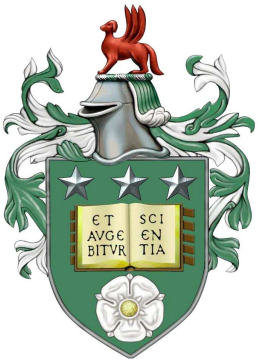
# The Anelastic Approximation: Magnetic Buoyancy and Magnetoconvection

Nikolai Andrei Berkoff




Submitted in accordance with the requirements for the degree of Doctor of Philosophy

The University of Leeds

Department of Applied Mathematics

July 2011

The candidate confirms that the work submitted is his own, except where work which has formed part of jointly-authored publications has been included. The contribution of the candidate and the other authors to this work has been explicitly indicated below. The candidate confirms that appropriate credit has been given within the thesis where reference has been made to the work of others.

# Joint Work

The work in Chapters 5 and 6 is based largely on work first published under the name *Comparison of the anelastic approximation with fully compressible equations for linear magnetoconvection and magnetic buoyancy*, authored by Berkoff, Kersale, & Tobias. The paper was published in the 2010 edition of Geophysical and Astrophysical Fluid Dynamics, volume 104, pages 545–563. The candidate was the sole developer of the computer program and obtained all the results presented. The co-authors advised on the development of the computer program, helped analyse the results, and edit the text.

# Acknowledgements

I would like to extend my greatest thanks to my supervisors, Steve Tobias and Evy Kersalé. They have been very patient and given plenty of support: they both had unwavering belief in me when I, at times, did not.

I would also like to thank, Chris Jones for the useful conversations I had with about everything anelastic; Gary Glatzmaier in offering advice; Krzysztof Mizerski for help with coding; Tamara Rogers for even more anelastic discussions; and all the above for help with explaining sections of their computer codes and papers.

My mother, Sarah, and who accpeted the task of proofreading, my sister, Ruth, has been very supportive, as have Tina and Rob who listened to my whingeing.

I have been supported in this work by a grant from STFC.

# Abstract

In this thesis I discuss a series of anelastic approximations and detail the assumptions used in the derivation. I derive an entropy and temperature formulation of the anelastic approximation along with a simplification to the entropy formulation introduced by Lantz (1992) and independently by Braginsky & Roberts (1995). I assess range of applicability of the anelastic approximation, which is often used in describing the dynamics of geophysical and astrophysical flows.

I consider two linear problems: magnetoconvection and magnetic buoyancy and compare the fully compressible solutions with those determined by solving the anelastic problem. I further compare the Lantz-Braginsky simplification with the full anelastic formulation which I find to work well if and only if the atmosphere is nearly adiabatic. I find that for the magnetoconvection problem the anelastic approximation works well if the departure from adiabaticity is small (as expected) and determine where the approximation breaks down. When the magnetic field is large then the anelastic approximation produces results which are markedly different from the fully compressible results. I also investigate the effects of altering the boundary conditions from isothermal to isentropic and the effect of stratification on how some of the parameters scale with the Chandrasekhar number. The results for magnetic buoyancy are less straight-forward, with the accuracy of the approximation being determined by the growth rate of the instability.

I argue that these results make it difficult to assess *a priori* whether the anelastic approximation will provide an accurate approximation to the fully compressible system for stably stratified problems. Thus, unlike the magnetoconvection problem, for magnetic buoyancy it is difficult to provide general rules as to when the anelastic approximation can be used. When the instability grows quickly or the magnetic field

is large the results do not compare well with the fully compressible equations. I outline a method for a two-dimensional non-linear time-stepping computer program and explain some problems with current non-linear programs.

# Contents

## List of figures

# List of figures

# Chapter 1

# Introduction

## 1.1 The Sun

The Sun has been studied for millennia with Babylonian astronomers taking some of the first early recordings of the Sun via the naked eye. Along with most other ancient peoples Babylonians studied the Sun to gain knowledge about the seasons and also due to their religious ceremonies. They applied mathematics to the collections of data recorded from water-clocks to deduce orbital periods. They deduced that the Earth was in orbit around the Sun and calculated the time for a sidereal year, the time for the Eath to rotate around the Sun, within 6 minutes of the current value (Leverington, 2003). Trying to understand our nearest star is still a challenge today due to the scales involved in the problem. It is of practical importance that we can predict solar phenomena, for instance space weather as this can damage satellites and other equipment, but it is of greater importance to science in investigating the natural world around us. To this end I will start by explaining some of the observations that motivate this work.

### 1.1.1 Observations

One of the earliest and easiest solar observations to make is that of sunspots. Sunspots are phenomena visible as dark spots compared with surrounding regions. Galileo Galilei published the first modern description of sunspots in 1613, in response to Christoph Scheiner who argued that sunspots were little planets. Galilei responded by tracking the motion of sunspots across the face of the Sun and proving that they rotate with the Sun's surface (Galilei et al., 1613). His 28-day observations, made with the new instrument of the time, the telescope, are shown in figure 1.1. Galilei even postulated that the spots, irregular in nature, were clouds of cool gas.

The number of sunspots at this time was decreasing as the Sun entered a period called the Maunder minimum during 1645 - 1715 (Eddy, 1976). The Maunder minimum was a period when sunspot numbers decreased dramatically. The sunspots that did appear during the minimum, 1680-1710, were confined to the southern hemisphere (Ribes & Nesme-Ribes, 1993). This shows that the mechanism which generated the spots was altered and although it was still capable of developing spots it did so in far fewer numbers and in only one hemisphere.

The sunspot cycle was discovered later. In his search for inter-mercurial planets Schwabe (1843) discovered that every ten years the number of sunspots reached a maximum. This observation has later been corrected to the eleven year sunspot cycle. Evidence for these cycles can be inferred from the effect the solar wind had on the Earth in the past. During normal solar activity the Sun produces a stream of magnetic particles, the solar wind, which deflects many galactic cosmic rays from reaching the Earth's heliosphere. Galactic cosmic rays are high energy particles and when these enter the terrestrial atmosphere they produce unstable radioactive isotopes, namely $^{10}\mathrm{Be}$ and $^{14}\mathrm{C}$ which can be measured in polar ice core records and tree fossil records respectively (Beer, 2000). An increase in radioactive isotopes indicates a decrease in the solar wind. This suggests events similar to the Maunder minimum, called Grand minima, occur aperiodically throughout the records. The radioactive record for $^{10}\mathrm{Be}$ shows evidence of

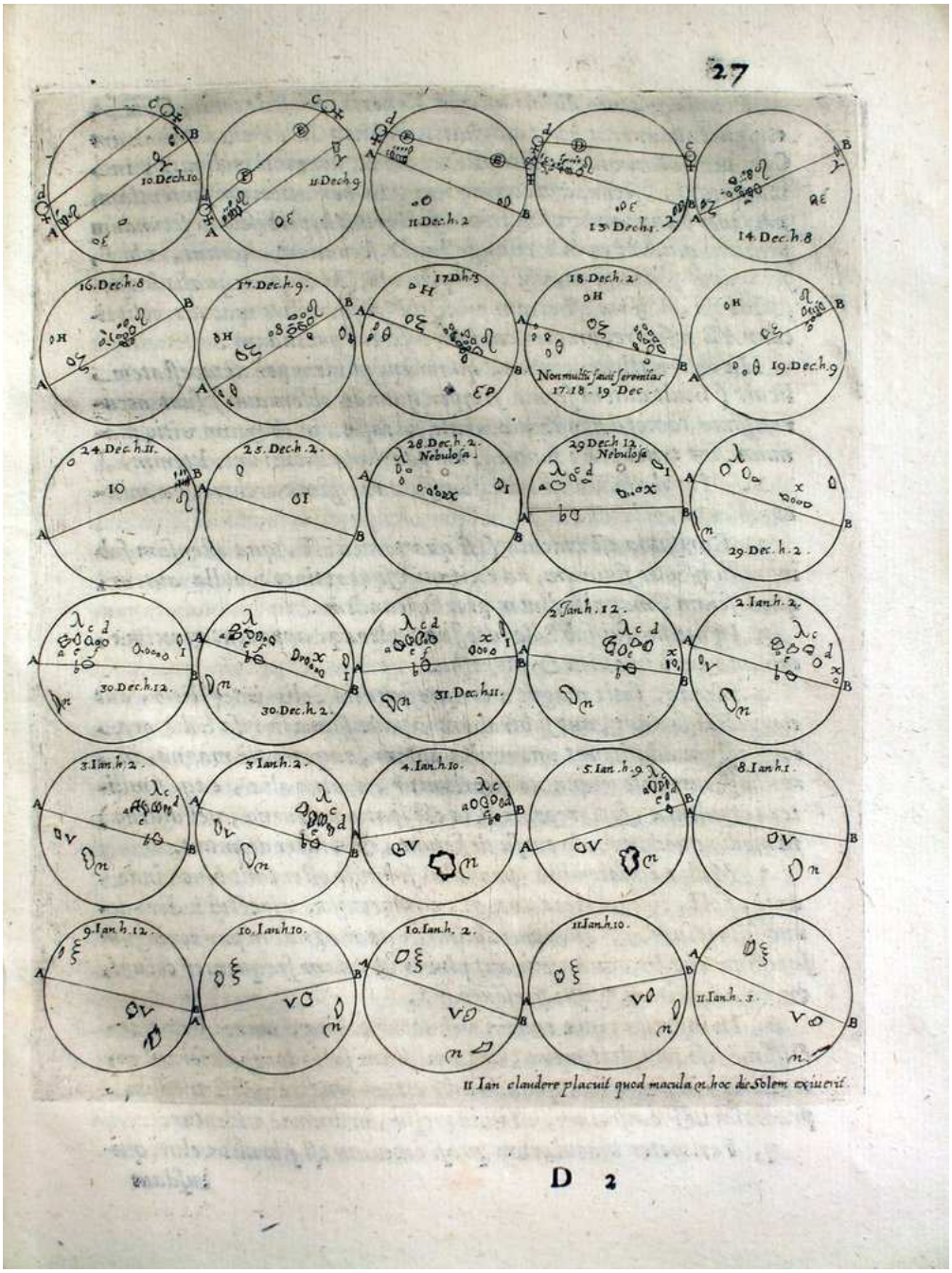



Figure 1.1: Galileo's drawings of sunspots from 'Letters on Sunspots' (1613) where Galileo, along with his student Benedetto Castelli, used a telescope to show that sunspots moved on the surface of the Sun.

the eleven-year cycle going back over 9 000 years (Vonmoos et al., 2006).

The magnetic nature of sunspots was discovered by measuring the Sun's spectral lines using the phenomena of Zeeman splitting. The splitting into different polarisation states of spectral lines, or widening, in the presences of magnetic fields is known as Zeeman splitting and was discovered in 1896. Hale (1908) used data from Mt. Wilson observatory, along with an understanding of Zeeman splitting, and concluded that sunspots possessed a magnetic field. He also noticed that sunspots typically appear in pairs of opposite polarity with the leading sunspot, i.e. the most Eastward, having a different polarity in the northern and southern hemispheres as in figure 1.3. This indicates that the Sun has a magnetic field which is coherent on a global scale.

A sunspot marks a region on the solar surface where a nearly vertical magnetic flux tube emerges from the solar interior. The radial field in the centre can be measured on magnetograms with a typical strength of the spot around $3$ kG. Large sunspots occur in pairs of opposite polarity, see figure 1.2 for an emerging pair. The polarity of the leading spot is the same for all pairs in the same hemisphere. The total solar radiance increases when sunspot activity is highest as active regions are on average brighter due to the additional faculae, even though they contain darker sunspots. The variation of solar irradiance between solar minimum and maximum is about $0.1\%$ which would lead to an expected effect on the Earth's temperature of $\sim 0.1\,^{\circ}\mathrm{C}$. For a review on sunspots see Thomas & Weiss (2008).

Along with the temporal variation, sunspots also display a spatial evolution. At the beginning of the sunspot cycle a small number of sunspots form at mid-latitudes $\sim 30^{\circ}$ then, slowly increasing in number, the regions of sunspot generation migrate to the equator. The number of sunspots then decreases and the cycle starts again. When the cycle starts again it does so with a reversed magnetic field so what was the north pole becomes the south pole. This magnetic field reversal means that although the sunspot cycle is over eleven-years the magnetic solar-cycle is over approximately twenty-two years. The polarity of the leading sunspot in each hemisphere is reversed after the each

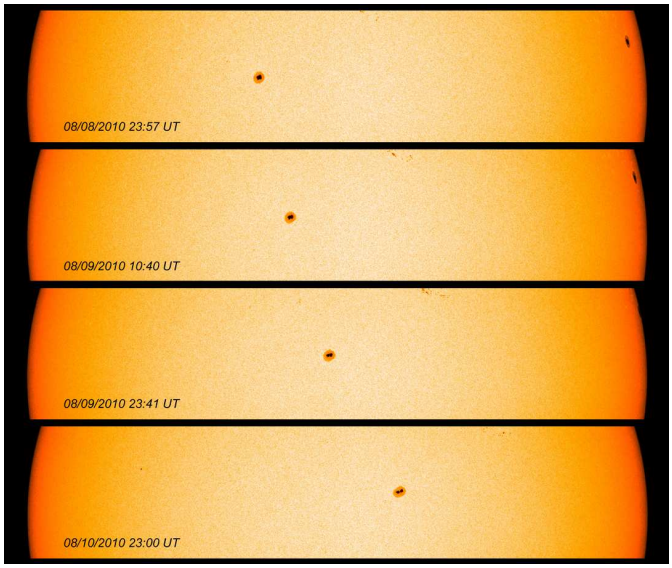



Figure 1.2: The emerging sunspot divided and became two spots over a two-day period. Each of the spots is about the size of Earth. Courtesy of NASA/SDO and the AIA, EVE, and HMI science teams.

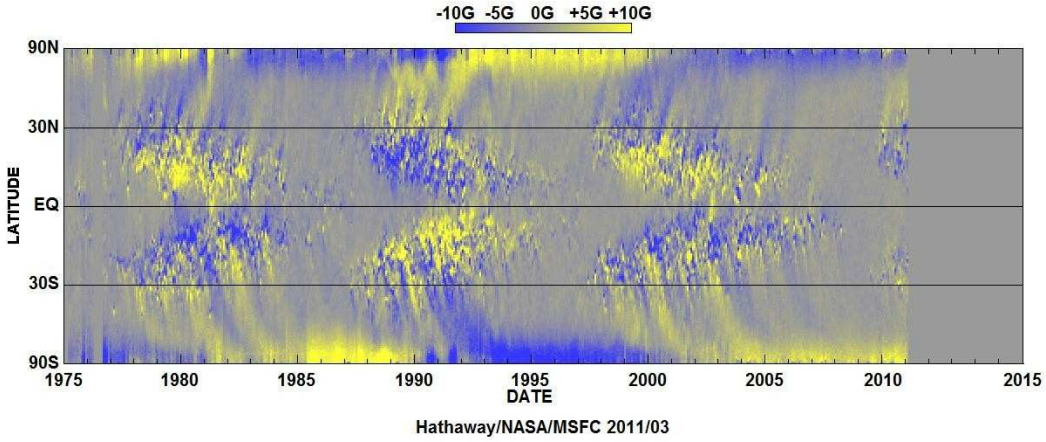



Figure 1.3: A magnetic butterfly diagram showing the polarity of the sunspots for each rotation of the Sun since 1975. Courtesy of David Hathaway at NASA.

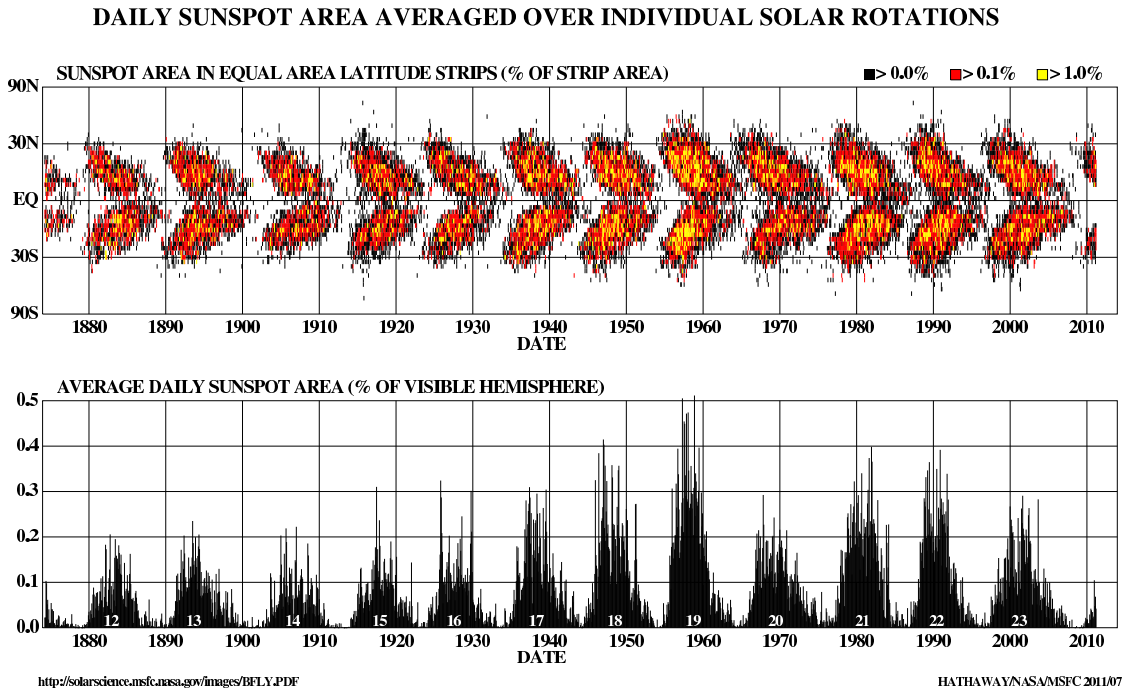



Figure 1.4: A butterfly diagram showing the positions of the sunspots for each rotation of the Sun since May 1874 shows that bands develop first from mid-latitudes, widen, and then move toward the equator as each cycle progresses. Courtesy of David Hathaway at NASA.

sunspot cycle, indicating the field reversal. If this process is mapped out on a time-latitude plot of sunspots this produces the well known 'butterfly diagram', first exhibited by Maunder (1904) and shown in figure 1.4. The cycles are shown in this 'butterfly diagram' where modulations can be seen in cycle strength and the latitude at which sunspots appear, but there is near symmetry between the two hemispheres. The solar coronal field can be used as a measure of the radial component of magnetic field and the coronal field reverses in the middle of the eleven-year sunspot cycle when numbers of sunspots is at a maximum.

It is often helpful to decompose magnetic fields into two components: poloidal and toroidal. In axisymmetric fields the toroidal component is in the same plane as the long dimension of a torus, or the component of magnetic field parallel to latitudinal lines. The poloidal field is outwards from the poles, and is the part of the magnetic field that contains a radial component. The coronal, or poloidal, field is out of phase with the magnetic field that creates sunspots, the toroidal field. This means that at solar maximum, when the number of sunspots is greatest, the coronal field is weakest, as in figure 1.3.

Today observations can be carried out using advanced satellites such as, Solar Terrestrial Relations Observatory (STEREO, `http://stereo.gsfc.nasa.gov/`), Solar and Heliospheric Observatory (SOHO, `http://sohowww.nascom.nasa.gov/` ), and Hinode (`http://solarb.msfc.nasa.gov/`) capturing many different wavelengths, high resolution measurements of the magnetic field, and even producing so-called 'three-dimensional images'. Although satellites remove the errors introduced from the Earth's atmosphere there are many ground based telescopes equipped with adaptive optics that are trained on the Sun, such as the Swedish 1-m Solar Telescope, which produces high resolution images, or the BiSON and GONG networks that provide an almost continual monitor of oscillations in the Sun from locations across the Earth.

Modern telescopes have allowed a much clearer imaging of the Sun and this has nurtured the study of coronal loops, fine structure in sunspots and other solar surface phenomena.

The telescopes have also allowed a more advanced prediction of space weather from monitoring events such as coronal mass ejections and gamma ray bursts. Observations of the Sun show the magnetic field is not diffuse but is instead concentrated into regions of intense field, ranging from sunspots (diameters $\sim 20$ Mm and field strengths of 3 kG) to magnetic knots ($200$ km, 1 kG) (see e.g. Thomas & Weiss, 2008). Images reveal a complicated pattern of fine structures in both the penumbra and in the umbra, as shown in figure 1.5. The umbra is the dark centre of a sunspot and it is surrounded by the penumbra, consisting of linear bright and dark radial elements. There are bright points that can be seen in the penumbra amongst downflows and these are clear indicators of magnetoconvection, convection in a magnetic medium. Some of these phenomena, such as umbral dots, can be explained by current theory and can be used to test that computer models are capable of producing results seen in the Sun. Umbral dots have diameters of 100-200 km and bright dots can last for days. Thomas & Weiss (2008) has a review of the features in sunspots and umbral dots can be seen in figure 1.5. Some sunspots are symmetrical but many are highly irregular.

As well as sunspots, which are a manifestation of the large solar magnetic field, the Sun also exhibits a small scale magnetic field on the surface. This is often referred to as the "magnetic carpet" and can be seen in the granulation patterns which occurs at two discrete scales: granules and supergranules. Granules are hot plumes of gas and have a size of $\sim 1,500$ km and time-scales of 15 minutes and supergranules are 20 times as large with time-scales of $\sim 24$ hours (Thomas & Weiss, 2008).

The final piece of observational data is from measurements of the solar rotation profile. The Sun is not a solid body. At the equator the Sun rotates in about 25.6 days and slower at the poles. The rotation rate observed at the surface is a monotonic smooth transition from the poles to the equator. It rotates counter-clockwise if viewed from a position above the Earth's northern pole. The internal rotation profile of the Sun is discussed in more depth later in this chapter.

The observational data fits together to build up a picture of the Sun but a theoretical

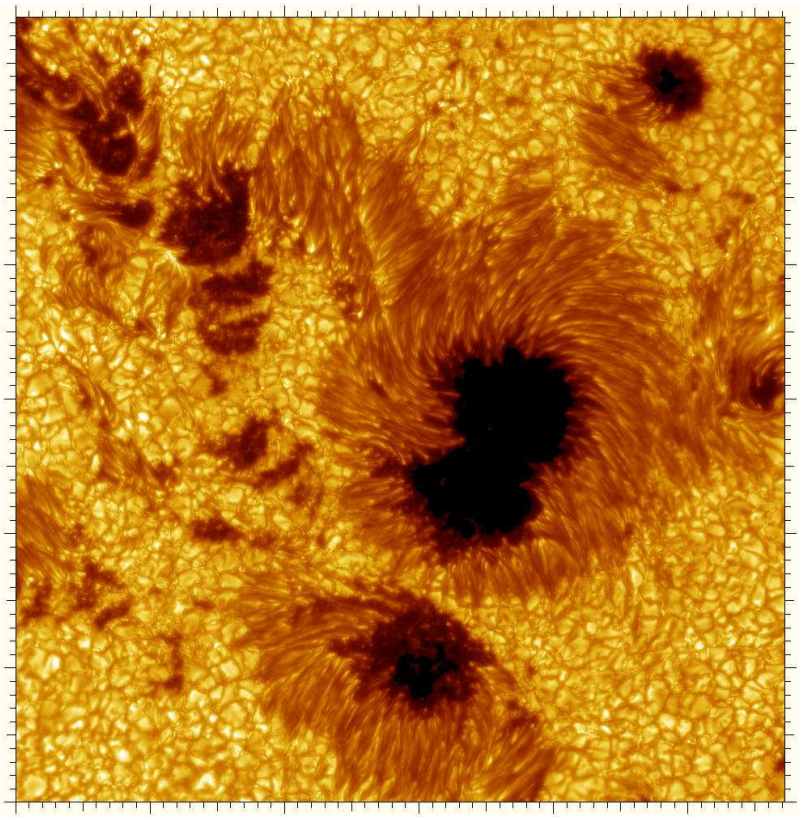


Figure 1.5: Image in visible light from the Swedish 1-m Solar Telescope showing dark filaments around a sunspot. The dark regions are the umbra and they are surrounded by the penumbra with a background of granulation. Penumbra filaments with dark cores can be seen protruding into the umbra. Credit: Royal Swedish Academy of Sciences.

model is also required, this is where solar magnetohydrodynamics is required. The aim of solar magnetohydrodynamics is to explain these observations and develop a coherent picture of the physical processes which take place.

### 1.1.2 Solar interior

In the solar core, where temperatures are $\sim 13.6 \times 10^6$ K, nuclear fusion combines hydrogen ions together in proton-proton reactions to form helium ions. There is a small amount of mass loss involved in nuclear fusion and that is carried away by high energy photons, neutrinos and gamma rays. Heat is radiated out from the core into the radiative interior of the Sun. At about $\sim 0.7$ $R_\odot$, where $R_\odot$ is the radius of the Sun, the opacity increases so that radiative processes are no longer capable of transferring the heat and the interior becomes unstable to convection. The convection zone is unstably stratified with highly time-dependent flows driven by vigorous thermal convection. The radiation and convection zones can be seen in figure 1.6 along with the transition between the two, which is discussed in more depth later. The convection zone is surrounded by the concentric zones of the photosphere, chromosphere and the corona respectively. A more detailed study of the internal solar structure is found in Priest (1984).

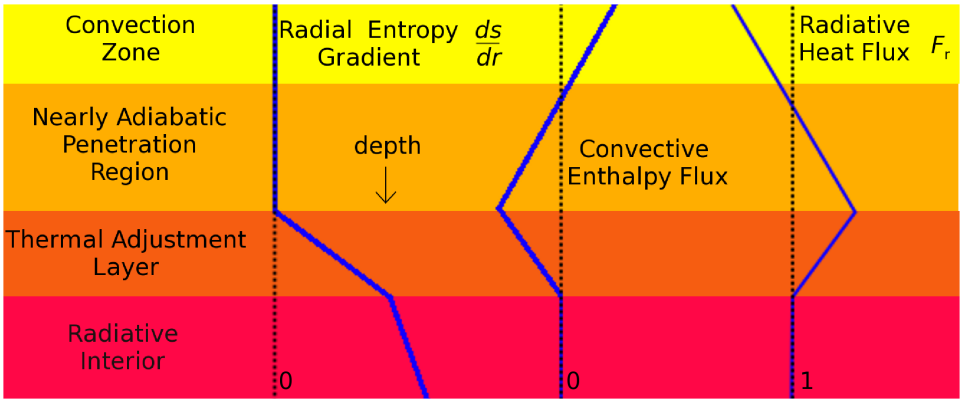



Figure 1.6: A schematic of the radial entropy gradient, $\mathrm{d}s/\mathrm{d}r$, convective enthalpy flux, and radiative heat flux $F_\mathrm{r}$. Courtesy of Miesch (2005).

The convection zone is highly turbulent. This turbulence can be seen on the surface in the form of solar granulation. The small scale magnetic dynamo, which is the self-generation of magnetic field on a small-scale, is thought to result from the interplay between magnetic fields and turbulent convection (see e.g. Cattaneo et al., 2003; Vögler & Schüssler, 2007) which gives rise to the "magnetic carpet", as previously discussed.

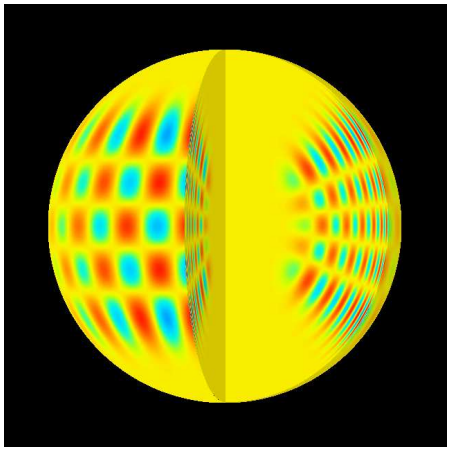


Figure 1.7: Computer generated image of a p-mode solar oscillation ($l = 20$, $m = 16$ and $n = 14$). Courtesy of Kosovichev et al. (1997).

Until the advent of helioseismology little was known about the internal rotation profile of the Sun. As discussed earlier, it was known that the surface rotation was non-uniform with faster rotation at the equator than at the poles. Some results could be deduced from the oblateness of the Sun but, beyond this, the internal rotation profile was guesswork, with suggestions that it obeyed the Taylor-Proudman theorem of constant rotation on cylinder (Pedlosky, 1987). There was also some debate as to how the Sun shed its angular momentum during the initial phase of gravitational collapse before it reached the main-sequence phase. This issue is solved, in part, by the solar wind torque which provides a mechanism to explain the loss of solar angular momentum. How deep this profile extends into the Sun was unknown until the mid-eighties when the field of helioseismology started producing results.

Helioseismology is the study of 'sun-quakes', i.e. of globally resonant oscillations inside the Sun. In the Sun there are many globally resonant modes and these can be analysed, using Fourier transforms, and decomposed into spherical harmonics. The turbulent convection creates a range of oscillations at the surface of the Sun. The acoustic or p-modes then travel towards the centre of the Sun but are refracted when the horizontal phase speed is equal to the sound speed. In the Sun the sound speed increases as a function of depth so high frequency oscillations are trapped in a thin acoustic cavity near the surface whereas low frequency oscillations can penetrate much deeper into the solar interior, see figure 1.7.

The frequencies of the different modes depends on $l$, $m$ and $n$, which are the angular order, azimuthal order, and absolute number of nodes in the radial direction respectively. The azimuthal wave number would exhibit degeneracy in a non-rotating case but the Sun is rotating so this degeneracy is removed. Thousands of these modes are measured by ground-based telescopes with varying frequencies and modes. The frequency depends on all three wave numbers and the frequency splits between modes with different radial order $n$. From the frequency splitting data information can be obtained about the internal rotation profile as a function of depth by solving the inverse problem. The current rotation profile is shown in figure 1.8; there is a lack of data at the poles as acoustic modes can give no information at high latitudes. An overview of this technique applied to the Sun is given in Christensen-Dalsgaard (1988) and Di Mauro (2008).

It is clear from the rotation profile shown in figure 1.8 that the variation seen at the surface between the equator and the poles extends into the solar interior with a sharp transition at the base of the convection zone. The rotation profile is not constant on cylinders, as previously suggested, but is more constant on radial lines. The dashed line, which extends from the 'Tachocline' label in figure 1.8, marks the base of the convection zone. There is a sharp transition from the stably stratified radiative zone, which rotates as a solid body, to the unstably stratified turbulent convection zone, which is rotating differentially. The transitional zone, or shear-layer, is called the tachocline which has

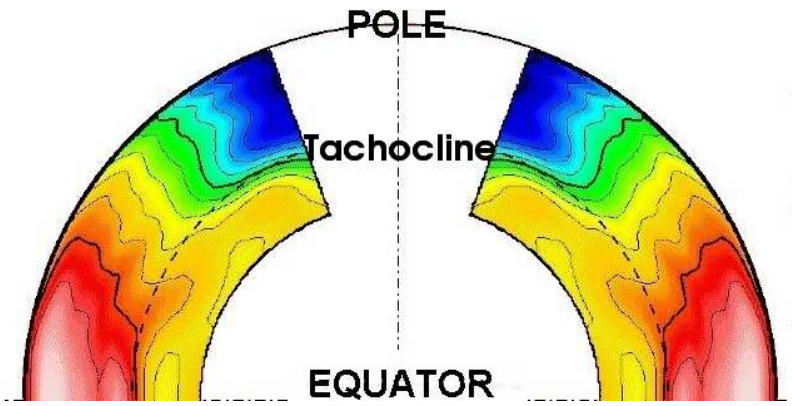



Figure 1.8: The internal rotation rate of the Sun with red for fast and blue for slow rotation rate. Image from M. J. Thompson.

a rotation profile which matches onto solid body rotation in the radiative zone. The tachocline is thin, helioseismology can only give an upper estimate of the thickness of this region but it is thought to be around $0.04$ $R_\odot$ (Charbonneau et al., 1999). The change from radiation to convection as the dominant heat transfer mechanism occurs at a similar place as the change from solid body rotation to differential rotation, see figure 1.6. The tachocline is the penetrative region and the thermal adjustment layer. In the thermal adjustment layer (the slow tachocline) the atmosphere is not nearly adiabatic, unlike the regions above it.

## 1.2 Theoretical background

Sunspots can be used as an indicator for large-scale magnetic field and the sunspot cycle exhibits a large-scale pattern which suggests a dynamic solar magnetic field of a global scale. Small scale magnetic field generation is generated by the interaction between the turbulent convection with the magnetic field and is reasonably well understood (see e.g. Cattaneo & Hughes, 2001; Cattaneo et al., 2003; Vögler & Schüssler, 2007). The large-scale field can be inferred from the solar cycle. The decay rate of the solar magnetic field due to diffusion alone is $\sim 10^{10}$ years which is on a similar time-scale to the lifetime of

the Sun. The time-scales do not rule out the possibility that the large-scale magnetic field is a fossil field: the decaying remnants of an initial field. Although not ruled out by time-scales it is unlikely that the solar magnetic field is the result of a fossil field as it is dynamic, as shown by the 'Butterfly diagram' in figure 1.4 along with surface magnetic field and Grand Minima. Along with time-variations in the solar differential rotation there are also oscillations called torsional oscillations, which are bands of faster or slower rotation which migrate in latitude on the solar surface (Proctor, 2006). All these features suggest that the solar magnetic field is not a simple oscillator.

Discarding the possibility of a fossil field the solar magnetic field must be sustained by converting kinetic energy into magnetic energy. Understanding the processes which drive the magnetic field generation is the goal of solar dynamo theory. Cowling (1933) proved that no steady axisymmetric magnetic field could be maintained by dynamo action. This means it is impossible to capture dynamo action in simple systems, unless additional assumptions are made e.g. in mean field theory (Krause & Rädler, 1980).

