ exactly. Also, the $\mathcal{K}_\perp$ term is zero for the state $\mathbf{X}_U$. Therefore,

$$\mathcal{A}(\mathbf{X}_U;\mathbf{X}) = \int_0^1\!\!\int_{-1}^1 \rho_0 \left[\tfrac{1}{2} C_{mm}(M_U - M)^2 + C_{m\theta}(M_U - M)(\Theta_U - \Theta) + \tfrac{1}{2} C_{\theta\theta}(\Theta_U - \Theta)^2\right] \mathrm{d}y\,\mathrm{d}z. \tag{4.4.5}$$

Substituting (4.2.32), $\mathbf{X}_U$, and $\mathbf{X}$ into (4.4.5) gives, after some cancellation,

$$\begin{aligned}
\mathcal{A}(\mathbf{X}_U;\mathbf{X}) &= \int_0^1\!\!\int_{-1}^1 \rho_0 \left\{ \frac{\beta_\delta}{2b}\left[-\tfrac{1}{2}(b_U - b)y^2 + \lambda_U y\right]^2 - \frac{\beta_\delta}{b} M_0 \left[-\tfrac{1}{2}(b_U - b)y^2 + \lambda_U y\right] \right\} \mathrm{d}y\,\mathrm{d}z \\
&\quad + \int_0^1\!\!\int_{-1}^1 \rho_0 \left[\frac{1}{2}\left(\frac{\gamma_\alpha^2}{\Gamma} + \frac{\beta_\delta}{b}\right) M_0^2 + \frac{\gamma_\alpha}{\epsilon\Gamma} M_0 \Theta_0 + \frac{1}{2}\left(\frac{1}{\epsilon^2\Gamma}\right)\Theta_0^2 \right. \\
&\qquad \left. - \frac{1}{\Gamma}\left(\gamma_\alpha M_0 + \frac{1}{\epsilon}\Theta_0\right)(\Gamma_U - \Gamma) z + \frac{1}{2}\frac{(\Gamma_U - \Gamma)^2}{\Gamma} z^2 \right] \mathrm{d}y\,\mathrm{d}z.
\end{aligned} \tag{4.4.6}$$

After integration,

$$\begin{aligned}\mathcal{A}(\mathbf{X}_U;\mathbf{X}) &= I_0\left(\frac{\beta_\delta}{b}\right)\left\{\frac{1}{20}(b_U-b)^2+\frac{1}{3}\left[\lambda_U^2+M_0(b_U-b)\right]+M_0^2\right\}\\ &\quad+I_0\left(\frac{1}{\Gamma}\right)\left(\gamma_\alpha M_0+\frac{1}{\epsilon}\Theta_0\right)^2-2I_1\left(\gamma_\alpha M_0+\frac{1}{\epsilon}\Theta_0\right)\left(\frac{\Gamma_U-\Gamma}{\Gamma}\right)\\ &\quad+I_2\left[\frac{(\Gamma_U-\Gamma)^2}{\Gamma}\right],\end{aligned} \tag{4.4.7}$$

where

$$I_k \equiv \int_0^1 z^k \rho_0(z)\,\mathrm{d}z. \tag{4.4.8}$$

We seek values of b, Γ, M_0, and Θ_0 which minimize $\mathcal{A}$. To that end, we calculate the partial derivatives of $\mathcal{A}$ with respect to each parameter and set them equal to zero. Since Θ_0 only appears in (4.4.7) in the grouping $\Pi_0 \equiv \gamma_\alpha M_0 + (1/\epsilon)\Theta_0$, it is convenient to differentiate $\mathcal{A}$ with respect to Π_0 rather than Θ_0:

$$\frac{\partial \mathcal{A}}{\partial \Pi_0} = \frac{2I_0}{\Gamma}\Pi_0 - \frac{2I_1}{\Gamma}(\Gamma_U-\Gamma), \tag{4.4.9}$$

whence the optimum choice of Π_0 is

$$\Pi_0^{(\mathrm{min})} = \frac{I_1}{I_0}(\Gamma_U-\Gamma). \tag{4.4.10}$$

Differentiating with respect to Γ gives

$$\frac{\partial \mathcal{A}}{\partial \Gamma} = -\frac{1}{\Gamma^2}I_0\Pi_0^2 + 2I_1\Pi_0\left(\frac{1}{\Gamma}+\frac{\Gamma_U-\Gamma}{\Gamma^2}\right) - 2I_2\left[\frac{\Gamma_U-\Gamma}{\Gamma}-\frac{(\Gamma_U-\Gamma)^2}{2\Gamma^2}\right]. \tag{4.4.11}$$

Inserting $\Pi_0^{(\mathrm{min})}$ for Π_0 and setting $\partial\mathcal{A}/\partial\Gamma = 0$ gives a minimizing value of Γ of

$$\Gamma^{(\mathrm{min})} = |\Gamma_U|, \tag{4.4.12}$$

where the absolute value is used because for stability of $\mathbf{X}$, Γ must be positive. Evidently, if Γ_U is positive (statically stable), then the minimizing choice of Γ is Γ_U itself. This is perhaps not surprising. Continuing, we calculate

$$\frac{\partial \mathcal{A}}{\partial M_0} = 2\beta_\delta I_0\left[\frac{1}{6}\left(\frac{b_U-b}{b}\right)+\frac{M_0}{b}\right], \tag{4.4.13}$$

$$\begin{aligned}\frac{\partial \mathcal{A}}{\partial b} &= 2\beta_\delta I_0\left\{-\frac{1}{b^2}\left[\frac{1}{40}(b_U-b)^2+\frac{1}{6}\left(\lambda_U^2+M_0(b_U-b)\right)+\frac{1}{2}M_0^2\right]\right.\\ &\qquad\left.-\frac{1}{b}\left[\frac{1}{20}(b_U-b)-\frac{1}{6}M_0\right]\right\},\end{aligned} \tag{4.4.14}$$

whence

$$b^{(\min)} = -|b_U|\sqrt{1+15\left(\frac{\lambda_U}{b_U}\right)^2}, \tag{4.4.15}$$

$$M_0^{(\min)} = -\frac{1}{6}(b_U - b^{(\min)}). \tag{4.4.16}$$

Combining (4.4.10) and (4.4.12) gives a minimizing value of Θ_0 of

$$\Theta_0^{(\min)} = \epsilon\left[\frac{I_1}{I_0}(\Gamma_U - |\Gamma_U|) + \frac{\gamma_\alpha}{6}(b_U - b^{(\min)})\right]. \tag{4.4.17}$$

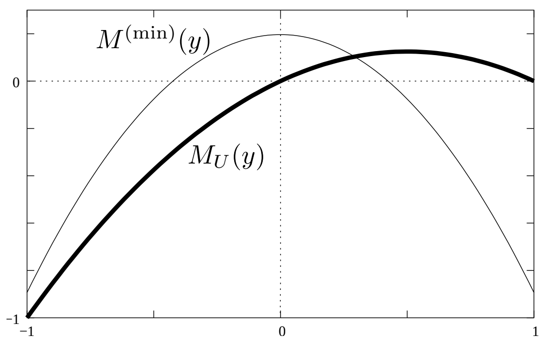
4.4.1 Examples

Example 1: Inertial instability

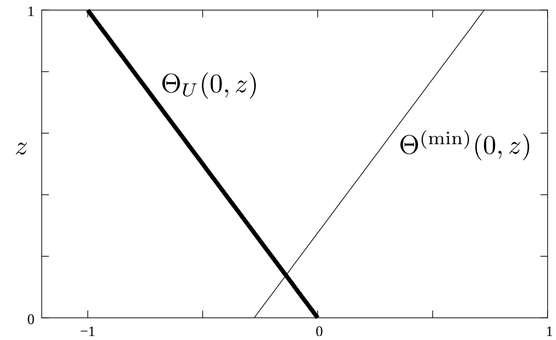
Consider a statically stable ($\Gamma_U > 0$), inertially unstable ($\lambda_U > 0$) state $\mathbf{X}_U$. The z dependence of the minimizing stable state $\mathbf{X}^{(\min)}$ is the same as that of $\mathbf{X}_U$ by virtue of (4.4.12). The y dependence of $\mathbf{X}^{(\min)}$, including the shifts M_0 and Θ_0 is given by (4.4.15)-(4.4.17). $M^{(\min)}(y)$ for an example case is plotted in Figure 4.9a. Substituting (4.4.12), (4.4.15), (4.4.16), and (4.4.17) into (4.4.7) gives a minimum value of $\mathcal{A}$ of

$$\begin{aligned}\mathcal{A}^{(\min)}(\lambda_U, b_U) &\equiv \mathcal{A}(\mathbf{X}_U;\mathbf{X}^{(\min)}) \\ &= \frac{\beta_\delta I_0 b_U}{\sqrt{1+15\left(\frac{\lambda_U}{b_U}\right)^2}}\left[\frac{2}{45}\left(1-\sqrt{1+15\left(\frac{\lambda_U}{b_U}\right)^2}\right)+\frac{2}{3}\left(\frac{\lambda_U}{b_U}\right)^2\right].\end{aligned} \tag{4.4.18}$$

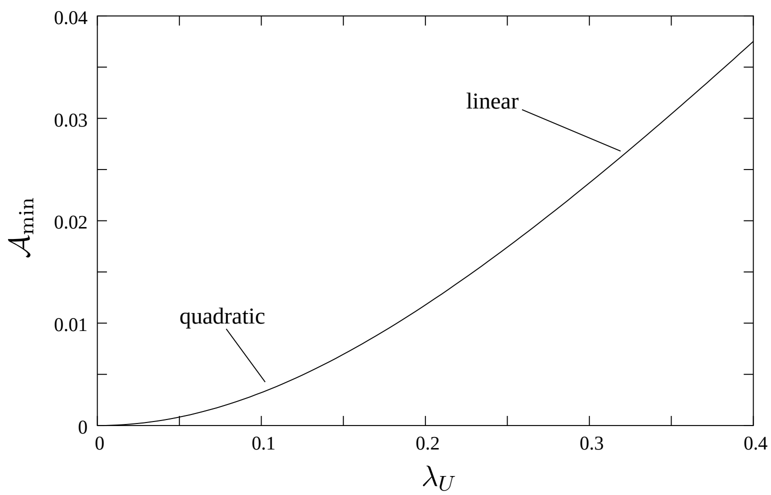
In Figure 4.9c, $\mathcal{A}^{(\min)}$ is plotted as a function of λ_U for fixed b_U. Notice in particular that $\mathcal{A}^{(\min)}$ approaches zero as λ_U approaches zero, i.e. as the stability threshold is approached.



(a)



(b)



(c)

Figure 4.9: (a) Stable $M^{\min}(y)$ (thin curve) which minimizes $\mathcal{A}$ for $M_U(y) = -\frac{1}{2}y^2 + \frac{1}{2}y$ (thick curve) for inertially unstable case. (b) Stable $\Theta^{(\min)}(0,z)$ (thin) which minimizes $\mathcal{A}$ for convectively unstable $\Theta_U(0,z) = -z$ (thick). The point of intersection depends on $\rho_0(z)$, which in turn depends on the domain depth parameter B. The case shown is for a choice of $B = 0.5$. (c) Upper bound on kinetic energy in released into overturning flow during inertial adjustment, $\mathcal{A}_{\min}$ as a function of velocity shear λ_U for fixed $b_U = 0.1$.

Example 2: Convective instability

Now let $\mathbf{X}_U$ be inertially stable ($\lambda_U = 0$) but statically unstable ($\Gamma_U < 0$).

Substituting (4.4.12) and (4.4.17) into (4.4.7) gives a saturation bound of

$$\mathcal{A}^{(\min)} = \frac{4}{I_0}(I_0 I_2 - I_1^2)|\Gamma_U|, \tag{4.4.19}$$

It must be true from the theory previously presented, but note anyway that $\mathcal{A}^{(\min)}$ in (4.4.19) is necessarily positive because $\rho_0(z) > 0$, and

$$I_0 I_2 - I_1^2 = \frac{1}{2}\int_0^1\int_0^1 (x-y)^2 \rho_0(x)\rho_0(y)\,\mathrm{d}x\,\mathrm{d}y > 0. \tag{4.4.20}$$

See Figure 4.9b for a plot of the stable potential temperature profile which minimizes $\mathcal{A}^{(\min)}$ for a given convectively unstable potential temperature profile Θ_U.

4.5 Summary and discussion

Using the energy-Casimir method, conditions for linear equatorial symmetric stability in the anelastic system were derived, and in the case of steady states that are even symmetric about the equator, were shown to imply nonlinear stability. The conditions, that potential vorticity have the sign of latitude and that potential temperature increase (decrease) in the direction of the local planetary rotation vector in the northern (southern) hemisphere, are formally identical with the corresponding conditions derived in Chapter 3 for the Euler equations.

Unlike in the Euler equations case, however, in the anelastic case, the "inertial stability" condition — that angular momentum increase towards the equator — implied by the other two conditions through thermal wind balance, does *not* depend on the equilibrium pressure field. Because of this difference the steady states considered in Chapter 3 that were inertially and statically stable but, in the vicinity of the equator, failed the potential vorticity condition with respect to the Euler equations, are stable with respect to the anelastic equations.

We solved an anelastic version of the problem of Dunkerton (1981), namely the computation of the normal mode solution of the equations linearized about the inertially unstable state with constant meridional shear in the zonal velocity at the equator. The solution in the anelastic case is qualitatively very similar to the hydrostatic case considered by Dunkerton, differing only for small values of aspect ratio $\alpha = L/H$, for which the hydrostatic approximation becomes unrealistic (see Chapter 2).

Following Mu et al. (1996), we used a class of nonlinearly stable states (which are necessarily even symmetric functions of y) to calculate a rigorous upper bound on the kinetic energy released when an unstable state is perturbed slightly from equilibrium, the so-called saturation bound. We showed that for the unstable state of nonzero meridional shear in the zonal velocity at the equator, the saturation bound approaches zero as the shear at the equator approaches zero.

The Lyapunov (nonlinear) stability result does not apply to steady states that are not even symmetric with respect to the equator. In fact, we saw that the pseudoenergy $\mathcal{A}(\mathbf{x};\mathbf{X})$, which for stability must be positive definite, can be negative if $\mathbf{X}$ is asymmetric even if the linear stability criteria are satisfied. The minimum (most negative) value of $\mathcal{A}$ is apparently bounded in terms of $\mathbf{X}$, suggesting that if the symmetric part of $\mathcal{A}$, i.e. $\mathcal{A}-\mathcal{A}_a$, is positive definite, then a norm $||\mathbf{x}-\mathbf{X}||$ is bounded in terms of its initial value, but not arbitrarily close to zero. The bound on $||\mathbf{x}-\mathbf{X}||$ depends on $\sqrt{|\mathcal{A}_a^{(\min)}|}$, where $\mathcal{A}_a^{(\min)}$ is the most negative value of the asymmetric part of $\mathcal{A}$, a lower bound on which is given by (4.2.26a). Explicitly, for any $\varepsilon > \sqrt{|\mathcal{A}_a^{(\min)}|}$, $||\mathbf{x}(t)-\mathbf{X}|| < \varepsilon$ for all t provided

$$||\mathbf{x}(0) - \mathbf{X}|| < \delta \equiv k_\lambda \sqrt{\varepsilon^2 - |\mathcal{A}_a^{(\min)}|}, \tag{4.5.1}$$

where $k_\lambda > 0$ depends on the ratio of the minimum to the maximum eigenvalue of the coefficient matrix of the integrand of the symmetric part of $\mathcal{A}$ (see (4.2.28)). Note that this is not a statement of Lyapunov stability because ε cannot be arbitrarily small.

Asymmetric basic states can also be used to calculate saturation bounds, with the meridional-vertical kinetic energy $\mathcal{K}_\perp$ bounded by

$$\mathcal{K}_\perp(t) \leq \mathcal{A}(\mathbf{X}_U;\mathbf{X}) + |\mathcal{A}_a^{(\min)}(\mathbf{X})|, \tag{4.5.2}$$

for all unstable equilibria $\mathbf{X}_U$ and *linearly* stable $\mathbf{X}$. Since $|\mathcal{A}_a^{(\min)}(\mathbf{X})|$ is probably a gross overestimation of the actual minimum value that $\mathcal{A}$ is likely to attain, it seems unlikely that (4.5.2) can have practical application to, for example, a parameterization scheme.