The simplistic theoretical picture for the solar dynamo is that an initial poloidal field is wound up, due to differential rotation, into a toroidal field. Parker (1975) showed that the large-scale magnetic field cannot be generated within the convection zone because it will rise to the surface too quickly, thus not allowing sufficient time for amplification. Before the discovery of the tachocline Parker predicted a dynamo, deep-seated in the Sun, where the solar interior is stable so that magnetic fields could remain stored for a time longer than convective time-scale. This would allow amplification by dynamo action before the fields rise, due to the magnetic buoyancy instability, through the turbulent convection zone and arrive at the surface as a pair of sunspots. For a review of rising flux in the convection zone see Fan (2009). To complete this theoretical picture the toroidal field needs then to be turned into poloidal field via another mechanism, such as proposed by Parker (1955a) and later by Steenbeck & Krause (1966), but this is beyond the scope of this thesis. For reviews on dynamo theory see (Ossendrijver, 2003; Tobias & Weiss, 2007; Charbonneau, 2010). Part of this theoretical picture is the winding-up of magnetic

field lines, this requires a strong shear flow. From the internal rotation profile the region with the largest shear is the tachocline. As the tachocline is a likely site for generation of the large-scale toroidal magnetic field modelling this region correctly is important for the understanding of the dynamo.

The reason the tachocline is so thin has been the subject of research for some time with early suggestions from Spiegel & Zahn (1992) who pose the problem by setting an initial condition then, as time advances, allow the system to spin-down. They maintain that a meridional flow, driven by baroclinic instabilities, would carry the differential rotation into the radiative zone over the lifetime of the Sun. In order to counteract this flow they suggest that stratified turbulence will mix the angular momentum in such a way as to make movements on a sphere more efficient than radial ones. The nature of the turbulence then plays a large role in keeping the tachocline so thin as a strong preference for horizontal motions can keep radiative effects from diffusing the tachocline. Spiegel & Zahn claim that the stratification in the tachocline will engender a two-dimensional turbulence over a three-dimensional turbulence.

Spiegel & Zahn (1992) did not consider the effects of magnetic fields in the horizontal turbulence but Gough & McIntyre (1998) suggest it is inevitable that the radiative interior has a large-scale magnetic field. A field as low as $10^{-2}$G is all that is required for the radiative zone to rotate as a solid body. In the limit where viscous effects are ignored Ferraro's law of isorotation states that fluid angular velocity is constant on a surface mapped by rotating a magnetic field line. Although this goes some way to explaining the thin tachocline it leaves the question of how the magnetic field in the interior does not transport the differential rotation from the convection zone inwards through magnetic coupling. For a review of the tachocline see Tobias (2005); Gough (2007); McIntyre (2007).

There are two processes which are fundamental to a solar dynamo model: magnetoconvection and magnetic buoyancy. I will give a brief introduction to both these processes.

### 1.2.1 Background to Magnetoconvection

Convection is the process of transferring fluid heat by motion. Warm fluid is less dense and, in a gravitationally stratified atmosphere, becomes buoyant. A parcel of warm rising fluid will occupy an increasing volume as it rises which will *smooth out* any perturbations present in the parcel. Conversely, a parcel of cold fluid is denser and so it will descend and occupy a decreasing volume, which will *intensify* any variations present in the parcel. Magnetoconvection is the study of how a convecting hydrodynamic system responds in the presence of magnetic fields, and how the two interact. The reviews of magnetoconvection by Hughes & Proctor (1988); Proctor & Weiss (1982); Proctor (2005) give a wide and detailed introduction to the field.

To understand magnetoconvection requires an understanding of how magnetic fields evolve in a conducting fluid. Magnetic fields obey Maxwell's equations and, assuming that the displacement current can be neglected, the evolution of magnetic fields is governed by the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) - \nabla \times (\eta \nabla \times \mathbf{B}) \,, \tag{1.1}$$

where $\eta$ is the magnetic diffusivity, and $\mathbf{B}$, $\mathbf{u}$ are the magnetic field and velocity field respectively. The magnetohydrodynamic equations will be further developed in §2.1.2. A non-dimensional number for measuring the importance of the magnetic advection to magnetic diffusion is the magnetic Reynolds number $R_m = u_s d/\eta$ where $u_s$, $d$ are relevant velocity and length scales. In the astrophysical context $R_m \gg 1$ so advection is dominant therefore the magnetic field lines move with the fluid and are wound up until they are on a length scale where diffusion is again important. For the tightly wound-up field the effective length-scale decreases so the local magnetic Reynolds number becomes smaller allowing reconnection to occur locally. Reconnection is where magnetic field lines from different domains are spliced together changing the magnetic topology of the system (see e.g. Priest & Forbes, 2000). Magnetic fields also have a force on the fluid flow, called the Lorentz force.

Early numerical studies of magnetoconvection were carried out using the Boussinesq approximation, where density stratification is ignored. For a longer discussion of the Boussinesq approximation see §2.2.1. The numerical results by Weiss (1966) show that in rising convective cells flux is expelled to the edge of the cell. Inside a cell the scale of flux variations decreases and resistive effects dominate, this phenomena is called flux expulsion. The total pressure, that is the thermal and magnetic pressure combined, must be nearly equal or a flow would develop to equalise the pressure imbalance. In flux tubes near the surface of the Sun the magnetic pressure is almost equal to the un-magnetised surroundings so the flux tube must be almost evacuated. The flux seen at the surface exceeds equipartition and around the tube is a strong downwelling of fluid as the tube emerges on the surface (Galloway et al., 1977). The strong magnetic fields impede convection via the Lorentz force, so the temperature in the flux tube will be less than the surrounding temperature. Simulations and observations show broad diffuse up-flows surrounded by cool dense flux tubes, with the minimum size of the downwellings limited by dynamical effects of the magnetic field. In a system with convective flows steadily overturning then magnetic fields will be advected into converging regions. This leads to flux stretching, which transfers energy from the velocity to magnetic field. This is important for dynamos.

Nonlinear Boussinesq magnetoconvection has been studied in great detail, with an overview in Proctor & Weiss (1982). There are limitations with using the Boussinesq approximation for magnetoconvection such as modelling evacuated regions or a changing plasma $\beta$. Magnetic pressure is not correctly modelled in Boussinesq magnetoconvection as evacuated regions cannot develop. The Boussinesq approximation has an up-down-symmetry which tends to favour hexagonal convective cells.

Using a fully compressible linear stability analysis Cattaneo (1984) studied two-dimensional modes. He found that when the plasma beta $\beta$, which is a ratio of the thermal pressure to the magnetic pressure, was $\beta \approx 1$ then fast and slow modes conspire

to make atmospheres, which are stable when un-magnetised, become convectively unstable. As this thesis is focused on the anelastic approximation the differences in fully compressible and Boussinesq convection are of interest to see where the anelastic approximation fairs better at modelling these phenomena.

### 1.2.2 Background to Magnetic Buoyancy

The Lorentz force can be split into a magnetic pressure and a curvature force. For an element of flux to be in dynamical equilibrium with its surroundings the sum of the thermal pressure and the magnetic pressure must be equal both inside and outside the element. If the element is magnetised and the surroundings are not then the external thermal pressure will exceed the internal thermal pressure. If the element is sufficiently thin, or for other reasons, the temperature in the element is equal to the surrounding temperature then the element will be less dense than its surroundings. In a gravitationally stratified atmosphere lighter elements will rise.

The term magnetic buoyancy was coined by Parker (1955b) to explain the formation of sunspots. In the review by Hughes (2007) he noted there are three different mechanisms referred to as magnetic buoyancy.

(i) In an isolated magnetic flux tube there is internal magnetic pressure so the internal thermal pressure will be less than the external pressure and so the density is lower inside the tube than out. Gravity will thus cause the tube to become buoyant and rise. This is more a lack of equilibrium than an instability.

(ii) A similar case where there is an isolated flux tube cooler than its surroundings so the pressure is in equilibrium but not overall mechanical equilibrium.

(iii) The buoyancy effect of a magnetic field can act as an instability mechanism in a magnetized atmosphere in equilibrium. The simplest being an atmosphere with

the horizontal field component only dependent on height as studied by Newcomb (1961).

I will only investigate the last mechanism and from now onwards, this is the mechanism I will be referring to by the term ‘magnetic buoyancy’.

In magnetic buoyancy two-dimensional instabilities, where the magnetic field lines are not bent, are called interchange modes because one magnetic field line exchanges position with another. Early theoretical work by Newcomb (1961) looked at the stability of interchange and three-dimensional modes in ideal plasmas using the energy principle of Bernstein et al. (1958).

An interesting feature of magnetic buoyancy is that three-dimensional modes can be more destabilising than interchange modes. The extra work done by three-dimensional modes against magnetic tension is often less than the extra work interchange modes must do to overcome thermal pressure and magnetic pressure to create a density perturbation (Hughes & Cattaneo, 1987).

The effect of shear flows on magnetic buoyancy was investigated by Tobias & Hughes (2004) who looked at the stability of a flow with $\mathbf{B} = B(z)\hat{\mathbf{x}}, \mathbf{U} = U(z)\hat{\mathbf{x}}$. Their analysis found that shear ultimately had a stabilising effect, but for certain modes the effect could be initially destabilising. Vasil & Brummell (2008) extended the work to the non-linear regime and found a very strong shear was required to cause a layer to become unstable to magnetic buoyancy although it is not shown if this is still true for sufficiently long time-scales, a recent update of this work is in Silvers et al. (2011).

It is still not clear if the concentrated flux that appears at the surface can be a result of the magnetic buoyancy instability alone, as non-linear effects of the Kelvin–Helmholtz instability occur on strongly buoyant fields wrapping them up so no large-scale flux escapes. Cattaneo & Hughes (1988) consider the nonlinear evolution of an interchange mode of a uniform magnetic field in an otherwise non-magnetic atmosphere. The instability develops as a usually Rayleigh-Taylor ‘mushroom’ instability but then there is

a second stage where the motion is dominated by interactions of vortices of opposite sign from neighbouring mushrooms. This leads to a fairly rapid break-up of the rising layer and so little flux escapes. Wissink et al. (2000) extended the two-dimensional model and investigated three-dimensional modes and noted the importance of twisting the flux tube to retain coherence as the tube rises.

# Chapter 2

# Mathematical Modelling

In this section I discuss the sets of equations that are used to model the physical system in this thesis. I also introduce the dimensionless parameters that are used. The physical systems of interest are compressible electrically conducting plasmas that can be modelled using a single fluid version of the magnetohydrodynamic equations.

## 2.1 Mathematical formulation of the problem

One set of equations that I will use to model the magnetic buoyancy and the magnetoconvection instabilities in the Sun are the magnetohydrodynamic equations. The magnetohydrodynamic equations are a combination of Maxwell's equations and Ohm's law along with the equations used in hydrodynamics. These equations are commonly used for solar modelling. A complete derivation is lengthy and unnecessary here but can be found in most text books on the subject, such as Davidson (2001).

### 2.1.1 Maxwell's Equations

It is useful to begin with Maxwell's equations and explain the set of assumptions that are required to derive the magnetohydrodynamic equations. The Maxwell's equations

describing the evolution of electro-magnetic fields are

$$\text{Gauss' law:} \quad \nabla \cdot \mathbf{E} = \varrho \mu_0 c^2, \tag{2.1a}$$

$$\text{Gauss' law of magnetic fields:} \quad \nabla \cdot \mathbf{B} = 0, \tag{2.1b}$$

$$\text{Faraday's law:} \quad \nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}, \tag{2.1c}$$

$$\text{Ampère's law:} \quad \nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \frac{1}{c^2}\frac{\partial \mathbf{E}}{\partial t}, \tag{2.1d}$$

where $\mathbf{E}$ is the electric field, $\varrho$ the charge density, $\mu_0$ is the permittivity of free space, and $c$ is the speed of light in a vacuum, $\mathbf{B}$ is the magnetic field, $t$ is time, and $\mathbf{J}$ is the electric current density.

Along with Maxwell's equations another relation is required to derive the magnetohydrodynamic equation: Ohm's law,

$$\mathbf{E} + \mathbf{u} \times \mathbf{B} = \frac{\mathbf{J}}{\sigma_e}, \tag{2.2}$$

where $\mathbf{u}$ is the velocity relative to the magnetic field $\mathbf{B}$ and $\sigma_e$ is the electrical conductivity of the medium. This is only valid in the reference frame of the plasma.

When the particles are moving at non-relativistic speeds, i.e. $u \ll c$ where $u$ is a typical velocity in the system, then the $\partial \mathbf{E}/\partial t$ in Ampère's law (2.1d) is negligible with respect to the other terms. This can be seen from a simple dimensional analysis. This allows equations to be combined and simplified given that

(i) the phenomena under consideration are slow compared to the plasma frequency so that the plasma is quasi-neutral, i.e. the number of electrons and ions in a volume is equal;

(ii) the plasma is collision dominated so it obeys a Maxwell-Boltzmann distribution of energy;

(iii) the length-scales must be large compared with the Debye length, which is the length over which a charge particle is screened, and the Larmor radius, which is the radius of the helical motions an electron makes about a field line.

With these assumptions in place the Navier-Stokes equations can be combined with the Maxwell equations to give the magnetohydrodynamic equations. In the Sun the scales of interest are large and the fluid is relatively slow and dense so these assumptions (i)-(iii) are met. This allows (2.1) to be combined to produce an equation describing the time evolution of the magnetic field and (2.2) give the Laplace / Lorentz force.

### 2.1.2 The Magnetohydrodynamic Equations

I consider a plasma governed by the evolution equations for the mass density $\rho$, velocity field $\mathbf{u} = (u, v, w)$, magnetic field $\mathbf{B}$, thermal pressure $p$ and either temperature $T$ or entropy $s$. The fully compressible equations of magnetohydrodynamics (see e.g. Hurlburt et al., 1996) are given by

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

$$\rho \left[ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla)\, \mathbf{u} \right] = -\nabla p + \rho \mathbf{g} + \frac{1}{\mu_0} (\nabla \times \mathbf{B}) \times \mathbf{B} + \mu \nabla \cdot \boldsymbol{\tau}, \tag{2.3b}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + \eta \nabla^2 \mathbf{B} \quad \text{with} \quad \nabla \cdot \mathbf{B} = 0, \tag{2.3c}$$

$$c_v \rho \left[ \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right] = -p \nabla \cdot \mathbf{u} + k \nabla^2 T + \mu \frac{\partial u_i}{\partial x_j} \tau_{ij} + \frac{\eta}{\mu_0} (\nabla \times \mathbf{B})^2, \tag{2.3d}$$

$$\text{where} \qquad \tau_{ij} = \left( \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \frac{\partial u_k}{\partial x_k} \delta_{ij} \right), \tag{2.3e}$$

$$p = (c_p - c_v) \rho T, \tag{2.3f}$$

where $c_p$ and $c_v$ are the specific heats at constant pressure and volume respectively; $\mathbf{g}$ is gravity; $\mu$ is the dynamic viscosity; $\eta = 1/\left(eta_0 \sigma_e\right)$ is the magnetic resistivity; $k$ is the coefficient of thermometric conductivity (or thermal conductivity) and is related to thermal diffusivity by $\kappa = k/(\rho c_p)$; $\mu_0$ is the magnetic permeability of free space. I

will assume that all of these parameters are constant. This assumption is valid only if the height of the domain of interest is much smaller than the gravity length scale and diffusion length scales, the variations in the non-radial (or in a Cartesian representation, horizontal) direction in the Sun are suitably small, and the Sun behaves as an ideal gas. The assumption of constant g means that self-gravity is ignored so that any motion of fluid should not be so large as to alter the local gravity. The assumption that the Sun behaves as an ideal gas is valid even at high pressures experienced in the solar interior; the high temperatures ensure that the gas is a charged plasma of H and He ions, which are orders of magnitude smaller in volume than H and He atoms, and so the gas can still be modelled as ideal point-like ideal gas.

### 2.1.3 The Lorentz Force

The momentum equation (2.3b) shows how the magnetic field effects the fluid. The magnetic field interacts with the velocity field via the Lorentz force, which can be written as

$$\frac{(\nabla \times \mathbf{B}) \times \mathbf{B}}{\mu_0} = \frac{\mathbf{B} \cdot \nabla \mathbf{B}}{\mu_0} - \nabla \left( \frac{|\mathbf{B}|^2}{2\mu_0} \right), \tag{2.4}$$

where $\mu_0$ is the magnetic permeability of free space. From this it is clear the Lorentz force has two distinct parts: a curvature force, $\mathbf{B} \cdot \nabla \mathbf{B}$ and a magnetic pressure $\frac{1}{2\mu_0}|\mathbf{B}|^2$. The curvature force acts to straighten curves in magnetic field lines by exerting a force with a component normal to the direction of the curvature of the field. This is what gives the magnetic field lines a restoring force so that the field lines can act like a string. It is this restoring force that allows the propagation of Alfvén waves. These waves are dispersionless wave that propagates in a similar manner to sound waves in that they cause no flow perpendicular to the direction they travel in. The magnetic pressure acts in a similar way to the thermal pressure but has a magnetic, rather than thermal, origin. For two neighbouring elements to be in total pressure equilibrium the magnetic and thermal pressure must sum to the same value in both elements.

### 2.1.4 A Discussion on Modelling Approaches

In the later chapters I shall predominantly be using equations in dimensionless form in which the dimensionless numbers show the relative dominance of various terms, e.g. the Reynolds number $\mathrm{Re} = UL/\mu$ is the ratio of inertial forces to the viscous forces, where $U$ is a typical speed and $L$ is a typical length. At low values of the Reynolds number the flow is laminar and small scale motions are heavily damped by viscosity so that disturbances to the mean flow will dissipate over time. This is the case if the flow is in a small domain, slowly moving or highly viscous. At high values of the Reynolds number the flow is turbulent and this is a much harder flow to model. The small scale motions are not heavily damped, so, to accurately model the flow small scales must be included. These small scale motions may effect the large scale flow through non-linear interactions, an inverse cascade, or through creating a type of anisotropy in the flow.

The Reynolds number is a standard dimensionless number in the dimensionless momentum equation and in magnetohydrodynamics there is a magnetic equivalent: the magnetic Reynolds number $\mathrm{Re}_m = UL/\eta$, which is in the dimensionless induction equation. This dimensionless number measures the effect of advection compared to magnetic diffusivity. In the case of a large magnetic Reynolds number the magnetic field is advected along with the fluid motions with minimal diffusion. At infinite magnetic Reynolds number the field is 'frozen in' which means that the field lines move with the fluid and do not diffuse (Alfvén, 1943). In the Sun's atmosphere the magnetic Reynolds number is large and so without a dynamo the magnetic field would take $\sim 10^{10}$ years to decay away. The range of parameters required to model the Sun is given in table 2.1. From the Reynolds number it is possible to estimate the range of physically relevant length scales from Kolmogorov's dissipation length to global oscillations on the scale of the Sun. This estimate gives a range of length-scales $\sim 10^{10}$ and three-dimensional computer modelling would require $\sim 10^{30}$ grid points, which is far greater than $\sim 10^{11}$ points that is around the current upper bound available in state-of-the-art super-computers. Making progress on computational modelling of the

Sun requires a tactic other than developing more efficient algorithms for the currently available computers.

Attempts have been made to parametrise the effects of the small scales or use mean field theory so they do not need to be explicitly resolved. Reynolds (1895) developed an important theory for modelling turbulence by averaging the Navier-Stokes equations. The problem with this approach is that averaging and parametrising motions introduces more unknowns than there are equations. The additional unknowns will require additional assumptions in an attempt to close the system. Subgrid-scale closure is a way of parametrising effects that are thought to exist on a scale unresolved in the simulation. The earliest paper using a subgrid-scale closure is Smagorinsky (1963) which relates the stresses to strains using an eddy viscosity. All these parametrisations rely on various additional assumptions which will be correct in certain instances but the range of validity is not known.

For this reason I have used relatively simple models so I can alter parameters computationally cheaply. Here simple means not including the whole range of physical processes in order to focus on relevant or essential mechanisms. The aim was to allow insight into how each parameter affects the system. This runs the risk of ignoring a physical process that may be fundamental to understanding the system e.g. a depth dependent diffusivity, but has the benefit that the equations I am solving are the correct equations to use e.g. there are no terms from a turbulence closure model. The benefits of having a simple system is that many simulations can be run cheaply to survey parameter space so that robust features are clearly identified. If the diffusivities were set to values expected from observations of the Sun then small structures could develop, which would be under-resolved in most models, so for the most part the values of the dimensionless numbers used in this thesis are not values expected in the Sun, this is to make the problem computationally tractable. This leads to a possible compromise where is it possible to use the correct equations with the wrong parameters or the wrong equations but at the observed parameter values. In this thesis I am using the correct equations but with the

wrong parameter values with the hope of finding trends that can be extrapolated into the correct parameter value regimes.

### 2.1.5 Non-dimensional form

The equations (2.3a-d) can be written in non-dimensional form. It is convenient to work with the equations in non-dimensional form as it reduces the number of parameters to a minimum. I will assume that the region of plasma that is to be modelled has a depth, $d$, which will be the unit of length and the dynamical time, $\sqrt{d/g}$, is the unit of time; the units of magnetic field, temperature and mass density, $B_0$, $T_0$ and $\rho_0$ respectively, are their values at the top boundary of the domain. The unit of pressure from dimensional analysis is then $p_0 \sim g\rho_0 d$. The dynamics of the problem is set not by the value of the individual diffusivities and flows but by their relative importance in the equations. From the relative importance of some parameters one can see when terms in the equations are dominant or not which may give an insight into what is driving certain instabilities.

Using similar dimensionless numbers as in Tobias et al. (1998) I define the following set of non-dimensional numbers

$$C_k = \frac{k}{d^2\rho_0 c_p}\sqrt{\frac{d}{g}}, \quad \mathrm{Pr} = \frac{\mu c_p}{k}, \quad \zeta = \frac{\eta c_p \rho_0}{k}, \quad R = \frac{g c_p d^3 \rho_0^2}{k\mu}, \quad \text{and } \mathcal{F} = \frac{B_0^2}{g d \rho_0 \mu_0}, \tag{2.5}$$

where $C_k$ is the non-dimensional thermal conduction of the system — it is also the ratio of the thermal relaxation time ($d^2\rho c_p/k$) to the sound crossing time ($d/\sqrt{p_0/\rho_0}$); the Prandtl number, $\mathrm{Pr}$, is the ratio of momentum diffusivity to thermal conductivity; the inverse of the Roberts number, $\zeta$, is the ratio of magnetic diffusivity to thermal conductivity; $R$ is a non-dimensional measure of the strength of gravity $g$ to diffusion, and $\mathcal{F}$ is the non-dimensional strength of the magnetic field — it is proportional to the inverse of the plasma $\beta$. (The magnetic Prandtl number can be obtained by $\mathrm{Pr}_m = \mathrm{Pr}/\zeta$.) The range of these parameters in a solar context is given in table 2.1 where the Mach number $M = u/u_s$ appears, with $u$ representing a typical velocity

| | Base of convection zone | Photosphere |
|---|---|---|
| $R = \frac{g c_p d^3 \rho_0^2}{k \mu}$ | $10^{20}$ | $10^{16}$ |
| $\zeta$ | $10^{-4}$ | $10^{-1}$ |
| Pr | $10^{-7}$ | $10^{-7}$ |
| $\mathcal{F}$ | $10^{-5} \ldots 10^{-7}$ | $10^{0}$ |
| $C_k$ | $10^{-11}$ | $10^{-5}$ |
| $M$ | $10^{-4}$ | $10^{0}$ |

Table 2.1: Values of relevant non-dimensional numbers from Ossendrijver (2003).

and $u_s$ the sound speed. Relations between these dimensionless numbers and other commonly utilised ones is given in Appendix C .

Using these non-dimensional numbers the governing equations become

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \mathbf{u}) = 0, \tag{2.6a}$$

$$\rho \left[ \frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u} \cdot \nabla) \mathbf{u} \right] = -\nabla(\rho T) + \rho \hat{\mathbf{g}} + \mathcal{F} (\nabla \times \mathbf{B}) \times \mathbf{B} + \left( \frac{\mathrm{Pr}}{R} \right)^{1/2} \nabla \cdot \boldsymbol{\tau}, \tag{2.6b}$$

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{u} \times \mathbf{B}) + C_k \zeta \nabla^2 \mathbf{B} \quad \text{with} \quad \nabla \cdot \mathbf{B} = 0, \tag{2.6c}$$

$$\rho \left[ \frac{\partial T}{\partial t} + (\mathbf{u} \cdot \nabla) T \right] = -(\gamma - 1) \rho T \nabla \cdot \mathbf{u} + (\gamma - 1) C_k^2 \left( \mathrm{Pr}^3 R \right)^{1/2} \frac{\partial u_i}{\partial x_j} \tau_{ij} + \frac{\gamma}{(\mathrm{Pr}\, R)^{1/2}} \nabla^2 T + (\gamma - 1) C_k^2 (\mathrm{Pr} R)^{1/2} \zeta \mathcal{F} (\nabla \times \mathbf{B})^2, \tag{2.6d}$$

$$p = \frac{1}{C_k^2 \, \mathrm{Pr}\, R} \rho T, \tag{2.6e}$$

where $\gamma = c_p / c_v$ is the ratio of the specific heats.

## 2.2 Convective Approximations

In many astrophysical and geophysical situations, although some effects of compressibility are essential for modelling the dynamics appropriately, not all are relevant. In particular, it is often the case that the dynamics of sound waves (and fast magneto-acoustic waves in magnetised domains) are secondary to the evolution of the system. It is also the case that focusing on the small fluctuations only and ignoring a static background state can make interpreting results from analytical and computational simulations easier. In the Sun estimates of sound speeds are $\sim 0.1$ Mm.s$^{-1}$ in the solar interior (Christensen-Dalsgaard, 1985) and so sound waves are the fastest waves in the system. The convective approximations were developed to make the analysis of the system easier. In fully compressible codes most of the computational power is going into correctly modelling the fast magneto-acoustic and sound waves, both of which are not thought to play a role in many instabilities in the Sun's interior. This is inefficient and it would be preferable to filter such fast waves.

A way to remove the fast waves and therefore simplify the system is to use a convective approximation. Both the Boussinesq and anelastic approximations are convective approximations and are not applicable to all systems. They both require velocities to be far smaller than the speed of sound. The system modelled must also be steadily driven, i.e. not fast (on convective time-scale) changes in the boundary conditions or from any heat sources. Convective approximations treat the pressure and buoyancy forces in a linear manner, whilst advection is still non-linear. They also remove a large stationary state and focus on the fluctuations which can make the analysis of the system easier.

### 2.2.1 Background to the Boussinesq equations

The Boussinesq approximation is the oldest convective approximation to the magnetohydrodynamic equations (see Spiegel & Veronis (1960) or Chandrasekhar (1961)). The Boussinesq approximation assumes that the typical depth of the layer

modelled $d$ is small compared with the pressure scale height of the fluid, $d_p$ where

$$d \ll d_p = \left| \frac{1}{p} \frac{\mathrm{d}p}{\mathrm{d}z} \right|^{-1}, \tag{2.7}$$

where $z$ is depth. The other assumption is that density fluctuations are due to temperature changes, with the pressure remaining relatively unchanged. In general, relative density and temperature fluctuations will be of the same order. In a gas the density fluctuation is the driver of a convective instability and therefore this term is retained when it is coupled with the acceleration due to gravity. This term must remain or the approximation will remove the driver of the system. These assumptions have been used by Rayleigh to study Bénard convection and in many systems since (Rayleigh, 1916).

It is helpful to lay some mathematical ground-work for why the pressure fluctuations can be ignored. I will express each variable $\xi$ as

$$\xi\left(\boldsymbol{x}, t\right) = \overline{\xi}\left(z\right) + \xi^{*}\left(\boldsymbol{x}, t\right), \tag{2.8}$$

where $\overline{\xi}$ is the horizontally averaged quantity, and $\xi^{*}$ is a small fluctuation to that state. Here I am using a Cartesian geometry. The horizontal average used is time independent and so the decomposition of $\xi$ above will not be able to model time dependent vertical boundary conditions. Starting with hydrostatic balance

$$\frac{\partial \overline{p}}{\partial z} = g\overline{\rho}, \tag{2.9}$$

where gravity is pointing in the positive $\hat{\mathbf{z}}$-direction so that $z$ increases with depth. The pressure fluctuations are driven by the flow, $|\mathbf{u}|$

$$|\mathbf{u}|^2\overline{\rho} \approx p^{*}, \tag{2.10}$$

whereas for the density fluctuation the kinetic energy of the flow can be balanced with the gravitational energy

$$|\mathbf{u}|^2\overline{\rho} \approx gd\rho^{*},$$

where the layer depth is $d$. This leads to

$$\frac{p^{*}}{\overline{p}} \approx \left(\frac{d}{d_p}\right)\frac{\rho^{*}}{\overline{\rho}},$$

which from (2.7) shows that the relative pressure fluctuations are far smaller than the relative density fluctuations.

If only linear calculations are being done then the only assumption required is for the layer depth to be much smaller than the stratification scale, but for non-linear work a further assumption is required namely that the density fluctuations do not become larger than the variations in the horizontally averaged density (Spiegel & Veronis, 1960). The assumptions made about density and temperature are that they vary only slightly in the fluid but the flow is essentially buoyancy driven. This is the case in an almost incompressible fluid but, unlike in an incompressible fluid, the density fluctuations are retained, as previously mentioned. Spiegel & Veronis (1960) showed that the Boussinesq approximation is formally equivalent to the incompressible system when the temperature gradient is replaced with the departure from the adiabatic temperature gradient. Since density perturbations are only kept in the buoyancy term then the conservation of mass (2.3a) is reduced to

$$\nabla \cdot \mathbf{u} = 0, \tag{2.11}$$

and, as pointed out by Lilly (1996), this means that the system conserves volume rather than mass. The energetics of the Boussinesq approximation is discussed in Chandrasekhar (1961).

When it is possible to separate the fast processes (e.g. acoustic) from the slow processes (e.g. convection), the Boussinesq approximation filters out the fast processes. The aim of the approximation is to retain only the minimum required complexity in the system but to still capture the essential physics.

**Magneto-Boussinesq equations**

Spiegel & Weiss (1982) extended the Boussinesq approximation to include varying magnetic fields which allows magnetic buoyancy to be investigated. They found that the magnetic buoyancy instability is captured within this magneto-Boussinesq

approximation. In the standard Boussinesq approximation the fluctuation pressure is neglected in the equation of state and so if a magnetic field is added then it must not create a large magnetic pressure or this would alter the thermal pressure. In the standard Boussinesq equations the pressure fluctuation is ignored which, if magnetic fields are introduced, would mean that the field profile must be slowly altering at most or the field to be sufficiently weak. In the magneto-Boussinesq equations magnetic fields are allowed to exert a significant magnetic pressure. The total pressure fluctuations, made from the magnetic pressure fluctuations and thermal pressure fluctuations, are still considered small in the magneto-Boussinesq approximation. The thermal pressure fluctuations can no longer be ignored, as is done in the standard Boussinesq approximation, and must be modelled. Also the way the induction equation is altered in the Boussinesq approximation means that the magnetic field is not kept exactly solenoidal. This can be seen by substituting $\nabla \cdot \mathbf{u} = 0$ into the induction equation. When the divergence of the induction equation is taken it is clear that the condition $\nabla \cdot \mathbf{B} = 0$ is not satisfied exactly. This should not be a problem for modelling magnetic buoyancy as Spiegel & Weiss (1982) show that the departure of the magnetic field from being solenoidal is proportional to the scale variations of the vertical velocity parallel to magnetic field. For magnetic buoyancy the scale variations of the vertical velocity remain small so the magnetic field will remain mostly solenoidal. When the length scale of the variations is of the order of the layer depth then the magneto-Boussinesq approximation is no longer valid.

### 2.2.2 Background to the Anelastic Approximation

In early computational simulations of convection the Boussinesq approximation was used for simplicity (see e.g. Durney, 1970; Deardorff, 1964). In many situations of geophysical and astrophysical interest it is important to include some effects of stratification and compressibility, which are outside the scope of the Boussinesq approximation, which neglects both of these features. These situations include the

interior of stars and giant planets, characterised by a large number of density scale heights (Glatzmaier, 2005). The strong variation in density is believed to play an important role in determining the dynamics of these systems and there is interest in modelling situations where in some regions the fluid is stable to convection while in other regions the fluid is not. For example, modelling of the deep solar interior often focuses on the behaviour of both the unstably stratified convective zone and the stably stratified tachocline with the radiative zone below (and sometimes the transition between the two) (see e.g. Gough & McIntyre, 1998; Tobias et al., 2001; Brummell et al., 2002; Brun & Toomre, 2002; Garaud, 2002; Rogers & Glatzmaier, 2005; Miesch, 2005; Rogers et al., 2006). Not including effects of stratification may introduce symmetries such as an up-down symmetry i.e. the fluid in the lower half of the domain is equivalent to the fluid in the upper half on average, provided that the fluid identities are reversed and that the boundary conditions at the top and bottom are equivalent. Clearly the Boussinesq approximation is not ideal in modelling the transition between stably and unstably stratified regions as the background atmosphere is fixed.

An intermediate approximation between the fully compressible equations and the Boussinesq approximation is then appropriate. The anelastic approximation is exactly such an approximation; it retains some effects of compressibility whilst filtering out the sound waves and in the magnetic case fast-magnetoacoustic waves. In the Boussinesq equations there was an assumption that the typical scale height was small compared to the pressure scale height; this constraint is relaxed in the anelastic approximation. The anelastic approximation has been heavily used to model astrophysical and geophysical fluids (see e.g. Glatzmaier & Roberts, 1996; Miesch et al., 2000; Fan, 2001; Anufriev et al., 2005; Rogers & Glatzmaier, 2005; Clune et al., 1999). Various forms of the anelastic approximation have been derived, using different small parameters for the asymptotic expansion leading to the constitutive equations and different physical assumptions to filter out the sound waves. The features that hold these various forms of the anelastic approximations in common is that they attempt to retain stratification whilst removing sound waves. I will look at the validity of some of these anelastic

approximations in later chapters.

The anelastic equations were originally derived by Batchelor (1953) in a meteorological context where the pressure and density are assumed to be close to their adiabatic values. This was motivated by an investigation of flows with a low Richardson number, which is the ratio of potential to kinetic energy, and resulted from the consideration of dynamical similarities in such flows. Low Richardson number means the fluid is weakly stratified and buoyancy is unimportant in the flow. A more formal scale analysis was performed by Ogura & Phillips (1962). They constructed an asymptotic expansion using a small parameter defined as the departure from adiabaticity and used a time-scale built upon the Brunt-Väisälä frequency in order to separate the dynamics of gravity and acoustic waves. The Brunt-Väisälä frequency is the characteristic rising time of a convective element or the frequency of a gravity wave. It is the frequency which separates high frequency acoustic waves from low frequency waves. Ogura & Phillips also assumed adiabatic motion and state the advantage of filtering sound waves in numerical computations. This filtering is a consequence of choosing a time-scale based on the Brunt-Väisälä frequency; from this choice in time-scale the fluid velocity must be small compared to the sound speed.

A complementary approach was used by Gough (1969), who derived the anelastic equations using a small parameter based on temperature fluctuations from the convective heat flux. This results in equations that are the same as those derived by Ogura & Phillips only when the atmosphere is perfectly adiabatic. Gough also allowed for the possibility of external forcing and a time-dependent reference state. Gough's small parameter arises from assuming that the temperature flux throughout the stratified atmosphere is the convective flux minus the lateral temperature flux. Gough argues that this allows the assumption that the atmosphere remains close to adiabatic to be relaxed; however the assumption about the dynamical time-scale, and so small Mach number, remains. Although not mentioned in Gough's paper the assumption about the atmosphere being allowed to depart from adiabatic stratification is only slightly relaxed or this would

introduce inconsistencies. Magnetic fields can be included in the anelastic equations if the local Alfvén speed scales as the convective speed (Glatzmaier, 1984). The equations in Gilman & Glatzmaier (1981) extend those in Ogura & Phillips (1962) to a spherical geometry and with dissipative effects included. The spherical magnetohydrodynamics code ASH (Clune et al., 1999; Miesch et al., 2000) is based upon the set of anelastic equations in Gilman & Glatzmaier (1981) but evolves the reference state. The ASH code is in wide use e.g. convection in B-type stars (Augustson et al., 2010), wreath-building dynamos (Brown, 2009), and rapidly rotating stars (Brown et al., 2007).

An additional common assumption made in anelastic models, originating from Glatzmaier (1984) and Glatzmaier (1985), is that the diffusion of heat is proportional to the entropy gradient. This assumption comes from sub-grid-scale turbulent-eddy arguments. The molecular temperature diffusion acts on a length scale derived from the mean free particle path but, as the systems of interest are turbulent, it is argued that the mixing of eddies is more important. Simulations do not extend down to molecular level so molecular thermal diffusion is already modelled incorrectly with inflated diffusion parameters; entropy diffusion is an attempt to parametrise the diffusion in a turbulent domain more realistically. This is explained more in §3.2.

The fully compressible and anelastic equations have non-linearities, such as the magnetic non-linearities as well as the more familiar hydrodynamic non-linearity of advection, which require extra computational resources to be solved correctly. For computational simplicity the anelastic approximation aims to make the system as simple as possible whilst retaining the complexity required to capture the dynamics of interest. The hydrodynamic and magnetic advection terms are the cause of much of the interesting dynamics and should not be removed. There are, however, thermodynamic non-linearities which are removed in the anelastic approximation with the consequence that the thermodynamics will be less accurately modelled in some cases, such as when the fluctuation is too large or the atmosphere is far from being adiabatically stratified. Removal of the non-linearities is done in the same way as in the Boussinesq

approximation, namely that an average (reference) state is taken and that thermodynamic fluctuations are taken about that reference state. As the thermodynamic terms are all related by an equation of state then a term such as a temperature multiplying a pressure is a non-linear term. The thermodynamic non-linearity is removed by replacing most instances of density with this slowly, if at all, varying reference state density. For example when (2.3d) is divided by $\rho$ then there are pressure and temperature terms that are divided by the density, from the equation of state these are linked and this is the type of non-linearity that the anelastic approximation removes. This allows the thermodynamic relations to be linearised but other non-linearities in the system to remain.

Finally, Lantz (1992) and, independently, Braginsky & Roberts (1995) made a further simplification by writing the thermodynamic variables in terms of entropy and pressure, with the pressure term becoming negligible when the atmosphere is nearly adiabatic. This allows for the thermodynamics to be written in terms of entropy alone, leading to further computational savings. Even though the anelastic approximation relies on the atmosphere being nearly adiabatic the Lantz-Braginsky simplification will only be equivalent to the anelastic approximation derived in Gilman & Glatzmaier (1981) in the limit of a perfectly adiabatic atmosphere (Berkoff et al., 2010).

**Assumptions used in the Anelastic Approximation**

The anelastic approximation is formulated using a number of assumptions. The dynamics are treated as the superposition of a reference state atmosphere and fluctuations about that reference state. In order for the approximation to be valid the reference state may only evolve at a slower rate than the convective time-scale. Moreover the fluctuations of the thermodynamic variables of temperature, pressure and density must be small compared with the reference state. The relative density, temperature, and pressure fluctuations are of the same order of magnitude, unlike in the Boussinesq approximation. Lantz & Fan (1999) suggest similarities between mixing length theory

and the anelastic approximation, in so much that both are derived from a reference state background and in both the reference state is adiabatic, which assumes only a weakly super-adiabatic atmosphere in the Sun. To produce the luminosity observed at the surface of the Sun only a weakly super-adiabatic solar atmosphere is required.

In the anelastic approximation the dynamical time-scale of motion is the inverse of the Brunt-Väisälä frequency. For this time-scale to capture the dynamics of the problem the Mach number of the flow must be small and the fastest wave motion relevant to the problem to be slow compared to the Brunt-Väisälä frequency; any dynamics which occur due to high frequency motions will not be included. For the system to be accurately modelled then the high frequency modes that are filtered out must not be physically important to the dynamics. If the reference atmosphere is not adiabatic then it is also difficult to ensure that the flows remain strongly subsonic, as argued in Lantz & Fan (1999). In a strongly convective atmosphere, even if buoyancy is inhibited by diffusion, the wave motion has a high frequency; conversely in a very stable atmosphere gravity waves again have high frequencies. In either the highly sub-adiabatic or super-adiabatic stratification the anelastic approximation would be inappropriate. The problem will be illustrated with the temperature equation in the anelastic approximation in §3.7.2.

When magnetic fields are included the Alfvén waves must evolve on the slow dynamical time-scale, and the magnetic field must be weak enough so as not to upset hydrostatic balance. Although it may be possible to formulate magnetic fields into leading order the Alfvénic frequency associated with large magnetic fields would not then be captured in the approximation.

In the remaining chapters I shall derive the anelastic approximation and I shall examine the range of applicability of the anelastic approximation in both stably stratified and convectively unstable atmospheres. The arguments in the formulation of the anelastic approximation were developed with convection in mind so the relevance to magnetic buoyancy in stably stratified atmospheres is less clear. In Rogers & Glatzmaier (2005) the anelastic approximation was used in a stable atmosphere but they experienced some

problems with an inward turbulent heat flux in stable regions. The limit of how far from stably stratified an atmosphere can be and be accurately modelled within the anelastic approximation is still a matter of debate.

# Chapter 3

# The Mathematics of the Anelastic Approximation: Formal Scale Analysis

## 3.1 Physical Balances

In the derivation of the anelastic approximation I chiefly follow the procedure described in Gough (1969) and in Lantz & Fan (1999). To make the anelastic approximation I decompose all variables into a reference state and fluctuations about the reference state, i.e.

$$\xi(x, y, z, t) = \overline{\xi}(z) + \xi^*(x, y, z, t).$$

The choice of what qualifies as a reference state is an issue, Ogura & Phillips (1962) used an isentropic state, Gough (1969) used a non-adiabatic stratification and Clune et al. (1999) used a reference state that is a slowly varying spherically averaged mean. Nordlund (1982) appears not to have used a reference state at all. I am using a Cartesian geometry and take gravity to point in the positive $\hat{\mathbf{z}}$-direction so that $z$ increases with depth, as in the Boussinesq discussion. Using Cartesian coordinates, when the Sun itself is an oblate sphere, is valid only if the curvature force can be neglected, which imposes some limits on layer width and aspect ratio of the domain.

From the first law of thermodynamics

$$T\,\mathrm{d}s = c_p\,\mathrm{d}T - \frac{\delta_p}{\rho}\,\mathrm{d}p, \tag{3.1}$$

where the thermal expansion coefficient is $\delta_p = -(\partial \ln \rho/\partial \ln T)_p$ which is unity for an ideal gas. One can see that the adiabatic temperature gradient is

$$\left.\frac{\mathrm{d}T}{\mathrm{d}z}\right|_{\text{adb}} = \frac{\delta_p}{\rho c_p}\frac{\mathrm{d}p}{\mathrm{d}z}.$$

For now I will also use the entropy formulation of the energy equation

$$\rho T\left[\frac{\partial s}{\partial t} + (\mathbf{u}\cdot\nabla)\,s\right] = \nabla\cdot k\nabla T + \mu\frac{\partial u_i}{\partial x_j}\tau_{ij} + \frac{\eta}{\mu_0}\left(\nabla\times\mathbf{B}\right)^2, \tag{3.2}$$

$$\text{with } s = c_v \ln\left(p\rho^{-\gamma}\right) \tag{3.3}$$

in this derivation, instead of the temperature formulation.

Assuming the gravity length scale is much larger than the layer depth, then hydrostatic balance gives

$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} = \overline{\rho}g.$$

It is now time to introduce a small parameter $\epsilon$. The standard method is for the small parameter to measures the departure from adiabaticity of the reference state, i.e.

$$\epsilon = \frac{d}{H}\left|\frac{\mathrm{d}\ln\overline{T}}{\mathrm{d}\ln\overline{p}} - \left.\frac{\mathrm{d}\ln\overline{T}}{\mathrm{d}\ln\overline{p}}\right|_{\text{adb}}\right| = \frac{d}{\overline{T}}\left|\left.\frac{\mathrm{d}\overline{T}}{\mathrm{d}z}\right|_{\text{ref}} - \frac{g}{c_p}\right| = \frac{d}{c_p}\left|\left.\frac{\mathrm{d}\overline{s}}{\mathrm{d}z}\right|_{\text{ref}}\right. \tag{3.4}$$

where subscript $_{\text{ref}}$ means evaluated at a reference point in the layer, in this case taken to be the top ($z = 0$); $d$ is the layer depth; subscript $_{\text{adb}}$ means evaluated for adiabatic values; and $H$ is the pressure scale height defined as

$$H = \frac{\mathrm{d}z}{\mathrm{d}\ln\overline{p}} = \frac{\overline{p}}{g\overline{\rho}}. \tag{3.5}$$

As previously mentioned, I assume that the relative density, temperature and pressure fluctuations are of similar order

$$\frac{|\rho^*|}{\overline{\rho}} \approx \frac{|T^*|}{\overline{T}} \approx \frac{|p^*|}{\overline{p}} \approx |\epsilon|, \tag{3.6}$$

which maintains the assumption that fluctuation terms are of much smaller amplitude than in the reference state. Unlike in the Boussinesq approximation where the pressure is ignored, in the anelastic approximation it is retained. In a similar ansatz to the Boussinesq approximation, pressure fluctuations are balanced with the vertical flow (2.10), it is the vertical motion in the fluid that causes a change in pressure. Gough (1969) introduced an additional parameter to his anelastic expansion

$$\delta = \begin{cases} \dfrac{d}{H} & \text{if } d < H, \\ 1 & \text{if } d \geq H, \end{cases} \tag{3.7}$$

which will be used to relate the anelastic approximation to the Boussinesq approximation. To leading order (i.e. neglecting viscous forces) the characteristic parcel speed can be estimated by balancing the gravitational energy stored due to the buoyancy forces acting over the pressure scale height with the kinetic energy

$$\overline{\rho}|\mathbf{u}|^2 \approx -g\delta H\rho^*. \tag{3.8}$$

The buoyancy force approximation is done over $\delta H$ and not simply the layer depth as when the layer depth is much greater than the pressure scale height then pressure forces overwhelm the buoyancy forces. If, on the other hand, the pressure scale height is larger than the layer depth then (3.8) will be true except that now it is the layer depth which is a typical length scale. Typical velocities in a layer are driven by the buoyancy forces over a length-scale of: the layer if $d < H$ or the pressure scale height if $d > H$. The cumulative effect of the pressure fluctuations is to reduce vertical motion. Therefore it is assumed that the relative pressure fluctuations are of the same order as the relative density and temperature fluctuations, see (3.6). In the Boussinesq case the layer depth $d$ is much smaller that the pressure scale height $H$. If the layer is much larger than $H$ then pressure fluctuations will be very efficient at inhibiting the vertical flow.

It is useful to relate (3.8) to the Mach speed $M = u/c_s$ where the sound speed is

$$c_s^2 = \left(\frac{\partial p}{\partial \rho}\right)_s = \gamma\frac{p}{\rho} = \gamma g H.$$

so the Mach number can be expressed as

$$M^2 \approx \frac{\rho^*}{\overline{\rho}\gamma} \approx \frac{\epsilon}{\gamma}. \tag{3.9}$$

The difference between the actual temperature gradient and the adiabatic temperature gradient is the super-adiabatic (or sub-adiabatic if the atmosphere is stable) temperature gradient $\beta$ defined to be

$$c_p\beta = -\overline{T}\frac{\mathrm{d}\overline{s}}{\mathrm{d}z}. \tag{3.10}$$

When a parcel rises due to convection it will transfer heat to its environment. The amount of heat it transports will be proportional to the super-adiabatic temperature gradient across the layer depth $d$

$$\hat{\theta} = \int_0^d |\beta| \,\mathrm{d}z \approx d\left|\frac{\delta_p}{c_p\,\overline{\rho}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} - \frac{\mathrm{d}\overline{T}}{\mathrm{d}z}\right| = d\left|\frac{\overline{T}}{c_p}\frac{\mathrm{d}s}{\mathrm{d}z}\right|.$$

The convective heat flux will then be

$$F_{\mathrm{cv}} \approx \overline{\rho}c_p\hat{\theta}w. \tag{3.11}$$

There are other ways the small parameter $\epsilon$ can be defined. Gough (1969) argues that the heat flux will be the convective heat flux minus temperature fluctuations. He uses this, combined with the equation (3.8), to build a small parameter. I will explore this more in §3.7.1.

## 3.2 Energy Diffusion

I will make a quick digression into entropy diffusion models with the aid of Jones et al. (2009). In many astrophysical bodies, such as in stars, small-scale turbulence will lead to a diffusion of entropy which will normally be much larger than the molecular conductivity. The diffusion parameters are artificially high in most astrophysical fluid simulations to stop the flow developing a structure too fine to be resolved, given the

limited ability to model many spatial scales in one simulation. Prandtl's mixing-length theory states that turbulent elements travel over a mixing-length before releasing their entropy and joining the ambient background state. This suggests that the turbulent entropy flux is proportional to the entropy gradient, not the temperature gradient. In Boussinesq convection the turbulent thermal diffusion is often modelled as similar to the molecular thermal diffusion, but with a much larger diffusivity, in a compressible flow this is not the case. The sub-grid model proposed by Gilman & Glatzmaier (1981) assumes the diffusive flux is due to small-scale eddies in the superadiabatic convection zone and so proportional to $\nabla T - \nabla T_{\text{adb}}$ , i.e. proportional to potential temperature rather than actual temperature. The essential assumption is that there is a turbulent velocity $u_T$ which gives rise to a turbulent entropy fluctuation $s_T$, where these have averages $\langle \xi_T \rangle$ of zero on a short length-scale, but where $\langle \rho u_T s_T \rangle$ has a non-zero result. The turbulent entropy diffusion is then assumed proportional to $\nabla s$. This turbulent entropy diffusion creates a source term in the entropy formulation of the energy equation so that (3.3) becomes

$$\rho T \left[ \frac{\partial s}{\partial t} + (\mathbf{u} \cdot \nabla) s \right] = \nabla \cdot k \nabla T + \nabla \cdot T \frac{k_T}{c_p} \nabla s + \mu \frac{\partial u_i}{\partial x_j} \tau_{ij} + \frac{\eta}{\mu_0} (\nabla \times \mathbf{B})^2 , \quad (3.12)$$

where the new term contains $k_T$, the turbulent thermal conductivity. Equation (3.12) is taken from Braginsky & Roberts (1995).

The source of entropy is chosen so that it is not a source of energy, which can be seen as it appears as a divergence. Defining $\lambda_T = k_T/k$ and taking $\lambda_T \gg 1$ then the turbulent entropy diffusion alone has been used in numerous papers (see e.g. Glatzmaier, 1984; Braginsky & Roberts, 1995; Clune et al., 1999; Lantz & Fan, 1999), as both assumptions have been used in previous works for now I will keep both terms in this work.

The turbulent diffusive flux is from unresolved eddies. Glatzmaier (1984) pointed out that having turbulent diffusive energy flux is preferable in cases where the anelastic approximation is modelling the base of the convection zone. Here the superadiabatic temperature gradient, $(\nabla T - \nabla T_{\text{adb}})$ jumps from a small positive value in the convection zone to a large negative value below. As turbulent diffusivities are used then when

turbulent convection penetrates into the subadiabatic region below the convection zone then the turbulent eddies will cause a smoother transition. It can be argued that having a diffusion based on the atmosphere being turbulent, when in computer models the atmosphere is often initially quiescent, is somewhat dubious. The atmosphere cannot diffuse entropy via turbulent eddies when it is quiescent. It may be safer to start with molecular temperature diffusion at high resolution, show that an instability exists and then repeat the calculation with turbulent entropy diffusion predicted from the molecular diffusion results but this would be outside the scope of this work. The models used in this work only account for isotropic diffusion and so anisotropic diffusion is outside the scope of this thesis also. In this work the coefficients of diffusion $k$ and $k_T$ are constant but this is an oversimplification, Glatzmaier (1984) cited private communication with Gilman for suggesting that $k_T$ should decrease with depth. The argument is that $k_T$ is a model for sub-scale diffusion by unresolved eddies. Near the top of convection zone the pressure scale height is very small suggesting that eddies will also be small. Lower down the pressure scale height increases so more of the turbulent eddies will be resolved, so the sub-grid is modelling less unresolved motion and the diffusion should decrease.

There is another reason that the diffusion of entropy has been preferred in the literature. In the appendix of Lantz & Fan (1999) is a discussion as to why entropy diffusion is preferred as when temperature diffusion is used then the reference state temperature is overdetermined. This can be seen from the energy equation at leading order,

$$\nabla \cdot k \nabla \overline{T} = 0, \tag{3.13}$$

and, from the first law of thermodynamics (3.1) applied to an adiabatic process in hydrostatic equilibrium, the other condition on $\overline{T}$ is

$$\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} = \frac{g_r}{c_{p,r}} \frac{d}{T_r}. \tag{3.14}$$

Parts of this thesis are looking at when the reference atmosphere departs from being perfectly adiabatic so although the equation (3.13) will always hold, the same cannot be said about equation (3.14). The overdetermined system has one important solution, that of a polytrope.

It is possible to look at how the two different diffusivity terms are related using (3.12)

$$\nabla^* s^* - \frac{\nabla^* T^*}{\overline{T}} \propto \frac{1}{\overline{p}} \left( \nabla^* p^* - \rho^* \hat{\mathbf{z}} \right), \tag{3.15}$$

which is interaction between the pressure and density fluctuations and is an important driving term for large scale convective motions.

## 3.3 Preliminary Scalings

I decompose the variables into a steady, stationary, non-magnetised reference state plus fluctuations, denoted by a superscript $^*$, which may or may not have a mean component. Guided by this we can write the preliminary scalings where subscript $_s$ denotes a scaling factor, which may depend on $\epsilon$,

$$\rho = \rho_s \left( \overline{\rho} + \epsilon \rho^* \right), \quad T = T_s \left( \overline{T} + \epsilon T^* \right), \quad p = p_s \left( \overline{p} + \epsilon p^* \right)$$

$$\mathbf{u} = u_s \mathbf{u}^*, \quad \mathbf{B} = B_s \mathbf{B}^*,$$

$$g = g_s, \quad c_p = c_{p,s}, \quad c_v = c_{v,s}, \quad \mu_0 = \mu_{0,s},$$

$$k = k_s, \quad k_T = k_{T,s}, \quad \eta = \eta_s, \quad \mu = \mu_s.$$

The aim of this is to develop the anelastic scalings $\xi_r$ which are independent of $\epsilon$ and $\delta$, where $\xi$ represents any variable, these scaling terms are therefore $\mathcal{O}(1)$. The $k_T$ term is a turbulent diffusion term which is explained in §3.2. The turbulent diffusivity term appears as

$$\nabla \cdot k_T T \nabla s$$

in the energy equation, see equation (3.12).

The reference state is nearly adiabatic so the reference and fluctuating entropy will enter in at the same order

$$s = s_s c_{p,s} + s_s c_{p,s} \left( \overline{s} + s^* \right),$$

where the $c_{p,s}$ is included to be dimensionally consistent. The small parameter $\epsilon$ is based on the departure of adiabaticity of the atmosphere and must be small. The first term on the right-hand-side, $s_s c_{p,s}$, does not enter into any equations so will be neglected. The time and velocity scale are given by (3.8)

$$u_s = \sqrt{\epsilon \delta H_r g_r}, \quad \frac{\partial}{\partial t} = \sqrt{\frac{\epsilon g_r}{H_r \delta}} \frac{\partial}{\partial t^*},$$

which shows the velocity is scaled using the Brunt-Väisälä speed.

If the reference state is allowed to evolve it must do so on a time scale much larger than the convective time-scale. Gough (1969) derived an equation for a time-varying reference state although it is more common to take a time-independent reference state. In the following derivation the reference state is taken to be time-independent so the time derivatives of the reference state are zero. There are two length scales, the layer depth $d$ and the pressure scale height $H$. The Boussinesq approximation assumes that $\delta = d/H \ll 1$ but in the anelastic case this restriction is not required.

The Boussinesq approximation can be obtained by taking $\epsilon,\ \delta \ll 1$. To show how the two approximations are related $\delta$ will be retained. The anelastic approximation can be derived by using the layer depth as the only length scale - ignoring the difference between layer depth and the pressure scale will include unnecessary terms when $\delta \ll 1$ but will not neglect important terms. In this derivation $d_s = H_r \delta$, the layer depth or the pressure scale height depending on whether the layer depth is smaller or larger than the pressure scale height.

Some terms must enter at leading order which makes thermodynamic scalings: $\rho_s = \rho_r$ and $T_s = T_r$. Any change in gravity, including self-gravity, is ignored so with hydrostatic equilibrium at leading order this suggests $g_s = g_r$. The pressure scaling must be consistent with (3.5) which suggests the pressure scale height $p_s = H_r g_r \rho_r$ but this scale is only relevant to the reference state pressure; for the fluctuating pressure the scale is over the layer depth and so $p = H_r g_r \rho_r \left(\overline{p} + \epsilon \delta p^*\right)$. The gas constant does not fluctuate so $c_{p,s} = c_{p,r}$ and $c_{v,s} = c_{v,r}$. The permittivity of free space $\mu_0$ is also treated as a constant that should not scale on $\epsilon$ or $\delta$ so $\mu_{0,s} = \mu_{0,r}$.

The vertical derivatives will have different dependencies on $\delta$ for the reference and fluctuating states. It is natural to use the pressure scale height (3.5) when considering derivatives of the reference state pressure, density or temperature. The vertical entropy derivative $(\mathrm{d}s/\mathrm{d}z)_s$ scales with a $\delta$ dependence from thermodynamic relationships or the definition of the small parameter (3.4). This leads to the scaling factors on the derivatives to be

$$\nabla = \begin{cases} 1/H_r \nabla^* & \text{operating on } \overline{p},\, \overline{\rho},\, \overline{T},\, \overline{s}, \\ \dfrac{1}{\delta H_r}\nabla^* & \text{otherwise.} \end{cases}$$

The $\nabla$ could be replaced by the vertical derivative of the reference state as in this work the reference state is a function of $z$ only.

The compressible equations are expanded using these preliminary expansions to determine how the diffusive terms scale.

Mass conservation becomes

$$\epsilon \frac{\partial \rho^*}{\partial t^*} = -\left[\overline{\rho}\nabla^* \cdot \mathbf{u}^* + \delta \mathbf{u}^* \cdot \nabla^* \overline{\rho}\right] - \epsilon \nabla^* \cdot \left(\rho^* \mathbf{u}^*\right),$$

where the Boussinesq limit of $\nabla \cdot \mathbf{u}$ can be recovered by taking $\epsilon, \delta \to 0$. The momentum equation (2.6b) with the preliminary scaling becomes

$$\begin{aligned} \epsilon \rho_r g_r \left(\overline{\rho} + \epsilon \rho^*\right) \left[\frac{\partial \mathbf{u}^*}{\partial t^*} + \left(\mathbf{u}^* \cdot \nabla^*\right)\mathbf{u}^*\right] &= -\rho_r g_r \nabla^* \left(\overline{p} + \epsilon p^*\right) \\ &+ \rho_r g_r \left(\overline{\rho} + \epsilon \rho^*\right)\hat{\mathbf{z}} + \frac{B_s^2}{H_r \delta}\frac{\left(\nabla^* \times \mathbf{B}^*\right) \times \mathbf{B}^*}{\mu_{0,r}} + \left[\frac{\mu_s \left(\epsilon g_r \delta H_r\right)^{1/2}}{\delta^2 H_r^2}\right] \nabla^* \cdot \boldsymbol{\tau}^*. \end{aligned}$$

When the layer is shallow the $\delta^{-1}$ factor in front of the Lorentz force would unbalance the equation. The Lorentz force must also not upset hydrostatic balance but should be included at the next order suggesting the scaling

$$B_s = \sqrt{\epsilon\delta} B_r.$$

A similar argument for the viscous forces not to upset hydrostatic balance gives $\mu_s = \left(\epsilon\delta^3\right)^{1/2} \mu_r$.

Expanding the induction equation (2.6c), using the preliminary scalings, results in

$$\epsilon\frac{\partial \mathbf{B}^*}{\partial t^*} = \epsilon\nabla^* \times (\mathbf{u}^* \times \mathbf{B}^*) + \eta_s \left(\frac{\epsilon}{g_r \delta^3 H_r^3}\right)^{1/2} \nabla^{*2}\mathbf{B}^*, \tag{3.16}$$

and, if magnetic diffusion is to be included in this equation, then $\eta_s = (\epsilon\delta^3)^{1/2}\,\eta_r$ which is similar to the viscosity scale dependence on $\epsilon$ and $\delta$.

Finally the energy equation when expanded using the scalings developed is

$$\begin{aligned}\rho_r\left(\overline{\rho}+\epsilon\rho^*\right)T_r\left(\overline{T}+\epsilon T^*\right)\left(\frac{\epsilon g_r}{\delta H_r}\right)^{1/2} c_{p,r}s_s\left[\frac{\partial s^*}{\partial t^*}+\left(\mathbf{u}^*\cdot\nabla^*\right)\left(\delta\overline{s}+s^*\right)\right]\\ = \frac{k_s T_r}{H_r^2}\nabla^{*2}\left(\overline{T}+\frac{\epsilon}{\delta^2}T^*\right)+\frac{k_{T,s}T_r c_{p,r}}{H_r^2}\nabla^*\cdot\left(\overline{T}+\frac{\epsilon}{\delta}T^*\right)\nabla^*\epsilon\left(\overline{s}+\frac{s^*}{\delta}\right)\\ +\frac{\epsilon^{3/2}\eta_r\delta^{1/2}B_r^2}{\mu_{0,r}H_r^2}\left(\nabla^*\times\mathbf{B}^*\right)^2+\frac{\mu_r}{2}\frac{\epsilon^{3/2}\delta^{1/2}g_r}{H_r}\boldsymbol{\tau}^{*2}.\end{aligned}$$

For the $\epsilon\delta^{-2}\nabla^{*2}T^*$ and $\epsilon\nabla^*\cdot\overline{T}\nabla^*\left(\overline{s}+\delta^{-1}s^*\right)$ terms to effect the evolution of entropy in a shallow layer then $k_s = (\epsilon\delta^3)^{1/2}\,k_r$ and $k_{T,s} = (\epsilon\delta^3)^{1/2}\,k_{T,r}$, which is the same dependence as the other diffusivities. The dependence on $\delta$ may seem arbitrary but this means that entropy diffuses in the low $\delta$ limit and that both $k$ and $k_T$ have the same dependence on $\epsilon$ and $\delta$. For the left-hand-side to be in balance $s_s = \epsilon$; there is no need for a $s_r$ scaling factor as that can be incorporated into $c_{p,r}$. The $s_s$ dependence on $\epsilon$ is expected from the definition of $\epsilon$ in equation (3.4) due to the reference state atmosphere being nearly adiabatic.

## 3.4 The Anelastic Scalings

From these preliminary scalings $\xi_s$ I now develop the anelastic equations which are independent of $\epsilon$, where $\xi$ represents any variable. I now use

$$\rho = \rho_r\left(\overline{\rho}+\epsilon\rho^*\right), \quad T = T_r\left(\overline{T}+\epsilon T^*\right), \quad p = H_r g_r \rho_r\left(\overline{p}+\epsilon\delta p^*\right)$$

$$s = \epsilon c_{p,r}\left(\overline{s}+s^*\right), \quad \mathbf{u} = \sqrt{\epsilon\delta H_r g_r}\,\mathbf{u}^*, \quad \frac{\partial}{\partial t} = \sqrt{\frac{\epsilon g_r}{H_r\delta}}\frac{\partial}{\partial t^*},$$

$$\mathbf{B} = \epsilon^{1/2} B_r \mathbf{B}^*, \quad g = g_r,$$

$$k = \left(\epsilon\delta^3\right)^{1/2} k_r, \quad k_T = \left(\epsilon\delta^3\right)^{1/2} k_{T,r}, \quad \eta = \left(\epsilon\delta^3\right)^{1/2} \eta_r,$$

$$\mu = \left(\epsilon\delta^3\right)^{1/2} \mu_r, \quad \nabla = \begin{cases} 1/H_r \nabla^*, & \text{operating on } \overline{p}, \overline{\rho}, \overline{T}, \text{ or } \overline{s} \\ \dfrac{1}{\delta H_r} \nabla^*, & \text{otherwise} \end{cases}$$

where again $\overline{\xi}$ denotes a stationary, non-magnetised reference state and $\xi^*$ denotes a fluctuating term which may or may not have a mean. The scalings $\xi_r$ are now independent of $\epsilon$ and $\delta$. Where relevant the scalings $\xi_r$ are defined at the top of the domain $z = 0$.

I will start from the equations (2.3a-c), together with the evolution equation (3.12) for the entropy $s = c_v \ln(p\rho^{-\gamma})$.