The above statements about asymmetric basic states assume that $|\mathcal{A}_a|$ can be bounded in terms of the basic state $\mathbf{X}$, which assumes that the Casimir density functions $C^-(m,\theta)$ and $C^+(m,\theta)$ can always be defined such that the absolute value of their difference has a maximum value, which we can compute. We showed how this can be done for a simple case in Example 3 of Section 4.2.3, in which the Θ dependences of C^- and C^+ were identical, but we did not show that it can be done in the general case.

4.A Derivation of anelastic equations

Exact equations

Consider the Euler equations for adiabatic flow of an ideal gas on the equatorial β-plane with first-order representation of both the $\sin\phi$ and $\cos\phi$ Coriolis acceleration terms (Equations (2.3.18)):

$$u_t + uu_x + vu_y + wu_z = \beta yv - \gamma w - \tfrac{1}{\rho}p_x, \tag{4.A.1a}$$

$$v_t + uv_x + vv_y + wv_z = -\beta yu - \tfrac{1}{\rho}p_y, \tag{4.A.1b}$$

$$w_t + uw_x + vw_y + ww_z = \gamma u - g - \tfrac{1}{\rho}p_z, \tag{4.A.1c}$$

$$\rho_t + u\rho_x + v\rho_y + w\rho_z = -\rho(u_x + v_y + w_z), \tag{4.A.1d}$$

$$\theta_t + u\theta_x + v\theta_y + w\theta_z = 0, \tag{4.A.1e}$$

where θ is potential temperature. The equation of state is the ideal gas law, $p = \rho RT$, where T is temperature[1]. It will prove convenient to replace p with the non-dimensional *Exner pressure* π, defined by

$$\pi \equiv \left(\frac{p}{p_{00}}\right)^{\kappa}, \tag{4.A.2}$$

where p_{00} is a constant reference pressure, and $\kappa \equiv R/c_p$, with c_p being the heat capacity with respect to constant pressure processes. ρ and T can be written as functions of θ and π according to

$$\rho = \left(\frac{p_{00}}{R}\right)\frac{\pi^{\frac{1}{\kappa}-1}}{\theta}, \qquad T = \pi\theta. \tag{4.A.3}$$

We non-dimensionalize the equations by introducing characteristic length, time and potential temperature scales. We temporarily adopt the convention that symbols followed by an asterisk are non-dimensional, and let

$$(x, y) = L(x^*, y^*), \qquad z = Hz^*, \qquad t = \tau t^*, \qquad \theta = \Theta\theta^*. \tag{4.A.4}$$

[1] Do not confuse T with equilibrium temperature, as it is in Chapter 3. The symbol τ is used in this chapter for the characteristic time scale.

The associated velocity, density and temperature scales are defined by

$$(u, v) = (L/\tau)(u^*, v^*), \qquad w = (H/\tau)w^*, \tag{4.A.5a}$$

$$\rho = (p_{00}/R\Theta)\rho^*, \qquad T = \Theta T^*. \tag{4.A.5b}$$

The Coriolis parameters are scaled in a trivial way:

$$\beta = \frac{2\Omega}{a}\beta^*, \qquad \gamma = 2\Omega\gamma^*, \tag{4.A.6}$$

with $\gamma^* = \beta^* = 1$. We leave γ^* and β^* in the equations so that we can monitor the effects of rotation on the conditions for equilibrium and stability.

To reduce the clutter in the equations, the asterisk notation is suppressed unless there is possibility of confusion. Henceforth, all variables are assumed to be dimensionless unless otherwise stated. Equations (4.A.1) become

$$\begin{aligned}
\frac{L^2}{\tau^2}\frac{Du}{Dt} &= \frac{2\Omega L^2}{\tau}\left(\frac{L}{a}\beta y v - \frac{H}{L}\gamma w\right) - (c_p\Theta)\theta\frac{\partial \pi}{\partial x}, && \text{(4.A.7a)}\\
\frac{L^2}{\tau^2}\frac{Dv}{Dt} &= \frac{2\Omega L^2}{\tau}\left(-\frac{L}{a}\beta y u\right) - (c_p\Theta)\theta\frac{\partial \pi}{\partial y}, && \text{(4.A.7b)}\\
\frac{H^2}{\tau^2}\frac{Dw}{Dt} &= \frac{2\Omega H L}{\tau}\gamma u - Hg - (c_p\Theta)\theta\frac{\partial \pi}{\partial z}, && \text{(4.A.7c)}\\
\left(1 - \frac{1}{\kappa}\right)\frac{D}{Dt}(\ln \pi) &= \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} + \frac{\partial w}{\partial z}, && \text{(4.A.7d)}\\
\frac{D\theta}{Dt} &= 0, && \text{(4.A.7e)}
\end{aligned}$$

The system (4.A.7) is exact insofar as (4.A.1) is exact. Of course, the scale factors have been introduced with the understanding that the nondimensional variables are of order unity. The specification of the scale factors is tantamount to choosing a regime of solution space for which certain approximations are justifiable.

Classical anelastic approximation

To obtain the classical anelastic equations, we choose the characteristic time scale τ to be

$$\tau = \frac{L}{H}N^{-1} = \sqrt{\frac{L^2}{gH\epsilon}}, \tag{4.A.8}$$

where $\epsilon \equiv \Delta\theta/\Theta = \Delta\theta^*$, and $N^2 \equiv (g/H)(\Delta\theta/\Theta)$, where $\Delta\theta$ is the width of the range of values to which θ is presumed to be limited.

Inserting (4.A.8) into the equations (4.A.7a-c) gives

$$\left(\frac{Hg}{c_p\Theta}\right)\epsilon\left[\frac{Du}{Dt} - \left(\frac{2\Omega}{N}\right)\left(\frac{L}{H}\right)\left(\frac{L}{a}\beta yv - \frac{H}{L}\gamma w\right)\right] + \theta\frac{\partial\pi}{\partial x} = 0, \quad (4.A.9a)$$

$$\left(\frac{Hg}{c_p\Theta}\right)\epsilon\left[\frac{Dv}{Dt} + \left(\frac{2\Omega}{N}\right)\left(\frac{L}{H}\right)\left(\frac{L}{a}\beta yu\right)\right] + \theta\frac{\partial\pi}{\partial y} = 0, \quad (4.A.9b)$$

$$\left(\frac{H}{L}\right)^2\left(\frac{Hg}{c_p\Theta}\right)\epsilon\left[\frac{Dw}{Dt} - \left(\frac{2\Omega}{N}\right)\left(\frac{L}{H}\right)^2(\gamma u)\right] + \frac{Hg}{c_p\Theta} + \theta\frac{\partial\pi}{\partial z} = 0, \quad (4.A.9c)$$

where one factor of τ in the Coriolis terms is deliberately left in terms of N instead of ϵ to emphasize that the decision to neglect those terms is dependent on the size of $(2\Omega L)/(NH)$ and not on the smallness of ϵ.

Define the dimensionless parameters

$$B \equiv \frac{Hg}{c_p\Theta}, \qquad S \equiv \frac{H}{L}\left(\frac{N}{2\Omega}\right), \quad (4.A.10a)$$

$$\alpha \equiv \frac{L}{H}, \qquad \delta \equiv \frac{a}{l}. \quad (4.A.10b)$$

α is the aspect ratio of the flow, and δ is the *planetary* aspect ratio. S is a form of the *Burger number*. B may be interpreted in terms of the thickness of an atmosphere with flat bottom topography at $z = 0$ and uniform potential temperature Θ (i.e. $\theta^* = 1$) in hydrostatic balance. This follows from the nondimensional hydrostatic balance condition

$$B + \theta\frac{\partial\pi}{\partial z} = 0, \quad (4.A.11)$$

whence, for $\theta = 1$,

$$\pi_{\text{hydrostatic}}(z) = 1 - Bz. \quad (4.A.12)$$

The pressure $\pi_{\text{hydrostatic}}$ vanishes at $z = B^{-1}$ so we may write $B = H/H_{\max}$, where $H_{\max}$ is the maximum thickness of the corresponding isentropic atmosphere. Obviously, B must be less than unity for the equations to have any meaning — otherwise, there would be negative pressure in the upper part of the domain.

Perturbation expansion

The (exact) dimensionless momentum equations as they now stand are

$$\epsilon\left[\frac{Du}{Dt} - \beta_\delta yv + \gamma_\alpha w\right] + \frac{1}{B}\theta\frac{\partial\pi}{\partial x} = 0, \tag{4.A.13a}$$

$$\epsilon\left[\frac{Dv}{Dt} + \beta_\delta yu\right] + \frac{1}{B}\theta\frac{\partial\pi}{\partial y} = 0, \tag{4.A.13b}$$

$$\epsilon\left[\frac{1}{\alpha^2}\frac{Dw}{Dt} - (\gamma_\alpha u)\right] + 1 + \frac{1}{B}\theta\frac{\partial\pi}{\partial z} = 0, \tag{4.A.13c}$$

where $\beta_\delta \equiv \beta/(S\delta)$ and $\gamma_\alpha \equiv \gamma/(S\alpha)$.

Consider the perturbation expansion in ϵ of all dependent variables:

$$\boldsymbol{v} = \boldsymbol{v}_0 + \epsilon\boldsymbol{v}_1 + \epsilon^2\boldsymbol{v}_2 + \ldots, \tag{4.A.14a}$$

$$\pi = \pi_0 + \epsilon\pi_1 + \epsilon^2\pi_2 + \ldots, \tag{4.A.14b}$$

$$\theta = 1 + \epsilon\theta_1 + \epsilon^2\theta_2 + \ldots, \tag{4.A.14c}$$

where in (4.A.14c), we have followed Ogura and Phillips (1962) and set the leading order term of θ equal to unity because of (4.1.2). This is appropriate since the departure from $\theta^* = 1$ is assumed to be of order ϵ.

Substituting (4.A.14) into (4.A.13) and (4.A.7d) yields the zeroth order equations

$$\frac{\partial\pi_0}{\partial x} = \frac{\partial\pi_0}{\partial y} = 0, \tag{4.A.15a}$$

$$\frac{\partial\pi_0}{\partial z} = -B, \tag{4.A.15b}$$

$$\frac{\partial\rho_0}{\partial t} = -\left[\frac{\partial}{\partial x}(\rho_0 u_0) + \frac{\partial}{\partial y}(\rho_0 v_0) + \frac{\partial}{\partial z}(\rho_0 w_0)\right], \tag{4.A.15c}$$

where

$$\rho_0 = \pi_0^{\frac{1}{\kappa}-1}. \tag{4.A.16}$$

We can solve (4.A.15a,b) for π_0 (and hence for ρ_0):

$$\pi_0(z,t) = \pi_0(0,t) - Bz, \tag{4.A.17}$$

and, assuming mass conserving boundary conditions, integrating (4.A.15c) over the entire domain gives

$$\iiint \frac{\partial \rho_0}{\partial t}\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = -\iiint \nabla\cdot(\rho_0 \boldsymbol{v}_0)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z = 0. \tag{4.A.18}$$

Substitution of (4.A.17) into (4.A.18) using (4.A.16) leads to, after some manipulation,

$$\left[\frac{\partial}{\partial t}\pi_0(0,t)\right]\int \pi_0^{\frac{1}{\kappa}-2}\mathrm{d}z = 0. \tag{4.A.19}$$

Since $\pi_0 > 0$ everywhere in the domain, it follows that $\pi_0(0,t)$ is a constant. If we choose p_{00} to be the surface pressure of the hydrostatic isentropic atmosphere with $\theta \equiv 1$, then we have that $\pi_0|_{z=0} \equiv 1$. It follows that $\partial\rho_0/\partial t = 0$, and the zeroth order continuity equation becomes

$$\frac{\partial}{\partial x}(\rho_0 u_0) + \frac{\partial}{\partial y}(\rho_0 v_0) + \frac{\partial}{\partial z}(\rho_0 w_0) = 0. \tag{4.A.20}$$

That is, the divergence of the $\mathcal{O}(1)$ mass flux vector $\rho_0\boldsymbol{v}_0$ vanishes.

The $\mathcal{O}(\epsilon)$ equations are

$$\begin{aligned}
\frac{D_0 u_0}{Dt} - \beta_\delta y v_0 + \gamma_\alpha w_0 + \frac{1}{B}\frac{\partial \pi_1}{\partial x} &= 0, && \text{(4.A.21a)}\\
\frac{D_0 v_0}{Dt} + \beta_\delta y u_0 + \frac{1}{B}\frac{\partial \pi_1}{\partial y} &= 0, && \text{(4.A.21b)}\\
\frac{D_0 w_0}{Dt} + \alpha^2\left[-\gamma_\alpha u_0 + \frac{1}{B}\frac{\partial \pi_1}{\partial z} - \theta_1\right] &= 0, && \text{(4.A.21c)}\\
\frac{D_0 \theta_1}{Dt} &= 0, && \text{(4.A.21d)}
\end{aligned}$$

where

$$\frac{D_0}{Dt} = \frac{\partial}{\partial t} + u_0\frac{\partial}{\partial x} + v_0\frac{\partial}{\partial y} + w_0\frac{\partial}{\partial z}. \tag{4.A.22}$$

The final equation required is obtained by noting that the $\mathcal{O}(1)$ anelastic continuity equation (4.A.20) must hold for all time. Equivalently, the time derivative of the left hand side must vanish. Using (4.A.21a-c), we arrive at the elliptic (which is to say readily solved numerically) diagnostic equation for the pressure perturbation π_1

$$\frac{1}{B}\left[\frac{\partial}{\partial x}\left(\rho_0\frac{\partial\pi_1}{\partial x}\right)+\frac{\partial}{\partial y}\left(\rho_0\frac{\partial\pi_1}{\partial y}\right)+\alpha^2\frac{\partial}{\partial z}\left(\rho_0\frac{\partial\pi_1}{\partial z}\right)\right]$$

$$= \alpha\frac{\partial}{\partial z}(\rho_0\theta_1) - \nabla\cdot[\rho_0(\boldsymbol{v}_0\cdot\nabla)\boldsymbol{v}_0]$$

$$+\left\{\frac{\partial}{\partial x}\left[\rho_0\left(\beta_\delta y v_0-\gamma_\alpha w_0\right)\right]+\frac{\partial}{\partial y}\left[\rho_0(-\beta_\delta y u_0)\right]+\alpha\frac{\partial}{\partial z}[\rho_0(\gamma_\alpha u_0)]\right\}. \tag{4.A.23}$$

We now have a closed set of equations for $\boldsymbol{v}_0$, π_1 and θ_1, with ρ_1, p_1 and T_1 obtained from

$$\frac{\rho_1}{\rho_0} = \left(\frac{1-\kappa}{\kappa}\right)\frac{\pi_1}{\pi_0}-\theta_1, \tag{4.A.24a}$$

$$\frac{p_1}{p_0} = \frac{1}{\kappa}\frac{\pi_1}{\pi_0}, \tag{4.A.24b}$$

$$\frac{T_1}{T_0} = \frac{\pi_1}{\pi_0}+\theta_1, \tag{4.A.24c}$$

where, from (4.A.2) and (4.A.3), $p_0=\pi_0^{\frac{1}{\kappa}}$ and $T_0=\pi_0$.

Energy in anelastic system

The anelastic system as derived above conserves a form of the energy functional that is almost the $\mathcal{O}(\epsilon)$ approximation to the energy conserved by the fully compressible equations. We may discover the form that the energy functional takes by the following educated trial and error approach.

The total energy conserved by the compressible Euler equations is

$$E_{\text{comp}} = \iiint \rho\left(\tfrac{1}{2}|\boldsymbol{v}|^2+gz+c_vT\right)\,\mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.A.25}$$

where all of the dependent and independent variables are dimensional. In terms of nondimensional variables,

$$E_{\text{comp}} = \left(\frac{p_{00}}{\kappa}\right)L^2HB\iiint\rho^*\left\{\frac{\epsilon}{2}\left[\ u^{*2}+v^{*2}+\frac{1}{\alpha^2}w^{*2}\right]\right.$$
$$\left.+z^*+\frac{1}{B}(1-\kappa)T^*\right\}\mathrm{d}x^*\mathrm{d}y^*\mathrm{d}z^* \tag{4.A.26}$$

Naturally, we define the dimensionless energy E^*_{comp} to be the integral in (4.A.26). Dropping the asterisks once more, we have

$$E_{\text{comp}} = \iiint \rho \left\{ \frac{\epsilon}{2}\left[u^2 + v^2 + \frac{1}{\alpha^2}w^2\right] + z + \frac{1}{B}(1-\kappa)T \right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.A.27}$$

where all variables are dimensionless.

If we now substitute the series (4.A.14) into (4.A.27), we find

$$\begin{aligned} E_{\text{comp}} &= \iiint \rho_0 \left[\left(\frac{1-\kappa}{B}\right) + \kappa z\right] \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &\quad + \epsilon \iiint \rho_0 \left\{ \tfrac{1}{2}\left[u_0^2 + v_0^2 + \frac{1}{\alpha^2}w_0^2\right] - \theta_1 z \right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z \\ &\quad + \epsilon \iiint \left(\frac{1-\kappa}{B}\right)\left(\rho_0 \frac{p_1}{p_0}\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z + \mathcal{O}(\epsilon^2). \end{aligned} \tag{4.A.28}$$

The first integrand in (4.A.28) is a function of z and as such the integral is trivially a constant. We may expect, therefore, that the anelastic equations (4.A.21) conserve the $\mathcal{O}(\epsilon)$ terms in (4.A.28). However, it happens that the equations conserve the second integral exactly and not the third,

$$E_{\text{elastic}} \equiv \epsilon \iiint \left(\frac{1-\kappa}{B}\right)\left(\rho_0 \frac{p_1}{p_0}\right) \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.A.29}$$

the so called *elastic* energy term. Sound wave solutions, which we want to exclude, involve the oscillating exchange of elastic and kinetic energies. The omission of this term from the energy is what gives the system its name. Hence, we define the *anelastic* energy by

$$E \equiv \iiint \rho_0 \left\{ \tfrac{1}{2}\left[u_0^2 + v_0^2 + \frac{1}{\alpha^2}w_0^2\right] - \theta_1 z \right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z. \tag{4.A.30}$$

It may be verified that E is conserved by the equations subject to "conservative" boundary conditions (e.g. no normal flow on the boundaries of a closed domain).

Modified anelastic system

It was decided (Wilhelmson and Ogura, 1972) that the anelastic system of Ogura and Phillips (1962) was too coarse an approximation for the horizontal momentum equations for some purposes.

The horizontal pressure gradient force in (4.A.21a,b) depends only on the mean potential temperature Θ (dimensional). For quasihorizontal motions which do not significantly alter the potential temperature profile from a slowly varying basic state $\theta_0(z)$, it is preferable to include an $\mathcal{O}(\epsilon^2)$ term that accounts for that behaviour.

The resulting modified anelastic equations cannot be derived in as clean a manner as the classic anelastic equations. We argue somewhat qualitatively below.

Consider the alternative expansion for the (dimensionless) potential temperature

$$\begin{aligned} \theta &= 1 + \epsilon(\bar{\theta_1}(z) + \theta_1) + \epsilon^2\theta_2 + \dots \\ &\equiv \theta_0(z) + \epsilon\theta_1 + \epsilon^2\theta_2 + \dots, \end{aligned} \tag{4.A.31}$$

in which θ_1 has been redefined implicitly to be the departure of the potential temperature perturbation from the specified profile $\bar{\theta_1}(z)$.

The equations of motion are adjusted in order to accomodate the extra precision in the horizontal momentum equations while still conserving energy (Lipps and Hemler, 1982). This involves changing both the $\mathcal{O}(1)$ and $\mathcal{O}(\epsilon)$ vertical momentum equations. The $\mathcal{O}(1)$ equation (4.A.15b) becomes

$$B + \theta_0(z)\frac{\mathrm{d}\pi_0}{\mathrm{d}z} = 0. \tag{4.A.32}$$

The $\mathcal{O}(\epsilon)$ equation (4.A.21c) becomes

$$\frac{D_0 w_0}{Dt} + \alpha^2\left\{-\gamma_\alpha u_0 + \frac{1}{B}\left[\frac{\partial}{\partial z}(\theta_0\pi_1) + \frac{\mathrm{d}\pi_0}{\mathrm{d}z}\theta_1\right]\right\} = 0, \tag{4.A.33}$$

where the new term $(\pi_1/B)\mathrm{d}\theta_0/\mathrm{d}z$ is added solely to ensure conservation of energy. The new term is $\mathcal{O}(\epsilon)$ smaller than the other terms. Evidently, we have moved some of the precision from the vertical to the horizontal equations.

We summarize the modified anelastic equations:

$$\frac{D_0 u_0}{Dt} - \beta_\delta y v_0 + \gamma_\alpha w_0 + \frac{1}{B}\theta_0\frac{\partial \pi_1}{\partial x} = 0, \tag{4.A.34a}$$

$$\frac{D_0 v_0}{Dt} + \beta_\delta y u_0 + \frac{1}{B}\theta_0\frac{\partial \pi_1}{\partial y} = 0, \tag{4.A.34b}$$

$$\frac{D_0 w_0}{Dt} + \alpha^2 \left\{ -\gamma_\alpha u_0 + \frac{1}{B} \left[\frac{\partial}{\partial z}(\theta_0 \pi_1) + \frac{\mathrm{d}\pi_0}{\mathrm{d}z}\theta_1 \right] \right\} = 0, \tag{4.A.34c}$$

$$\frac{D_0 \theta_1}{Dt} + \frac{w_0}{\epsilon}\frac{\mathrm{d}\theta_0}{\mathrm{d}z} = 0, \tag{4.A.34d}$$

$$\frac{\partial}{\partial x}(\rho_0 u_0) + \frac{\partial}{\partial y}(\rho_0 v_0) + \frac{\partial}{\partial z}(\rho_0 w_0) = 0. \tag{4.A.34e}$$

Notice the thermodynamic equation (4.A.34d) has changed formally only (because of the introduction of $\theta_0(z)$ and the redefinition of θ_1). The equation for π_1 is again

$$\frac{\partial}{\partial t}\left[\nabla \cdot (\rho_0 \boldsymbol{v}_0)\right] = 0. \tag{4.A.35}$$

The other thermodynamic variables ρ, p, and T change in the obvious ways, satisfying

$$\frac{\rho_1}{\rho_0} = \left(\frac{1-\kappa}{\kappa}\right)\frac{\pi_1}{\pi_0} - \frac{\theta_1}{\theta_0}, \tag{4.A.36a}$$

$$\frac{p_1}{p_0} = \frac{1}{\kappa}\frac{\pi_1}{\pi_0}, \tag{4.A.36b}$$

$$\frac{T_1}{T_0} = \frac{\pi_1}{\pi_0} + \frac{\theta_1}{\theta_0}, \tag{4.A.36c}$$

where now $\pi_0(z)$ is determined from (4.A.32), $\rho_0 = (1/\theta_0)\pi_0^{\frac{1}{\kappa}-1}$, and $T_0 = \theta_0 \pi_0$.

The anelastic energy in this case is

$$E = \iiint \rho_0 \left\{ \frac{1}{2}\left[u_0^2 + v_0^2 + \frac{1}{\alpha^2}w_0^2\right] + \frac{1}{B}\pi_0\theta_1) \right\} \mathrm{d}x\,\mathrm{d}y\,\mathrm{d}z, \tag{4.A.37}$$

which no longer has the neat interpretation as the total energy less the pressure perturbation term, but it differs from the exact anelastic energy only by $\mathcal{O}(\epsilon)$.

4.B Functional derivatives of $\mathcal{H}$

Let $\mathbf{w} = (\Delta m, \Delta\zeta, \Delta\theta)^T$ and $\varepsilon \in \mathbb{R}$. Then

$$\begin{aligned} \mathcal{H}(\mathbf{x} + \varepsilon\mathbf{w}) = & \iint \Big\{ \rho_0 \left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(m + \varepsilon\Delta m) \\ & + \frac{1}{2\rho_0}\left[\left(\frac{\partial}{\partial z}(\psi + \varepsilon\Delta\psi)\right)^2 + \frac{1}{\alpha^2}\left(\frac{\partial}{\partial y}(\psi + \varepsilon\Delta\psi)\right)^2 \right] \\ & + \frac{1}{\epsilon B}\rho_0\pi_0(\theta + \varepsilon\Delta\theta) \Big\} \mathrm{d}y\,\mathrm{d}z, \end{aligned} \tag{4.B.1}$$

where $\Delta\psi$ is the change in ψ induced by the change $\Delta\zeta$. From (4.1.20),

$$\Delta\zeta = -\left[\frac{\partial}{\partial z}\left(\frac{1}{\rho_0}\frac{\partial}{\partial z}(\Delta\psi)\right) + \frac{1}{\alpha^2}\frac{\partial}{\partial y}\left(\frac{1}{\rho_0}\frac{\partial}{\partial y}(\Delta\psi)\right)\right]. \tag{4.B.2}$$

Expanding (4.B.1),

$$\begin{aligned}\mathcal{H}(\mathbf{x}+\varepsilon\mathbf{w}) \;=\; & \iint \Bigg\{\rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(m+\varepsilon\Delta m) \\ & + \frac{1}{2\rho_0}\left[\left(\frac{\partial\psi}{\partial z}\right)^2 + \frac{1}{\alpha^2}\left(\frac{\partial\psi}{\partial y}\right)^2\right] \\ & + \frac{\varepsilon}{\rho_0}\left[\left(\frac{\partial\psi}{\partial z}\right)\left(\frac{\partial\Delta\psi}{\partial z}\right) + \frac{1}{\alpha^2}\left(\frac{\partial\psi}{\partial y}\right)\left(\frac{\partial\Delta\psi}{\partial y}\right)\right] \\ & + \frac{1}{\epsilon B}\rho_0\pi_0(\theta+\varepsilon\Delta\theta)\Bigg\}\,\mathrm{d}y\,\mathrm{d}z + \mathcal{O}(\varepsilon^2),\end{aligned} \tag{4.B.3}$$

Differentiating with respect to ε,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\mathcal{H}(\mathbf{x}+\varepsilon\mathbf{w}) \;=\; & \iint \Bigg\{\rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(\Delta m) \\ & + \frac{1}{\rho_0}\left[\left(\frac{\partial\psi}{\partial z}\right)\left(\frac{\partial\Delta\psi}{\partial z}\right) + \frac{1}{\alpha^2}\left(\frac{\partial\psi}{\partial y}\right)\left(\frac{\partial\Delta\psi}{\partial y}\right)\right] \\ & + \frac{1}{\epsilon B}\rho_0\pi_0(\Delta\theta)\Bigg\}\,\mathrm{d}y\,\mathrm{d}z + \mathcal{O}(\varepsilon).\end{aligned} \tag{4.B.4}$$

Integrating the middle term by parts, and, finally, setting $\varepsilon = 0$,

$$\begin{aligned}\left.\frac{\mathrm{d}}{\mathrm{d}\varepsilon}\right|_{\varepsilon=0}\mathcal{H}(\mathbf{x}+\varepsilon\mathbf{w}) \;=\; & \iint \Bigg\{\rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)(\Delta m) \\ & + \psi(\Delta\zeta) + \frac{1}{\epsilon B}\rho_0\pi_0(\Delta\theta)\Bigg\}\,\mathrm{d}y\,\mathrm{d}z.\end{aligned} \tag{4.B.5}$$

Comparing (4.B.5) with (4.1.24), the functional gradient of $\mathcal{H}$ has components

$$\frac{\delta\mathcal{H}}{\delta m} = \rho_0\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), \tag{4.B.6a}$$

$$\frac{\delta\mathcal{H}}{\delta\zeta} = \psi, \tag{4.B.6b}$$

$$\frac{\delta\mathcal{H}}{\delta\theta} = \frac{1}{\epsilon B}\rho_0\pi_0. \tag{4.B.6c}$$

4.C Proof of conservation of $\mathcal{H}_L$ by linearized dynamics

In Section 4.2.1, the Hamiltonian for the linearized dynamics was defined

$$\begin{aligned}\mathcal{H}_L \;=\; & \iint\limits_{\mathcal{D}} \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}(\psi'_y)^2 + (\psi'_z)^2\right] \mathrm{d}y\,\mathrm{d}z \\ & + \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \rho_0 \left[C^{(i)}_{mm}(M,\Theta)m'^2 + 2C^{(i)}_{\theta m}(M,\Theta)\theta' m' + C^{(i)}_{\theta\theta}(M,\Theta)\theta'^2\right] \mathrm{d}y\,\mathrm{d}z\end{aligned} \tag{4.C.1}$$

We now show that $\mathcal{H}_L$ is conserved by the linearized equations. Taking its time derivative, we find

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\mathcal{H}_L \;=\; & \iint\limits_{\mathcal{D}} \frac{1}{\rho_0}\left[\frac{1}{\alpha^2}\psi'_y\psi'_{yt} + \psi'_z\psi'_{zt}\right] \mathrm{d}y\,\mathrm{d}z \\ & + \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \rho_0 \left[\left(C^{(i)}_{mm}(M,\Theta)m' + C^{(i)}_{\theta m}(M,\Theta)\theta'\right) m'_t\right. \\ & \left. + \left(C^{(i)}_{\theta\theta}(M,\Theta)\theta' + C^{(i)}_{m\theta}(M,\Theta)m'\right)\theta'_t\right] \mathrm{d}y\,\mathrm{d}z.\end{aligned} \tag{4.C.2}$$

Substituting (4.2.5):

$$\begin{aligned}\frac{1}{2}\frac{d}{dt}\mathcal{H}_L \;=\; & \iint\limits_{\mathcal{D}} \psi'\left\{\partial\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), m'\right] + \partial\left(\frac{\pi_0}{\epsilon B}, \theta'\right)\right\} \mathrm{d}y\,\mathrm{d}z \\ & + \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \left[\left(C^{(i)}_{mm}(M,\Theta)m' + C^{(i)}_{\theta m}(M,\Theta)\theta'\right)\partial(\psi', M)\right. \\ & \left. + \left(C^{(i)}_{\theta\theta}(M,\Theta)\theta' + C^{(i)}_{m\theta}(M,\Theta)m'\right)\partial(\psi', \Theta)\right] \mathrm{d}y\,\mathrm{d}z.\end{aligned} \tag{4.C.3}$$

Expanding the integrals in the sum by parts:

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\mathcal{H}_L &= \iint\limits_{\mathcal{D}} \psi' \left\{ \partial\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), m'\right] + \partial\left(\frac{\pi_0}{\epsilon B}, \theta'\right)\right\} \mathrm{d}y\,\mathrm{d}z \\
&\quad - \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \psi' \Big[\partial(C^{(i)}_{mm}(M,\Theta)m', M) + \partial(C^{(i)}_{m\theta}(M,\Theta)\theta', M) \\
&\qquad + \partial(C^{(i)}_{m\theta}(M,\Theta)m', \Theta) + \partial(C^{(i)}_{\theta\theta}(M,\Theta)\theta', \Theta)\Big] \mathrm{d}y\,\mathrm{d}z \\
&\quad + \sum_{i,j=1}^{n} \int\limits_{\partial\mathcal{D}^{(i)}\cap\partial\mathcal{D}^{(j)}} \psi' \left\{ \hat{\imath} \times \left[\left((C^{(i)}_{mm} - C^{(j)}_{mm})m' + (C^{(i)}_{m\theta} - C^{(j)}_{m\theta})\theta'\right)\nabla M\right.\right. \\
&\qquad \left.\left. + \left((C^{(i)}_{m\theta} - C^{(j)}_{m\theta})m' + (C^{(i)}_{\theta\theta} - C^{(j)}_{\theta\theta})\theta'\right)\nabla\Theta\right]\right\} \cdot \hat{\nu}^{(i)} \mathrm{d}l^{(i)}(y,z). \qquad (4.C.4)
\end{aligned}
$$

The boundary integrals vanish if the $C^{(i)}$ are such that their second partial derivatives match on the boundaries of the $\mathcal{D}^{(i)}$ at equilibrium. Expanding the Jacobians using the product rule for derivatives, and substituting from (4.2.15):

$$
\begin{aligned}
\frac{1}{2}\frac{d}{dt}\mathcal{H}_L &= \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \psi' \left\{ \partial\left[\left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right), m'\right] + \partial\left(\frac{\pi_0}{\epsilon B}, \theta'\right)\right. \\
&\qquad + \frac{1}{\rho_0 Q}\left(\partial\left[\Theta, \left(\tfrac{1}{2}\beta_\delta y^2 - \gamma_\alpha z\right)\right]\partial(m', M) + \partial\left(\frac{\pi_0}{\epsilon B}, M\right)\partial(\theta', \Theta)\right. \\
&\qquad\quad \left.\left. + \partial\left(\Theta, \frac{\pi_0}{\epsilon B}\right)\left[\partial(\theta', M) + \partial(m', \Theta)\right]\right)\right\} \mathrm{d}y\,\mathrm{d}z \qquad (4.C.5) \\
&= \sum_{i=1}^{n} \iint\limits_{\mathcal{D}^{(i)}} \psi' \left\{ m'_y \left[\gamma_\alpha + \frac{1}{\rho_0 Q}\left(M_z(-\gamma_\alpha \Theta_y - \beta_\delta y \Theta_z) + \frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\Theta_z\Theta_y\right)\right]\right. \\
&\qquad + m'_z \left[\beta_\delta y + \frac{1}{\rho_0 Q}\left(-M_y(-\gamma_\alpha\Theta_y - \beta_\delta y\Theta_z) - \frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\Theta_y\Theta_y\right)\right] \\
&\qquad + \theta'_y \left[-\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z} + \left(\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\right)\frac{1}{\rho_0 Q}\left(-M_y\Theta_z + M_z\Theta_y\right)\right] \\
&\qquad \left. + \theta'_z \left[\frac{1}{\epsilon B}\frac{\mathrm{d}\pi_0}{\mathrm{d}z}\frac{1}{\rho_0 Q}\left(-\Theta_y(-M_y) - M_y\Theta_y\right)\right]\right\} \mathrm{d}y\,\mathrm{d}z \qquad (4.C.6) \\
&= 0,
\end{aligned}
$$

where (4.2.1) and (4.2.6) are used to see that each term vanishes.