The equations (2.3a-c) expanded using the anelastic scalings yield

$$\epsilon\frac{\partial\rho^*}{\partial t^*}+\overline{\rho}\nabla^*\cdot\mathbf{u}^*+\delta\mathbf{u}^*\cdot\nabla^*\overline{\rho}+\epsilon\nabla^*\cdot\left(\rho^*\mathbf{u}^*\right)=0, \tag{3.17a}$$

$$\begin{aligned}\epsilon\rho_r g_r\left(\overline{\rho}+\epsilon\rho^*\right)\left[\frac{\partial\mathbf{u}^*}{\partial t^*}+\left(\mathbf{u}^*\cdot\nabla^*\right)\mathbf{u}^*\right] &= +\rho_r g_r\left(\overline{\rho}+\epsilon\rho^*\right)\hat{\mathbf{z}}\\ &\quad+\epsilon\frac{B_r^2}{H_r}\frac{\left(\nabla^*\times\mathbf{B}^*\right)\times\mathbf{B}^*}{\mu_{0,r}}\\ &\quad-\rho_r g_r\nabla^*\left(\overline{p}+\epsilon p^*\right)+\mu_r\epsilon\left(\frac{g_r}{H_r^3}\right)^{1/2}\nabla^*\cdot\boldsymbol{\tau}^*,\end{aligned} \tag{3.17b}$$

$$\frac{\partial\mathbf{B}^*}{\partial t^*}=\nabla^*\times\left(\mathbf{u}^*\times\mathbf{B}^*\right)+\eta_r\left(\frac{1}{g_r H_r^3}\right)^{1/2}\nabla^{*2}\mathbf{B}^*, \tag{3.17c}$$

$$\begin{aligned}\sqrt{\frac{g_r}{H_r}}\left(\overline{\rho}+\epsilon\rho^*\right)\left(\overline{T}+\epsilon T^*\right)\epsilon\left[\frac{\partial s^*}{\partial t^*}+\left(\mathbf{u}^*\cdot\nabla^*\right)\left(\delta\overline{s}+s^*\right)\right] &= \frac{\delta^2}{\rho_r c_{p,r}H_r^2}\left[k_r\nabla^{*2}\left(\overline{T}+\frac{\epsilon}{\delta^2}T^*\right)\right.\\ &\quad\left.+k_{T,r}\nabla^*\cdot\left(\overline{T}+\frac{\epsilon}{\delta}T^*\right)\nabla^*\epsilon\left(\overline{s}+\frac{s^*}{\delta}\right)\right]\\ &\quad+\frac{\epsilon\delta}{c_{p,r}}\left[\frac{\eta_r B_r^2}{\mu_r H_r^2}\left(\nabla^*\times\mathbf{B}^*\right)^2+\frac{\mu_r g_r}{2H_r}\boldsymbol{\tau}^{*2}\right].\end{aligned} \tag{3.17d}$$

With the introduction of the scaling factors the dimensionless numbers defined in (2.5), along with some related dimensionless numbers, are

$$\mathrm{Pr} = \frac{\mu_r c_{p,r}}{k_F}, \quad \zeta = \frac{\eta_r c_{p,r}\rho_r}{k_F}, \quad \tilde{\mathcal{F}} = \epsilon^{-1}\mathcal{F} = \frac{{B_r}^2}{g_r H_r \rho_r \mu_r}$$
$$\tilde{C}_k = \epsilon^{-1/2} C_k = \frac{k_F}{H_r c_{p,r}\rho_r \sqrt{(c_{p,r} - c_{v,r})T_r}}, \quad \text{and} \quad |\tilde{R}| = \frac{g_r c_{p,r} H_r^3 \rho_r^2}{k_F \mu_r}, \quad (3.18)$$

where

$$k_F = \begin{cases} k_{T,r} & \text{if } k_r < k_{T,r} \\ k_r & \text{if } k_r \geq k_{T,r} \end{cases}$$

From (3.17a) the low $\delta$ limit recovers the Boussinesq approximation at leading order. It is now possible to define equations for the reference state. As previously mentioned the reference state considered here is time independent so even if the fluctuations generate a mean then the reference state will not be updated. This should not present a problem as the assumptions about scaling factors and the atmosphere should prevent this. The equations to be satisfied by an anelastic reference state are

$$0 = \overline{\rho}\nabla^* \cdot \mathbf{u}^* + \delta \mathbf{u}^* \cdot \frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}, \quad (3.19a)$$
$$\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} = \overline{\rho}, \quad (3.19b)$$
$$0 = k_r \frac{\mathrm{d}^2\overline{T}}{\mathrm{d}z^2} + k_{T,r}\frac{\mathrm{d}}{\mathrm{d}z}\left(\overline{T}\epsilon\frac{\mathrm{d}\overline{s}}{\mathrm{d}z}\right). \quad (3.19c)$$

The energy equation reference state (3.19c) has two terms that are of a different order in $\epsilon$. This is not inconsistent as entropy diffusion will not be used in conjunction with temperature diffusion in this thesis for modelling purposes. For clarity the superscript $^*$ will be dropped on the fluctuating terms. Now that the reference state is clear the equations (3.17) can be non-dimensionalised so that fluctuations about the reference

state are given as

$$\overline{\rho}\nabla\cdot\mathbf{u}+\delta\mathbf{u}\cdot\nabla\overline{\rho}=0, \tag{3.20a}$$

$$\overline{\rho}\left[\frac{\partial\mathbf{u}}{\partial t}+(\mathbf{u}\cdot\nabla\mathbf{u})\right]=-\nabla p+\rho\hat{\mathbf{z}}+\tilde{\mathcal{F}}\left[(\nabla\times\mathbf{B})\times\mathbf{B}\right] \tag{3.20b}$$

$$+\left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}(\nabla\cdot\boldsymbol{\tau}), \tag{3.20c}$$

$$\frac{\partial\mathbf{B}}{\partial t}=\nabla\times(\mathbf{u}\times\mathbf{B})+\tilde{C}_k\zeta\nabla^2\mathbf{B} \tag{3.20d}$$

$$\text{with}\quad \nabla\cdot\mathbf{B}=0, \tag{3.20e}$$

$$\overline{\rho}\overline{T}\left[\frac{\partial s}{\partial t}+(\mathbf{u}\cdot\nabla)(\delta\overline{s}+s)\right]=\frac{1}{(\mathrm{Pr}\,\tilde{R})^{1/2}}\left(\nabla^2T+\nabla\cdot\overline{T}\nabla s\right)$$
$$+\frac{\gamma-1}{\gamma}\delta\tilde{C}_k^2\left(\mathrm{Pr}\tilde{R}\right)^{1/2}\tilde{\mathcal{F}}\zeta\left(\nabla\times\mathbf{B}\right)^2$$
$$+\frac{\gamma-1}{\gamma}\delta\tilde{C}_k^2\left(\mathrm{Pr}^3\tilde{R}\right)^{1/2}\frac{\partial u_i}{\partial x_j}\tau_{ij}, \tag{3.20f}$$

$$\text{with}\quad \frac{p}{\overline{p}}=\frac{\rho}{\overline{\rho}}+\frac{T}{\overline{T}}, \tag{3.20g}$$

$$\text{and}\quad s=\frac{1-\gamma}{\gamma}\frac{p}{\overline{p}}+\frac{T}{\overline{T}}. \tag{3.20h}$$

The equations (3.19) and (3.20) are the non-linear anelastic equations that describe a reference state and the fluctuations about that state.

### 3.4.1 A note on Non-dimensional Numbers

The dynamical time and the resultant non-dimensional numbers are chosen so they have any factor of $\epsilon$ explicitly shown in them. The explicit nature of $\epsilon$ can act as a warning as to when the values of the parameters will invalidate the approximation. For example, the leading order balance in the momentum equation is hydrostatic equilibrium, which is still true if $\epsilon$ is small and $\mathcal{F}$ large but when $\epsilon$ and $\mathcal{F}$ are both large, i.e. $\tilde{\mathcal{F}}$ is large, then the leading order balance may not be valid. This may make diagnosing when the parameters are liable to invalidate the anelastic assumptions easier to notice *a priori*.

That said, whether $\epsilon$ is implicit or otherwise it is not known, and it may not be possible to know, when the values of the parameters will invalidate the approximation.

### 3.4.2 Reference State

The reference state must satisfy (3.19). In this thesis the reference state is not considered to be time dependent and is a function of depth only. As mentioned Clune et al. (1999) considered a reference state that is updated. This can lead to problems unless the state is not allowed to depart from being almost adiabatic as although updating the reference state makes the fluctuating terms small it can mean that $\epsilon$, i.e. the departure from an adiabatic atmosphere, is large so invalidating neglecting higher order $\epsilon$ terms, such as $\partial\rho/\partial t$.

#### Dimensional Reference State

The reference state that will be used in this thesis is a polytrope. It is useful to go back to dimensionful equations so that the value of gravity $g$ can be followed explicitly. Assuming the atmosphere to be in hydrostatic equilibrium equation and an ideal gas then

$$\frac{\mathrm{d}p}{\mathrm{d}z} = g\rho \quad \text{and} \quad p = (c_p - c_v)\,\rho T.$$

A simple solution to this is to use the polytrope ansatz $p = K\rho^{1+1/m}$, where $m$ is the polytropic index and $K$ is an arbitrary constant. Putting this into the hydrostatic equilibrium results in

$$K\left(\frac{m+1}{m}\right)\rho^{1/m-1}\mathrm{d}\rho = g\mathrm{d}z$$

$$\rho = \left(\frac{g}{K\,(m+1)}z + c\right)^{m}.$$

From the polytrope ansatz the pressure must therefore be

$$p = K\left(\frac{g}{K\,(m+1)}z + c\right)^{m+1}.$$

I am assuming the plasma is an ideal gas so from the equation of state the temperature is

$$K\left(\frac{g}{K\left(m+1\right)}z+c\right)^{m+1}=\left(c_p-c_v\right)\left(\frac{g}{K\left(m+1\right)}z+c\right)^{m}T$$
$$\frac{1}{c_p-c_v}\left(\frac{g}{m+1}z+Kc\right)=T.$$

It is useful to define the thermal gradient at the bottom of the layer as $\mathrm{d}T/\mathrm{d}z\big|_{z=1}=\beta$ where $\beta$ is consistent with the definition given in equation (3.10).

**Dimensionless Reference State**

It now makes sense to use dimensionless numbers so

$$\frac{\mathrm{d}\left(\overline{T}+\epsilon T\right)}{\mathrm{d}z}=\theta \quad \text{at } z=0, \tag{3.21}$$

where $\theta = \beta H_r/T_r$ is the dimensionless thermal gradient. The introduction of new dimensionless numbers, $m$ and $\theta$, leads to the relation

$$\theta=\frac{H_r}{T_r}\frac{g_r}{\left(c_{p,r}-c_{v,r}\right)\left(m+1\right)}. \tag{3.22}$$

Expressing the reference state in dimensionless form results in

$$\overline{T}=\left(1+\theta z\right) \tag{3.23a}$$
$$\overline{\rho}=\left(1+\theta z\right)^{m} \tag{3.23b}$$
$$\overline{p}=\left(1+\theta z\right)^{m+1} \tag{3.23c}$$
$$\epsilon\overline{s}=\frac{\epsilon}{\theta}\ln\left(1+\theta z\right), \quad \text{where } \theta\neq 0 \tag{3.23d}$$

where in (3.23d) the definition of $\epsilon$ from (3.4) is used, along with the thermodynamic relation

$$\overline{s}=\frac{1}{\gamma}\ln\left(\overline{p}\,\overline{\rho}^{-\gamma}\right). \tag{3.24}$$

Some of the dimensionless numbers defined in (3.18) can be expressed in terms of the new dimensionless numbers including, in particular, the Rayleigh number, which can be written as

$$\tilde{R}=\epsilon R=\frac{\theta^2(m+1)^2}{\gamma C_k^2\mathrm{Pr}}\left(1-\frac{m\gamma}{m+1}\right), \tag{3.25}$$

using equation (3.4) together with the relation

$$C_k^2 \mathrm{Pr} R = (m+1)\theta, \tag{3.26}$$

which again shows that the non-dimensional parameters of the model are not all independent. Note that, for $\gamma = 5/3$, $\tilde{R}$ is positive if $m < 3/2$ (for example in convective instabilities) and negative if $m > 3/2$ (as it is in the magnetic buoyancy instability).

## 3.5 Layer Width

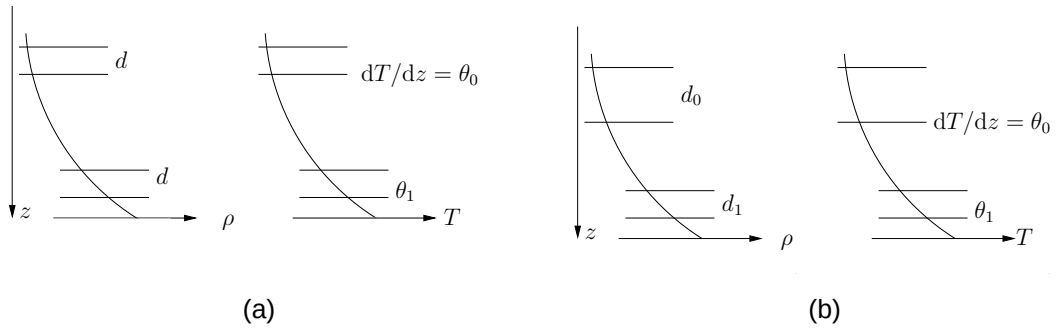



Figure 3.1: The effect of varying $\theta$, or the depth of the layer within the Sun, with a) constant depth and b) constant mass.

To be consistent in defining the basic state at different depths within the Sun an additional assumption is required about the layer width $d$. When modelling layers of plasma at different depths inside the Sun nothing has been said about how the layer width should be chosen. To model layers at different depths inside the Sun $\theta$ is altered which also alters the total mass within the layer. In figure 3.1 $\theta_0$ is a value corresponding to a temperature flux near the surface and $\theta_1$ corresponds to a temperature flux deeper in the Sun so $\theta_1 > \theta_0$. There are many ways of imposing an additional condition on $d$ and I will discuss two with strong physical motivations. One additional condition is to set $\rho(0) = 1$, which corresponds to a layer of constant width as in figure 3.1 (a). In this case to model a layer close to the surface of the Sun a small value of $\theta$ is chosen, and this also sets the layer width. It is also equally valid to have the additional condition of

$\int_0^1 \rho.\mathrm{d}z = 1$ which corresponds to fixing the total mass, as in figure 3.1 (b), where the layer width changes when $\theta$ changes. I choose to keep the width of the layer constant to compare with results from other works. The algorithm developed in Chapter 4 is capable of both constant width and constant mass.

### 3.5.1 Energetics

It can be shown that the anelastic equations are energetically consistent. This means that the total energy balance (the kinetic, internal plus potential) can be written in a closed form so that it does not depend on terms of higher order than the system of equations in (3.20). Gilman & Glatzmaier (1981) produced an equation similar to that derived by Gough (1969). Wilhelmson & Ogura (1972) points out that Gough's energetics calculations cannot be written in closed form without another further approximations as Gough's reference state was not necessarily isentropic, whereas Gilman's was. I will not derive the total energy balance but it is worth noting that there are possible inconsistencies if the reference state is allowed to depart from being adiabatic.

## 3.6 Different Formalisms

In the literature there have been many ways to derive the anelastic equations. This can involve different small parameters about which the asymptotic expansion is made. There is also a difference depending on whether the temperature or entropy formulation of the energy equation is used. First I will describe the differences in the small parameters used. As mentioned in the previous section most asymptotic expansions take the small parameter to be the departure from adiabaticity. Lantz & Fan (1999) argue that if the reference state atmosphere departs strongly from being adiabatic then it is not clear *a priori* that the velocities will remain small enough for the approximation to be valid. This seems reasonable in a convective atmosphere and also for magnetic buoyancy in

a strongly sub-adiabatic atmosphere. If the atmosphere is strongly sub-adiabatic then a slow convective instability can be expected but the gravity-wave speed would be very high which may negate the slow convective instability growth.

Gough (1969) uses a different small parameter. He assumes the total heat flux $F_{\mathrm{t}}$ is the the convective heat flux minus fluctuations. These fluctuations will seek to minimise any temperature gradients laterally. The actual heat flux will then be the convective flux minus diffusion

$$\begin{aligned} \overline{\rho} c_p w T &\approx \overline{\rho} c_p w \theta - k \left| \nabla T \right| \\ c_p \overline{\rho} w \left( \theta - T \right) &\approx \frac{kT}{\delta H}. \end{aligned} \tag{3.27}$$

Gough (1969) defines the small parameter $\epsilon$ from (3.27) and (3.8) as

$$\begin{aligned} \frac{\epsilon k_r T_r}{\delta H_r} &= \rho_r c_{p,r} \left( g_r \delta H_r \epsilon \right)^{1/2} \left( \theta - \epsilon T_r \right) \\ \implies \qquad \epsilon \left( \frac{k_r T_r}{\delta H_r} \right)^2 \frac{1}{g \delta H_r} \left( \frac{1}{\rho_r c_{p\,r}} \right)^2 &= \left( \theta - \epsilon T_r \right)^2 \\ \implies \qquad T_r^2 \epsilon^2 - \left( 2 \theta T_r + \frac{T_r \theta}{S} \right) \epsilon + \theta^2 &= 0 \\ \implies \qquad \epsilon &= \frac{\theta}{4 S T_r} \left( \sqrt{4S+1} - 1 \right)^2 \qquad (3.28) \\ &= \frac{\theta}{T_r} \psi(S), \qquad (3.29) \end{aligned}$$

where

$$S = \frac{g \delta^3 H_r^3 \theta \rho_r^2 c_{pr}}{k^2 T_r} = \mathcal{S} \frac{\theta}{T_r},$$

and the negative root of the quadratic in $\epsilon$ was taken. This is plotted in figure 3.2 for two values of $\mathcal{S}$ and compared with the case where $\epsilon$ is the departure from an adiabatic atmosphere as in

$$\begin{aligned} \epsilon =& \frac{d}{c_p} \frac{\mathrm{d}\overline{s}}{\mathrm{d}z} \\ =& \frac{\theta}{T_r}. \end{aligned}$$

As the atmosphere departs from being adiabatic figure 3.2 shows that Gough's small parameter remains smaller than the standard small parameter but the expansion is only

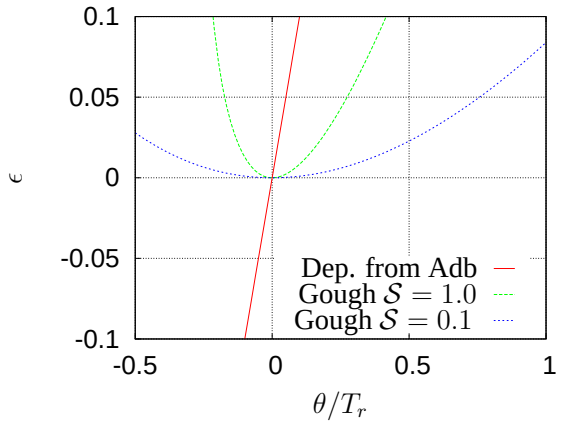



Figure 3.2: Comparison of different small parameters

formally valid as $\epsilon \rightarrow 0$. This does not fundamentally change equations but it may allow the anelastic approximation to be used in slightly less adiabatic atmospheres, which may be useful in the stably stratified limit. The only way to be sure that $\epsilon$ is small enough to give good agreement with the fully compressible results is to take an almost adiabatic atmosphere.

## 3.7 A Significant Simplification

Lantz (1992) and, independently, Braginsky & Roberts (1995) made a further simplification by writing the thermodynamic variables in the momentum equation in terms of entropy and pressure, with the pressure term becoming negligible when the atmosphere is nearly adiabatic.

The simplification essentially relies on the reference state atmosphere being nearly adiabatic whereas the anelastic approximation required the atmosphere to be nearly adiabatic, it is still a matter of debate whether these are equivalent or not. To show clearly

how this simplification can be done I will start with a thermodynamic relationship

$$\begin{aligned}\rho =& \left.\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}\overline{p}}\right|_s p + \left.\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}\overline{s}}\right|_p s,\\ \text{so}\qquad \frac{\rho}{\overline{\rho}} =& \frac{1}{\overline{\rho}}\left[\left(\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right)_s\left(\frac{\mathrm{d}z}{\mathrm{d}\overline{p}}\right)_s p + \left(\frac{\partial\overline{\rho}}{\partial\overline{T}}\right)_p\left(\frac{\partial\overline{T}}{\partial\overline{s}}\right)_p s\right]\\ =& \frac{p}{\overline{\rho}^2}\left(\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right)_s - s, \end{aligned} \tag{3.30}$$

where some of the partial differentials on the right-hand-side are unity for an ideal gas. I also require the definition of entropy (3.20h)

$$\frac{\rho}{\overline{\rho}} = \frac{p}{\gamma\overline{p}} - s. \tag{3.31}$$

Equating the equation (3.30) with (3.31) gives

$$\begin{aligned}\frac{p}{\overline{\rho}^2}\left(\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right)_s &= \frac{p}{\gamma\overline{p}}\\ \frac{\overline{p}}{\overline{\rho}^2}\left(\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right)_s &= \frac{1}{\gamma}. \end{aligned} \tag{3.32}$$

Equation (3.20b) divided by $\overline{\rho}$ contains the terms

$$-\frac{\nabla p}{\overline{\rho}} + \frac{\rho}{\overline{\rho}}\hat{\mathbf{z}}, \tag{3.33}$$

which involve $p$ and $\rho$. It would be far easier if terms which contain $p$ could be eliminated. Re-writing the equation (3.33) using equation (3.31) leads to

$$\begin{aligned}-\frac{\nabla p}{\overline{\rho}} + \frac{\rho}{\overline{\rho}}\hat{\mathbf{z}} =& -\frac{\nabla p}{\overline{\rho}} + \hat{\mathbf{z}}\left(\frac{p}{\gamma\overline{p}} - s\right)\\ =& -\nabla\left(\frac{p}{\overline{\rho}}\right) - s\hat{\mathbf{z}} + \frac{p\hat{\mathbf{z}}}{\overline{p}}\left(\frac{1}{\gamma} - \frac{\overline{p}}{\overline{\rho}^2}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right), \end{aligned} \tag{3.34}$$

where the final term in brackets is zero if the reference atmosphere is adiabatic, from equation (3.32), and small if the atmosphere is nearly adiabatic. If the final term in equation (3.34) is removed then this equation can be curled to remove the fluctuating pressure term. Another effect is that if entropy diffusion is used then removing the final term leaves entropy as the only remaining fluctuating thermodynamic variable in the system. This leads to fewer equations to solve and greatly simplifies the problem.

Braginsky noted that the resemblance of the equations to the Boussinesq equations is very strong but in this case the atmosphere is allowed to have strong density contrasts without violating any assumptions. The buoyancy force due to departure from the adiabatic reference atmosphere is now captured by entropy fluctuating term so that the Lantz-Braginsky momentum equation is

$$\left[\frac{\partial \mathbf{u}}{\partial t} + (\mathbf{u}\cdot\nabla\mathbf{u})\right] = -\nabla\left(\frac{p}{\overline{\rho}}\right) - s\hat{\mathbf{z}} + \frac{\tilde{\mathcal{F}}}{\overline{\rho}}\left[(\nabla\times\mathbf{B})\times\mathbf{B}\right] + \left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\left(\nabla\cdot\boldsymbol{\tau}\right). \quad (3.35)$$

Formally, the Lantz-Braginsky simplification is equivalent to the equations (3.20) when the atmosphere is perfectly adiabatic. As the atmosphere departs from being adiabatic both the anelastic approximation with and without the Lantz-Braginsky simplification will deviate from the fully compressible results; importantly both will do so in a different manner, as terms that are being neglected in one and not the other will become larger, but the difference will be of order $\epsilon^2$. There are many other higher order $\epsilon$ terms that are being neglected, namely the $\partial\rho/\partial t$ in the continuity equation, so including one and not including others is not increasing the accuracy of the equations. Likewise if one set of equations appears to be including more higher order $\epsilon$ terms than another it is not more accurate, unless it includes all the higher order $\epsilon$ terms as these may cancel each other out rather than increasing the accuracy by a small amount for each higher order term included.

In §3.2 I discussed the entropy diffusion, when this is combined with the Lantz-Braginsky simplification then the non-linear computation becomes far simpler as it removes the need for the pressure to be calculated. To calculate the pressure involves solving an elliptic equation with boundary conditions that are based on the previous time-step. Solving elliptic equations in a parallel environment is challenging as this is a global problem so all the processors need to combine solutions in order to solve global problems which can create a bottleneck. In massively parallel calculations the performance of the code is greatly reduced for global problems.

### 3.7.1 Gough's Energy Equation

Although in the temperature formulation the energy equations appear the same, there is a difference in the entropy formulation. As mentioned the only differences between Gough's equations and the more standard ones is in the energy equation. I will take Gilman & Glatzmaier (1981) as the comparison. To aid comparison I will ignore all the heating terms so the magnetic field is ignored. First I will rewrite Gilman & Glatzmaier (1981, equation (53))

$$\overline{\rho}\overline{\theta}\left[\frac{\partial s}{\partial t}+\mathbf{v}\cdot\nabla s\right]-w\overline{\rho}\Gamma=2FP\overline{\rho}\,\overline{\nu}\left(e_{ij}e_{ij}-\frac{\Delta^2}{3}\right)+\nabla\cdot(\overline{\rho}\,\overline{\kappa}\nabla\theta) \tag{3.36}$$

where in Gilman and Glatzmaier's notation $e_{ij}$ is the symmetric rate-of-strain tensor equal to $\partial u_i/\partial x_j$; $\theta$ is temperature equal to $T$; $\mathbf{v}$ the velocity equal to $\mathbf{u}$; $\Gamma$ is the superadiabatic temperature-gradient profile equal to $\beta$; $F=\kappa_0^2/(\epsilon c_p T_0 d^2)$ is the Froude number; $P$ is the Prandtl number $\mathrm{Pr}$; $\nu$ is the kinematic viscosity $\mu/\rho$; $\Delta$ is the divergence of the velocity field $\nabla\cdot\mathbf{u}$.

If the atmosphere is nearly adiabatic then $\beta=c_p\partial\overline{T}/\partial z$ so that, expressed in the notation used in the rest of this thesis, Gilman and Glatzmaier's equation is

$$\overline{\rho}\overline{T}\left(\frac{\partial s}{\partial t}+\mathbf{u}\cdot\nabla s\right)+w\overline{\rho}\left(c_p\frac{\partial\overline{T}}{\partial z}+g\right)=\text{R.H.S.} \tag{3.37}$$

where R.H.S. refers to equation (3.36). Equation (3.37) is similar to equation (3.20f). I will transform equation (3.37) into a form similar to the conservation of energy used in Gough (1969) by substituting (3.20h) into (3.37) and, along with the ideal gas law, this gives

$$\bar{\rho}\bar{T}\left[\frac{c_p}{\overline{T}}\frac{\partial T}{\partial t}-\frac{c_p-c_v}{\overline{p}}\frac{\partial p}{\partial t}+c_p\frac{\overline{T}\mathbf{u}\cdot\nabla T-T\mathbf{u}\cdot\nabla\overline{T}}{\bar{T}^2}\right.$$
$$\left.-(c_p-c_v)\,\frac{\overline{p}\mathbf{u}\cdot\nabla p-p\mathbf{u}\cdot\nabla\overline{p}}{\overline{p}^2}\right]+w\bar{\rho}\left(c_p\frac{\partial\overline{T}}{\partial z}+g\right)=\text{R.H.S.}$$

This can be further simplified using the assumption that the basic state atmosphere is in

hydrostatic balance so that

$$c_p\overline{\rho}\frac{\partial T}{\partial t}-\frac{\partial p}{\partial t}+w\overline{\rho}\left[c_p\left(1-\frac{T}{\overline{T}}\right)\frac{\mathrm{d}\overline{T}}{\mathrm{d}z}-\left(1-\frac{p}{\overline{p}}\right)\frac{1}{\overline{\rho}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z}\right] + \mathbf{u}\cdot(\overline{\rho}c_p\nabla T-\nabla p)=\text{R.H.S.} \quad (3.38)$$

This is now in a format where a comparison can be made with Gough (1969, equation (4.17)), except that $\hat{\mathbf{z}}$ is defined in the opposite manner in Gough's paper than in this thesis so I have removed a minus sign from the term involving $g$. Correcting for the definition of $\hat{\mathbf{z}}$ Gough's energy equation is

$$\overline{\rho}c_p\frac{\partial T_1}{\partial t}-\delta\frac{\partial p_1}{\partial t}-c_p\beta m_3+m_k\left(\frac{\partial h_1}{\partial x_k}-\frac{1}{\overline{\rho}}\frac{\partial p_1}{\partial x_k}\right)+\frac{g\rho_1 m_3}{\overline{\rho}}= \tau_{ik}\frac{\partial}{\partial x_k}\left(\frac{m_i}{\overline{\rho}}\right)+\overline{Q}+Q_1-\frac{\partial}{\partial x_k}\left(\overline{F}_k+F_{1k}\right) \quad (3.39)$$

where in Gough's notation $m$ is the momentum $\overline{\rho}u$ and $m_i$ is one of the three components of $\mathbf{m}$ so that $m_3$ is the vertical momentum; $p_1$ is the fluctuation pressure $p$; $\rho_1$ is the fluctuation density $\rho$; $h_1$ is the fluctuation enthalpy; $\delta=-(\partial\ln\rho/\partial\ln T)_p$ which is unity for an ideal gas; $\beta$ is the superadiabatic temperature gradient defined in (3.10); $\tau_{ik}$ is equal to

$$\mu\left(\frac{\partial u_i}{\partial x_k}+\frac{\partial u_k}{\partial x_i}-\frac{2}{3}\frac{\partial u_j}{\partial x_j}\delta_{ik}\right),$$

and so equal to $\tau\mu$ from equation (2.3e); $Q$ is any internal heat source (e.g. viscous or Ohmic heating); $F$ is the combined heat flux from radiation and conduction with $\overline{F}_k$ being a component of the reference state heat flux and $F_{1k}$ being a component of the fluctuating heat flux. Converting to the notation used elsewhere results in

$$c_p\overline{\rho}\frac{\partial T}{\partial t}-\frac{\partial p}{\partial t}+w\overline{\rho}\left(c_p\frac{\partial\overline{T}}{\partial z}-\frac{1}{\overline{\rho}}\frac{\partial\overline{p}}{\partial z}\right)+\mathbf{u}\cdot(\overline{\rho}c_p\nabla T-\nabla p)+g\rho w=\text{R.H.S. Gough,} \quad (3.40)$$

where R.H.S. Gough refers to equation (3.39). The difference between (3.38) and (3.40) is connected to how the two expansions treat $\mathrm{d}\overline{s}/\mathrm{d}z$ in different ways. In Gilman & Glatzmaier (1981), before any scaling assumptions are made, they have

$$-\epsilon^{3/2}\left(\overline{\rho}+\epsilon\rho\right)\left(\overline{T}+\epsilon T\right)\left[\frac{\partial s}{\partial t}+\mathbf{u}\cdot\nabla s+w\frac{\mathrm{d}\overline{s}}{\mathrm{d}z}\right]=\text{R.H.S.} \quad (3.41)$$

as the atmosphere is nearly adiabatic then $\mathrm{d}\overline{s}/\mathrm{d}z$ is of order $\epsilon$. If the term $\mathrm{d}\overline{s}/\mathrm{d}z$ was of order 1 then the terms

$$w\left(\overline{T}\rho + T\overline{\rho}\right)\frac{\partial \overline{s}}{\partial z} = w\left(\overline{T}\rho + T\overline{\rho}\right)\left(\frac{c_p}{\overline{T}}\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} - \frac{c_p - c_v}{\overline{p}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z}\right)$$
$$= w\rho c_p\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} + w\overline{\rho}c_p\frac{T}{\overline{T}}\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} - w\frac{p}{\overline{p}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z},$$

should be included in (3.38). If they are included then the equation (3.38) becomes

$$c_p\overline{\rho}\frac{\partial T}{\partial t} - \frac{\partial p}{\partial t} + w\overline{\rho}\left(c_p\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} - \frac{1}{\overline{\rho}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z}\right) + \mathbf{u}\cdot(\overline{\rho}c_p\nabla T - \nabla p) + w\rho c_p\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} = \text{R.H.S.} \quad (3.42)$$

but it also is now inconsistent as $\overline{\rho}\overline{T}\mathrm{d}\overline{s}/\mathrm{d}z$ would be un-balanced. C. A. Jones (personal communication) showed that (3.38) and (3.40) are equivalent if the atmosphere is adiabatic. For the two equations to be equivalent then

$$\frac{p}{\overline{p}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} - \frac{\overline{\rho}c_pT}{\overline{T}}\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} = g\rho. \quad (3.43)$$

This can be shown by starting with the entropy formulation of the energy equation

$$\frac{\mathrm{d}\overline{s}}{\mathrm{d}z} = \frac{c_v - c_p}{c_p\overline{p}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} + \frac{1}{\overline{T}}\frac{\mathrm{d}\overline{T}}{\mathrm{d}z}, \quad (3.44)$$

but taking an adiabatic reference state $\mathrm{d}\overline{s}/\mathrm{d}z = 0$ so (3.43) can be expressed as

$$\left(\frac{p}{\overline{p}} - \frac{\overline{\rho}(c_v - c_p)T}{\overline{p}}\right)\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} = g\rho$$
$$\left(\frac{p}{\overline{p}} - \frac{T}{\overline{T}}\right)\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} = g\rho$$
$$\frac{\rho}{\overline{\rho}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z} = g\rho$$

which is simply stating the reference state is in hydrostatic balance. It is clear that the two formalisms of Gough's and Gilman's are equivalent in the adiabatic limit but it is not clear if one would perform better when the atmosphere departs from being adiabatically stratified.

### 3.7.2 The Temperature Formulation

The use of the energy or temperature formulation of the energy equation is equivalent in the fully compressible magnetohydrodynamic equations. In the compressible case this

can be done by combining the first law of thermodynamics (3.1) and the time derivative of the equation of state (2.3f),

$$\frac{\partial T}{\partial t} = \frac{T}{c_p}\frac{\partial s}{\partial t} + \frac{c_p - c_v}{c_p \rho} T \frac{\partial \rho}{\partial t}. \tag{3.45}$$

and then using conservation of mass (2.3a) the pressure time derivative can be written in terms of temperature. Also required is a combination of the conservation of mass in the fully compressible case (2.3a) and the gradient of (3.12) so that

$$\nabla s = c_v \frac{\nabla T}{T} + (c_v - c_p) \frac{\nabla \rho}{\rho}. \tag{3.46}$$

I can now use equations (3.45) and (3.46) to rewrite the L.H.S. of the energy equation

$$c_v \rho \left[ \frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T \right] + p \nabla \cdot \mathbf{u} = k \nabla^2 T + \mu \frac{\partial u_i}{\partial x_j} \tau_{ij} + \frac{\eta}{\mu_0} \left( \nabla \times \mathbf{B} \right)^2. \tag{3.47}$$

The energy equation can then be written in terms of entropy

$$\rho T \left[ \frac{\partial s}{\partial t} + (\mathbf{u} \cdot \nabla) s \right] = k \nabla^2 T + \mu \frac{\partial u_i}{\partial x_j} \tau_{ij} + \frac{\eta}{\mu_0} \left( \nabla \times \mathbf{B} \right)^2$$

where the diffusion can be temperature or energy diffusion even though only temperature is shown, see §3.2 for a discussion on diffusion terms.

In the anelastic case this is not possible as the conservation of mass (3.17a) has no time derivative. The anelastic scalings derived in §3.4 used an entropy formulation of the energy equation but there is an equivalent temperature formulation in the compressible case. To be able to compare and contrast with the entropy formulation an anelastic form of the temperature formulation of the energy equation will be derived.

Starting again with the fully compressible temperature formulation of the energy equation (2.3d) and using an asymptotic expansion similar to that used in the anelastic scaling previously except now dimensions are kept so that $T = \overline{T} + \epsilon T$, $\rho = \overline{\rho} + \epsilon \rho$, $p = \overline{p} + \epsilon p$, $\mathbf{B} = \epsilon^{1/2}\mathbf{B}$, and $\mathbf{u} = \epsilon^{1/2}\mathbf{u}$. The diffusion coefficients scale as $\epsilon^{1/2}$. The possibility of a thin layer is ignored for simplification. Keeping the dimensions in the

energy equation (3.47), but using this expansion, gives

$$c_v\epsilon^{3/2}\frac{\partial T}{\partial t} + c_v\epsilon^{1/2}\mathbf{u}\cdot\nabla\left(\overline{T}+\epsilon T\right) = \\ -\frac{\overline{p}+\epsilon p}{\overline{\rho}+\epsilon\rho}\epsilon^{1/2}\nabla\cdot\mathbf{u} + \epsilon^{1/2}\frac{k}{\overline{\rho}+\epsilon\rho}\nabla^2\left(\overline{T}+\epsilon T\right) \\ +\epsilon^{3/2}\frac{\mu}{\overline{\rho}+\epsilon\rho}\frac{\partial \mathbf{u}_i}{\partial \mathbf{x}_j}\tau_{ij} + \frac{\epsilon^{1/2}\eta}{\mu_0\left(\overline{\rho}+\epsilon\rho\right)}\left(\nabla\times\epsilon^{1/2}\mathbf{B}\right)^2. \quad (3.48)$$

It is also useful to note that

$$\frac{1}{\overline{\rho}+\epsilon\rho} = \left(\frac{1}{\overline{\rho}} - \frac{\epsilon\rho}{\overline{\rho}^2}\right).$$

In (3.48) most of the terms are similar to the terms in the entropy formalism except there are two extra terms at leading order $\epsilon^{1/2}$, namely

$$c_v\mathbf{u}\cdot\nabla\overline{T} + \frac{\overline{p}}{\overline{\rho}+\epsilon\rho}\nabla\cdot\mathbf{u} - \frac{k}{\overline{\rho}+\epsilon\rho}\nabla^2\overline{T}. \quad (3.49)$$

From the reference state $\nabla^2\overline{T}=0$ and, using $\overline{\rho}\nabla\cdot\mathbf{u} = -w\,\mathrm{d}\overline{\rho}/\mathrm{d}z$, it is possible to show the other two terms from (3.49) should enter the equations at higher order. Dropping the $\epsilon^{1/2}$ the terms in (3.49) can be rewritten as

$$\begin{aligned} c_v\mathbf{u}\cdot\nabla\overline{T} + \frac{\overline{p}}{\overline{\rho}+\epsilon\rho}\nabla\cdot\mathbf{u} &= c_v\overline{\rho}w\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} - \frac{\overline{p}w}{\overline{\rho}}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\left(\frac{1}{\overline{\rho}} - \frac{\epsilon\rho}{\overline{\rho}^2}\right) \\ &= wc_p\left(\frac{\mathrm{d}\overline{T}}{\mathrm{d}z} - \frac{1}{c_p\overline{\rho}}\frac{\mathrm{d}\overline{p}}{\mathrm{d}z}\right) + \left(c_p - c_v\right)\epsilon\frac{\overline{T}\rho w}{\overline{\rho}^2}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z} \\ &= wc_p\beta + \left(c_p - c_v\right)\epsilon\frac{\overline{T}\rho w}{\overline{\rho}^2}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}, \end{aligned}$$

where the final term on the right then combines with $p\nabla\cdot\mathbf{u}$ and can be simplified. To scale $\beta$ I look to the definition (3.2) which suggests $\beta = \overline{\beta}\theta/d$ and $\epsilon = \theta/T_r$ which is an equivalent definition to (3.4). At leading order $\nabla^2\overline{T}=0$ and at $\epsilon^{3/2}$

$$c_v\left(\frac{\partial T}{\partial t} + \mathbf{u}\cdot\nabla T\right) + wc_p\overline{\beta} = \left(c_p - c_v\right)\frac{Tw}{\overline{\rho}}\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z} + k\frac{1}{\overline{\rho}}\nabla^2 T \\ +\frac{\mu}{\overline{\rho}}\frac{\partial \mathbf{u}_i}{\partial \mathbf{x}_j}\tau_{ij} + \frac{\eta}{\mu_0\overline{\rho}}\left(\nabla\times\mathbf{B}\right)^2, \quad (3.50)$$

which is similar to Gough's anelastic energy equation. In the literature it is less common to use the temperature formalism with the anelastic approximation but Rogers

& Glatzmaier (2005) take this approach. In Rogers and Glatzmaier's paper equation (3), the energy equation, is

$$\frac{\partial T}{\partial t} + (\mathbf{v}\cdot\nabla)T = -v_z\left(\frac{\partial \overline{T}}{\partial z} - (\gamma-1)\right)\overline{T}h_\rho v_z + \gamma\overline{\kappa}\left[\nabla^2 T + (h_\rho + h_\kappa)\frac{\partial T}{\partial z}\right] + \gamma\overline{\kappa}\left[\nabla^2\overline{T} + (h_\rho + h_\kappa)\frac{\partial \overline{T}}{\partial z}\right] + \frac{\bar{Q}}{c_v}, \quad (3.51)$$

where in Rogers and Glatzmaier's notation $\mathbf{v}$ in the velocity field equal to $\mathbf{u}$ and $v_z$ is the velocity in the vertical direction $w$; $h_\rho$ is the density scale hight $h_\rho = \mathrm{d}/\mathrm{d}z \ln\overline{\rho}$; $h_\kappa$ is the thermal diffusivity scale hight $h_\kappa = \mathrm{d}/\mathrm{d}z \ln\overline{\kappa}$; and $\overline{Q}$ is the heating rate which maintains the reference state profile.

When ignoring variations in thermal diffusivity, $h_\kappa = 0$, and removing the reference state, which is represented in the last two terms, then in the notation used in this thesis equation (3.51) becomes

$$c_v\overline{\rho}\left[\frac{\partial T}{\partial t} + (\mathbf{u}\cdot\nabla)T\right] + wc_p\overline{\rho}\overline{\beta} = wT\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}(c_p - c_v) + k\nabla^2 T'. \quad (3.52)$$

The two equations (3.50) and (3.52) are the same as Gough's energy equation. This suggests that the temperature formulation of the energy equation is independent of the small parameter used but this is not fully clear as in Rogers & Glatzmaier (2005) the equations are not justified by a full asymptotic expansion.

However, it is clear that the anelastic conservation of energy written in terms of temperature (3.50) is not equivalent to when it is written in terms of entropy (3.20f). It is possible that if higher order terms were kept in the anelastic conservation of mass then there would be a way to express the density time derivative and so a conversion from an entropy formulation of the energy equation to a temperature formulation would be possible. The anelastic conservation of mass with higher-order-terms included is

$$\epsilon^{3/2}\frac{\partial \rho_1^*}{\partial t} = \epsilon^{1/2}\nabla\cdot\mathbf{u}_1^*\overline{\rho} + \epsilon^{3/2}\nabla\cdot\mathbf{u}_2^*\overline{\rho} + \epsilon^{3/2}\nabla\cdot\mathbf{u}_1^*\rho_1^*, \quad (3.53)$$

where the decomposition $\mathbf{u} = \sqrt{\epsilon\delta H_r g_r}\,(\mathbf{u}_1 + \epsilon\mathbf{u}_2)$ and $\rho = \rho_r\,(\overline{\rho} + \epsilon\rho^* + \epsilon^2\rho_2^*)$ and $\mathbf{u}_1^*$, $\rho_1^*$ are the first order fluctuations, the same as dealt with in §3.4 on the anelastic

scalings, and $\mathbf{u}_2^*$, $\rho_2^*$ are an order of $\epsilon$ smaller. Keeping higher-order-terms would stop the system from being closed. One important thing to note about this derivation is that when the atmosphere is not close to being adiabatic then the terms with $\beta$ will not be small and there is no reason that they should be excluded from the leading order balance. This would lead to a reference state that depended on the fluctuating velocity and so invalidate the approximation. In a state far from adiabatic then this would clearly make the anelastic equations not energetically consistent. Put in a different way, when the atmosphere is not close to being adiabatic then $\theta/T_r$ is not small and the $\beta$ term in the temperature formulation of the energy equation (3.50) will act as a spurious source of energy, as noted by Durran (1989).

# Chapter 4

# Linear Algorithm

Insight into how the anelastic approximation deals both with the magnetic buoyancy and magnetoconvection instabilities can be gained by investigating linear problems. It is possible to look at small perturbations to a stable basic state of the anelastic equations (3.20) and then linearise these equations. Linearising means finding an equilibrium, perturbing the variables, and then, as the perturbations are small, neglecting non-linear perturbation terms. This requires a reference state, which is described in §3.4.2, and will also require a basic state about which to perturb the equations. The reference state is the state about which the fluctuating anelastic equations were derived. The basic state is equivalent to an initial condition if the system were solved as an initial value problem.

I have studied the same problem for anelastic and for compressible cases with the aim of comparing and contrasting the results. This was done for both the magnetic buoyancy and magnetoconvection instabilities. I have then attempted to find the range of parameters for which the anelastic approximation works and where it does not.

In all cases described henceforth, I follow a similar ansatz to determine the stability of a plane parallel layer. In the anelastic case the fluctuating variables are all expanded as

$$\xi^*(x, y, z, t) = \xi_b^*(z) + \hat{\xi}(z) \exp(\sigma t + \mathrm{i}k_x x + \mathrm{i}k_y y), \tag{4.1}$$

where $\xi_b^*(z)$ is the anelastic basic state; $\hat{\xi}(z)$ is a perturbation to the basic state with a $z$

dependence only; $\sigma$ is the (possibly complex) growth rate of the instability; and $k_x$ and $k_y$ are the wavenumbers in the $\hat{\mathbf{x}}$ and $\hat{\mathbf{y}}$ horizontal directions respectively. The perturbation is much smaller than the basic state (unless the basic state is identically zero) so $\hat{\xi} \ll \xi_b^*$. In the compressible case the full variables are expanded in a similar manner except the variables were not split into reference and fluctuating parts so

$$\xi(x, y, z, t) = \xi_b(z) + \hat{\xi}(z) \exp(\sigma t + \mathrm{i}k_x x + \mathrm{i}k_y y), \tag{4.2}$$

where $\xi_b(z)$ is the compressible basic state. The basic state was calculated numerically for the magnetic buoyancy problem in both the compressible and anelastic cases. The linear equations for the perturbation variables are given in Appendix A and B.

## 4.1 Finite Difference Scheme

The continuous problem was solved by describing the system of ordinary differential equations using $N$ evenly spaced grid points. The differential operators acting on the variables were replaced with the fourth-order accurate finite difference representations with appropriate representations at each boundary, given by the boundary conditions.

A finite difference scheme is a way of approximating derivatives in a domain divided into discrete points and is based on the Taylor expansion of a well behaved function

$$f(x_0 + \Delta x) = f(x_0) + \sum_{i=1}^{n} \frac{f^{(i)}(x_0)}{i!} (\Delta x)^i + R_n(x), \tag{4.3}$$

where $f$ is the function, $x_0$ the point about which the function is being expanded, $\Delta x$ is the distance between discrete points, and $R_N$ the remainder if the function is expanded to the $n^{\text{th}}$ derivative. It is then possible to write the $n^{\text{th}}$ derivative in terms of the function at different grid points with the accuracy of this given by the reminder function.

My domain had uniform grid width $\Delta x$ and my finite difference stencils were fourth-order accurate in space. The representation of an $n^{\text{th}}$ derivative of any variable will have

identical internal matrix points, e.g. the matrix representing the first derivative of $f$ is

$$\begin{pmatrix} \text{B.P} & \cdots & & & & & \\ \text{B.P} & \cdots & & & & & \\ \frac{1}{12(\Delta x)} & \frac{-2}{3(\Delta x)} & 0 & \frac{2}{3(\Delta x)} & \frac{-1}{12(\Delta x)} & 0 & 0 \\ & \ddots & \ddots & \ddots & \ddots & \ddots & \\ 0 & 0 & \frac{1}{12(\Delta x)} & \frac{-2}{3(\Delta x)} & 0 & \frac{2}{3(\Delta x)} & \frac{-1}{12(\Delta x)} \\ & & & & & \cdots & \text{B.P} \\ & & & & & \cdots & \text{B.P} \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ f_3 \\ \vdots \\ f_{N-3} \\ f_{N-2} \\ f_{N-1} \end{pmatrix} = \begin{pmatrix} Df_1 \\ Df_2 \\ Df_3 \\ \vdots \\ Df_{N-3} \\ Df_{N-2} \\ Df_{N-1} \end{pmatrix} \tag{4.4}$$

where B.P. are the boundary points which will be discussed later, $f_i$ is the value of the function at the $i^{\text{th}}$ grid point, $N-2$ is the total number of internal points, and $Df_i$ is the first derivative of the function at the $i^{\text{th}}$ grid point. The matrix term on the left-hand-side of equation (4.4) is an example of a fourth-order accurate matrix representation of the first derivative. The central difference stencils, which represent the internal points, for the function itself and the first four derivatives are

$$\begin{pmatrix} \mathrm{d}^0/\mathrm{d}x^0 \\ \mathrm{d}^1/\mathrm{d}x^1 \\ \mathrm{d}^2/\mathrm{d}x^2 \\ \mathrm{d}^3/\mathrm{d}x^3 \\ \mathrm{d}^4/\mathrm{d}x^4 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & \frac{1}{12(\Delta x)} & \frac{-2}{3(\Delta x)} & 0 & \frac{2}{3(\Delta x)} & \frac{-1}{12(\Delta x)} & 0 \\ 0 & \frac{-1}{12(\Delta x)^2} & \frac{4}{2(\Delta x)^2} & \frac{-5}{2(\Delta x)^2} & \frac{4}{3(\Delta x)^2} & \frac{-1}{12(\Delta x)^2} & 0 \\ \frac{1}{8(\Delta x)^3} & \frac{-1}{(\Delta x)^3} & \frac{13}{8(\Delta x)^3} & 0 & \frac{-13}{8(\Delta x)^3} & \frac{1}{(\Delta x)^3} & \frac{-1}{8(\Delta x)^3} \\ 0 & \frac{2}{(\Delta x)^4} & \frac{-19}{3(\Delta x)^4} & \frac{28}{3(\Delta x)^4} & \frac{-19}{3(\Delta x)^4} & \frac{2}{(\Delta x)^4} & 0 \end{pmatrix}, \tag{4.5}$$

where the third and fourth derivatives were only used for testing purposes.

To be able to model the third derivative requires seven grid points in the central difference scheme. This means there will be three rows at the top and bottom of the matrix that will depend on the boundary condition. In the boundary regions to keep the fourth-order accuracy then eight grid points are required. The boundary finite difference stencils are given for a free boundary condition, i.e. one where nothing is imposed on the function, where $l$ represents which row of the matrix the boundary stencil refers to, for example in equation (4.4) the first row would be $l = 1$ and $l < 0$ represents the boundary stencils

for the rows at the bottom of the matrix, with $l = N - 1$ being the bottom row.

$l = 1$

$$\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-1089}{420(\Delta x)} & \frac{7}{(\Delta x)} & \frac{-441}{42(\Delta x)} & \frac{35}{3(\Delta x)} & \frac{-35}{4(\Delta x)} & \frac{21}{5(\Delta x)} & \frac{-49}{42(\Delta x)} & \frac{3}{21(\Delta x)} \\ \frac{469}{90(\Delta x)^2} & \frac{-669}{30(\Delta x)^2} & \frac{7911}{180(\Delta x)^2} & \frac{-949}{18(\Delta x)^2} & \frac{369}{9(\Delta x)^2} & \frac{-603}{30(\Delta x)^2} & \frac{1019}{180(\Delta x)^2} & \frac{-21}{30(\Delta x)^2} \\ \frac{-967}{120(\Delta x)^3} & \frac{5104}{120(\Delta x)^3} & \frac{-11787}{120(\Delta x)^3} & \frac{15560}{120(\Delta x)^3} & \frac{-12725}{120(\Delta x)^3} & \frac{6432}{120(\Delta x)^3} & \frac{-1849}{120(\Delta x)^3} & \frac{232}{120(\Delta x)^3} \\ \frac{28}{3(\Delta x)^4} & \frac{-166}{3(\Delta x)^4} & \frac{426}{3(\Delta x)^4} & \frac{-609}{3(\Delta x)^4} & \frac{528}{3(\Delta x)^4} & \frac{-277}{3(\Delta x)^4} & \frac{82}{3(\Delta x)^4} & \frac{-10}{3(\Delta x)^4} \end{pmatrix} \tag{4.6a}$$

$l = 2$

$$\begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-60}{420(\Delta x)} & \frac{-609}{420(\Delta x)} & \frac{1260}{420(\Delta x)} & \frac{-1050}{420(\Delta x)} & \frac{70}{42(\Delta x)} & \frac{-315}{420(\Delta x)} & \frac{84}{420(\Delta x)} & \frac{-1}{42(\Delta x)} \\ \frac{126}{180(\Delta x)^2} & \frac{-70}{180(\Delta x)^2} & \frac{-486}{180(\Delta x)^2} & \frac{855}{180(\Delta x)^2} & \frac{-670}{180(\Delta x)^2} & \frac{324}{180(\Delta x)^2} & \frac{-90}{180(\Delta x)^2} & \frac{11}{180(\Delta x)^2} \\ \frac{-232}{120(\Delta x)^3} & \frac{889}{120(\Delta x)^3} & \frac{-1392}{120(\Delta x)^3} & \frac{1205}{120(\Delta x)^3} & \frac{-680}{120(\Delta x)^3} & \frac{267}{120(\Delta x)^3} & \frac{-64}{120(\Delta x)^3} & \frac{7}{120(\Delta x)^3} \\ \frac{10}{3(\Delta x)^4} & \frac{-56}{3(\Delta x)^4} & \frac{127}{3(\Delta x)^4} & \frac{-162}{3(\Delta x)^4} & \frac{125}{3(\Delta x)^4} & \frac{-60}{3(\Delta x)^4} & \frac{16}{3(\Delta x)^4} & \frac{-2}{3(\Delta x)^4} \end{pmatrix} \tag{4.6b}$$

$l = 3$

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{12(\Delta x)} & \frac{-2}{3(\Delta x)} & 0 & \frac{2}{3(\Delta x)} & \frac{-1}{12(\Delta x)} & 0 & 0 & 0 \\ \frac{-1}{12(\Delta x)^2} & \frac{16}{12(\Delta x)^2} & \frac{-30}{12(\Delta x)^2} & \frac{16}{12(\Delta x)^2} & \frac{-1}{12(\Delta x)^2} & 0 & 0 & 0 \\ \frac{-1}{8(\Delta x)^3} & \frac{-1}{(\Delta x)^3} & \frac{35}{8(\Delta x)^3} & \frac{-48}{8(\Delta x)^3} & \frac{29}{8(\Delta x)^3} & \frac{-1}{(\Delta x)^3} & \frac{1}{8(\Delta x)^3} & 0 \\ \frac{2}{3(\Delta x)^4} & \frac{-3}{(\Delta x)^4} & \frac{10}{3(\Delta x)^4} & \frac{-2}{3(\Delta x)^4} & \frac{-4}{3(\Delta x)^4} & \frac{1}{(\Delta x)^4} & 0 & 0 \end{pmatrix} \tag{4.6c}$$

$l = N-3$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{12(\Delta x)} & \frac{-2}{3(\Delta x)} & 0 & \frac{2}{3(\Delta x)} & \frac{-1}{12(\Delta x)} \\ 0 & 0 & 0 & \frac{-1}{12(\Delta x)^2} & \frac{4}{3(\Delta x)^2} & \frac{-5}{2(\Delta x)^2} & \frac{4}{3(\Delta x)^2} & \frac{-1}{12(\Delta x)^2} \\ 0 & \frac{-1}{8(\Delta x)^3} & \frac{1}{(\Delta x)^3} & \frac{-29}{8(\Delta x)^3} & \frac{6}{(\Delta x)^3} & \frac{-35}{8(\Delta x)^3} & \frac{1}{(\Delta x)^3} & \frac{1}{8(\Delta x)^3} \\ 0 & 0 & \frac{1}{(\Delta x)^4} & \frac{-4}{3(\Delta x)^4} & \frac{-2}{3(\Delta x)^4} & \frac{10}{3(\Delta x)^4} & \frac{-9}{3(\Delta x)^4} & \frac{2}{3(\Delta x)^4} \end{pmatrix} \tag{4.6d}$$

$l = N-2$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ \frac{1}{42(\Delta x)} & \frac{-21}{60(\Delta x)} & \frac{63}{84(\Delta x)} & \frac{-35}{21(\Delta x)} & \frac{105}{42(\Delta x)} & \frac{-3}{(\Delta x)} & \frac{609}{420(\Delta x)} & \frac{1}{7(\Delta x)} \\ \frac{11}{180(\Delta x)^2} & \frac{-1}{2(\Delta x)^2} & \frac{27}{15(\Delta x)^2} & \frac{-67}{18(\Delta x)^2} & \frac{171}{36(\Delta x)^2} & \frac{-27}{10(\Delta x)^2} & \frac{-7}{18(\Delta x)^2} & \frac{21}{30(\Delta x)^2} \\ \frac{-7}{120(\Delta x)^3} & \frac{8}{15(\Delta x)^3} & \frac{-267}{120(\Delta x)^3} & \frac{17}{3(\Delta x)^3} & \frac{-241}{24(\Delta x)^3} & \frac{174}{15(\Delta x)^3} & \frac{-889}{120(\Delta x)^3} & \frac{29}{15(\Delta x)^3} \\ \frac{-2}{3(\Delta x)^4} & \frac{16}{3(\Delta x)^4} & \frac{-20}{(\Delta x)^4} & \frac{125}{3(\Delta x)^4} & \frac{-54}{(\Delta x)^4} & \frac{127}{3(\Delta x)^4} & \frac{-56}{3(\Delta x)^4} & \frac{10}{3(\Delta x)^4} \end{pmatrix} \tag{4.6e}$$

$l = N-1$

$$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ \frac{-1}{420(\Delta x)} & \frac{4}{420(\Delta x)} & \frac{-1764}{420(\Delta x)} & \frac{3675}{420(\Delta x)} & \frac{-40}{420(\Delta x)} & \frac{44}{420(\Delta x)} & \frac{-29}{420(\Delta x)} & \frac{89}{420(\Delta x)} \\ \frac{126}{180(\Delta x)^2} & \frac{19}{180(\Delta x)^2} & \frac{-3618}{180(\Delta x)^2} & \frac{73}{180(\Delta x)^2} & \frac{-94}{180(\Delta x)^2} & \frac{7911}{180(\Delta x)^2} & \frac{-14}{180(\Delta x)^2} & \frac{938}{180(\Delta x)^2} \\ \frac{-232}{120(\Delta x)^3} & \frac{1849}{120(\Delta x)^3} & \frac{-6432}{120(\Delta x)^3} & \frac{12725}{120(\Delta x)^3} & \frac{-155}{120(\Delta x)^3} & \frac{11787}{120(\Delta x)^3} & \frac{-54}{120(\Delta x)^3} & \frac{967}{120(\Delta x)^3} \\ \frac{-1}{6(\Delta x)^4} & \frac{82}{3(\Delta x)^4} & \frac{-277.5}{3(\Delta x)^4} & \frac{528}{3(\Delta x)^4} & \frac{-9.5}{3(\Delta x)^4} & \frac{426}{3(\Delta x)^4} & \frac{-166.5}{3(\Delta x)^4} & \frac{28}{3(\Delta x)^4} \end{pmatrix} \tag{4.6f}$$

The other boundary condition such as Dirichlet and Neumann boundary conditions have different boundary stencils which are not shown here. The Dirichlet boundary condition stencil is similar to the free boundary condition stencils shown previously but where the value of the function at the point just outside the domain is zero.

Using the notation $\mathrm{D}n_l^{\mathrm{bc}}(m)$ where $n$ is the derivative; $l$ is the same $l$ as described previously; bc is the type of boundary i.e. f(ree), D(irichlet), or N(eumann); and $m$ is the column in the matrix. An example would be $\mathrm{D}4_2^{\mathrm{f}}(3) = 127/3$, from the stencil in equation (4.6b).

The Dirichlet condition can be derived from the free condition where $l < 4$ by

$$\mathrm{D}n_l^D(m) = Dn_{l+1}^f(m+1), \tag{4.7}$$

and when $l > N - 4$ by

$$\mathrm{D}n_l^D(m) = Dn_{l-1}^f(m-1). \tag{4.8}$$

For example with $l = 1$, the first row in the derivative matrix, the fourth derivative of a function with a Dirichlet boundary is

$$\mathrm{D}4_1^D = \begin{pmatrix} \frac{-56}{3(\Delta x)^4} & \frac{127}{3(\Delta x)^4} & \frac{-162}{3(\Delta x)^4} & \frac{125}{3(\Delta x)^4} & \frac{-60}{3(\Delta x)^4} & \frac{16}{3(\Delta x)^4} & \frac{-2}{3(\Delta x)^4} & 0 \end{pmatrix}. \tag{4.9}$$

Neumann boundary conditions are more involved but make use of the first row $l = 1$ from the first order derivative of the function

$$\mathrm{D}1_1^f = \begin{pmatrix} \frac{-1089}{420(\Delta x)} & \frac{7}{(\Delta x)} & \frac{-441}{42(\Delta x)} & \frac{35}{3(\Delta x)} & \frac{-35}{4(\Delta x)} & \frac{21}{5(\Delta x)} & \frac{-49}{42(\Delta x)} & \frac{3}{21(\Delta x)} \end{pmatrix}. \tag{4.10}$$

A Neumann condition means that when the stencil in equation (4.10) is applied to the function at the boundary grid point, which is just outside the domain, the result will be zero by definition. The Neumann boundary condition stencils for $l < 4$ are created by

$$\mathrm{D}n_l^{\mathrm{N}}(m) = -\frac{\mathrm{D}n_{l+1}^{\mathrm{f}}(1)}{\mathrm{D}1_1^{\mathrm{f}}(1)}\mathrm{D}1_1^{\mathrm{f}}(m+1) + \mathrm{D}n_{l+1}^{\mathrm{f}}(m+1),$$

with a similar expression for the $l > N - 4$ stencils.

## 4.2 Numerical Procedure

As can be seen from the previous section the result of the finite difference representation of the system is a banded matrix $A$, and the algorithm takes advantage of this. The $N$ eigenvalues $\sigma_N$ and corresponding eigenvectors $\boldsymbol{x}_N$ of matrix $A$ were found using the inverse iteration method described by Fearn (1985). An initial guess for the eigenvector $\boldsymbol{y}$ was made, which can be expressed as $\boldsymbol{y} = \sum_{i=1}^{N} \alpha_i \boldsymbol{x}_i$, where $\alpha_i$ are the coefficients of the eigenvectors. An initial guess $\varpi$ for the eigenvalue was also made. It is then possible to formulate an expression that converges to $\sigma_j$, the closest eigenvalue to $\varpi$,

$$(A - \varpi I)^{-m} \boldsymbol{y} = \sum_{i=1}^{N} \alpha_i \left(\sigma_i - \varpi\right)^{-m} \boldsymbol{x}_i, \tag{4.11}$$

where the right hand side will tend towards the $j^{\text{th}}$ eigenvalue as $m$ increases. This can be expressed in an iterative scheme

$$\boldsymbol{y}_{m+1} = \frac{(A - \varpi I)^{-1} \boldsymbol{y}_m}{y_{\text{max}}}. \tag{4.12}$$

At each iteration the guess for the eigenvector was normalised to have the element with the largest modulus $y_{\text{max}}$ set equal to one. The convergence of $\boldsymbol{y}$ on the closest actual eigenvector depends on the distance between $\varpi$ and $\sigma_j$ relative to the distance between $\varpi$ and the other eigenvalues. The operation to find the inverse is computationally expensive so although it is possible to recalculate $(A - \varpi I)$ with an improved guess for the eigenvalue $\varpi_2$ this is not efficient. The process was restarted with $\varpi_2$ as an improved guess for the eigenvalue only if $|\varpi_1 - \varpi_2| > 0.75$ and then the algorithm recalculated the matrix inverse from the updated eigenvalue guess $\varpi_2$. The inverse was calculated using LU decomposition which takes advantage of the banded form of the matrix. The iteration performed in equation (4.12) was considered complete when $\boldsymbol{y}_m$ and $\boldsymbol{y}_{m+1}$ were parallel to within a tolerance of $10^{-9}$ or smaller. The eigenvalue was

then found from

$$\frac{|\boldsymbol{y}_m|}{|\boldsymbol{y}_{m+1}|} = \frac{\alpha_j \left(\sigma_j - \varpi\right)^{-m} |\boldsymbol{x}_j|}{\alpha_j \left(\sigma_j - \varpi\right)^{-(m+1)} |\boldsymbol{x}_j|},$$
$$\sigma_j = \varpi + \frac{|\boldsymbol{y}_m|}{|\boldsymbol{y}_{m+1}|}.$$

In the compressible case then temperature rather than entropy was used. In the compressible case there is an evolution equation for the pressure so the variables were set-up using

$$A \begin{pmatrix} u_0 \\ v_0 \\ w_0 \\ b_{x0} \\ b_{y0} \\ b_{z0} \\ \rho_0 \\ T_0 \\ u_1 \\ \vdots \\ T_N \end{pmatrix} = \begin{pmatrix} \sigma & 0 & \cdots & & & & & & & & 0 \\ 0 & \sigma & 0 & \cdots & & & & & & & 0 \\ \vdots & 0 & \sigma & 0 & \cdots & & & & & & 0 \\ & \vdots & 0 & \sigma & 0 & \cdots & & & & & 0 \\ & & \vdots & 0 & \sigma & 0 & \cdots & & & & 0 \\ & & & \vdots & 0 & \sigma & 0 & \cdots & & & 0 \\ & & & & \vdots & 0 & \sigma & 0 & \cdots & & 0 \\ & & & & & \vdots & 0 & \sigma & 0 & \cdots & 0 \\ & & & & & & \vdots & 0 & \sigma & \ddots & 0 \\ & & & & & & & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \\ w_0 \\ b_{x0} \\ b_{y0} \\ b_{z0} \\ \rho_0 \\ T_0 \\ u_1 \\ \vdots \\ T_N \end{pmatrix}$$

so that $A$ is in banded form.

In the anelastic case instead of a $\sigma$ multiplying the pressure perturbation there is a zero as the pressure responds instantaneously to changes in the fluid. This can be seen when the divergence of the anelastic momentum equation (3.20b) is taken to get another relationship which is based on pressure

$$\nabla \cdot [\overline{\rho} \left(\mathbf{u}^* \cdot \nabla^* \mathbf{u}^*\right)] = -\nabla^{*2} p^* + \tilde{\mathcal{F}} \nabla \cdot [(\nabla^* \times \mathbf{B}^*) \times \mathbf{B}^*] + \left(\frac{\mathrm{Pr}}{R}\right)^{1/2} \nabla \cdot \left[\frac{1}{\overline{\rho}} \left(\nabla^* \cdot \boldsymbol{\tau}^*\right)\right], \quad (4.13)$$

which shows that pressure responds instantly to changes in $\mathbf{u}^*$ and $\mathbf{B}^*$. The anelastic continuity equation (3.20a) can now be seen as an diagnostic equation for the pressure,

as in Kersalé et al. (2004). This is equivalent to sound waves travelling instantaneously over all space so the pressure effects are also instantaneous as equation (4.13) shows. This leads to a generalised eigenvalue problem of

$$A\begin{pmatrix} u_0 \\ v_0 \\ w_0 \\ b_{x0} \\ b_{y0} \\ b_{z0} \\ p_0 \\ s_0 \\ u_1 \\ \vdots \\ s_N \end{pmatrix} = \begin{pmatrix} \sigma & 0 & \cdots & & & & & & & & 0 \\ 0 & \sigma & 0 & \cdots & & & & & & & 0 \\ \vdots & 0 & \sigma & 0 & \cdots & & & & & & 0 \\ & \vdots & 0 & \sigma & 0 & \cdots & & & & & 0 \\ & & \vdots & 0 & \sigma & 0 & \cdots & & & & 0 \\ & & & \vdots & 0 & \sigma & 0 & \cdots & & & 0 \\ & & & & \vdots & 0 & 0 & 0 & \cdots & & 0 \\ & & & & & \vdots & 0 & \sigma & 0 & \cdots & 0 \\ & & & & & & \vdots & 0 & \sigma & \ddots & 0 \\ & & & & & & & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \sigma \end{pmatrix} \begin{pmatrix} u_0 \\ v_0 \\ w_0 \\ b_{x0} \\ b_{y0} \\ b_{z0} \\ p_0 \\ s_0 \\ u_1 \\ \vdots \\ s_N \end{pmatrix}$$

so the problem can be formulated as

$$A\boldsymbol{x} = \sigma B\boldsymbol{x}. \tag{4.14}$$

The generalised eigenvalue problem can be solved in much the same way except with the equation (4.11) being replaced with

$$(A - \varpi B)^{-m} B\boldsymbol{y} = \sum_{i=1}^{N} \alpha_i \left(\sigma_i - \varpi\right)^{-m} \boldsymbol{x}_i, \tag{4.15}$$

so that the iteration scheme (4.12) is now

$$\boldsymbol{y}_{m+1} = \frac{(A - \varpi B)^{-1} B\boldsymbol{y}_m}{y_{\max}}. \tag{4.16}$$

### 4.2.1 Validation of Anelastic Algorithm

The anelastic code was validated by first comparing with the analytical results given in Chandrasekhar (1961). The validation also gives a justification for the number of

grid points by comparison to the Boussinesq case where a result can be analytically derived, as shown in Chapter 2. It is possible to recover the Boussinesq limit in the anelastic equations by taking $\theta = 0$ which sets all the thermodynamic reference states to be constant. In the Boussinesq case Chandrasekhar has an analytical equation for vertical magnetoconvection (for a more detailed introduction to Boussinesq magnetoconvection see §5.1)

$$\tilde{R}_{\mathrm{c}} = \left[(\pi^2 + k^2)^2 + \pi^2 Q\right] \tag{4.17}$$

where $k^2 = k_x^2 + k_y^2$; $Q$ the Chandrasekhar number $Q = B_0^2 d^2/(\mu_0 \mu \eta)$, see appendix C for details on how $Q$ relates to the other dimensionless numbers; and $\tilde{R}_{\mathrm{c}}$ is the critical Rayleigh number. The critical Rayleigh number $\tilde{R}_{\mathrm{c}}$ is the lowest Rayleigh number required for the system to be unstable, i.e. at $R = \tilde{R}_{\mathrm{c}}$ then $\sigma = 0$. The critical Rayleigh number can be minimised over $k$ to find the minimum critical Rayleigh number $\tilde{R}_{\mathrm{c}}^m$ which happens when

$$Q = \pi^2 \left( 2 \left( \frac{k}{\pi} \right)^6 + 3 \left( \frac{k}{\pi} \right)^4 - 1 \right). \tag{4.18}$$

This can then be compared to the value that the computer program finds in the anelastic case to give a fractional difference. This fractional difference in the Rayleigh number for the analytical Boussinesq and numerical anelastic equations is defined as $|\tilde{R}_{\mathrm{Bouss.}} - \tilde{R}_{\mathrm{an.}}|/|\tilde{R}_{\mathrm{Bouss.}}|$; a similar definition is used for the fractional differences in other parameters. Figure 4.1 shows the fractional difference as a function of resolution for the anelastic code. Figure 4.1 shows there is a steep decrease in the difference up to $600$ points, where the difference is under $0.001\%$. Then at numbers of grid points higher than $1500$ the difference begins to increase. This increase is due to fourth derivative terms which, when represented in the fourth-order accurate finite difference scheme, involve division by the grid-spacing to the fourth power. When the number of grid points increases above $1500$ then the numbers involved are small enough for machine precision to become a limitation. This suggests that the number grid points should be between $800 - 1200$ which is the range used in this thesis. The $\tilde{R}_{\mathrm{c}}^m$ values are independent of

the diffusion parameters $\mathrm{Pr}$ and $\zeta$ as well as being independent of the magnetic field $\mathcal{F}$. I will still give the parameters used in the following figures for completeness, $\mathrm{Pr} = 1.0$, $\zeta = 1.0$, $\gamma = 5/3$, $\theta = 0$, and $m$ is not used so could remain undefined. Figure 4.2 shows the difference of the numerical results in the Boussinesq case for a range of $k$ values and shows that the difference remains relativity unchanged and are less than $0.005\%$. This shows the fractional difference is independent of the wavenumber, as expected.