Chapter 5

Nonlinear stability of inviscid homogeneous Taylor-Couette flow

In this chapter, we consider the stability of steady flow of a homogeneous, incompressible fluid confined between two rotating solid cylinders. The subject is known as *Taylor-Couette* flow, after M. Couette, who used a stationary inner cylinder inside a rotating outer cylinder to measure the viscosity of liquids, and G. I. Taylor, who made a comprehensive theoretical and experimental study of the instability of the flow at high rotation rates of the inner cylinder with stationary outer cylinder (Taylor, 1923). The experiment has an obvious connection with the atmospheric problem of symmetric instability.

We focus on the case of both cylinders rotating. By increasing the rate of rotation of the inner cylinder compared to the outer, a stable laminar azimuthal flow can be made unstable. The resulting axisymmetric pattern of helical rolls is known as *Taylor vortex flow*. Further increase of the ratio of rotation rates takes the system through a series of bifurcations corresponding to different classes of stable flow and eventually into turbulence (Andereck et al., 1986). Experimentally, the flow can be made visible by injecting dye into the fluid or by suspending reflective particles that align with the direction of flow. See Tagg (1994) for a review of experimental and numerical results, Chandrasekhar (1961) and Chossat and Iooss (1994) for mathematical treatments, and Koschmieder (1993) for an overview of the general subject.

In the absence of viscosity, the instability is explained by Rayleigh's centrifugal stability theorem (Rayleigh, 1916), which holds that an axisymmetric rotating fluid is stable as long as the magnitude of angular momentum increases with distance from the axis of rotation. If *no-slip* boundary conditions are assumed, then Rayleigh's criterion is always violated by the case of the inner cylinder rotating with the outer cylinder at rest, but viscosity delays the onset of instability. Taylor calculated the neutral stability curve for the linearized viscous problem, and his result agrees well with experiments, including his own. Rayleigh's theorem is the zero viscosity limit of Taylor's result.

The only steady solutions to the viscous Navier-Stokes equations with no-slip boundary conditions on the surfaces of the cylinders have the azimuthal *Couette* velocity profile $V(r) = Ar + B/r$, where the constants A and B depend on the radii and rotation rates of the cylinders. Although the uniqueness of the Couette solution depends on viscosity, the flow is itself independent of the Reynolds number, with $\mathrm{Re} = \infty$ being a singular limit. This suggests that the behaviour of the instability, at least for short times, might be described by inviscid dynamics, and we therefore think it reasonable to consider the stability of the Couette solution under inviscid conditions. Also, since the observed flow for weakly unstable axisymmetric initial conditions is axisymmetric, i.e. Taylor vortex flow, we restrict our attention to axisymmetric disturbances.

The viscous problem has been studied extensively. Joseph (1976) describes an energy method with which it can be shown that flow sufficiently close to solid-body rotation (small $|B|$) is asymptotically stable[1]. If only axisymmetric perturbations are considered, as in our case, then satisfying the Rayleigh criterion implies asymptotic stability when the inner cylinder is rotating faster than, but in the same sense as, the outer. See Figure

[1] The inviscid problem that we consider cannot exhibit asymptotic stability because it would violate energy conservation.

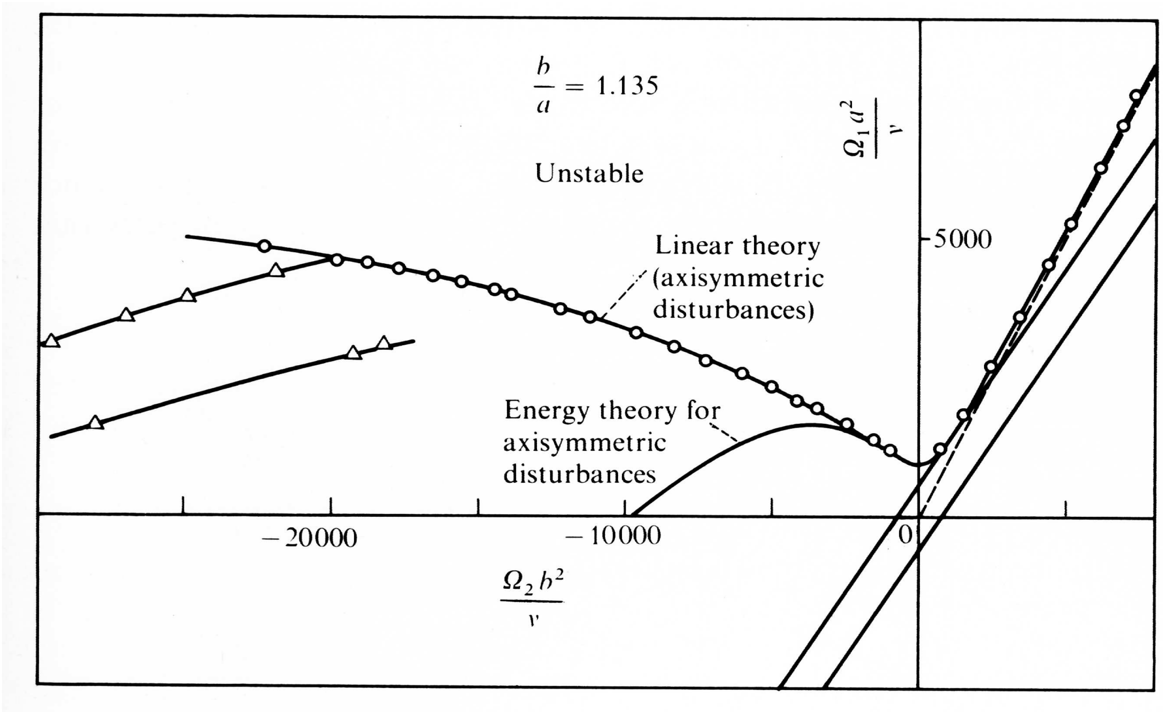


Figure 5.1: Stability regions for viscous Taylor-Couette flow. Ω_1 and Ω_2 are the angular speeds of, respectively, the inner cylinder of radius a and the outer cylinder of radius b. ν is the kinematic viscosity. The dashed line through the origin is the inviscid criterion of Rayleigh (1916). Below the uppermost curve, viscous flow is linearly stable to axisymmetric disturbances (Taylor, 1923). Between the parallel diagonal lines straddling the origin, representing near solid-body rotation, the flow is asymptotically stable (Joseph, 1976). The circles and triangles represent points of instability in experiments of Coles (1965). From Joseph and Hung (1971).

5.1 for a summary of linear and nonlinear results.

Szeri and Holmes (1988) showed that Rayleigh's criterion is valid for finite but small amplitude disturbances to steady axisymmetric inviscid flows in an annular cylinder, of which the Couette profile is a particular case. We extend this result to arbitrarily large disturbances to stable Couette profiles with the restriction that the ratio of cylinder speeds not be too *high* and apply the saturation bound method of Shepherd (1988) to estimate the amount of energy that can be released into the overturning component of Taylor vortex flow.

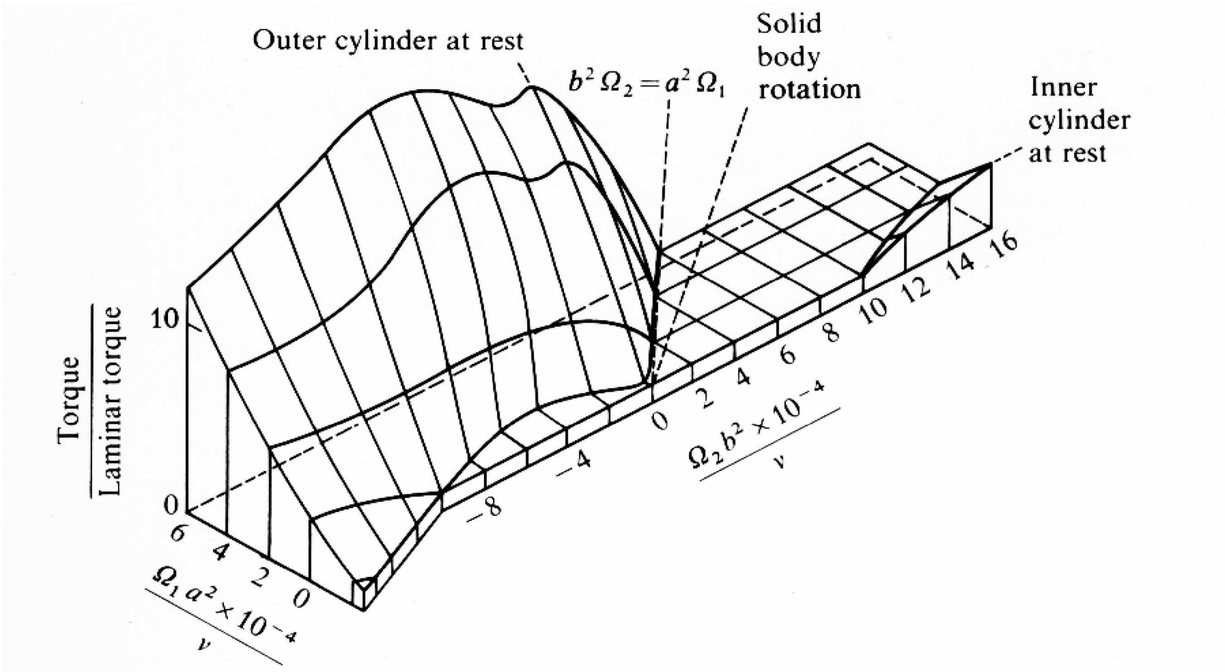


Figure 5.2: Torque required to maintain the cylinders rotating at Ω_1 and Ω_2, respectively, divided by the corresponding torque if the fluid were undergoing Couette flow. Large values indicate instability of Couette flow. Note, in particular, instability at high $\Omega_2 b^2/\Omega_1 a^2$. From Joseph (1976), and see references therein.

Our failure to demonstrate nonlinear stability at high values of the rotation rate of the outer cylinder is qualitatively consistent with experiments. Measurements of the torque required to maintain given cylinder rotation rates compared to the torque required if the fluid between the cylinders were undergoing Couette flow show that Couette flow breaks down at high rotation rates of the outer cylinder (see Figure 5.2).

5.1 Stability of Couette profile

The Euler equations in cylindrical coordinates for axisymmetric motion of an unstratified, incompressible fluid are

$$u_t + uu_r + wu_z \;=\; \tfrac{1}{r}\,v^2 - p_r, \tag{5.1.1a}$$

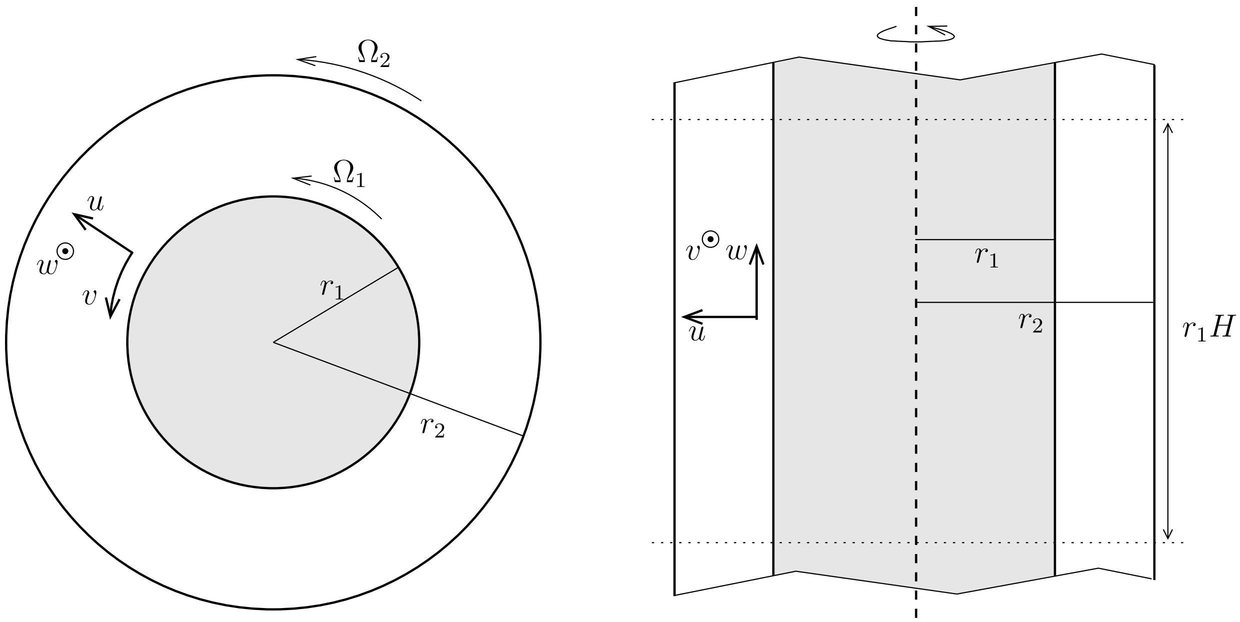


Figure 5.3: Schematic diagram of Taylor-Couette apparatus: top view on left and cross-section through rotation axis on right. The flow is assumed to be periodic in the vertical direction with some large period $r_1 H$.

$$v_t + uv_r + wv_z = -\tfrac{1}{r}\, uv, \tag{5.1.1b}$$

$$w_t + uw_r + ww_z = -p_z, \tag{5.1.1c}$$

$$(ru)_r + (rw)_z = 0, \tag{5.1.1d}$$

where u, v and w are the radial, azimuthal and vertical components of the velocity, p is the fluid pressure, defined so as to incorporate the uniform density and gravity, r is distance from the axis of rotation, and z is height measured from an arbitrary reference point. (5.1.1b) can be rewritten

$$m_t + um_r + wm_z = 0, \tag{5.1.2}$$

where $m \equiv rv$ is the vertical component of angular momentum about the rotation axis. Thus m is a materially conserved quantity. We consider the flow between two infinitely long coaxial cylinders rotating with constant, but independent, angular speeds. Let r_1 be the radius of the inner cylinder and Ω_1 its angular speed, and let r_2 and Ω_2 be the radius and angular speed respectively of the outer cylinder (Figure 5.3). We assume that all variables are periodic in z with period $r_1 H$, and that there is no normal flow at the surfaces of the cylinders (i.e. $u(r = r_1) = u(r = r_2) = 0$). For brevity of notation, we introduce the dimensionless constants $\eta \equiv r_2/r_1$ and $\mu \equiv \Omega_2 r_2^2/\Omega_1 r_1^2$.

Let the equations be nondimensionalized so that r_1 is the unit of length, $\Omega_1 r_1$ the unit of velocity and Ω_1^{-1} the unit of time.

Solutions to (5.1.1) conserve total energy

$$\mathcal{H}(u, v, w) = \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{1}^{\eta} \left[\tfrac{1}{2}(u^2 + v^2 + w^2)\right] r \,\mathrm{d}r \,\mathrm{d}z \tag{5.1.3}$$

and all integrals of the form

$$\mathcal{C}(u) = \int_{-\frac{H}{2}}^{\frac{H}{2}} \int_{1}^{\eta} C(m) r \,\mathrm{d}r \,\mathrm{d}z, \tag{5.1.4}$$

where $C(m)$ is an arbitrary function of angular momentum. $\mathcal{C}$ is a Casimir invariant in the Hamiltonian formulation of the dynamics. Note that in all subsequent integrations, the limits are as in (5.1.3) and (5.1.4) and will be omitted.