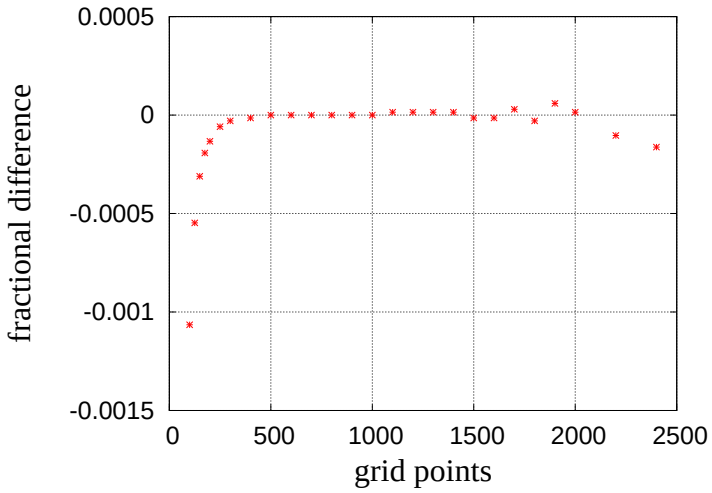



Figure 4.1: The fractional difference between the computational and the analytical Boussinesq results showing the fractional difference as function of the number of grid points for a fixed $R = 1761.8$ and $k = 3.27$.

### 4.2.2 Validation of Compressible Algorithm

The compressible code was adapted from a code written by Evy Kersalé and tested against that code. This validation was only done for cases with a constant magnetic field as in magnetoconvection. The validation that the anelastic code was correctly modelling magnetic buoyancy was made through comparisons with the compressible code.

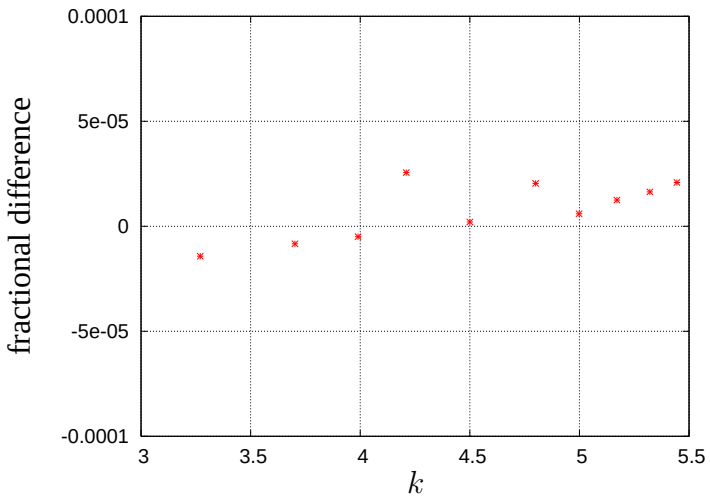



Figure 4.2: The fractional difference between the computational anelastic and the analytical Boussinesq results for a range of wavenumbers. The fractional difference is the difference between the computed and analytical $\tilde{R}_c$ both where $800$ grid points were used in the computational case.

# Chapter 5

# Linear Magnetoconvection Results

## 5.1 An Introduction to Boussinesq Magnetoconvection

Some magnetoconvection linear theory is useful to help understand the computational results presented later in this chapter.

From Proctor & Weiss (1982) and Chandrasekhar (1961), I will consider a Boussinesq plasma between two horizontal layers at a distance $d$ apart. Energy lost to magnetic dissipation leads to a source term in the energy equation (2.3d) in the form of Ohmic heating, but this is small and will be ignored for rest of the discussion on convection. The constituent equations of magnetohydrodynamics in the Boussinesq approximation are then non-dimensionalised using the length scale $d$, thermal relaxation time $d^2/(\rho_0 c_p k)$, velocity $\mathbf{u}$ with $\rho_0 c_p k/d$, and pressure $p$ with $\mu c_p k/d^2$. Then the magnetic field is written $\mathbf{B} = B_0\,(\hat{\mathbf{z}} + \mathbf{B}^*)$ and temperature $T$ as $T = T_0\,(1 + \theta(1-z) + T^*)$ where $B_0, \theta$ are the dimensionless magnetic and thermal gradients respectively and $T_0$ is the background temperature. Dropping the superscript $^*$ the magnetohydrodynamic equations, in non-

dimensional form, are then

$$\frac{1}{\mathrm{Pr}}\left(\frac{\partial \mathbf{u}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{u}\right)=-\nabla p+RT\hat{\mathbf{z}}+Q\zeta\left(\frac{\partial \mathbf{B}}{\partial z}+\mathbf{B}\cdot\nabla\mathbf{B}\right)+\nabla^2\mathbf{u}, \tag{5.1a}$$

$$\frac{\partial T}{\partial t}+\mathbf{u}\cdot\nabla T=\theta\mathbf{u}\cdot\hat{\mathbf{z}}+\nabla^2 T, \tag{5.1b}$$

$$\frac{\partial \mathbf{B}}{\partial t}+\mathbf{u}\cdot\nabla\mathbf{B}=\frac{\partial \mathbf{u}}{\partial z}+\mathbf{B}\cdot\nabla\mathbf{u}-\zeta\nabla\times(\nabla\times\mathbf{B}), \tag{5.1c}$$

$$\nabla\cdot\mathbf{u}=\nabla\cdot\mathbf{B}=0, \tag{5.1d}$$

where $\mathrm{Pr}=\mu c_p/k$ is the Prandtl number; $\zeta=\eta c_p\rho_0/k$ is a diffusivity ratio (or inverse Roberts number); $R=gc_p\alpha\theta d^3\rho_0^2/(k\mu)$ is the Rayleigh number which is a ratio of the buoyancy forces to the viscous forces, where $\alpha$ is the thermal coefficient of expansion, and $Q=B_0^2d^2/(\mu_0\mu\eta)$ is the Chandrasekhar number, see Appendix C for how these related to the dimensionless numbers in (2.5). Unlike elsewhere in this thesis, here $z=0$ is the bottom of the domain and $z=1$ is the top. The boundary conditions at $z=0,1$ are isothermal, stress-free with a vertical magnetic field so $\mathbf{B}\cdot\hat{\mathbf{x}}=\mathbf{B}\cdot\hat{\mathbf{y}}=0$. It is often helpful to decompose fields into two components: poloidal and toroidal. In axisymmetric fields the toroidal component of field is parallel to latitudinal lines. The poloidal field is outwards from the poles. I now separate the fields into poloidal and toroidal components, $\mathbf{u}=\nabla\times(\nabla\times\phi\hat{\mathbf{z}})+\nabla\times\psi\hat{\mathbf{z}}$ and $\mathbf{B}=\nabla\times(\nabla\times\xi\hat{\mathbf{z}})+\nabla\times\chi\hat{\mathbf{z}}$. The components decouple with the toroidal component describing solutions that decay and so will be neglected. Substituting the poloidal terms into (5.1a-c), linearising, and taking the $\hat{\mathbf{z}}$ component of the curl leaves

$$\frac{1}{\mathrm{Pr}}\frac{\partial}{\partial t}\nabla^2\phi=-RT+Q\zeta\frac{\partial}{\partial z}\nabla^2\xi+\nabla^4\phi, \tag{5.2a}$$

$$\frac{\partial T}{\partial t}=-\nabla_H^2\phi+\nabla^2 T, \tag{5.2b}$$

$$\frac{\partial \xi}{\partial t}=\frac{\partial \phi}{\partial z}+\zeta\nabla^2\xi, \tag{5.2c}$$

where $\nabla_H$ is the horizontal derivative. On assuming a normal mode solution for the perturbations of the form $f(z)\exp(i\mathbf{k}\cdot\mathbf{x}+\sigma t)$ and its complex conjugate, where $\mathbf{k}=(k_x,k_y)$, the system reduces to one containing only ordinary differential equations.

This leads to the dispersion relation, cubic in the growth rate $\sigma$, of

$$\varpi^2\left(\sigma+\varpi^2\right)\left(\sigma+\mathrm{Pr}\,\varpi^2\right)\left(\sigma+\zeta\varpi^2\right) + \zeta\,\mathrm{Pr}\,Q\varpi^2\pi^2\left(\sigma+\varpi^2\right) - R\,\mathrm{Pr}\,k^2(\sigma+\zeta\varpi^2)=0, \quad (5.3)$$

where $\varpi^2 = k^2+\pi^2$ and $k = |\mathbf{k}|$. There is a steady-state bifurcation (exchange of stabilities) at $\sigma = 0$ where the critical Rayleigh number, the minimum Rayleigh number required for convection to onset, is

$$R_\mathrm{c} = k^{-2}\left(\varpi^6 + Q\pi^2\varpi^2\right). \quad (5.4)$$

The critical Rayleigh number can be minimised over $k$. Taking a step back to see how this relates to magnetoconvection is useful at this point. The minimum critical Rayleigh number $R_\mathrm{c}^m$ is an increasing function of $Q$. As magnetic fields become stronger then more energy is required to displace field lines and so convection onsets proportionally later for higher $R$. For sufficiently large $Q$ then $R_\mathrm{c}^m \sim \pi^2 Q$.

A Hopf bifurcation is also possible in this system when (5.3) has purely imaginary roots, $\sigma = \pm i\omega$, and occurs when

$$R_\mathrm{oc} = k^{-2}\left(A\varpi^6 + BQ\pi^2\varpi^2\right), \quad (5.5)$$

$$\text{where } A = 1+\frac{\zeta}{\mathrm{Pr}}\left(1+\mathrm{Pr}+\zeta\right),\ B = \frac{\zeta\left(\mathrm{Pr}+\zeta\right)}{1+\mathrm{Pr}}.$$

The frequency $\omega$ must satisfy

$$\omega^2 = -\zeta^2\varpi^4 + \frac{\mathrm{Pr}\,\zeta(1-\zeta)}{1+\mathrm{Pr}}\pi^2 Q. \quad (5.6)$$

No Hopf bifurcation is possible if $\zeta > 1$ as then $\omega^2 < 0$. The limiting case is $\omega = 0$ and substituting (5.6) into (5.5) and comparing with (5.4) shows that if a Hopf bifurcation is possible then it will occur at a lower Rayleigh number than the steady-state bifurcation. If the system is in the regime where a Hopf bifurcation will occur then, as the Rayleigh number is increased steadily, the system will transition smoothly from a stable state to oscillatory convection. The Takens-Bogdanov (TB) point is where the Hopf bifurcation

coincides with the steady state (or pitchfork) bifurcation and the frequency tends to zero. Near the TB point the frequency at the Hopf bifurcation is small. In liquid-metals $\zeta \gg 1$ but in astro- and geophysical situations $\zeta < 1$ so it initially appears likely that oscillatory convection sets in first. This is somewhat of an over-simplification as in the Sun $\zeta$ is proportional to density and passes through unity from the surface to the bottom of the convection zone.

## 5.2 Linear Code modelling Magnetoconvection

I will now explain the set-up of the linear problem. This was done for the anelastic case and then compared with the compressible case where I have assumed that in the compressible case the problem was solved exactly. It is then possible to compare and contrast the anelastic equation with and without the Lantz-Braginsky simplification to the compressible equations. The fully compressible equations are in equation (2.3) and the anelastic equations without the Lantz-Braginsky are given in (3.20) with the Lantz-Braginsky approximation given in §3.7.

Owing to the inherent symmetry of the linear problem the results do not depend on the separate horizontal wavenumbers and so I proceed by calculating the critical Rayleigh number as a function of $k$ where $k^2 = k_x^2 + k_y^2$. The critical Rayleigh number $\tilde{R}_{\rm c}$ is defined as that Rayleigh number for which $\Re\{\sigma\} = 0$ and $\tilde{R}_{\rm c}^m$ is then the lowest value of $\tilde{R}_{\rm c}$ when optimised over $k$. For simplicity, I denote the wavenumber at which this minimum occurs as $k$ in the figures following. This procedure is formally equivalent to solving an eigenvalue problem for $\tilde{R}_{\rm c}$, optimised over $k$, with all other parameters fixed apart from $C_k$, which is related to $\tilde{R}$ by equation (3.25); this in turn can be thought as an eigenvalue problem for $C_k$.

The systems of full anelastic, Lantz-Braginsky approximation, and fully compressible linear equations, together with the appropriate boundary conditions, were solved in the form of generalised eigenvalue problems.

### 5.2.1 Boundary Conditions

The boundary conditions are consistent with boundary conditions on full variables, but were imposed on the perturbation terms, and are that the top and bottom boundaries were impermeable, stress free and isothermal so

$$\hat{w} = \frac{\partial}{\partial z}\hat{u} = \frac{\partial}{\partial z}\hat{v} = \hat{T} = 0 \quad \text{at } z = 0, 1$$

where the velocity components are $\mathbf{u} = (u, v, w)$ and the terms such as $\hat{w}$ terms are from the expansion in (4.1) in the anelastic case and (4.2) in the compressible case. This was true even in the anelastic case where I used the entropy equation rather than the temperature equation and this will be further investigated in § 5.3.6. Stress free boundary conditions were chosen as these are easier to implement in the non-linear case so comparisons would be simpler. The temperature boundary condition is non-physical as there is no area in the Sun where the fluctuations to a time-independent profile are always zero and was chosen for mathematical convenience. Although it is non-physical it is relevant to the Sun and other temperature boundary conditions, such as fixed flux, have similar issues.

I also imposed the illustrative boundary condition suggested by Chandrasekhar (1961) on top and bottom of the magnetic field, i.e.

$$\hat{B}_x = \hat{B}_y = \frac{\partial \hat{B}_z}{\partial z} = 0, \tag{5.7}$$

where magnetic field components are $\mathbf{B} = (B_x, B_y, B_z)$. This guaranteed a vertical field whilst satisfying the solenoidal condition (2.1b).

### 5.2.2 Basic State

The basic state in the anelastic equations is not equivalent to the basic state in the compressible equations. In the anelastic case, the reference state is a non-magnetic polytrope and the magnetic field enters only in the basic state. In the compressible case

there is no reference state so the basic state is that of a polytrope, but in the presence of a weak magnetic field with a strength consistent with the anelastic approximation.

The basic state of the compressible equations (2.3) is denoted $\xi_b$, where $\xi$ can represent any variable. I consider a steady, stationary basic state given by a polytropic solution together with a uniform vertical magnetic field $\mathbf{B}_b = (0, 0, 1)$. Thus, $T_b = (1 + \theta z)$ and $\rho_b = (1 + \theta z)^m$ where $m$ is the polytropic index as explained in §3.4.2. For a polytropic atmosphere the density contrast is defined as

$$\chi = \frac{\rho_b(1)}{\rho_b(0)} = (\theta + 1)^m \,. \tag{5.8}$$

Although there is a magnetic field in the basic state it has no gradient and does not alter the polytrope. Figure 5.1 shows the compressible basic state for typical parameter values used in the computations later and it shows that the compressible basic state is a polytrope.

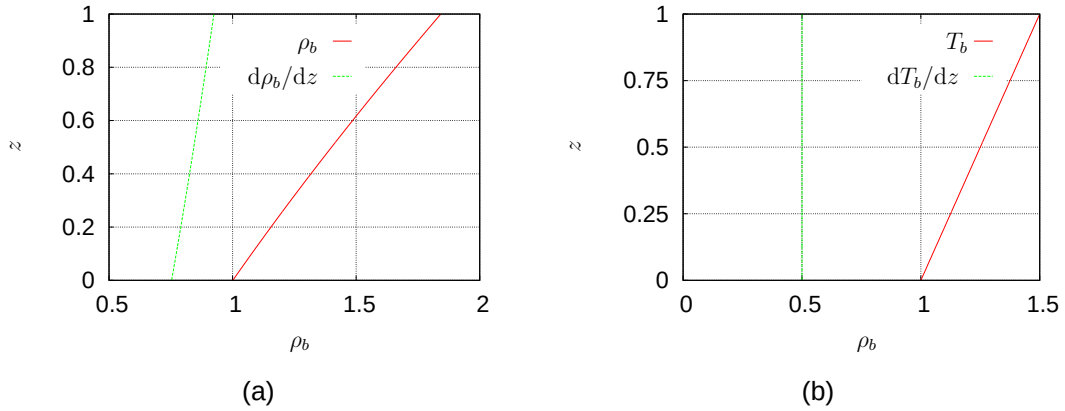



Figure 5.1: The compressible basic state density (a) and the temperature (b) for a typical case with $\gamma = 5/3$, $m = 1.495$, $\theta = 0.5$, $\mathrm{Pr} = 1.0$, $\zeta = 5 \times 10^{-2}$, and $\mathcal{F} = 5 \times 10^{-4}$.

In the anelastic approximation the basic state $\zeta_b^*$ is also simple. For magnetoconvection the field in the basic state is uniform and vertical $\mathbf{B}_b^* = (0, 0, 1)$. Imposing a constant vertical field causes no Ohmic heating and no Lorentz force. I consider a basic state which is a stationary and steady solution of the equations (3.20a-d), given by

$s_b^* = T_b^* = p_b^* = \rho_b^* = \mathbf{u}_b^* = 0$. This simplicity is in contrast to the magnetic buoyancy instability which is discussed later in §6.2.3.

## 5.3 Linear Magnetoconvection Results

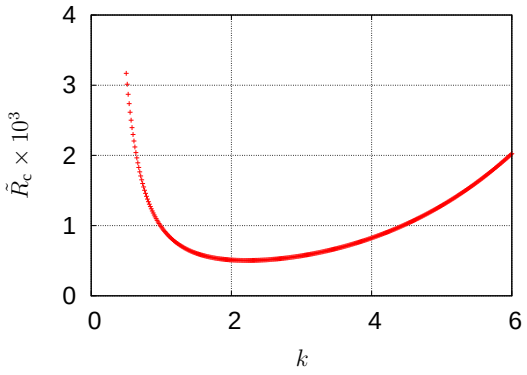
The numerical algorithm determines the growth rates (eigenvalues) of the perturbation terms as a function of the input parameters and the wavenumbers. For historical reasons the critical diagnostic in convection problems has been the marginal Rayleigh number (Chandrasekhar, 1961) and I choose to retain this diagnostic. Starting from a stable state I slowly increase the Rayleigh number from a starting point until the system becomes unstable. This bifurcation defines the critical Rayleigh number $\tilde{R}_{\rm c}$, where the growth rate is zero, as in the Boussinesq case.

In setting the polytropic index $m$ so that the atmosphere is super-adiabatic then convection may occur. For $\gamma = 5/3$ super-adiabaticity requires $m < 1.5$. I fix the following parameters as

$$\gamma = 5/3, \quad m = 1.495, \quad \mathrm{Pr} = 1.0, \quad \zeta = 5 \times 10^{-2} \quad \text{and} \quad \mathcal{F} = 5 \times 10^{-4},$$

unless otherwise stated in the figure captions. The value of $m$ was chosen so the atmosphere was unstable to convection but the departure from an adiabatic atmosphere was small. This is required so that $\epsilon$, defined in (3.4), is kept small. In computations where $m$ is not altered, if there are differences between the fully compressible and anelastic equations then it is not due to a large $\epsilon$. The fractional difference is defined with regard to the compressible equations for example the fractional difference in the wavenumber is defined as $|k_{\text{comp.}} - k_{\text{an.}}|/|k_{\text{comp.}}|$; a similar definition is used for the fractional differences in other parameters. When I compared the full anelastic with the Lantz-Braginsky simplification the fractional difference was defined as $|k_{\text{an.}} - k_{\text{Lantz-Braginsky.}}|/|k_{\text{an.}}|$ as this produced clearer plots. This has the disadvantage that when the full anelastic approximation produces a different result from the fully

compressible equations then it does not show if the Lantz-Braginsky simplification is increasing or decreasing the difference. It was always the case that the fractional difference between the Lantz-Braginsky simplification and the compressible case was larger than between the full anelastic and compressible case.




(a)

Figure 5.2: For $\chi = 1.5$, ($\theta = 0.311$) variations in the marginal or critical Rayleigh vs. the wavenumber.

Figure 5.2 shows a typical $\tilde{R}_c$ dependence on $k$, the lowest eigenvalue optimised over $k$ has already been defined as the minimum critical Rayleigh number $\tilde{R}_c^m$.

### 5.3.1 The effect of altering $\theta$

The temperature gradient $\theta$ is related to other dimensional numbers from equations (3.22) and (3.26). The anelastic equations are formally equivalent to the Boussinesq equations in the limit $\theta \to 0$; so the Boussinesq results from §5.1 should be recoverable.

Figure 5.3(a) shows the critical Rayleigh number dependence on $\theta$ (with $m$ fixed so that $\chi$ varies) for the fully compressible problem. As the temperature gradient is normalised in $\tilde{R}_c$, increasing $\theta$ stabilises the layer, leading to a larger $\tilde{R}_c^m$ in agreement with earlier studies (see e.g. Gough et al., 1976; van Ballegooijen, 1982). Increasing the stratification of the atmosphere causes a steep increase in the minimum critical

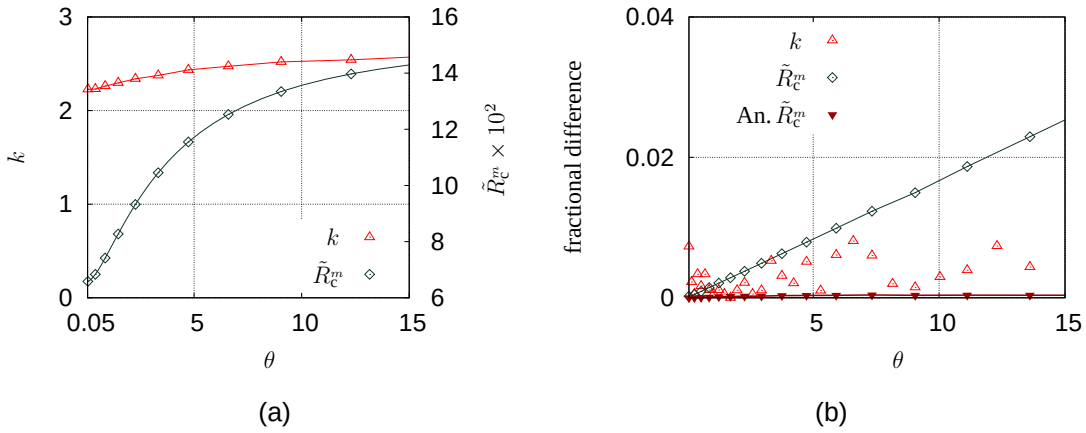
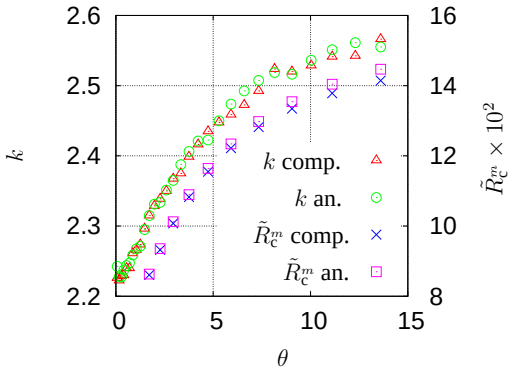



Figure 5.3: For $m$ fixed, variations vs. $\theta$ of (a) the minimum value of the critical Rayleigh number, $\tilde{R}_c^m$, and the corresponding wavenumber, $k$, for the compressible model; (b) the fractional differences in $\tilde{R}_c^m$ between compressible and anelastic models, and between anelastic models with the Lantz-Braginsky simplification and the anelastic equations, denoted An. $\tilde{R}_c^m$. The An. $\tilde{R}_c^m$ points denote the fractional difference between the full anelastic equations and the Lantz-Braginsky simplification. Here $\epsilon$ lies in the range $10^{-4} - 3 \times 10^{-2}$.

Rayleigh number but this increase then levels off at higher stratification. Figure 5.3(b) shows the fractional difference in the critical Rayleigh number and wavenumber for the anelastic approximation and the fully compressible equations. It is clear that the anelastic approximation gives an accurate representation of the fully compressible solutions, with errors of less than $3\%$, even for the most stratified cases where $\theta = 15$ and $\chi = 60$. This is not unexpected since, with $m = 1.495$, the atmosphere is chosen to be very close to adiabatic and so one might expect the anelastic approximation to perform well. Figure 5.3(b) also shows the fractional difference in $\tilde{R}_c^m$ between the anelastic equations solutions and the solutions calculated using the Lantz-Braginsky approximation. The fractional difference of the difference between the anelastic approximation with and without the Lantz-Braginsky simplification is also shown in figure 5.3(b). As the atmosphere is exceptionally close to being adiabatic, the Lantz-Braginsky approximation performs well. The Lantz-Braginsky approximation is exact for $m = 1.5$ and should be very good for small $m$. Figure 5.4 has both the full anelastic and compressible results which shows that the $\tilde{R}_c^m$ in the full anelastic equations is larger than in the compressible equations.




(a)

Figure 5.4: The same as figure 5.3(a) except now showing the compressible along with the anelastic results on the same plot.

Figure 5.5 also shows a similar plot as figure 5.3 but with a stronger magnetic field of $\mathcal{F} = 0.2$. At low values of $\theta$ the instability is oscillatory and at larger values it becomes

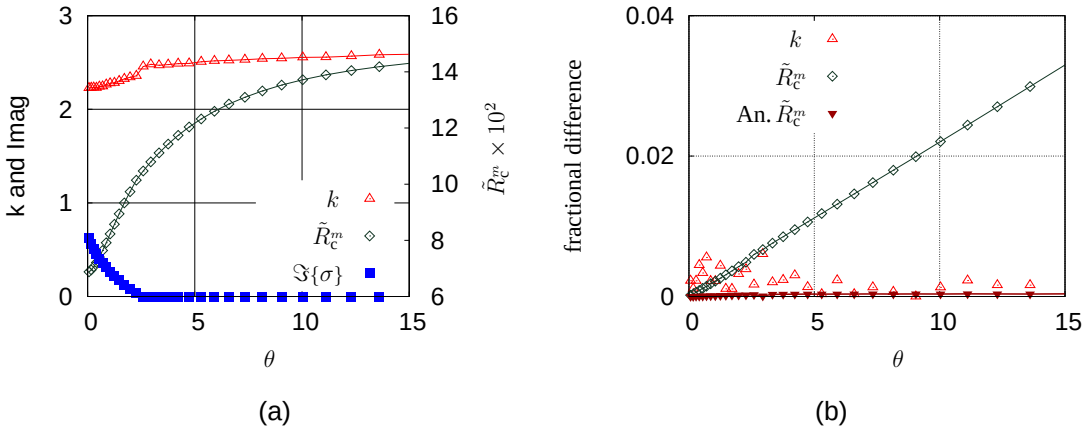



Figure 5.5: As for figure 5.3 but here $\mathcal{F} = 0.2$ and $\zeta = 2 \times 10^{-2}$. Also the imaginary part of the eigenvalue is plotted, denoted $\Im\sigma$.

steady. At the TB point the wavenumber is discontinuous, as the oscillatory instability and the steady instability occur at different wavenumbers, but $\tilde{R}_c^m$ is continuous. A sketch of how the critical Rayleigh number is altered around the TB point is given in figure 5.6. The fractional difference in figure 5.5 is slightly larger than in figure 5.3 even though $\epsilon$ is the same because of the larger magnetic field the Alfvénic time-scale decreases, upsetting the anelastic ordering.

### 5.3.2 The effect of altering $m$

The polytropic index $m$ measures the stratification of the atmosphere with an adiabatic atmosphere having $m = 1.5$. The anelastic equations are only valid when the atmosphere is nearly adiabatic and the Lantz-Braginsky simplification is equivalent to the full anelastic equations only if the atmosphere is adiabatic, see §3.7 and later discussions.

The effect of varying the polytropic index on the critical Rayleigh number and wavenumber is shown in figure 5.7(a). Here $m$ is varied along with $\theta$ to keep the density contrast constant. The equation (3.22) shows that altering the polytropic index also alters the gravity of the system. As $m$ is decreased the atmosphere becomes more

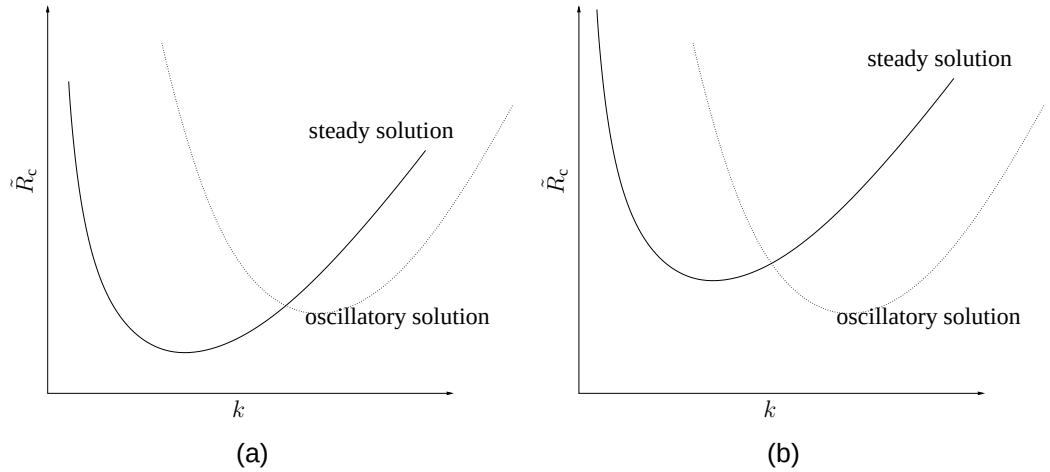



Figure 5.6: A sketch of the steady and oscillatory critical Rayleigh number for (a) parameter values below the TB point and (b) parameter values above the TB point.

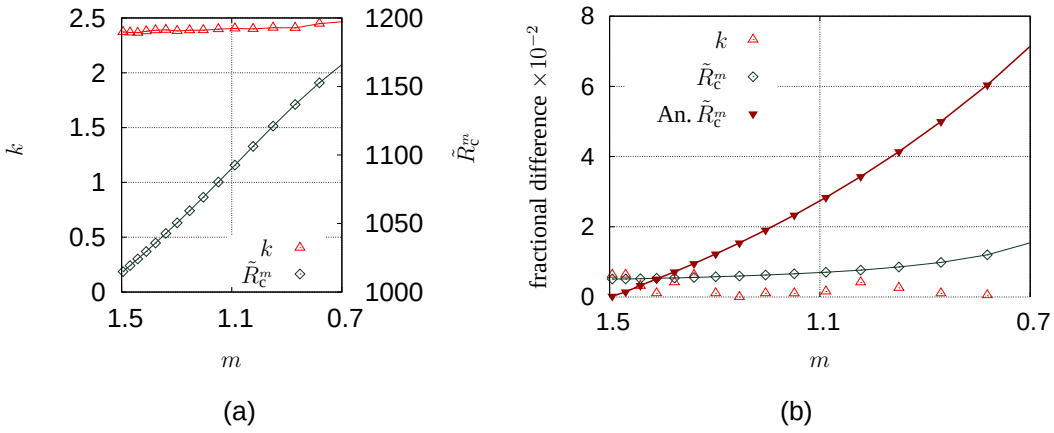



Figure 5.7: As for figure 5.3 but variations vs. $m$ for $\chi = 8$ fixed; $m$ is decreasing to the right. In addition (b) shows the fractional difference in $k$ between compressible and anelastic models. Here $\epsilon$ lies in the range $6 \times 10^{-3} - 10$.

unstable but, as $m$ is also in the gravity term, the minimum critical Rayleigh number increases, although the wavenumber at which the system becomes critical remains largely unaffected. Again, figure 5.7(b) shows how well the anelastic approximation reproduces the fully compressible results by presenting fractional differences. Somewhat surprisingly the anelastic approximation performs well on the linear problem even when $m$ differs significantly from $1.5$, the adiabatic value. As $m$ is decreased the anelastic approximation becomes less accurate, but it is still well within $2\%$ at $m = 0.7$ where $\epsilon \sim 10$. Berkoff et al. (2010) believed this accuracy was an artifact of the linear problem. The Lantz-Braginsky simplification is equivalent to the anelastic equations in an adiabatic atmosphere and the two differ in terms of order $\epsilon^2$. In figure 5.7(b) when $m$ becomes small then $\epsilon$ becomes large causing differences of up-to $7\%$ with the anelastic equations. The Lantz-Braginsky approximation is not capturing this instability as well as the anelastic equations, but this may be another artifact of the linear problem. When the density contrast is reduced from $\chi = 8$ to $\chi = 3$ (not shown) there is a slight decrease in the fractional differences. It therefore seems probable that the anelastic approximation performs better when the density contrast is lower.

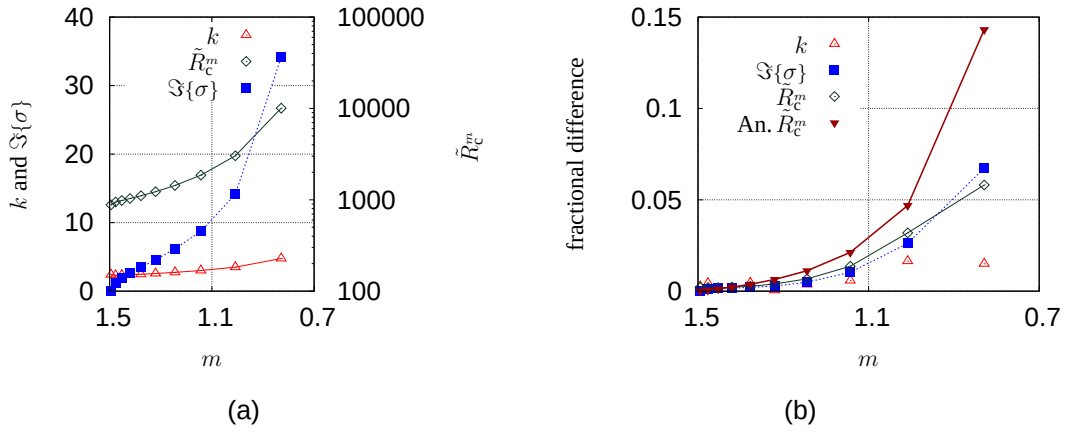


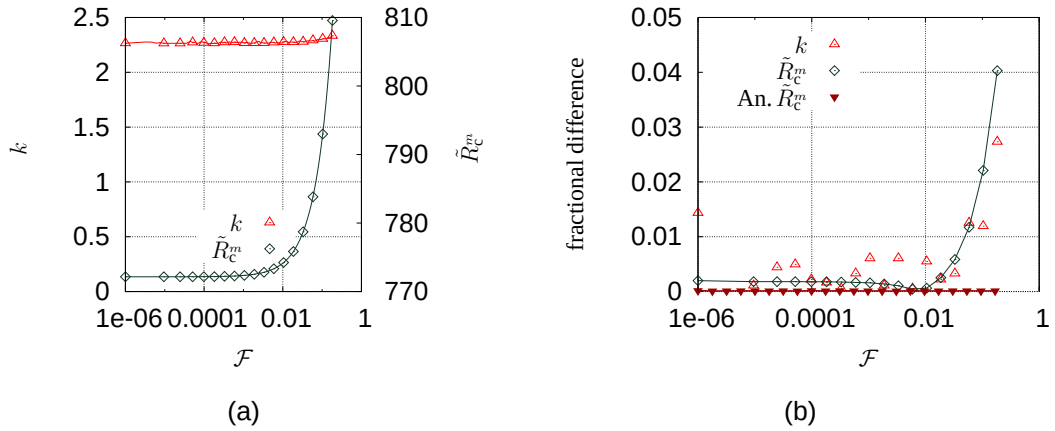

Figure 5.8: As for figure 5.7, for $\chi = 3$ and $\mathcal{F} = 0.5$. In addition, (a) shows the imaginary component of $\sigma$ vs. $m$ and (b) shows its fractional difference between compressible and anelastic models.

Figure 5.8 shows a similar comparison between the anelastic and compressible models when the imposed field is stronger, as the polytropic index $m$ is once again moved away from $1.5$. Interestingly, both the anelastic equations and the Lantz-Braginsky approximation perform worse than in figure 5.7, although the anelastic equations are still accurate to 7%. The Lantz-Braginsky simplification compared with the anelastic equations has a $15\%$ difference at $m = 0.8$ in contrast to when the magnetic field is weaker, where the difference is under 2% (not shown). This again suggests that when $m$ is far from its adiabatic value the Lantz-Braginsky anelastic approximation performs less well. The mode is oscillatory and the anelastic approximation is able to capture this accurately. The oscillations have a higher frequency when the atmosphere is further from adiabatic. The anelastic model also captures the oscillation frequency less accurately when the atmosphere is far from adiabatic.

### 5.3.3 The effect of altering $\mathcal{F}$

Altering the dimensionless parameter $\mathcal{F}$ changes the strength of the magnetic field. The anelastic approximation is only valid for a weak field i.e. where the Alfvén waves can be captured by the slow dynamical time-scale. Strong magnetic fields violate the assumptions used in Chapter 3 on the time-scale of the evolution of the fluctuation terms and so it is expected that instabilities with strong magnetic fields will be not be accurately captured by the anelastic approximation.

The effect of a strong magnetic field in reducing the accuracy of the anelastic approximation can be seen clearly in figure 5.9. Magnetic fields inhibit convection but this process only becomes noticeable for $\mathcal{F}$ larger than $\sim 0.01$. This threshold is sensitive to whether the fully compressible or anelastic approximation is used. As $m$ is fixed near to its adiabatic value, the anelastic approximation does perform well, but performs less well as $\mathcal{F}$ is increased; the magnetic field may become sufficiently large so that the Alfvénic time-scale approaches the dynamic time-scale, breaking one of the assumptions of the anelastic approximation. As expected for a nearly adiabatic




(a) (b)

Figure 5.9: As for figure 5.7 but variations vs. $\mathcal{F}$ with $\epsilon = 2 \times 10^{-3}$. $\mathcal{F}$ on the $x$ axis is scaled logarithmically.

atmosphere, the Lantz-Braginsky approximation is a good approximation of the anelastic equations.

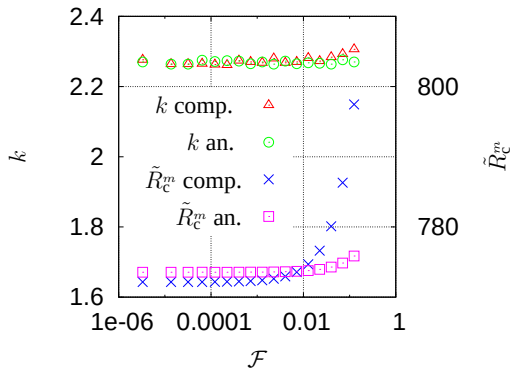



Figure 5.10: The same as figure 5.9(a) except now showing the compressible along with the anelastic results on the same plot.

It can be seen more clearly how the anelastic approximation differs from the fully compressible model in figure 5.10 which shows the anelastic results alongside the compressible results. In the compressible results the minimum critical Rayleigh number and the corresponding wavenumber increase when $\mathcal{F} > 0.08$ whereas in the anelastic results the minimum critical Rayleigh number and the corresponding wavenumber do

not show the same increase.

The figures in 5.11 show the absolute value of the eigenfunctions which correspond to the point $\mathcal{F} = 0.001$ in figure 5.9. At $\mathcal{F} = 0.001$ in figure 5.9 the fractional difference between the anelastic approximation and the compressible equations is small. In the compressible case the eigenfunction shown in figure 5.11 (a) for the horizontal magnetic field peaks just below the middle of the domain and is shown in $B_x$. The vertical magnetic field is not symmetric about the middle of the domain with a peak closer to the top. This is similar to the anelastic case without the Lantz-Braginsky simplification shown in figure 5.11 (b). The horizontal velocity, shown in figure 5.11 (c), passes through the origin in the lower half of the domain whereas the vertical velocity has its peak in the middle of the domain. The anelastic case, shown in figure 5.11 (d), is almost identical to the compressible case shown in 5.11 (c). The eigenfunctions for the thermodynamic variables are shown in figure 5.11 (e). $T$ is almost a sine curve but the other variables have more complex shapes. The anelastic case in figure 5.11 (f) shows that even when the eigenvalues agree well, as when $\mathcal{F} = 0.001$, then there are still differences in the eigenfunctions as maximum values of the thermodynamic variables are all different between the two cases by around $2\%$. The shapes of the thermodynamic eigenfunctions appears similar in the anelastic and compressible cases.

Slightly unexpectedly figure 5.12 gives almost the same results as figure 5.11 even though it has $\mathcal{F} = 0.05$ rather than $\mathcal{F} = 0.001$ . The compressible eigenfunctions shown on the left in figure 5.12 are very similar to those in figure 5.11, except that in figure 5.12 the thermodynamic variables have slightly larger maximum values. When $\mathcal{F} = 0.05$ then the fractional difference between the compressible and anelastic approximation is $3\%$, as shown in figure 5.9. Although the magnetic field is $50$ times stronger it is still a weak field and so not creating much difference in the eigenfunctions of the instability, as shown in the compressible case in the left panels of figures 5.11 and 5.12. The eigenfunctions in the anelastic approximation still compare very well to the compressible case with a fractional difference of $3\%$.

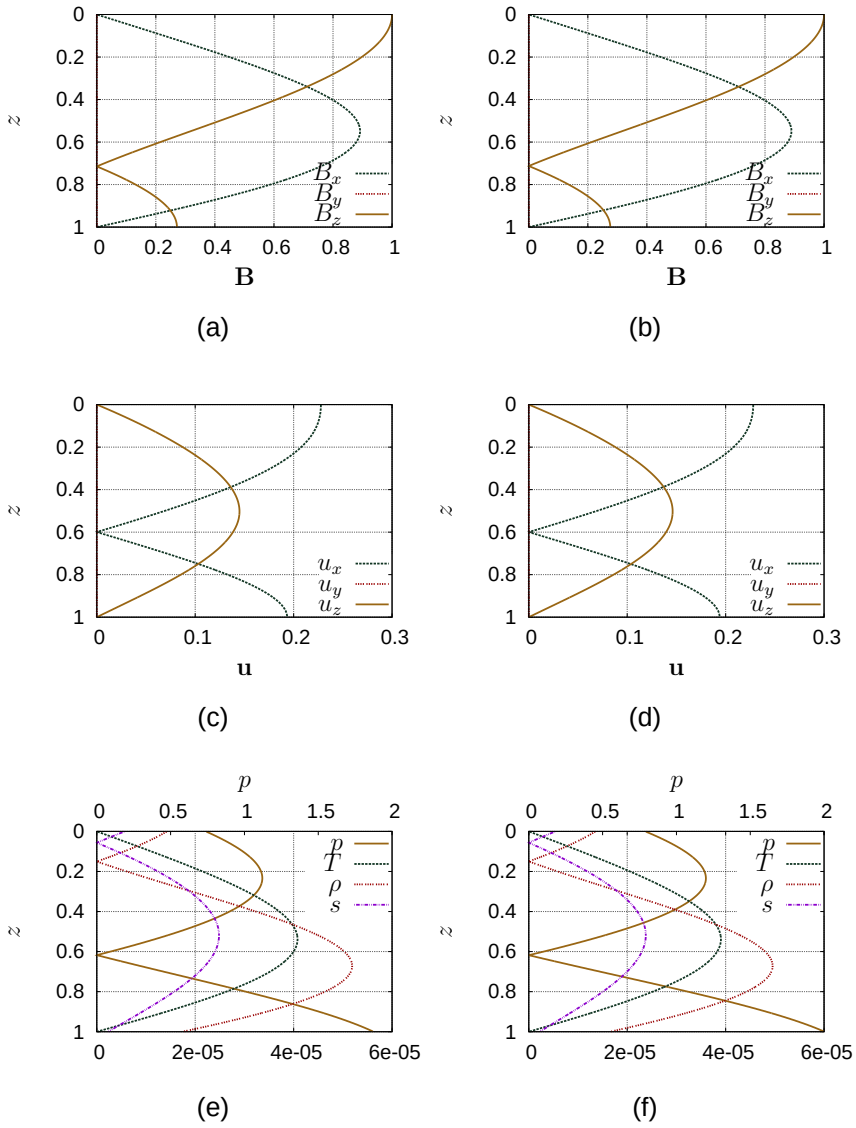



Figure 5.11: The absolute values of the eigenfunctions corresponding to minimum critical Rayleigh number shown in figure 5.9 with $\mathcal{F} = 0.001$. The left and right panels correspond to the compressible and anelastic (without the Lantz-Braginsky simplification) models respectively. The values in the compressible case are $k = 2.27$, $R = \tilde{R}_c^m = 772$ and in the anelastic case $k = 2.28$, $R = \tilde{R}_c^m = 773$. (a)-(b) Magnetic field ; (c)-(d) components of the fluid velocity $u_x$, $u_y$ and $u_z$; (e)-(f) thermodynamic variables $p$, $\rho$, $T$ and $s$.

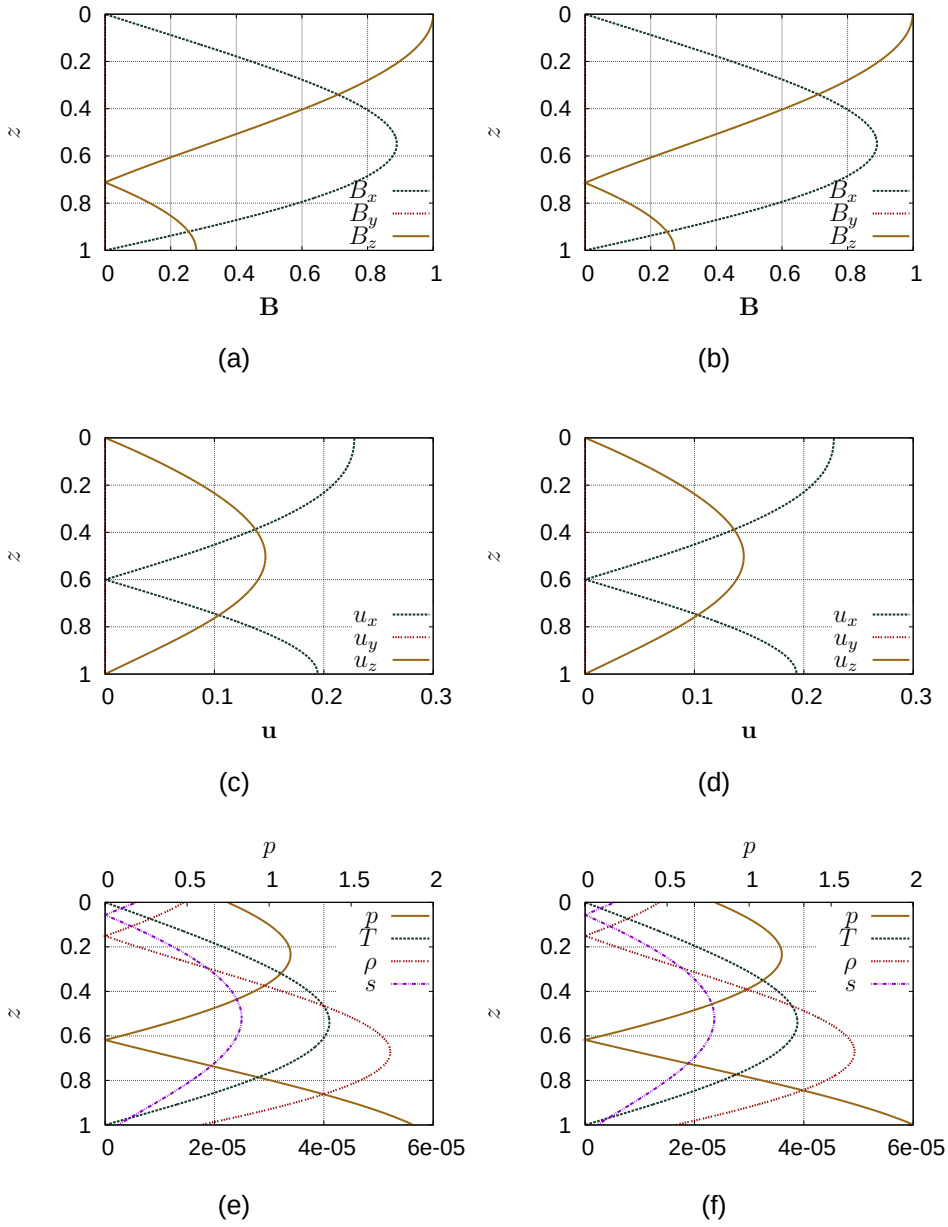



Figure 5.12: As for figure 5.11 except with $\mathcal{F} = 0.05$. The values in the compressible case are $k = 2.29$, $R = \tilde{R}_c^m = 783$ and in the anelastic case $k = 2.27$, $R = \tilde{R}_c^m = 775$.

### 5.3.4 The effect of altering $\zeta$

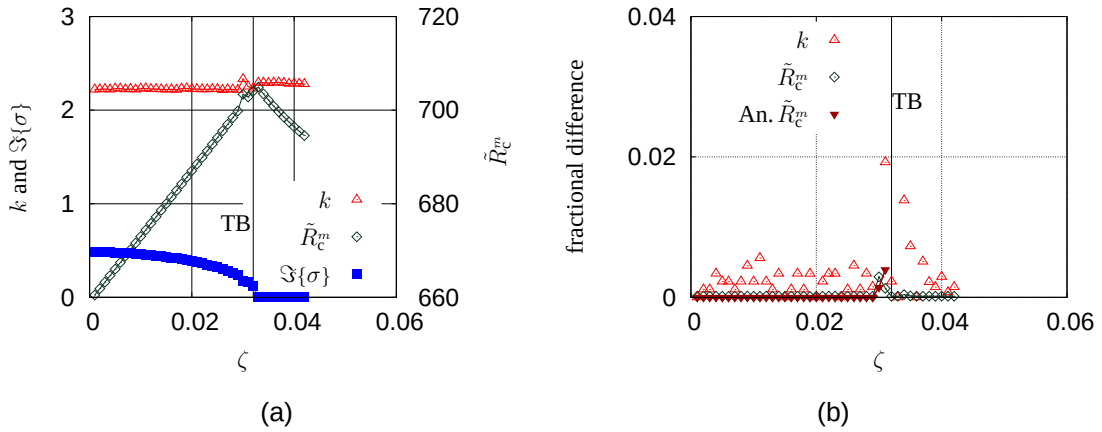



Figure 5.13: As for figure 5.3 with variations vs. $\zeta$ with $\mathcal{F} = 0.1$, $\chi = 1.1$, and $\theta = 0.0658$. The line marked TB is the Takens-Bogdanov point; for $\zeta$ values above this then $\Im\{\sigma\} = 0$.

The magnetic diffusivity proportional to the dimensionless number $\zeta$. From the assumptions made in Chapter 3 altering $\zeta$ should not effect the accuracy of the anelastic approximation.

When $\zeta$ is varied in figure 5.13 (a) a TB point occurs. For low values of $\zeta$ the solution which becomes unstable at the lower Rayleigh number has an eigenvalue with an imaginary component and so is an oscillatory instability. There may also be a steady instability with no imaginary component, but it must occur at a higher Rayleigh number. The oscillatory solution undergoes a Hopf bifurcation when the Rayleigh number passes through critical. This is expected from the Boussinesq theory shown in equation (5.3). As $\zeta$ is increased the oscillatory solution becomes unstable at larger Rayleigh numbers whereas the dependence of the steady solution on Rayleigh number appears constant for all $\zeta$ values in the results. When $\zeta$ is large enough then the steady solution will become unstable before the oscillatory solution. The steady mode bifurcates as the Rayleigh number passes through critical via a pitchfork bifurcation.

Capturing the TB point was numerically awkward as the solution had the tendency to jump onto the wrong branch, i.e. near the TB point there is a steady solution that is only slightly sub-critical (e.g. using an illustrative case then at $\tilde{R} = \tilde{R}_{\text{c}} - 10$ then $\sigma_{\text{st}} = -0.01 + 0i$ and at $\tilde{R} = \tilde{R}_{\text{c}} + 10$ then $\sigma_{\text{st}} = -0.01 + 0i$) and this has the largest real eigenvalue for most Rayleigh numbers below critical. There is also a second solution, which is oscillatory, with a real component which grows rapidly when the Rayleigh number increases (e.g. using an illustrative case again then at $\tilde{R} = \tilde{R}_{\text{c}} - 10$ then $\sigma_{\text{oc}} = -0.2 + 3i$ and at $\tilde{R} = \tilde{R}_{\text{c}} + 10$ then $\sigma = +0.1 + 3i$) making this oscillatory solution the first to become unstable. The algorithm often latched onto the steady solution even after the oscillatory solution became unstable. When the algorithm found an eigenvalue that decreased when the Rayleigh number increased then the algorithm restarted the inverse iteration but starting with an initial guess of the eigenvalue with a much larger real component, first with and then without an imaginary component. The algorithm then proceeded with the initial guess of the eigenvalue which was closest to the actual eigenvalue with the largest real component. Also when the algorithm had selected what might be the lowest minimum critical Rayleigh number the wavenumber was increased by large steps to check that there was no lower critical Rayleigh number at higher wavenumber, figure 5.6 shows how large changes in the wavenumber may find a lower critical Rayleigh number.

The value of $\mathcal{F}$ in figure 5.13 is higher than in other figures so that the oscillatory solution is visible for a large enough value of $\zeta$. For low values of $\zeta$ the diffusion reduces the scale of the instabilities and introduces boundary layers that cannot be accurately modelled without running into issues of round-off error.

In the anelastic case I have tracked the TB point over a range of values for the polytropic index, as done in figure 5.14, and showed that as $m$ increases, and so $\theta$ increases, the TB point occurs at lower $\zeta$ values and higher Rayleigh numbers.

Figure 5.15 shows the effect of increasing $\zeta$ at larger stratification than in figure 5.13. This figure shows a similar trend as figure 5.13 except that the oscillatory instability

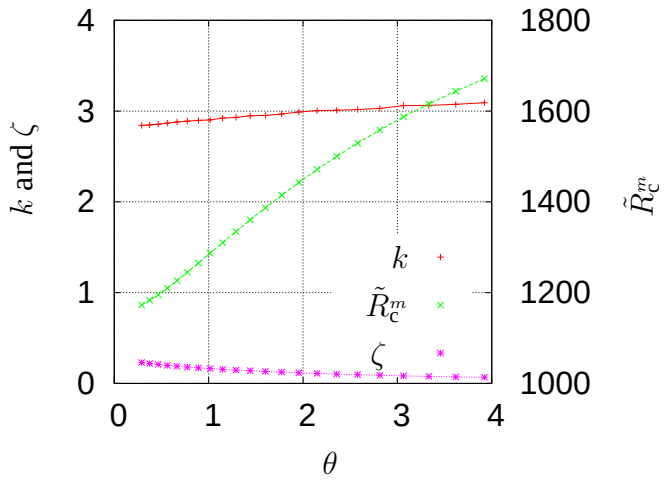



Figure 5.14: Tracking the TB bifurcation whilst altering $\theta$.

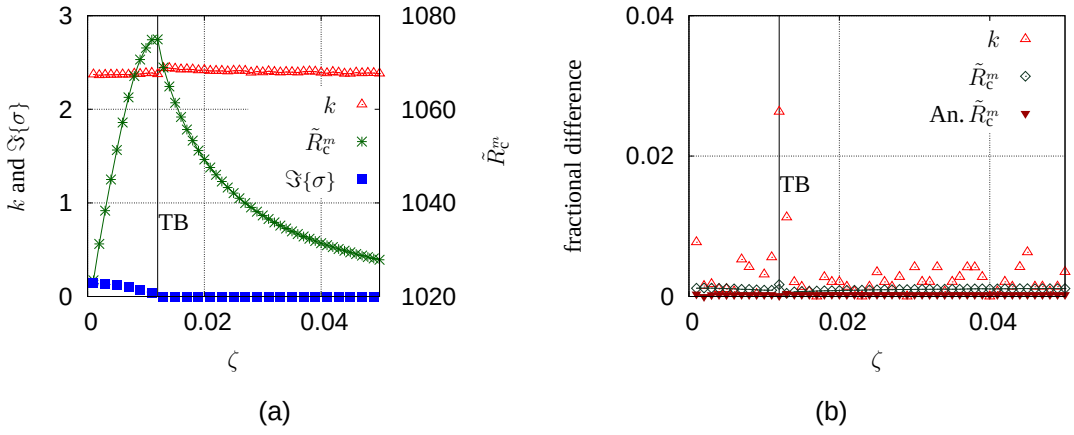



Figure 5.15: As for figure 5.13 with $\chi = 8.0$, and $\theta = 3.02$.

is more unstable than the steady instability for a smaller range of $\zeta$ values. The fractional differences between the anelastic and compressible models are slightly larger in figure 5.15 than in figure 5.13 which is expected from the results investigating when $\theta$ was altered as in figure 5.7.

### 5.3.5 The effect of altering $\mathrm{Pr}$

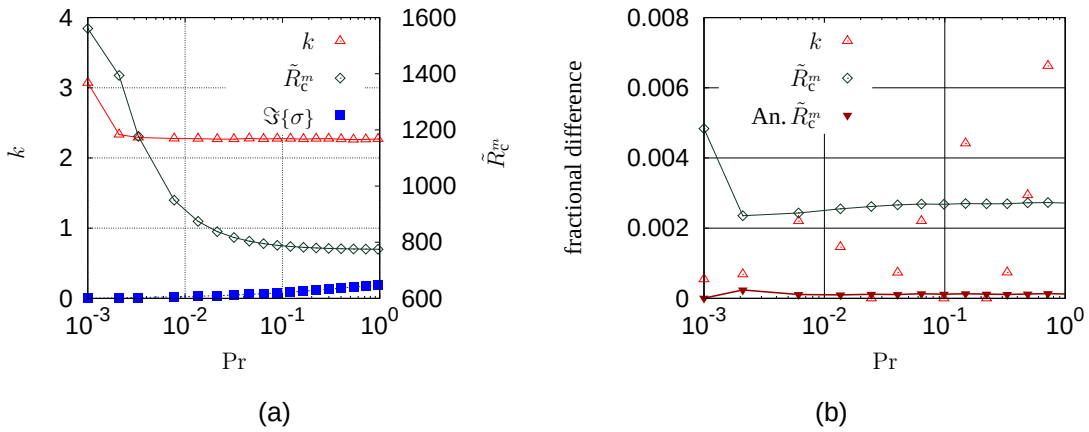



Figure 5.16: The parameter $\mathrm{Pr}$ is altered in (a) the compressible case and (b) the difference between the compressible and anelastic Lantz-Braginsky.

The viscosity is proportional to the parameter $\mathrm{Pr}$ and, as with magnetic diffusion, altering the thermal diffusion should not effect the accuracy of the anelastic approximation.

The effect of altering the Prandtl number is shown in figure 5.16. At very low Prandtl numbers the instability is hard to drive, but as the Prandtl number increases the minimum critical Rayleigh number decreases. The solution also passes through the TB point. The dimensionless numbers used in this section are not the same as used in the linear analysis of Chandrasekhar (1961) but equation (5.6) shows that the Prandtl number effects when an oscillatory mode can occur. The fractional difference shown in figure 5.16 (b) is below $0.5\%$ and, as the atmosphere is nearly adiabatic, it is expected that the anelastic

approximation will preform well. The large fractional difference on the first point is most likely due to the difficulties in calculating the critical Rayleigh number near the TB mentioned previously.

### 5.3.6 Isothermal to Isentropic Boundary Condition

The equations have so far been solved with an isothermal boundary condition. It is also equally valid to use an isentropic one as neither is particularly relevant physically however they are illustrative. Changing the boundary condition can be done using a continuation method and setting the thermodynamic boundary condition to be

$$0 = \lambda s + (1 - \lambda)\, T,$$

where $\lambda \in [0, 1]$ is a new parameter to allow continuity when changing from one boundary condition to another. In the anelastic case the entropy equation was used and so the isothermal boundary condition is obtained using equation (3.20h), that is

$$s = \frac{1 - \gamma}{\gamma} \frac{p}{\overline{p}} + \frac{T}{\overline{T}}. \tag{5.9}$$

This means the isothermal boundaries condition, in terms of the entropy, can be expressed as

$$s = \frac{1 - \gamma}{\gamma} \frac{p}{\overline{\rho} \overline{T}}, \tag{5.10}$$

on the boundary, with pressure being extrapolated so that its value can be calculated outside the domain.

Figure 5.17 shows how magnetoconvection is affected by $\lambda$. As the boundary condition changes from isothermal ($\lambda = 0$) to isentropic ($\lambda = 1$) then the growth-rate decreases and so the layer is becoming more stable. This means that it is harder to excite modes of the magnetoconvection instability in a layer with an isentropic boundary condition than a layer with an isothermal boundary condition. The effect the boundary conditions have on the Rayleigh number, corresponding wavenumber, and the imaginary component of

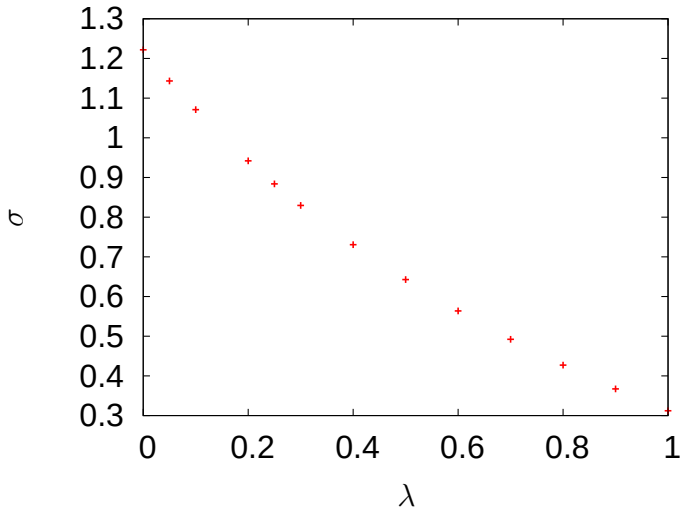



Figure 5.17: Showing the continuous change for $k = 3.02$ from the $s = 0$ to $T = 0$ boundary condition with $\zeta = 0.01$, $\theta = 3.01$, and $R = 1650$. These are the parameters at the most unstable mode in the isentropic case.

the eigenvalue are shown in figures 5.18 to 5.20, here $\mathrm{Pr} = 1.0$, $Q = 20$ and other dimensionless numbers are in the captions.

Figure 5.18 shows there is a fairly rapid change in the minimum critical Rayleigh number at low $\lambda$ values from $\tilde{R}_c^m = 1000$ at $\lambda = 0$ to $\tilde{R}_c^m = 1300$ at $\lambda = 0.2$. The rapid change at low $\lambda$ values is more pronounced at larger $\zeta$ values; figure 5.18 (a), with $\zeta = 0.5$, has a range $600$ of the minimum critical Rayleigh numbers whereas in 5.18 (b), with $\zeta = 0.01$, the range is only $300$. The wavenumber corresponding to the most unstable mode also has a rapid change at low $\lambda$ values. Increasing $\zeta$, whilst holding the other dimensionless numbers constant, is equivalent to increasing the magnetic diffusivity. At the higher $\zeta$ the solution is oscillatory and for the lower $\zeta$ value the solution is steady with the imaginary component of the eigenvalue $\Im\{\sigma\} = 0$ for all $\lambda$ values in figures 5.18 (a), 5.19 (a), and 5.20 (a).

In figure 5.19 the atmosphere is less stratified than in figure 5.18 and the effect of altering the boundary condition from isentropic to isothermal is reduced. For the two values

of $\zeta$, the difference in the minimum critical Rayleigh number between the isothermal and isentropic boundary condition is shown in figure 5.19 (a) and (b). Figure 5.20 shows a nearly Boussinesq case with $\chi \sim 1$ and $\theta \ll 1$ and the different boundary conditions have no effect on the imaginary component of the eigenvalue and almost no effect on the critical Rayleigh number and corresponding wavenumber. This is expected as in the Boussinesq case there are no pressure fluctuation so temperature and entropy are proportional so a change from isentropic to isothermal boundaries would make no difference to the equations.

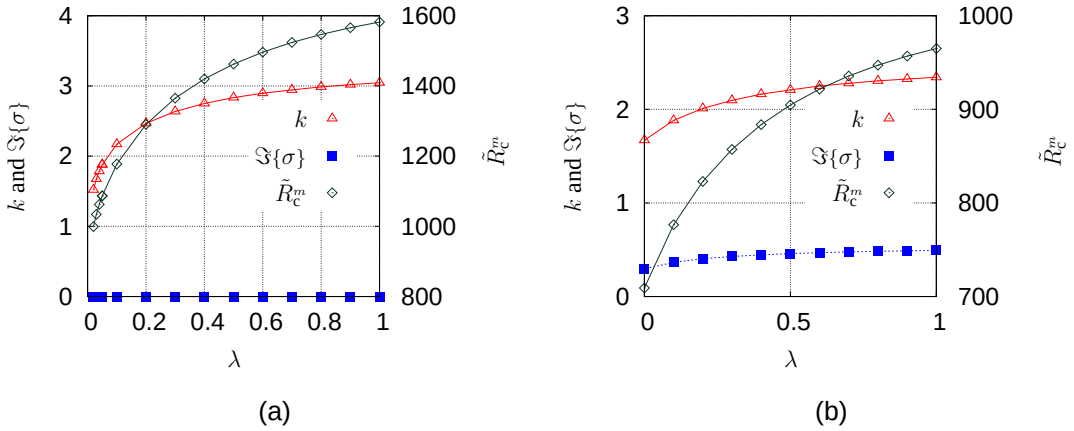
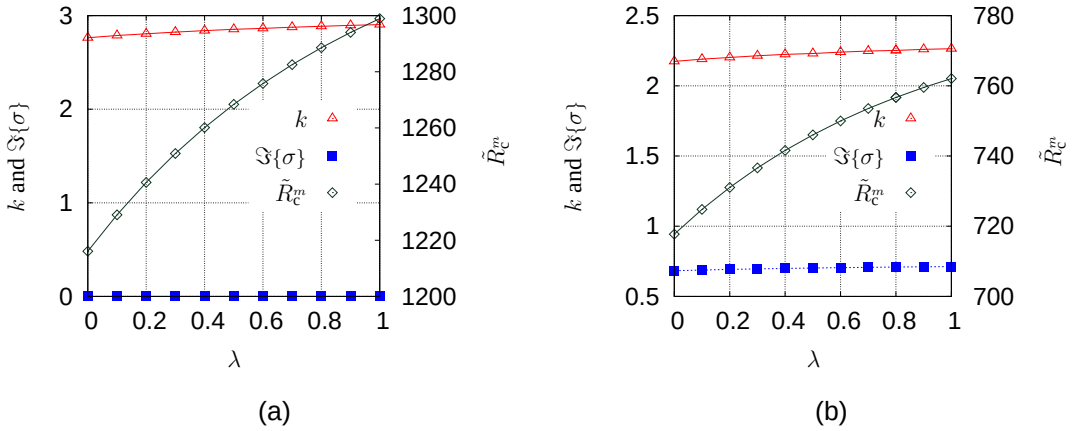



Figure 5.18: The change from the isentropic to isothermal boundary condition with $\chi = 8$, $\theta = 3.01$, and $\zeta =$ (a) $0.5$ and (b) $0.01$. The minimum critical Rayleigh number $\tilde{R}_c^m$ and corresponding wavenumber are shown along with the imaginary component of the eigenvalue $\Im\{\sigma\}$.