We seek conditions for the stability of steady state solutions with v having the form

$$V(r) = Ar + \frac{B}{r}, \tag{5.1.5}$$

and with $u = w = 0$. Assuming no-slip boundary conditions on the surfaces of the cylinders, the Couette coefficients in (5.1.5) are

$$A = \frac{\mu - 1}{\eta^2 - 1}, \qquad B = \frac{\eta^2 - \mu}{\eta^2 - 1}. \tag{5.1.6}$$

To prove that a steady state $\mathbf{X}$ is stable, we need to introduce a positive definite functional (a norm) on the phase space of disturbances to $\mathbf{X}$. Stability can then be defined in terms of bounds on the norm.

Let $||\Delta\mathbf{x}||$ be a norm defined on the disturbances $\Delta\mathbf{x} \equiv \mathbf{x} - \mathbf{X} = (u, v - V(r), w)^T$. The equilibrium $\mathbf{X}$ is stable if for every $\varepsilon > 0$, there exists a $\delta > 0$ such that

$$||\Delta\mathbf{x}(t=0)|| < \delta \Rightarrow ||\Delta\mathbf{x}(t)|| < \varepsilon \quad \forall t > 0. \tag{5.1.7}$$

We can show (5.1.7) by finding a conserved functional that can be bounded from above and below by comparable disturbance norms. The required functional is the disturbance *pseudoenergy*

$$\mathcal{A}(u, v, w; V) \equiv (\mathcal{H} + \mathcal{C})(u, v, w) - (\mathcal{H} + \mathcal{C})(0, V, 0) \tag{5.1.8}$$

with $\mathcal{C}$ defined so that $\mathcal{A}$ is stationary at $\mathbf{X}$. That is,

$$\delta\mathcal{A}(\delta\mathbf{x})|_{\mathbf{X}} = \delta(\mathcal{H} + \mathcal{C})|_{\mathbf{X}} \equiv \iint [(V + C'(M)r)\delta v]\, r\,\mathrm{d}r\,\mathrm{d}z = 0 \tag{5.1.9}$$

for arbitrary variations $\delta\mathbf{x} \equiv (\delta u, \delta v, \delta w)$. We therefore define the arbitrary function $C(\cdot)$ to be that satisfying

$$C'(M) = -\frac{V}{r} = -\frac{AM}{M - B}, \tag{5.1.10}$$

where $M \equiv rV = Ar^2 + B$ is the angular momentum of the steady state, and we have assumed that M is an invertible function of r^2. If $A \neq 0$, then M is invertible, and (5.1.9) can be satisfied at all points in the domain. On the other hand, if $A = 0$ ($\mu = 1$), then M is not invertible, (5.1.9) cannot be satisfied, and this method does not apply. As we will see, $\mu = 1$ is a bifurcation point between stable and unstable steady states. Henceforth, we will assume $A \neq 0$.

Explicitly,

$$\mathcal{A} = \iint \left\{ \tfrac{1}{2}[(\Delta u)^2 + (\Delta v)^2 + (\Delta w)^2] + V\Delta v + C(M + \Delta m) - C(M) \right\} r\,\mathrm{d}r\,\mathrm{d}z, \tag{5.1.11}$$

where $\Delta v \equiv v - V$, etc. We can simplify $\mathcal{A}$ by applying Taylor's remainder theorem to $C(m)$:

$$C(M + \Delta m) - C(M) = \Delta m C'(M) + \tfrac{1}{2}(\Delta m)^2 C''(\tilde{m}), \tag{5.1.12}$$

where $\tilde{m}(r) \in [M, M + \Delta m]$. Noting that $V\Delta v = -\Delta m C'(M)$, this gives

$$\mathcal{A} = \iint \tfrac{1}{2}\left[(\Delta u)^2 + (\Delta w)^2 + (1 + r^2 C''(\tilde{m}))(\Delta v)^2\right] r\,\mathrm{d}r\,\mathrm{d}z. \tag{5.1.13}$$

For infinitesimal disturbances, we may approximate the remainder term in (5.1.12) sufficiently well by evaluating C'' at M. From (5.1.10),

$$C''(M) = \frac{AB}{(M - B)^2}, \tag{5.1.14}$$

so $\mathcal{A}$ is strictly positive if

$$1 + r^2 C''(\tilde{m}) \approx \frac{M}{M - B} = \frac{Ar^2 + B}{Ar^2} = \frac{2}{r}\left(\frac{M}{\mathrm{d}M/\mathrm{d}r}\right) > 0, \tag{5.1.15}$$

and bounded if $A \neq 0$. Since $r \geq 1$, (5.1.15) is satisfied if $\mathrm{d}(M^2)/\mathrm{d}r > 0$ for all r, which is Rayleigh's criterion for stability. In terms of the coefficients A and B, (5.1.15) is satisfied if $B/A > -1$, which is true if and only if $\mu > 1$.

For arbitrarily large perturbations to $\mathbf{X}$, we must allow for all possible values of $\tilde{m}$. If the function

$$F(r, \tilde{m}) = 1 + r^2 C''(\tilde{m}) \tag{5.1.16}$$

can be bounded from above and below in the rectangle $(1 \leq r \leq \eta, 1 \leq \tilde{m} \leq \mu)$ by positive constants λ_+ and λ_-, then $\mathcal{A}$ can be bounded from above and below by norms satisfying

$$||\Delta\mathbf{x}||_\lambda^2 = \iint \tfrac{1}{2}\left[(\Delta u)^2 + (\Delta w)^2 + \lambda(\Delta v)^2\right] r\,\mathrm{d}r\,\mathrm{d}z, \tag{5.1.17}$$

where $\lambda = \lambda_-$ or $\lambda = \lambda_+$. Note that while it is permissible for m, and hence possible for $\tilde{m}$, to be outside of the interval $[1, \mu]$ (and we must bound $F(r, \tilde{m})$ for all possible values of $\tilde{m}$), the function $C(\cdot)$ can be extended beyond that interval in such a way that

$$C''(m) = C''(1) \quad \text{for} \quad m < 1, \tag{5.1.18a}$$

$$C''(m) = C''(\mu) \quad \text{for} \quad m > \mu \tag{5.1.18b}$$

(see Figure 5.4a), so we need only consider $C''(\tilde{m})$ for values of $\tilde{m}$ in the range of $M(r)$. Observe that the continuation of $C''(M)$ "blows up" at $m = B$ so a construction like

(5.1.18) is necessary. (5.1.18) may not be the best choice of extension in the sense that a different choice might lead to a stronger stability condition. The choice of how to define $C(m)$ outside the range of the equilibrium distribution of m is a freedom in the energy-Casimir procedure.

The function $C''(m)$ is obviously finite in the interval $[1, \mu]$ (and for all m because of (5.1.18)) and can only be negative if $A < 0$ ($\mu < 1$) or $B < 0$ ($\mu > \eta^2$). In either case, the minimum value of $F(r, \tilde{m})$ in the intervals $r \in [1, \eta]$ and $\tilde{m} \in [1, \mu]$ occurs for $r = \eta$ and $\tilde{m} = 1$. This represents a perturbation that transposes fluid from near the inner cylinder to near the outer cylinder, conserving its angular momentum. We have in that case

$$\min_{AB<0}[F(r, \tilde{m})] = F(\eta, 1) = 1 + \eta^2 \frac{B}{A} = (\eta^2 + 1 - \mu) \left(\frac{\eta^2 - 1}{\mu - 1} \right), \tag{5.1.19}$$

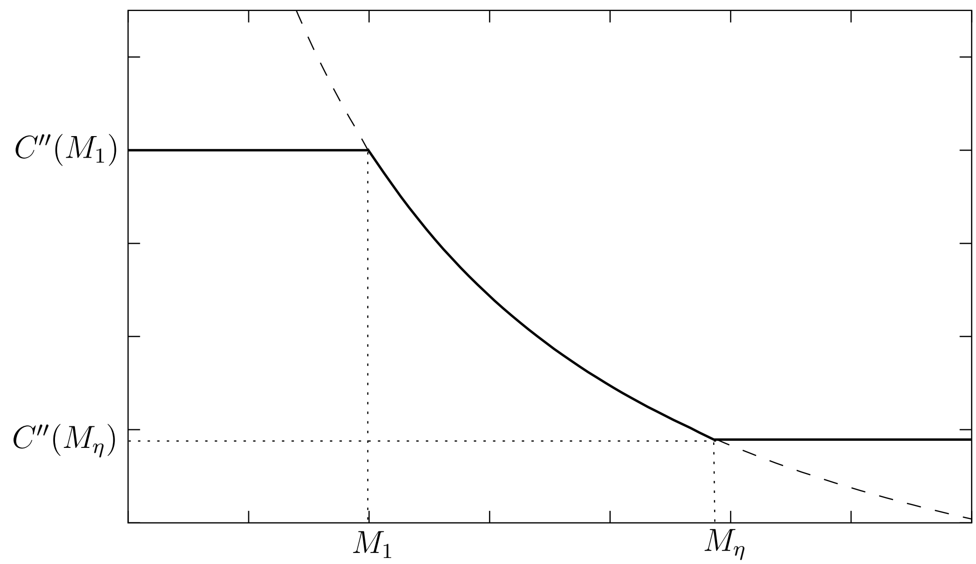
which is always negative in the case of $A < 0$ (this is the small amplitude condition) and in the case of $B < 0$, it is negative if $\mu > \eta^2 + 1$. We can conclude that the equilibrium profile $M(r)$ is nonlinearly stable if $1 < \mu < \eta^2 + 1$, since

$$||\Delta\mathbf{x}(t)||^2_{\lambda_-} \leq \mathcal{A}(t) = \mathcal{A}(0) \leq ||\Delta\mathbf{x}(0)||^2_{\lambda_+} < \frac{\lambda_+}{\lambda_-} ||\Delta\mathbf{x}(0)||^2_{\lambda_-} \tag{5.1.20}$$

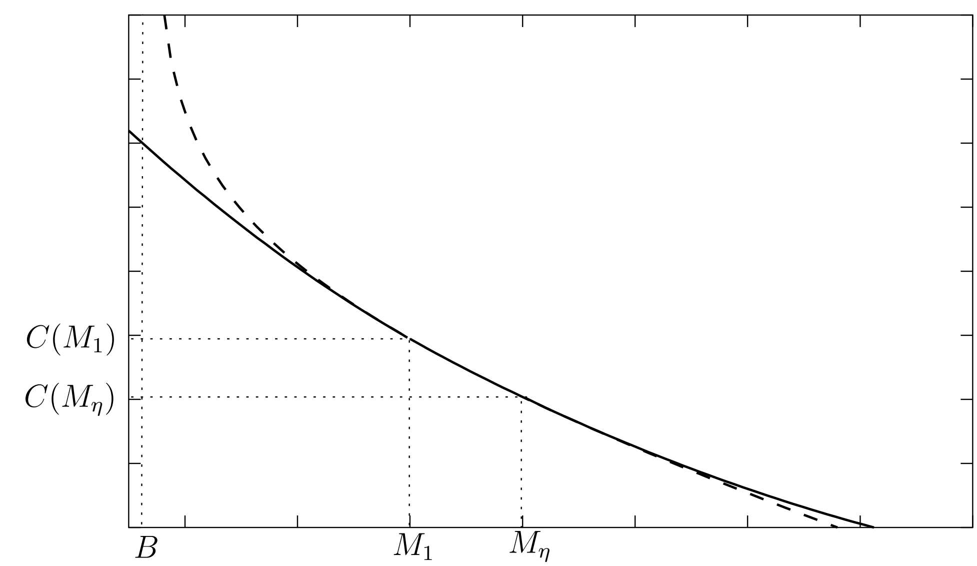
where $\lambda_- = \inf[F(r, \tilde{m})]$ and $\lambda_+ = \sup[F(r, \tilde{m})]$. (5.1.7) is then satisfied with $\delta = \sqrt{\lambda_-/\lambda_+}\varepsilon$.

Note that "small amplitude perturbation" and $||\Delta\mathbf{x}||_\lambda \ll 1$ are not the same. The former means that $\Delta\mathbf{x}(r, z, t)$ and its derivatives are small at all values of (r, z, t) while the latter means that the integral in (5.1.17) is small. The linear stability result (5.1.15) implies stability for suitably small amplitude perturbations, but it does not imply that there is a ball of radius δ about $\mathbf{X}$ in phase space (excluding $\mathbf{X}$ itself) inside which $\mathcal{A}$ is positive, and hence inside which the evolution of the system is contained if $||\mathbf{x}(t = 0)||$ is sufficiently small. Consider for example

$$m(r) = M(r) - G\sqrt{r} \exp\left[-\frac{(\eta - r)^2}{4\varepsilon_1} \right], \tag{5.1.21}$$



(a)



(b)

Figure 5.4: (a) $C''(m)$ and (b) $C(m)$ extended outside the range of $M(r)$. The function $C(m)$ is defined to satisfy (5.1.10) for the basic state with $\eta = 1.2$ and $\mu = 1.2$. The dashed curves in both plots are the extensions of the functional forms of $C''(M)$ and $C(M)$ over the range of $M(r)$ extended outside of the range of $M(r)$. Both dashed curves blow up at $m = B$.

where G and ε_1 are constants, and with $u = w = 0$. For fixed G, ε_1 can be made as small as we like, and for $\varepsilon_1 \ll \eta - 1$,

$$||\Delta \mathbf{x}||_\lambda^2 \approx \lambda H G^2 \sqrt{\frac{\pi \varepsilon_1}{8}}. \tag{5.1.22}$$

Therefore, any ball in the space of perturbations that includes the origin also includes perturbations of the form (5.1.21). If $\mu > \eta^2 + 1$, and if G is sufficiently large and ε_1 sufficiently small, then $\mathcal{A}(\mathbf{x}; \mathbf{X}) < 0$.

Hence, we cannot conclude nonlinear stability by this argument if the ratio of cylinder speeds is too high. In the early experiments of Couette (1890) and those of Schultz-Grunow (1959), instability was observed at high values of μ, but it was attributed to asymmetry or imperfections in the construction of the cylinders (see Koschmieder, 1993). Joseph notes the instability at high μ, claiming that in the case of $|\Omega_2|$ being sufficiently large, "the Rayleigh mechanism does not operate strongly, and more complicated time-dependent and subcritical motions are observed." (Joseph, 1976, p. 141)

5.2 Saturation bounds on Taylor-Couette instability

Observe that (5.1.20) is a stronger statement than (5.1.7), which simply requires that solutions are bounded near $\mathbf{X}$ provided their initial conditions are close enough to $\mathbf{X}$. But (5.1.20) implies that *every* solution is bounded relative to $\mathbf{X}$, no matter how far it is from $\mathbf{X}$ at $t = 0$.

In particular, the evolution of an initial condition in the neighbourhood of a linearly *unstable* basic state $\mathbf{X}_U$ is constrained by the initial value of $||\mathbf{X}_U - \mathbf{X}||$. This suggests an important application of the energy-Casimir method, namely the calculation of rigorous saturation bounds on an instability (Shepherd, 1988). By choosing an optimum nonlinearly stable basic state $\mathbf{X}$, an upper bound on, for example, the contribution to the kinetic energy by the radial and vertical components of the velocity that can be released into the flow during the evolution of the system from near an unstable equilibrium may be computed.

Consider an unstable equilibrium $\mathbf{X}_U = (0, V_U, 0)$, and a solution $\mathbf{x}(t)$ which begins near $\mathbf{X}_U$. The kinetic energy in the vertical-radial plane $\mathcal{K}_\perp(\mathbf{x})$ is always less than the pseudoenergy of $\mathbf{x}$ relative to *any* nonlinearly stable reference state $\mathbf{X} = (0, V, 0)^T$ (this is true for any $\mathbf{x}$):

$$\mathcal{K}_\perp(\mathbf{x}(t)) \equiv \iint \tfrac{1}{2}\left[(u(t))^2 + (w(t))^2\right] r\,\mathrm{d}r\,\mathrm{d}z \leq \mathcal{A}(\mathbf{x}(t); \mathbf{X}), \tag{5.2.1}$$

but conservation of $\mathcal{A}$ allows us to bound $\mathcal{K}(\mathbf{x})$ in terms of $\mathbf{x}(t{=}0) \approx \mathbf{X}_U$ according to

$$\mathcal{K}_\perp(\mathbf{x}(t)) \leq \mathcal{A}(\mathbf{X}_U; \mathbf{X}) = \iint \left[\tfrac{1}{2}(V_U - V)^2 + V(V_U - V) + C(M_U) - C(M)\right] r\,\mathrm{d}r\,\mathrm{d}z. \tag{5.2.2}$$

In this case, we can calculate $\mathcal{A}(\mathbf{X}_U; \mathbf{X})$ exactly. If we require that $C(m)$ and $C'(m)$ be continuous everywhere, then (5.1.10) and (5.1.18) imply (up to an additive constant)

$$C(m) = \begin{cases} -AB\ln|M_1 - B| - A(M_1 - B) \\ \qquad - \dfrac{AM_1}{M_1 - B}(m - M_1) + \dfrac{AB}{2(M_1 - B)^2}(m - M_1)^2, & m \leq M_1 \\[2ex] -AB\ln|m - B| - A(m - B), & M_1 \leq m \leq M_\eta \\[2ex] -AB\ln|M_\eta - B| - A(M_\eta - B) \\ \qquad - \dfrac{AM_1}{M_\eta - B}(m - M_\eta) + \dfrac{AB}{2(M_\eta - B)^2}(m - M_\eta)^2, & m \geq M_\eta \end{cases}, \tag{5.2.3}$$

where $M_1 = M(1)$ and $M_\eta = M(\eta)$ (see Figure 5.4*b*).

Consider an unstable equilibrium with angular momentum

$$M_U = \left(\frac{\mu_U - 1}{\eta^2 - 1}\right) r^2 + \left(\frac{\eta^2 - \mu_U}{\eta^2 - 1}\right) \equiv A_U r^2 + B_U, \tag{5.2.4}$$

corresponding to the fluid at the inner cylinder having angular momentum $M_U = 1$ and the fluid at the outer cylinder having $M_U = \mu_U < 1$. We look for the stable equilibrium with angular momentum of the form

$$M = M_1\left[\left(\frac{\mu - 1}{\eta^2 - 1}\right) r^2 + \left(\frac{\eta^2 - \mu}{\eta^2 - 1}\right)\right] \equiv Ar^2 + B, \tag{5.2.5}$$

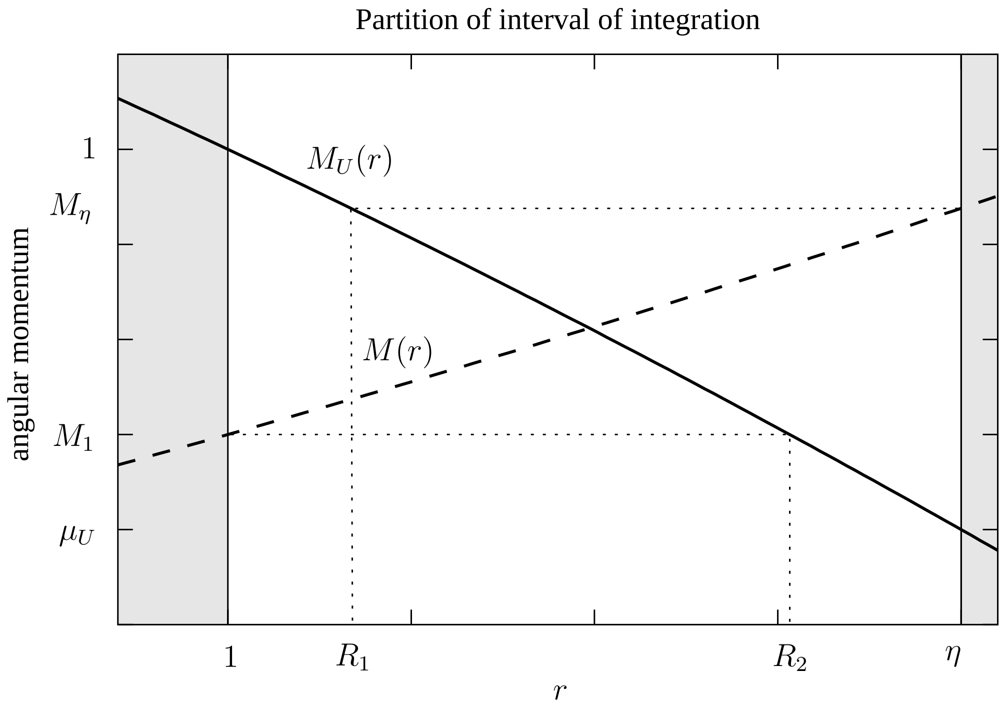


Figure 5.5: Partition of interval of integration for a typical stable reference state.

corresponding to a state in which the innermost fluid has $M = M_1$ and the outermost $M = M_\eta = \mu M_1$ (where $1 < \mu < \eta^2 + 1$), which minimizes $\mathcal{A}$.

Note that for this calculation, we have implicitly scaled m based on the unstable equilibrium $\mathbf{X}_U$ because it is fixed and we can compare $\mathcal{A}(\mathbf{X}_U; \mathbf{X})$ for stable equilibria $\mathbf{X}$ with different values of $M(1)$. Also note that $\mathbf{X}$ and $\mathbf{X}_U$ do not satisfy the same boundary conditions. This is not a problem because we did not assume that the perturbation satisfied any particular boundary conditions. Physically, we might think of the perturbation $\mathbf{X} \to \mathbf{X}_U$ as a quasistatic (which is to say that the fluid is always in a state of equilibrium) change in the rotation rates of the cylinders.

In evaluating $\mathcal{A}$, the integral over r is computed in three segments because of the partition of $C(m)$. Let R_1 be the point at which $M_U = M_1\mu$, if it exists and is in the interval $1 \le r \le \eta$, and let R_2 be the point at which $M_U = M_1$, if it exists and is in

$1 \le r \le \eta$ (see Figure 5.5). Then,

$$
\begin{aligned}
\frac{1}{H}\mathcal{A}(\mathbf{X}_U;\mathbf{X}) &= \int_1^\eta \left[\frac{1}{2}(V_U^2 - V^2) + C(M_U) - C(M)\right] r\,\mathrm{d}r && \text{(5.2.6a)}\\
&= \Bigg[\frac{A_U^2 + A^2}{8} r^4 + \frac{(A_U B_U - AB)}{2} r^2 \\
&\qquad + \frac{(B_U^2 - B^2)}{2}\ln|r| + \frac{AB}{2} r^2 (\ln|Ar^2| - 1)\Bigg]_1^\eta && \text{(5.2.6b)}\\
&\quad + \Bigg\{\frac{1}{2}\Big[-AB\ln|M_\eta - B| - A(M_\eta - B) \\
&\qquad + \frac{AM_\eta}{M_\eta - B}(M_\eta - B_U) + \frac{AB}{2(M_\eta - B)^2}(M_\eta - B_U)^2\Big] r^2 \\
&\qquad + \frac{1}{4}\left[-\frac{AM_\eta}{M_\eta - B}A_U - \frac{AB}{(M_\eta - B)^2}A_U(M_\eta - B_U)\right] r^4 \\
&\qquad + \frac{AB}{12(M_\eta - B)^2}A_U^2 r^6\Bigg\}_1^{R_1} && \text{(5.2.6c)}\\
&\quad + \Bigg[-\frac{A(B_U - B)}{2} r^2 - \frac{AA_U}{4} r^4 \\
&\qquad - \frac{AB}{2A_U}(A_U r^2 + B_U - B)\left(\ln|A_U r^2 + B_U - B|\right)\Bigg]_{R_1}^{R_2} && \text{(5.2.6d)}\\
&\quad + \Bigg\{\frac{1}{2}\Big[-AB\ln|M_1 - B| - A(M_1 - B) \\
&\qquad + \frac{AM_1}{M_1 - B}(M_1 - B_U) + \frac{AB}{2(M_1 - B)^2}(M_1 - B_U)^2\Big] r^2 \\
&\qquad + \frac{1}{4}\left[-\frac{AM_1}{M_1 - B}A_U - \frac{AB}{(M_1 - B)^2}A_U(M_1 - B_U)\right] r^4 \\
&\qquad + \frac{AB}{12(M_\eta - B)^2}A_U^2 r^6\Bigg\}_{R_2}^{\eta}. && \text{(5.2.6e)}
\end{aligned}
$$

If either of R_1 and R_2 is not in the interval $1 < r < \eta$, then the obvious adjustment to the integral limits is made (i.e. depending on how the range of M_U overlaps the interval $[M_1, M_\eta]$, one or two of the brackets (5.2.6c)-(5.2.6e) is omitted).

5.3 Numerical results

For a given unstable equilibrium $M_U(r)$, $\mathcal{A}/H$ was calculated relative to stable equilibria $M(r)$ with μ between 1 and $1+\eta^2$ and M_1 between 0 and 1.5. Some examples are plotted in Figure 5.6.

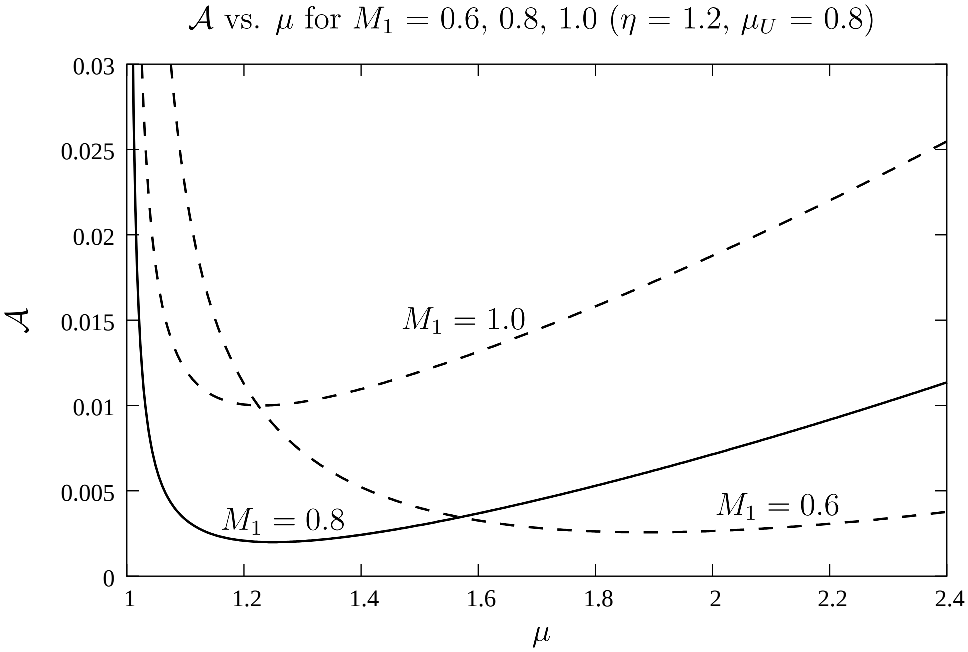
The minimum of $\mathcal{A}$ occurs when $M_1 = \mu_U$ and $\mu = \mu_U^{-1}$. Such a pair of $M_U(r)$ and $M(r)$ are shown in Figure 5.7. Such a basic state achieves an optimum balance between the azimuthal kinetic energy and Casimir contributions to $\mathcal{A}$. $V_U^2 - V^2$ is negative when $M > M_U$ (assuming both are positive) and, since $C(m)$ is a decreasing function in the range of $M(r)$, $C(M_U) - C(M)$ is positive when $M > M_U$. The reverse is true when $M < M_U$. We have shown that $\mathcal{A}$ is positive for every stable $M(r)$, so as the two profiles diverge, the contribution from the positive term must eventually dominate.

In Figure 5.8a, the minimum $\mathcal{A}$ is shown as a function of the angular momentum ratio of the unstable equilibrium μ_U. As the linear stability threshold $\mu_U = 1$ is approached from below, the maximum disturbance amplitude approaches 0. This is characteristic of supercritical bifurcations. Assuming the existence of a non-laminar equilibrium solution for $\mu_U < 1$ (which has been experimentally shown to be Taylor vortex flow near $\mu_U = 1$), its amplitude must be small for small $1 - \mu_U$.

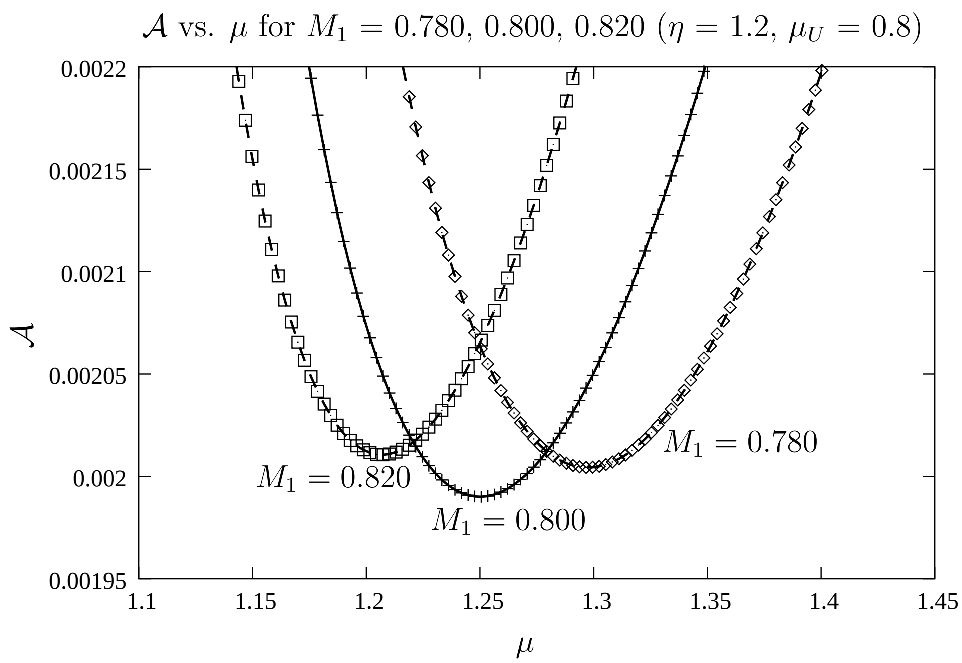
On the other hand, the threshold[1] $\mu_U = 1 + \eta^2$, which does not represent a critical point for linear stability, may be an indication of a subcritical bifurcation at some $\mu_U > 1 + \eta^2$. Joseph (1976) mentions an instability at high rotation rates of the outer cylinder, saying that in that regime, "the Rayleigh mechanism does not operate strongly, and more complicated time-dependent and subcritical motions are observed." In the case of a subcritical bifurcation, multiple stable equilibria exist for values of the stability parameter (μ_U in this case) on the stable side of the bifurcation point (see Figure 5.9 and, for example, Iooss and Joseph 1990).

Finally, in Figure 5.10, we show an upper bound on the *fraction* of the kinetic energy $\mathcal{K}(V_U)$ in the unstable equilibrium that can be converted to kinetic energy in the overturning components of the flow ($\mathcal{K}_\perp$) as a function of radius ratio. Evidently, a greater

[1] Actually, we do not claim that $\mu = 1 + \eta^2$ is necessarily a threshold of stability-instability, only that we cannot show stability for $\mu > 1 + \eta^2$.



(a)



(b)

Figure 5.6: Pseudoenergy of unstable equilibrium with angular momentum ratio $\mu_U = 0.8$ as a function of μ of stable state for multiple values of M_1. Radius ratio is $\eta = 1.2$. (a) values of M_1 widely spaced, (b) values of M_1 near the minimizing case.

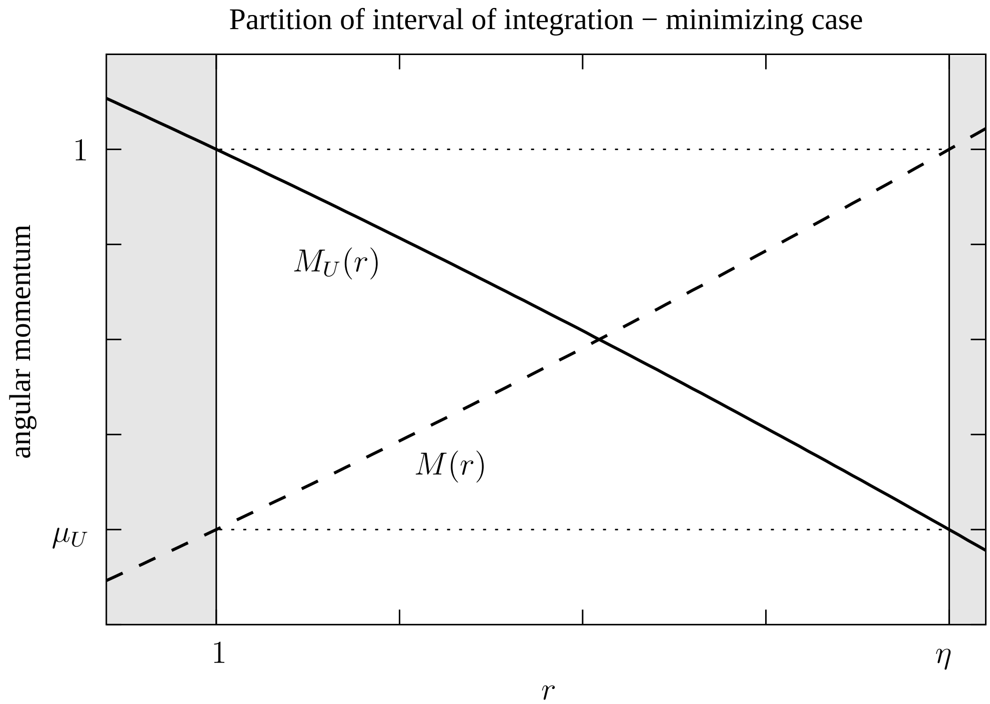


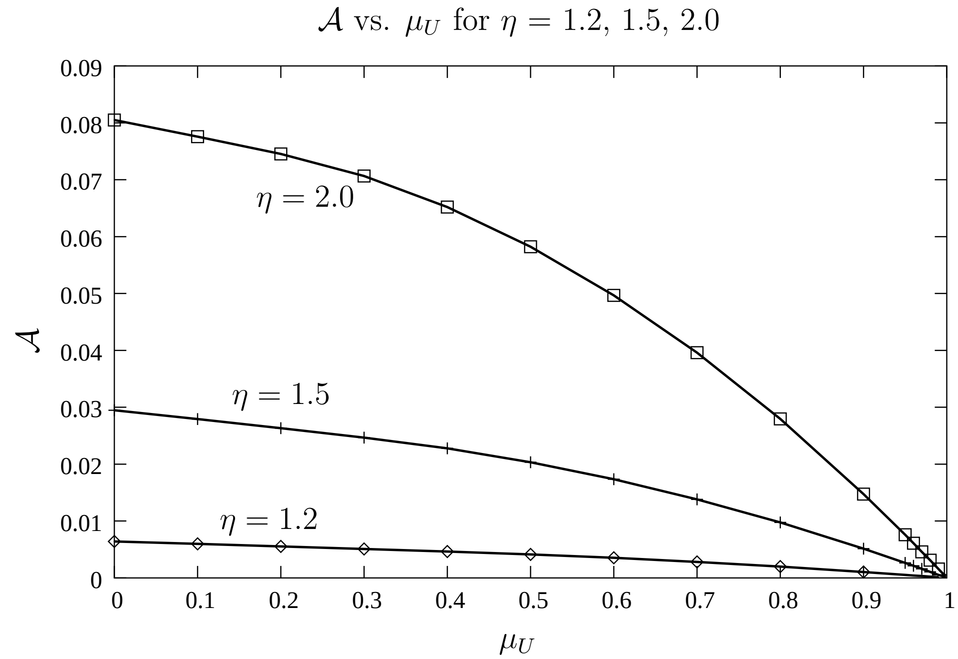
Figure 5.7: Partition of interval of integration for the stable reference state $M(r)$ which minimizes $\mathcal{A}(M_U(r); M(r))$ over the set of all stable Couette profiles.

fraction of the energy can be released for larger cylinder gaps. Eventually (not shown), the curve asymptotes to 1 as η gets very large. This does not imply that all of the energy can be released into the overturning flow, but rather that we cannot bound the energy released using this method. In any case, for extremely wide cylinder gaps, the notion of an inviscid fluid having the Couette velocity profile is unphysical.

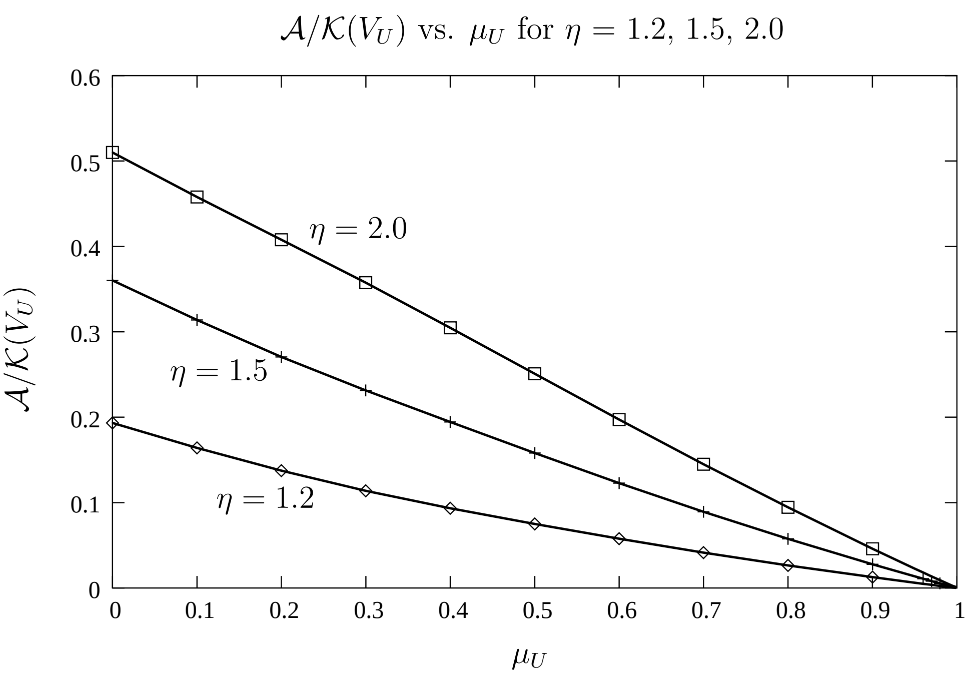
5.4 Summary

Using an energy-Casimir approach, we have shown that steady inviscid flow in the Taylor-Couette problem having the Couette velocity profile is nonlinearly stable if the Rayleigh criterion $\mu \equiv \Omega_2 r_2^2/\Omega_1 r_1^2 > 1$ is satisfied together with the further requirement that $\mu < \eta^2 + 1 \equiv (r_2/r_1)^2 + 1$.

We then applied the method of Shepherd (1988) to bound the amplitude at which



(a)



(b)

Figure 5.8: Pseudoenergy minimized over all stable Couette profiles as a function of angular momentum ratio μ_U of unstable equilibrium. Results for three values of radius ratio η are plotted. (a) Pseudoenergy. (b) Pseudoenergy divided by kinetic energy in unstable equilibrium, $\mathcal{K}(V_U)$.

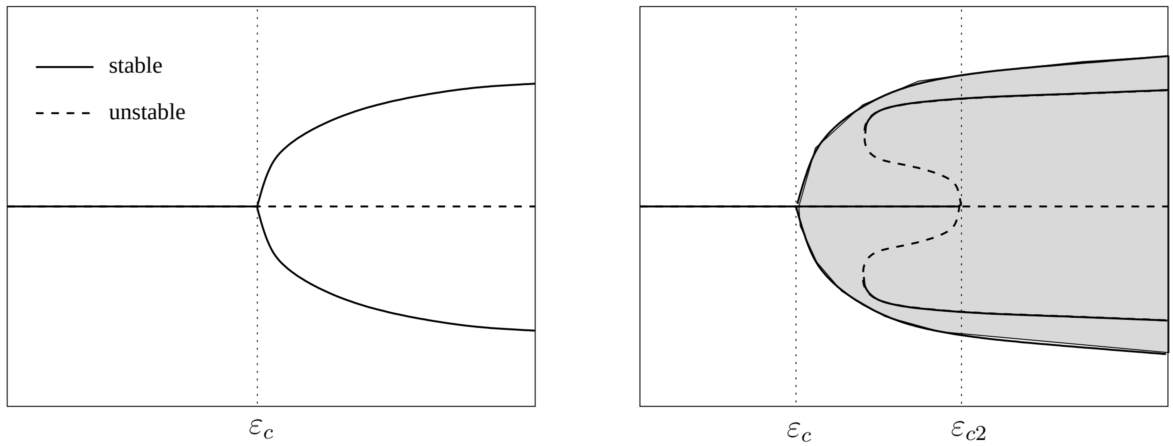


Figure 5.9: Bifurcation diagrams for supercritical (left) and subcritical (right) bifurcations. Plotted are amplitudes of equilibrium solutions as functions of some critical parameter ε. The dashed curves represent unstable equilibria and the solid curves represent stable equilibria. The pseudoenergy saturation bound curve in Figure 5.8 suggests that there is a supercritical bifurcation at the linear stability threshhold $\mu = 1$. The upper limit on nonlinear stability $\mu = \eta^2 + 1$ may indicate a subcritical bifurcation at some $\mu_2 > \eta^2 + 1$. The shaded region in the subcritical diagram indicates a nonlinear bound on the subcritical instability in terms of the "threshold" $\mu_c = \eta^2 + 1$.

disturbances to unstable equilibria must saturate. Bounds on saturation amplitude are obtained by calculating the pseudoenergy in the unstable equilibrium relative to any stable equilibrium. The stable equilibrium, considering only Couette velocity profiles for both the stable and unstable equilibria, that gives the least upper bound is that with angular momentum $M(r)$ satisfying $M(r_1) = M_U(r_2)$ and $M(r_2) = M_U(r_1)$, where $M_U(r)$ is the angular momentum profile of the unstable equilibrium. We have shown that the saturation amplitude in the neighbourhood of the linear stability threshold approaches zero, consistent with supercritical bifurcations.

A limitation on our saturation bound result is that we have only considered axisymmetric perturbations. While experiments on corotating cylinders (e.g. Andereck et al., 1986) suggest that the first bifurcation is into (axisymmetric) Taylor vortices, non-axisymmetric states emerge as the supercriticality is increased. The maximum su-

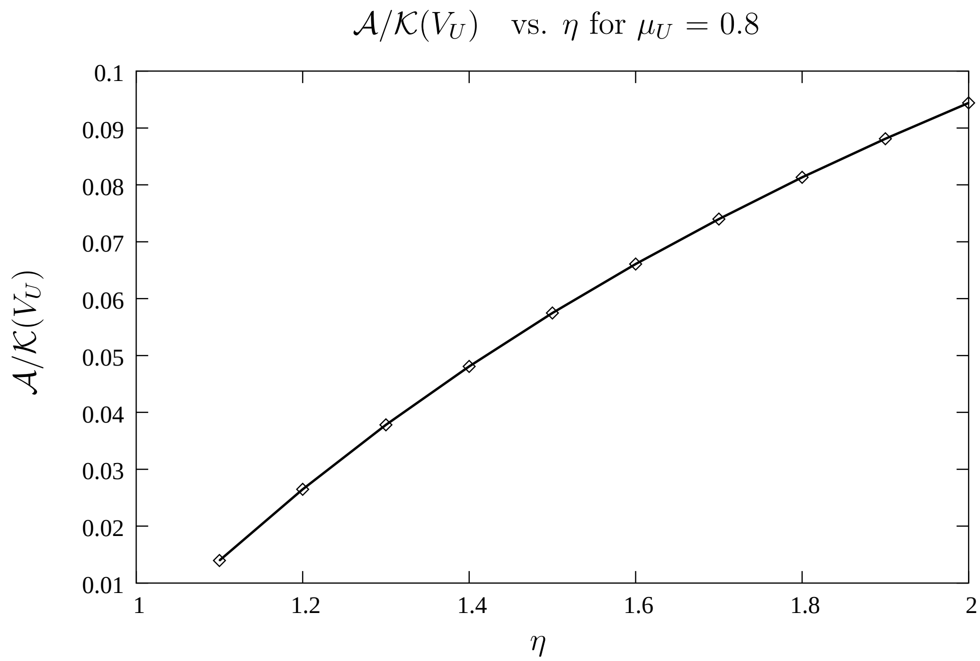


Figure 5.10: Pseudoenergy minimized over all stable Couette profiles for unstable angular momentum ratio $\mu_U = 0.8$ as a function of radius ratio η. Pseudoenergy divided by kinetic energy in unstable equilibrium, $\mathcal{K}(V_U)$ is plotted. For very large η, the curve approaches unity.

percriticality for which Taylor vortices prevail becomes ever smaller as the Reynolds number is increased. In that sense, the two assumptions of axisymmetry and inviscid flow are not particularly compatible.

We conjecture that the upper limit on μ for nonlinear stability is indirect evidence of a subcritical bifurcation at some higher value of μ.

5.A Uniqueness of Couette solution to steady viscous equations

In vector form, the Navier-Stokes equation is

$$\frac{\partial \boldsymbol{v}}{\partial t} + (\boldsymbol{v}\cdot\nabla)\boldsymbol{v} = -\nabla p + \nu\nabla^2\boldsymbol{v}, \tag{5.A.1}$$

where $\boldsymbol{v}$ is the fluid velocity, p is pressure, and ν the kinematic viscosity. The geometry of the Taylor-Couette problem suggests cylindrical coordinates (r,ϕ,z), with associated unit vectors $\hat{\mathbf{e}}_r$, $\hat{\mathbf{e}}_\phi$ and $\hat{\mathbf{e}}_z$. Let the velocity vector have components (u,v,w) so that

$$\boldsymbol{v} = u\hat{\mathbf{e}}_r + v\hat{\mathbf{e}}_\phi + w\hat{\mathbf{e}}_z. \tag{5.A.2}$$

The vectors $\hat{\mathbf{e}}_r$ and $\hat{\mathbf{e}}_\phi$ are themselves functions of ϕ:

$$\frac{\mathrm{d}}{\mathrm{d}\phi}\hat{\mathbf{e}}_r = \hat{\mathbf{e}}_\phi, \quad \frac{\mathrm{d}}{\mathrm{d}\phi}\hat{\mathbf{e}}_\phi = -\hat{\mathbf{e}}_r. \tag{5.A.3}$$

Therefore, the components of the advection term are

$$\begin{aligned}(\boldsymbol{v}\cdot\nabla)\boldsymbol{v} &\equiv u\frac{\partial \boldsymbol{v}}{\partial r} + \frac{v}{r}\frac{\partial \boldsymbol{u}}{\partial \phi} + w\frac{\partial \boldsymbol{v}}{\partial z} \\ &= \left(\boldsymbol{v}\cdot\nabla u - \frac{v^2}{r}\right)\hat{\mathbf{e}}_r + \left(\boldsymbol{v}\cdot\nabla v + \frac{uv}{r}\right)\hat{\mathbf{e}}_\phi + (\boldsymbol{v}\cdot\nabla w)\,\hat{\mathbf{e}}_z,\end{aligned} \tag{5.A.4}$$

and of the viscosity term are

$$\begin{aligned}\nabla^2\boldsymbol{v} &\equiv \frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial \boldsymbol{v}}{\partial r}\right) + \frac{1}{r^2}\frac{\partial^2 \boldsymbol{v}}{\partial \phi^2} + \frac{\partial^2 \boldsymbol{v}}{\partial z^2} \\ &= \left(\nabla^2 u - \frac{2}{r^2}\frac{\partial v}{\partial \phi} - \frac{u}{r^2}\right)\hat{\mathbf{e}}_r \\ &\quad + \left(\nabla^2 v + \frac{2}{r^2}\frac{\partial u}{\partial \phi} - \frac{v}{r^2}\right)\hat{\mathbf{e}}_\phi \\ &\quad + \nabla^2 w\hat{\mathbf{e}}_z.\end{aligned} \tag{5.A.5}$$

We seek steady solutions to (5.A.1) which are independent of ϕ and have $u = w = 0$. The components of the resulting equilibrium equation are

$$
\begin{aligned}
\frac{V^2}{r} - \frac{\partial P}{\partial r} &= 0, & \text{(5.A.6a)}\\
\nu\left[\frac{1}{r}\frac{\partial}{\partial r}\left(r\frac{\partial V}{\partial r}\right) - \frac{V}{r^2}\right] &= 0, & \text{(5.A.6b)}\\
\frac{\partial P}{\partial z} &= 0, & \text{(5.A.6c)}
\end{aligned}
$$

where $P(r,z)$ and $V(r,z)$ are the pressure and azimuthal velocities of the equilibrium solution. From (5.A.6c) and (5.A.6a) respectively, we have that $P(r,z)$ and $V(r,z)$ are independent of z. The general solution to (5.A.6b) is the Couette profile

$$V(r) = Ar + \frac{B}{r}, \tag{5.A.7}$$

where A and B are constants determined from boundary conditions, and from (5.A.6a), the pressure is (up to an arbitrary additive constant)

$$P(r) = \frac{A^2r^2}{2} + 2AB\ln r - \frac{B^2}{2r^2}. \tag{5.A.8}$$

Chapter 6

Conclusion

Symmetric instability is a fundamental phenomenon that can be observed in, for example, the Taylor-Couette experiment on flow of a liquid confined between coaxial rotating cylinders. It is also a significant part of the dynamics in the Earth's equatorial middle atmosphere during solstice seasons, when the heating by the sun is stronger in the summer hemisphere than in the winter. It contributes to the overturning circulation in the tropical middle atmosphere and the resulting smoothing of angular momentum and temperature gradients over the equator (Semeniuk and Shepherd, 2001), to the semiannual oscillation in the wind direction at the equatorial stratopause (Shepherd, 2000), and is one mechanism for exciting the two-day wave in the summer subtropical stratosphere (Limpasuvan et al., 2000).

The physics underlying the instability is simple. If a fluid is stably stratified parallel to an axis of symmetry, or homogeneous as is typical in the Taylor-Couette experiment, then instability sets in when the angular momentum somewhere decreases in the direction normal to the axis of symmetry. This is the Rayleigh (1916) inertial instability criterion. If the stratification is more complicated, then the inertial stability condition is (in a sense) mixed with the condition for convective stability.

In the Earth's atmosphere, the density stratification is approximately aligned with gravity (surfaces of constant density are approximately spherical), and the rest state with respect to the rotating Earth has cylindrical surfaces of constant ("planetary") angular momentum. Symmetric stability depends on the relative orientations of the planetary rotation vector $\mathbf{\Omega}$ and the gradients of entropy, angular momentum and pressure.

Previous studies (Dunkerton, 1981; Stevens, 1983; Bowman and Shepherd, 1995) have neglected the component of $\mathbf{\Omega}$ tangent to the surface, in which case the stratification is very nearly aligned with the rotation vector. The stability conditions that we have calculated do not differ significantly from those of the earlier studies for typical Earth velocity, length, time and temperature scales except in the immediate neighbourhood of the equator. This is due to the fact that the Coriolis force terms in the dynamical equations associated with the traditionally neglected component of $\mathbf{\Omega}$ are relatively small, although not negligibly so near the equator (White and Bromley, 1995). It is worth noting, however, that including the neglected Coriolis terms restricts the choices of velocity, density, and temperature (or entropy) fields that satisfy the conditions for equilibrium, which must obviously be satisfied before one can begin to discuss stability (see Section 3.1.6).

6.1 Summary of results

We have considered the problem of equatorial symmetric stability in the Earth's middle atmosphere, taking into account the effects of the often neglected Coriolis force terms associated with the component of the planetary rotation vector tangent to the surface. Using an energy-Casimir method based on the Hamiltonian structure of the governing equations, we derived conditions for the linear stability of a steady zonal solution to the adiabatic, compressible Euler equations on an equatorial β-plane. We showed that for stability, it is sufficient that the potential vorticity have the sign of latitude and that the entropy increase (decrease) in the direction of the local planetary rotation vector in the northern (southern) hemisphere. By explicitly solving the steady state equations, we showed that there are solutions that are stable under the "traditional" equations (without the neglected Coriolis terms) but unstable in the more general system.

We also looked at symmetric stability in the anelastic equations system. We defined an exact invariant functional $\mathcal{A}$, the pseudoenergy, that vanishes at the equilibrium state and showed that conditions for the positive definiteness of $\mathcal{A}$ are conditions for the nonlinear stability of the equilibrium. We were able to show that steady states that are *even* functions of latitude and that satisfy the conditions for linear stability are stable with respect to finite amplitude perturbations. We then applied the saturation bound method of Shepherd (1988) to calculate an upper bound on the kinetic energy that can develop in adjustment from an unstable steady state in terms of the pseudoenergy relative to a stable steady state. Surprisingly, it appears that for states that are asymmetric functions of latitude, perturbations for which $\mathcal{A}$ is negative always exist, although that in itself does not imply that the equilibrium is necessarily unstable.

We solved the anelastic equations linearized about a state of linear meridional shear in the zonal velocity, the anelastic counterpart to the problem of Dunkerton (1981). The solution is unstable in an interval of latitude between the equator and a latitude proportional to the strength of the shear. The normal mode solution exhibits Taylor vortices centred over the unstable region, stacked alternately signed anomalies in the zonal velocity on the equatorward side of the unstable region, and out of phase columns of temperature anomalies of alternating sign on either side of the unstable region. These are the features that, when identified in observational data or numerical simulations, are considered evidence of symmetric instability.

Lastly, we applied the energy-Casimir method to steady flows in the Taylor-Couette experiment. We showed that steady solutions having the Couette velocity profile are nonlinearly stable if they satisfy the Rayleigh criterion, that the magnitude of angular momentum increase with distance from the axis of rotation, and an upper bound on the gradient of angular momentum that depends on the width of the gap between the cylinders.

6.2 Future work

The linear stability criteria in Chapter 3 and Chapter 4 are *sufficient* for stability. To demonstrate that they are *necessary*, an approach such as that of Ooyama (1966) may be used to show that a positive definite measure of disturbance is always increasing with time if the stability conditions are not met.

Deriving conditions for stability in a spherical shell, as opposed to on the β-plane (Bowman and Shepherd, 1995, considered this case for the hydorstatic system), should also be possible. Any nonlinear calculation will have to take into account the position dependence of the volume element in the disturbance norm, as was necessary for the cylindrical geometry of the Taylor-Couette problem in Chapter 5.

If it can be shown to be comparable with "realistic" values (derived from, perhaps, high resolution, small domain simulations), the saturation bound on the kinetic energy in the transverse plane derived in Chapter 4 may be incorporated into a scheme for parameterizing the effects of unresolved inertial instability. Linear theory, such as was used in deriving the normal mode solution in Section 4.3, predicts that the most unstable modes have the smallest vertical scales, while observations and simulations show instability cells at slightly above or at the smallest vertical scale that can be resolved. Undoubtedly, there is energy in the unresolved scales. Furthermore, general circulation models that resolve equatorial inertial instability tend to exhibit overly strong vertical mixing, probably because all of the available potential energy is going into relatively large instability cells. An inertial adjustment scheme instantaneously redistributes angular momentum in an unstable state such that marginal stability is restored and total angular momentum is conserved. This is analogous to convective adjustment parameterization schemes that redistribute potential temperature to produce a marginally convectively stable state in such a way that total potential energy is conserved (see, e.g., Emanuel, 1994, Chapter 16). The saturation bound could be used to infer the mixing (of chemicals, for example) that would have occured during the adjustment, since the effective diffusivity is related to the kinetic energy in the overturning inertial instability cells (assuming that the mixing can be described by a diffusion model). The effective diffusivity also depends on the vertical scale of the cells, which would have to be inferred by another method.

Bibliography

Andereck, C. D., S. S. Liu, and H. L. Swinney (1986). Flow regimes in a cicular Couette system with independently rotating cylinders. *J. Fluid Mech. 164*, 155–183.

Andrews, D. G., J. R. Holton, and C. B. Leovy (1987). *Middle Atmosphere Dynamics.* Orlando: Academic Press.

Arfken, G. (1985). *Mathematical Methods for Physicists.* San Diego: Academic Press.

Arnold, V. I. (1989). *Mathematical Methods of Classical Mechanics.* New York: Springer.

Batchelor, G. K. (1967). *An Introduction to Fluid Dynamics.* Cambridge, U.K.: Cambridge University Press.

Bowman, J. C. and T. G. Shepherd (1995). Nonlinear symmetric stability of planetary atmospheres. *J. Fluid Mech. 296*, 391–407.

Chandrasekhar, S. (1961). *Hydrodynamic and Hydromagnetic Instability.* Oxford, U.K.: Clarendon Press.

Cho, H.-R., T. G. Shepherd, and V. A. Vladimirov (1993). Application of the direct Liapunov method to the problem of symmetric stability in the atmosphere. *J. Atmos. Sci. 50*, 822–836.

Chossat, P. and G. Iooss (1994). *The Couette-Taylor Problem.* New York: Springer-Verlag.

Clark, P. D. and P. H. Haynes (1996). Inertial instability on an asymmetric low-latitude flow. *Q. J. R. Meteorol. Soc. 122*, 151–182.

Coles, D. (1965). Transition in circular Couette flow. *J. Fluid Mech. 21*, 385–425.

Colin de Verdière, A. and R. Schopp (1994). Flows in a rotating spherical shell: the equatorial case. *J. Fluid Mech. 276*, 233–260.

Dunkerton, T. J. (1981). On the inertial stability of the equatorial middle atmosphere. *J. Atmos. Sci. 38*, 2354–2364.

Eckart, C. (1960). *Hydrodynamics of Oceans and Atmospheres.* New York: Pergamon Press.

Emanuel, K. A. (1994). *Atmospheric Convection.* Oxford: Oxford University Press.

Firing, E. (1987). Deep zonal currents in the central equatorial Pacific. *J. Marine Research 45*, 791–812.

Gill, A. (1982). *Atmosphere-Ocean Dynamics.* San Diego: Academic Press.

Griffiths, S. D. (2003). Nonlinear vertical scale selection in equatorial inertial instability. *J. Atmos. Sci. 60*, 977–990.

Grimshaw, R. H. J. (1975). A note on the β-plane approximation. *Tellus 27*, 352–357.

Hayashi, H., M. Shiotani, and J. C. Gille (1998). Vertically stacked temperature disturbances near the equatorial stratopause as seen in cryogenic limb array etalon spectrometer data. *J. Geophys. Res. 103*, 19,469–19,483.

Hitchman, M. H., C. B. Leovy, J. C. Gille, and P. L. Bailey (1987). Quasi-stationary zonally asymmetric circulations in the equatorial lower mesosphere. *J. Atmos. Sci. 44*, 2219–2236.

Holm, D. D., J. E. Marsden, T. Ratiu, and A. Weinstein (1985). Nonlinear stability of fluid and plasma equilibria. *Phys. Rep. 123*, 1–116.

Hua, B. L., D. W. Moore, and S. Le Gentil (1997). Inertial nonlinear equilibration of equatorial flows. *J. Fluid Mech. 331*, 345–371.

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

Iooss, G. and D. D. Joseph (1990). *Elementary Stability and Bifurcation Theory.* New York: Springer-Verlag.

Joseph, D. D. (1976). *Stability of Fluid Motions*, Volume 1. New York: Springer-Verlag.

Joseph, D. D. and W. Hung (1971). Contributions to the nonlinear theory of stability of viscous flow in pipes and between rotating cylinders. *Arch. Rational Mech. Anal. 44*, 1–22.

Koschmieder, E. L. (1993). *Bénard Cells and Taylor Vortices.* Cambridge, U.K.: Cambridge University Press.

Limpasuvan, V., C. B. Leovy, and Y. J. Orsolini (2000). Observed temperature two-day wave and its relatives near the stratopause. *J. Atmos. Sci. 57*, 1689–1701.

Lipps, F. B. and R. S. Hemler (1982). A scale analysis of deep moist convection and some related numerical calculations. *J. Atmos. Sci. 39*, 2192–2210.

Lorenz, E. N. (1967). *The Nature and Theory of the General Circulation of the Atmosphere.* Geneva: World Meteorological Organization.

Morrison, P. J. (1998). Hamiltonian description of the ideal fluid. *Rev. Mod. Phys. 70*, 467–521.

Mu, M., T. G. Shepherd, and K. Swanson (1996). On nonlinear symmetric stability and the nonlinear saturation of symmetric instability. *J. Atmos. Sci. 53*, 2918–2923.

Müller, R. (1989). A note on the relation between the "traditional approximation" and the metric of the primitive equations. *Tellus 41A*, 175–178.

Ogura, Y. and N. A. Phillips (1962). Scale analysis of deep and shallow convection in the atmosphere. *J. Atmos. Sci. 19*, 173–179.

Ooyama, K. (1966). On the stability of the baroclinic circular vortex: a sufficient criterion for instability. *J. Atmos. Sci. 23*, 43–53.

O'Sullivan, D. J. and M. H. Hitchman (1992). Inertial instability and Rossby wave breaking in a numerical model. *J. Atmos. Sci. 49*, 991–1002.

Perlis, S. (1952). *Theory of Matrices.* Reading, Mass., USA: Addison-Wesley.

Phillips, N. A. (1966). The equations of motion for a shallow rotating atmosphere and the "traditional approximation". *J. Atmos. Sci. 23*, 626–628.

Phillips, N. A. (1968). Reply to Veronis, 1968. *J. Atmos. Sci. 25*, 1155–1157.

Phillips, N. A. (1973). Principles of large scale numerical weather prediction. In P. Morel (Ed.), *Dynamic Meteorology*, pp. 1–96. Dordrecht, Holland: D. Reidel.

Rayleigh, Lord (1916). On the dynamics of revolving fluids. *Proc. R. Soc. Lond. A 93*, 148–154.

Salmon, R. (1982). Hamilton's principle and Ertel's theorem. *Am. Inst. Phys. Proc. 88*, 127–135.

Salmon, R. (1998). *Lectures on Geophysical Fluid Dynamics.* Oxford: Oxford University Press.

Semeniuk, K. and T. G. Shepherd (2001). The middle-atmosphere Hadley circulation and equatorial inertial adjustment. *J. Atmos. Sci. 58*, 3077–3096.

Shepherd, T. G. (1988). Rigorous bounds on the nonlinear saturation of instabilities to parallel shear flows. *J. Fluid Mech. 196*, 291–322.

Shepherd, T. G. (1990). Symmetries, conservation laws, and Hamiltonian structure in geophysical fluid dynamics. *Advances in Geophysics 32*, 287–338.

Shepherd, T. G. (2000). The middle atmosphere. *J. Atmos. Sol-Terr. Phy. 62*, 1587–1601.

Spiegel, E. A. and G. Veronis (1960). On the Boussinesq approximation for a compressible fluid. *Astrophys. J. 131*, 442–447.

Stevens, D. (1983). On symmetric stability and instability of zonal mean flows near the equator. *J. Atmos. Sci. 40*, 882–893.

Stommel, H. M. and D. W. Moore (1989). *An Introduction to the Coriolis Force.* New York: Columbia University Press.

Szeri, A. and P. Holmes (1988). Nonlinear stability of axisymmetric swirling flows. *Phil. Trans. R. Soc. Lond. A 326*, 327–354.

Tagg, R. (1994). The Couette-Taylor problem. *Nonlinear Science Today 4*, 1–25.

Taylor, G. I. (1923). Stability of a viscous liquid contained between two rotating cylinders. *Phil. Trans. R. Soc. Lond. A 223*, 289–343.

Veronis, G. (1963). On the approximations involved in transforming the equations of motion from a spherical surface to the β-plane. II.Baroclinic systems. *J. Marine Res. 21*, 199–204.

Veronis, G. (1968). Comments on Phillips' proposed simplification of the equations of motion for a shallow rotating atmosphere. *J. Atmos. Sci. 25*, 1154–1155.

White, A. A. (1999). Hydrostatic and quasi-hydrostatic versions of the basic "New Dynamics" equations. Technical Report 271, Meteorological Office NWP Division, Bracknell, United Kingdom.

White, A. A. (2002). A view of the equations of meteorological dynamics and various approximations. In J. Norbury and I. Roulstone (Eds.), *Large-Scale Atmosphere-Ocean Dynamics*, pp. 1–100. Cambridge, U.K.: Cambridge University Press.

White, A. A. and R. A. Bromley (1995). Dynamically consistent, quasi-hydrostatic equations for global models with a complete representation of the coriolis force. *Q. J. R. Meteorol. Soc. 123*, 399–418.

Wilhelmson, R. and Y. Ogura (1972). The pressure perturbation and the numerical modeling of a cloud. *J. Atmos. Sci. 29*, 1295–1307.

Zauderer, E. (1989). *Partial Differential Equations of Applied Mathematics.* New York: Wiley.