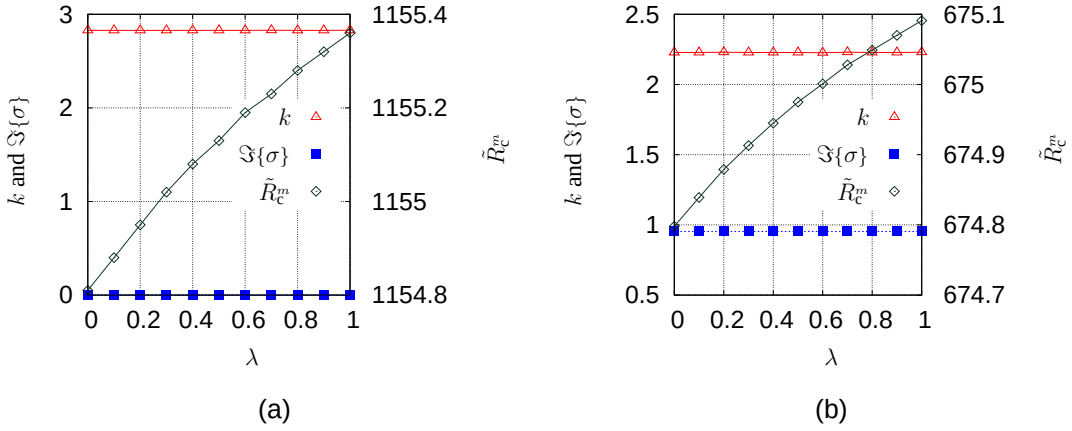
### 5.3.7 Chandrasekhar Exponent Dependences

I also investigated some of the parameters dependence on the Chandrasekhar number $Q = B_0^2 d^2/(\mu_0 \mu \eta)$ and how the exponents altered as the atmosphere became more stratified [see appendix C for the relation between commonly used dimensionless numbers in this thesis]. The power law scaling obtained in Chandrasekhar (1961) are




(a) (b)

Figure 5.19: As figure 5.18 but with $\chi = 3$, $\theta = 1.08$, and $\zeta =$ (a) $0.5$ and (b) $0.01$.




(a) (b)

Figure 5.20: As figure 5.18 but with $\chi = 1.1$ and $\theta = 0.065$ and $\zeta =$ (a) $0.5$ and (b) $0.01$.

for the Boussinesq limit. The analytical Boussinesq results for sufficiently large $Q$ are that $k \propto Q^{1/6}$, $\tilde{R}_{\mathrm{c}} \propto Q^{0.976}$, and $\Im\{\sigma\} \propto Q^{0.5}$. Figure 5.21 shows that with a low value of $\theta$ then the exponents are similar to the analytical results in Chandrasekhar (1961). To confirm the exponents an oscillatory and steady instability were studied, in practice this meant using a small value and large value of $\zeta$ so the TB point was in between the two $\zeta$ values. In figure 5.23, $\theta$ was larger than in figure 5.22 and the difference between the analytical and computed exponents became larger, which was expected as the analytical exponents are for $\theta = 0$. When $\theta$ increased the exponent of the wavenumber dependence on $Q$ increased and the exponent of the minimum critical Rayleigh number dependence on $Q$ decreased. As $\zeta$ increased then the exponents of both instability parameters $\tilde{R}_{\mathrm{c}}^{m}$ and $k$ dependence on $Q$ increased, e.g. in figure 5.22 (a) where $\zeta = 0.05$ the dependence was $k \propto Q^{0.175}$ but when $\zeta = 0.5$ as in figure 5.22 (b) then the exponent for the wavenumber dependence on $Q$ increased to $k \propto Q^{0.177}$. The line of best fit was only calculated when the data lay close to a straight line, for example in figure 5.22 (a) the best fit line was calculated from points $Q > 10^4$.

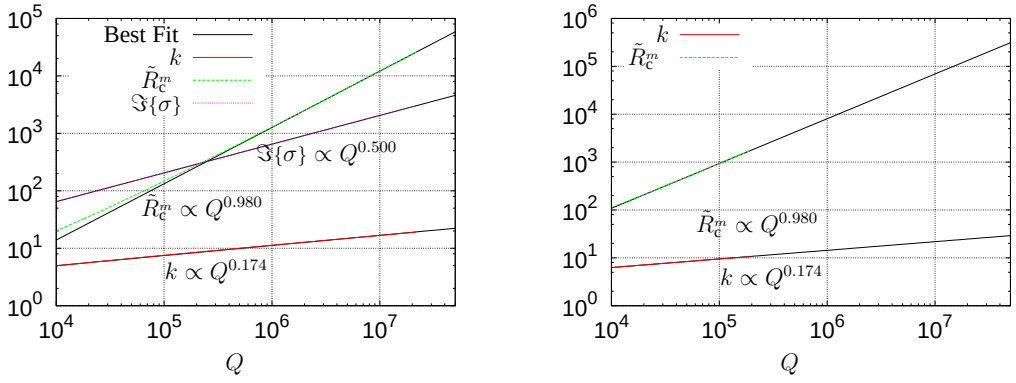



Figure 5.21: The dependencies on $Q$, the Chandrasekhar number. This is on a log-log plot with $\theta = 0.0658, (\chi = 1.1)$. The lines of best-fit and the exponents are also plotted. In (a) $\zeta = 0.1$ so the solution oscillated, and in (b) $\zeta = 0.5$ so the solution did not oscillate.

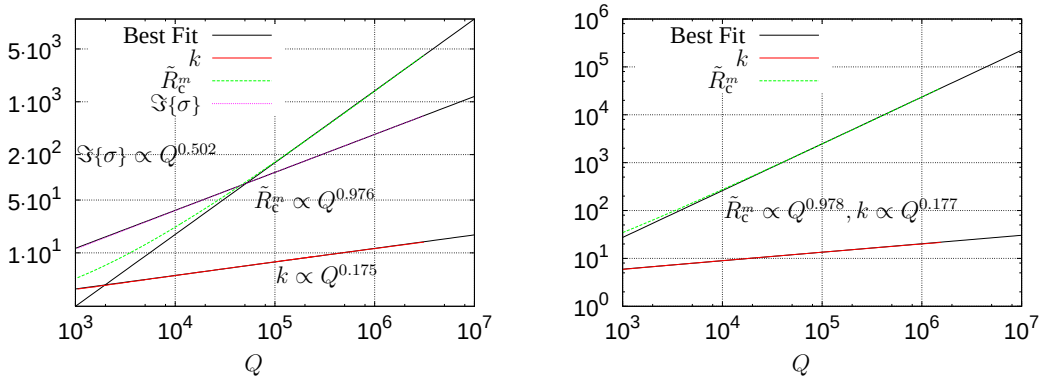



Figure 5.22: Same as figure 5.21 except with $\theta = 1.08$ ($\chi = 3.0$). In (a) $\zeta = 0.05$ so the solution oscillated, and on (b) $\zeta = 0.5$ so the solution did not oscillate.

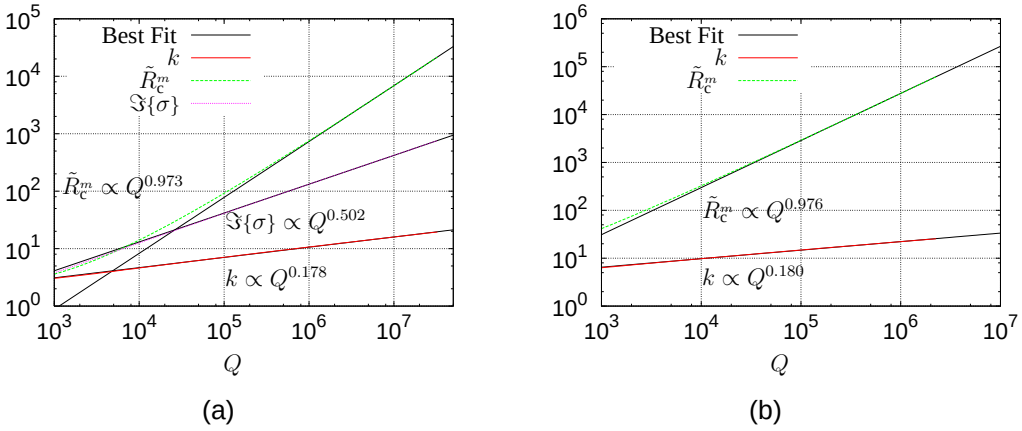



Figure 5.23: Same as figure 5.21 except with $\theta = 3.01, (\chi = 8.0)$. In (a) $\zeta = 0.01$ so the solution oscillates, and in (b) $\zeta = 0.5$ so the solution does not oscillate.

### 5.3.8 Tilted Field

So far I have only considered the case of a vertical magnetic field where $\mathbf{B}_b^* = (0, 0, 1)$. It is also possible to have a tilted field where $\mathbf{B}_b^* = (\sin\phi, 0, \cos\phi)$ and $\phi$ is the angle of the magnetic field from the vertical, with $\phi = 0$ corresponding to case being previously discussed in this chapter. The basic state is the same in the tilted field case as in §5.2.2 and the boundary conditions are the same as those in §5.2.1 except for the magnetic field boundary condition. The magnetic boundary condition becomes $B_z \sin\phi + B_x \cos\phi = 0$ at $z = 0, 1$ . This was only investigated in the anelastic case and the linear equations are given in Appendix B.

The interest in inclined fields is due to sunspots as it is thought that inclined fields may be responsible for some of the features seen. With an inclined field the symmetry of clockwise and anti-clockwise oscillations is broken. The perturbation terms, which result from a decomposition of the form $f(z)\exp(i\mathbf{k}\cdot\mathbf{x} + \sigma t)$, are calculated to find the eigenvalue. As time increases if $\Re\sigma > 0$ and $\Im\sigma \neq 0$ then the solution will increase and, if plotted on an Argand diagram, the instability will rotate around the origin with a clockwise or anti-clockwise as the solution spirals outwards. Figure 5.24 shows that instabilities which oscillate clockwise, where $\Im\sigma < 0$, occur at lower Rayleigh number and are more unstable than those which rotate anti-clockwise, where $\Im\sigma > 0$, when $\phi = 45°$. Figure 5.24 shows a plot of the critical Rayleigh number for the two rotation directions at $\phi = 45°$ and as the wavenumber is altered there is a minimum critical Rayleigh number at which the system becomes unstable.

Figure 5.24 shows how the critical Rayleigh number changes as the wavenumber is increased. This compares well with the compressible results in Matthews et al. (1992). The parameter values used are: $\mathrm{Pr} = 1,\ m = 1.495,\ \zeta = 0.05,\ \gamma = 5/3,\ \theta = 1.09$ (to 3 s.f.), $\chi = 3,\ Q = 20$.

An Argand diagram of eigenvalues as the Rayleigh number increases is shown in figure 5.25. In both the figures 5.25 (a) and (b) at a low Rayleigh there are two branches,

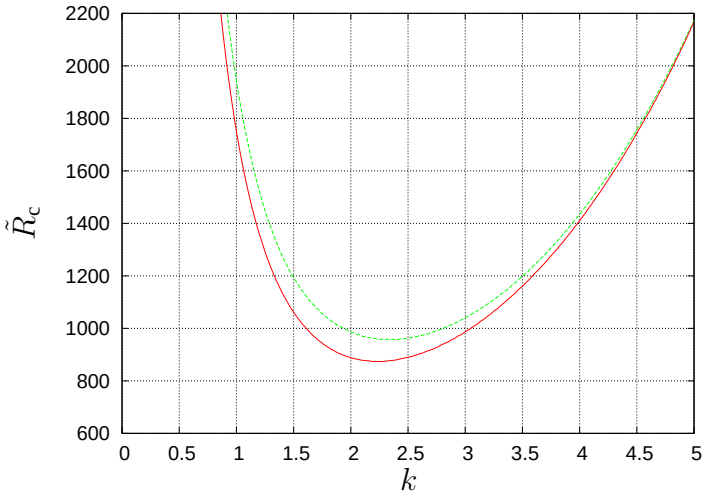



Figure 5.24: The critical Rayleigh number as a function of $x$ wavenumber with $\phi = 45°$, green dashed for $i > 0$ and red for $i < 0$

one with positive and the other with a negative imaginary part. As the Rayleigh number is increased then both branches in figure 5.25 (a) become unstable, in that the real part of both branches becomes positive. The positive branch then turns back and becomes stable again at higher Rayleigh numbers whereas the negative branch remains unstable. The branch which remains stable changes as $k$ is increased as shown in figure 5.25 (b) where now it is the positive branch which remains unstable. This change of stability in the branches is called 'stability reversal' and is discussed in Hurlburt et al. (1996) where they find that the stability reversal is very sensitive to boundary conditions. The results in Roxburgh (2007) are not the same as shown in figure 5.25 but this discrepancy has been put down to the sensitivity of the stability reversal. Roxburgh found that the stability reversal occurred at $k = 4.2$ whereas I found the reversal to occur at $k = 4.8$.

## 5.4 Summary

The anelastic approximation accurately captures the magnetoconvection instability when the atmosphere is nearly adiabatic, the magnetic field is weak, and the temperature flux

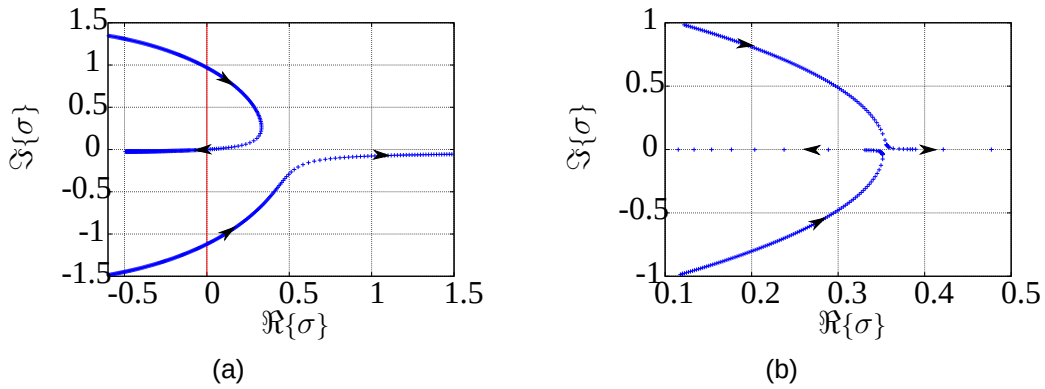



Figure 5.25: An Argand diagram of the eigenvalue (a) $k = 3.0$ and (b) $k = 4.9$ with $\phi = 60°$. The arrows point in the direction of increasing Rayleigh number.

gradient $\theta$ is small. This is true for the Lantz-Braginsky approximation as well as the full anelastic approximation.

When the atmosphere departs from being adiabatic then the full anelastic approximation is able to reproduce the instability more accurately than the Lantz-Braginsky approximation. The term that is neglected in the Lantz-Braginsky is small if and only if the atmosphere is nearly adiabatic. When the temperature flux gradient is large the anelastic approximation produces results with a larger fractional difference compared to the fully compressible results and the term neglected in the Lantz-Braginsky simplification is small so the difference between the two anelastic approximations also remains small. A strong magnetic field also causes large fractional differences between the anelastic and fully compressible cases.

Changing between isothermal and isentropic boundary conditions makes the magnetoconvection instability onset at higher Rayleigh numbers and wavenumbers. This effect is more pronounced when the temperature flux gradient is larger.

Various parameters dependencies on $Q$, obtained analytically in Chandrasekhar (1961) for the Boussinesq limit, were investigated in the anelastic approximation. As the temperature flux gradient increases, departing from the Boussinesq limit, then there are deviations from the analytical dependencies but these are small.

# Chapter 6

# Linear Magnetic Buoyancy Results

## 6.1 Linear Theory of Magnetic Buoyancy

It is possible to set out, in a general way, an argument based on density of how the magnetic buoyancy instability operates. If an element of fluid was lifted a small distance vertically in a stably stratified atmosphere then it would be heavier than its surroundings and would sink, overshooting and, subsequently, oscillating about its initial point at the Brunt-Väisälä frequency. In investigating the magnetic buoyancy instability it is helpful to use a simple parcel argument to understand the nature of the instability which was considered, in the absence of dissipation, by Acheson (1979) and later by Hughes (2007). I will consider a gravitationally stably stratified atmosphere in equilibrium with a horizontal magnetic field. I will simplify matters by considering a simple case where no field lines are twisted or bent, this is called an interchange mode which will be discussed in more depth later. A parcel is vertically displaced from $z$ to $z + \mathrm{d}z$, so that the parcel properties change from $\phi_{\text{int}}$ to $\phi_{\text{int}} + \delta\phi_{\text{int}}$ and the external properties change from $\phi_{\text{ext}}$ to $\phi_{\text{ext}} + \mathrm{d}\phi_{\text{ext}}$.

For a fully compressible fluid without magnetic diffusion it can be shown that $\mathbf{B}/\rho$ is

advected with the fluid

$$\left(\frac{\partial}{\partial t}+\mathbf{u}\cdot\nabla\right)\left(\frac{\mathbf{B}}{\rho}\right)=\left(\frac{\mathbf{B}}{\rho}\right)\cdot\nabla\mathbf{u}, \tag{6.1}$$

but this is not true in the anelastic equations or with the addition of diffusion. In the parcel argument the atmosphere is diffusionless and compressible so the mass per unit length and the flux are conserved, the quantity $B/\rho$ is thus conserved

$$\frac{B+\delta B}{\rho+\delta\rho}=\frac{B}{\rho},\ \text{i.e.}\ \frac{\delta B}{B}=\frac{\delta\rho}{\rho}. \tag{6.2}$$

Assuming that the parcel is adiabatically displaced then the specific entropy is conserved

$$\frac{\delta p}{p}=\gamma\frac{\delta\rho}{\rho}. \tag{6.3}$$

If the parcel is displaced slowly it will remain in pressure equilibrium with the surroundings

$$\delta\left(p+\frac{B^2}{2\mu_0}\right)=\mathrm{d}\left(p+\frac{B^2}{2\mu_0}\right)\Rightarrow\delta p+\frac{B\delta B}{\mu_0}=\mathrm{d}p+\frac{B\mathrm{d}B}{\mu_0} \tag{6.4}$$

For the magnetic buoyancy instability to occur the parcel density must be less than that of the new surroundings, $\delta\rho<d\rho$. This condition is combined with (6.2 - 6.4) to become an instability inequality

$$\left(\frac{B^2}{\mu_0\rho}+\frac{\gamma p}{\rho}\right)\delta\rho=\mathrm{d}p+\frac{B\mathrm{d}B}{\mu_0}, \tag{6.5}$$

which can be divided by $\rho\mathrm{d}z$ to get

$$\frac{\gamma p}{\rho^2}\frac{\delta\rho}{\delta z}-\frac{1}{\rho}\frac{\mathrm{d}p}{\mathrm{d}z}=\frac{B}{\mu_0\rho}\frac{\mathrm{d}B}{\mathrm{d}z}-\frac{B^2}{\mu_0\rho^2}\frac{\delta\rho}{\delta z} \tag{6.6}$$

or

$$\frac{B^2}{\mu_0\gamma p}\frac{\mathrm{d}}{\mathrm{d}z}\ln\left(\frac{B}{\rho}\right)>\frac{-g}{\gamma}\frac{\mathrm{d}}{\mathrm{d}z}\ln\left(p\rho^{-\gamma}\right)=N^2, \tag{6.7}$$

where $N$ is the Brunt-Väisälä frequency, and $\gamma=c_p/c_v$ is the ratio of the specific heats.

An important feature of (6.7) is that a magnetic field that decreases sufficiently rapidly with height can destabilise a convectively stable atmosphere.

The initial assumptions stipulated that the field lines were not bent so (6.7) is valid for two-dimensional modes called interchange modes occurring when one magnetic flux tube exchanges position with another. Another mode which may occur is a three-dimensional mode where the flux tubes bend. Early theoretical work by Newcomb (1961) looked at the stability of interchange and three-dimensional modes in ideal plasmas using the energy principle of Bernstein et al. (1958). He showed that a necessary and sufficient condition for the atmosphere to be stable to interchange modes was, after some manipulation,

$$\frac{c^2}{\rho}\frac{\mathrm{d}\rho}{\mathrm{d}z} - \frac{c^2}{\gamma p}\frac{\mathrm{d}p}{\mathrm{d}z} > \frac{a^2}{B}\frac{\mathrm{d}B}{\mathrm{d}z} - a^2\frac{\mathrm{d}\rho}{\mathrm{d}z} \qquad \text{where } a^2 = \frac{B^2}{\mu_0\rho},\ c^2 = \frac{\gamma p}{\rho} \tag{6.8}$$

or

$$-\frac{\mathrm{d}\rho}{\mathrm{d}z} > \frac{\rho g}{a^2 + c^2}. \tag{6.9}$$

which is an equivalent result to that obtained by the parcel argument.

An interesting feature of magnetic buoyancy is that three-dimensional modes can be more destabilising than interchange modes. Newcomb (1961) and, re-written explicitly showing the role of the magnetic field, Thomas & Nye (1975) showed that three-dimensional modes occurred if and only if the following inequality was satisfied somewhere in the fluid:

$$\frac{-g}{c^2}\frac{\mathrm{d}}{\mathrm{d}z}\ln B > k_x\left(1 + \frac{k_z}{k_y}\right) + \frac{N^2}{a^2}, \tag{6.10}$$

where $k_x$, $k_y$, and $k_z$ are wavenumbers in the $\hat{\mathbf{x}}$, $\hat{\mathbf{y}}$, and $\hat{\mathbf{z}}$ directions respectively. Simplistically, it would seem that three-dimensional modes must do extra work against magnetic tension but for interchange modes, work is done against thermal pressure and magnetic pressure to create a density perturbation. In three-dimensional instabilities the long variations in direction of the field allow the work done against magnetic pressure to be minimised. As the variations in the direction of the field are so long the benefits of minimising the work against the magnetic pressure outweigh the negligible extra work done against magnetic tension. The condition is necessary but not sufficient for instabilities to form (Hughes & Cattaneo, 1987).

Introducing diffusion adds complexity to the parcel argument. Acheson & Gibbons (1978) included the role of diffusivity which extends (6.10) to give

$$\frac{-ga^2}{c^2}\frac{\mathrm{d}}{\mathrm{d}z}\ln B > \frac{k_x}{\gamma}\left(1+\frac{k_z}{k_y}\right)+\zeta N^2, \tag{6.11}$$

where $\zeta = \eta c_p \rho_0 / k$ is the diffusivity ratio, the inverse of the Roberts number. By considering a parcel argument it appears beneficial to the instability to have small magnetic diffusion, $\eta$ (which helps maintain the destabilising field), and large thermal conductivity, $k$ (to reduce the stabilising entropy gradient). The laminar values in the Sun satisfy $\eta \ll k$ so the stability is greatly reduced, although it can be argued that the laminar values are inappropriate and the diffusion rates are all of order 1.

Diffusion can change the equations for interchange instabilities to be the same as the thermosolutal convection in the double diffusive case. I will treat the addition of diffusion in two cases: where the magnetic field decreases with height so that $\zeta \gg 1$, and the case where the field increases with height so that $\zeta \ll 1$.

In a decreasing field a rising parcel moves to a region with weaker magnetic field, assuming that $\zeta \gg 1$, then the magnetic field diffuses but the temperature does not. A decrease in flux means a decrease in magnetic pressure will be compensated by an increase in thermal pressure and density. The parcel is now denser than its surroundings so will sink and repeat a similar process. This can result in a situation where an increase of stabilising gradients can cause the instability to become stronger by matching the natural frequency of this overshooting process of repeated rising and sinking (see e.g. Hughes, 2007).

In the case where the field increases with height, and $\zeta \ll 1$, then a rising parcel is compressed by the pressure of the new background. The compression causes the element to be hotter than its surroundings. As the thermal diffusivity is much larger it loses its temperature, but not its magnetic field, so when it returns it will have lost heat and will consequentially be denser and so overshoot.

## 6.2 Linear Code Modelling Magnetic Buoyancy

Using the algorithm described in Chapter 4, I investigated the magnetic buoyancy in the fully compressible and anelastic (with and without the Lantz-Braginsky approximation) cases for a plasma in a layer using a Cartesian geometry. I then compared the anelastic cases with the compressible case to see in which parameter regimes the anelastic approximation fails to capture the compressible results. I am assuming that the plasma is modelled perfectly by the compressible system. The parameters used do not invalidate the assumptions used to derive the compressible equations and so the algorithm should be able to model this well. The fully compressible equations are in (2.3), the anelastic equations without the Lantz-Braginsky are given in (3.20), and with the Lantz-Braginsky approximation are given in §3.7.

### 6.2.1 Boundary Condition

The boundary condition in the magnetic buoyancy case is similar to that used in magnetoconvection described in §5.2.1 where the top and bottom boundaries were impermeable, stress free and isothermal so

$$\hat{w} = \frac{\partial}{\partial z}\hat{u} = \frac{\partial}{\partial z}\hat{v} = \hat{T} = 0 \quad \text{at } z = 0, 1$$

where the hat terms are from the expansion in (4.1). I also imposed the top and bottom magnetic field boundary conditions of

$$\frac{\partial \hat{B}_x}{\partial z} = \frac{\partial \hat{B}_y}{\partial z} = \hat{B}_z = 0, \tag{6.12}$$

which corresponds to a horizontal field whilst satisfying the solenoidal condition (2.1b). The basic state is different for the anelastic and fully compressible equations as mentioned in §5.2.2. The basic state used to study magnetic buoyancy instabilities is more involved then the basic state used in magnetoconvection.

### 6.2.2 Basic State for the fully Compressible Equations

In the compressible case there is no reference state so the basic state is that of a polytrope, but under the influence of a weak magnetic field. The basic state of the compressible equations (2.3) is denoted $\xi_b$, where $\xi$ can represent any variable. I consider a steady, stationary basic state given by a polytropic solution together with an imposed magnetic field, given by $\mathbf{B}_b = (B_{xb}(z), 0, 0)$ where $B_{xb} = 1 + H_b z$ and is a linear function of depth with gradient $H_b$. So, from equation (2.6d), ohmic heating leads to a departure of the temperature distribution from a polytrope, such that

$$T_b = -\frac{(\gamma - 1)}{\gamma}\mathcal{F}\zeta C_k^2 \,\mathrm{Pr}\, R \left(\frac{\mathrm{d}B_{xb}}{\mathrm{d}z}\right)^2 \left(z^2 - 2z\right) + \theta z + 1, \tag{6.13}$$

where $\theta$ is the temperature flux gradient at the bottom boundary, $z = 1$. Then, from the vertical momentum balance, see equation (2.6b), I obtain the basic state density; the solution of the equation

$$\frac{\mathrm{d}\rho_b}{\mathrm{d}z} + \left(\frac{\mathrm{d}T_b}{\mathrm{d}z} - 1\right)\frac{\rho_b}{T_b} = -\mathcal{F}\frac{B_{xb}}{T_b}\frac{\mathrm{d}B_{xb}}{\mathrm{d}z} \tag{6.14}$$

is computed numerically using a Runge-Kutta solver. It is assumed that the horizontal magnetic field is weak enough not to alter significantly the density stratification which is still accurately represented by equation (5.8).

In the magnetic buoyancy instability the basic state, for a typical case with $\mathcal{F} < 0.001$, is still very similar to that of a polytrope, shown in figure 6.1 (a). Figure 6.1 (b) shows that for stronger magnetic fields the basic state is distorted and it also shows that as the magnetic field increases there can be a situation where the basic state can be top-heavy. Top-heavy means that denser fluid is sitting atop lighter fluid. Top-heavy states were not included in the study of the magnetic buoyancy instability as the Rayleigh-Taylor instability would also be occurring which, although an interesting instability, is not relevant to the solar interior. The temperature profile changes only slightly from a polytrope due to ohmic heating as shown in figure 6.2 where, for even the top-heavy state in figure 6.2 (b), there are only minor deviations from a polytrope.

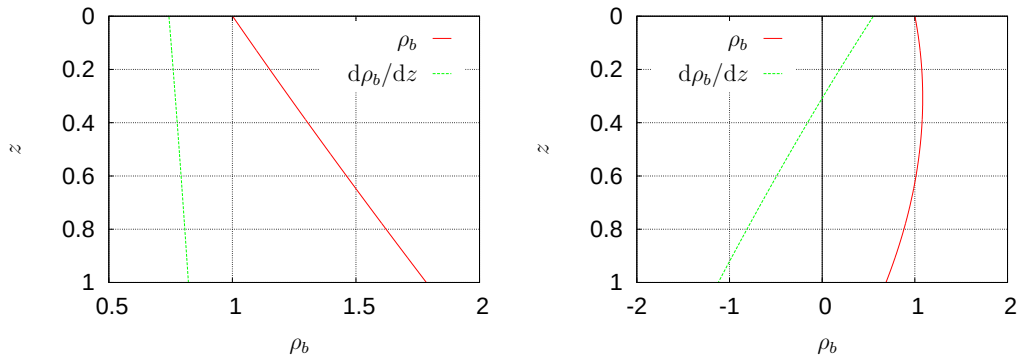



Figure 6.1: The compressible basic state density profile for a typical case where $\gamma = 5/3$, $m = 1.505$, $\theta = 0.5$, $C_k = 0.01$, $H_b = 10$, $\mathrm{Pr} = 0.5$ and (a) $\mathcal{F} = 0.001$ and (b) $\mathcal{F} = 0.02$ which is top-heavy.

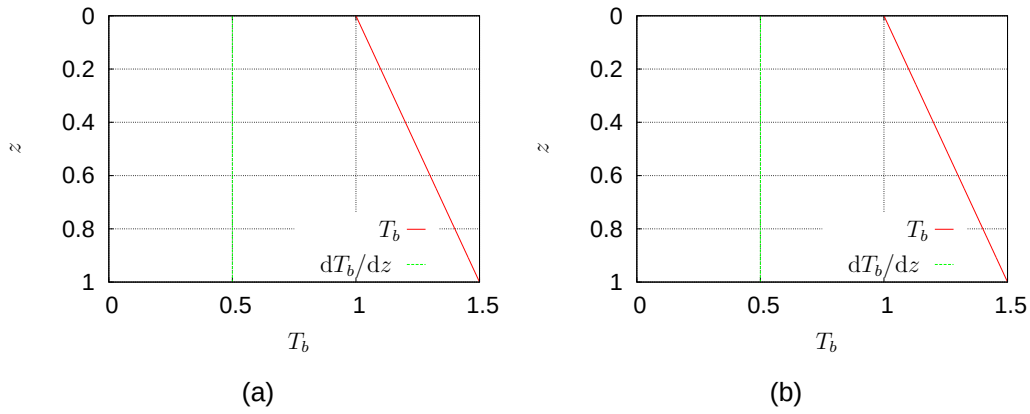



Figure 6.2: As for 6.1 but now showing the compressible basic state temperature profile.

### 6.2.3 Basic State for the Anelastic Equations

As there is no magnetic field in the anelastic reference state, the field only appears in the basic state. Imposing a field and not altering the entropy basic state would mean there would be no equilibrium so any results would be affected by the initial adjustment to the imposed field. For magnetic buoyancy the field $\mathbf{B}_b^* = (B_{xb}^*(z), 0, 0)$ is horizontal and varies linearly with depth as $B_{xb}^* = 1 + H_b z$. The solution of the entropy equation (3.20f) gives the basic state temperature

$$T_b^* = -\frac{\gamma - 1}{\gamma}\tilde{\mathcal{F}}\zeta\tilde{C}_k^2 \mathrm{Pr}\tilde{R}\left(\frac{\mathrm{d}B_{xb}^*}{\mathrm{d}z}\right)^2\left(\frac{z^2}{2} - z\right), \tag{6.15}$$

satisfying $T_b^* = 0$ at $z = 0$ and $\mathrm{d}T_b^*/\mathrm{d}z = 0$ at $z = 1$. For stationary solutions, which are functions of depth only, the $z$-component of the momentum equation (3.20b) reduces to

$$\frac{dp_b^*}{dz} - \frac{\overline{\rho}}{\overline{p}}p_b^* = -\frac{\overline{\rho}}{\overline{T}}T_b^* - \mathcal{F}B_{xb}^*\frac{dB_{xb}^*}{dz} \quad \text{with} \quad p_b^* = 0 \text{ at } z = 0. \tag{6.16}$$

The basic state pressure is obtained by integrating numerically the above equation using a Runge-Kutta solver. (From equation (3.20g) the boundary condition for $p_b^*$ is consistent with $\rho_b^* = T_b^* = 0$ at $z = 0$.) The basic state entropy, $s_b^*$, is found algebraically using equation (3.20h), with typical profiles shown in figure 6.3 where it can be seen that $s_b^* \propto \mathcal{F}$.

## 6.3 Linear Magnetic Buoyancy Results

The algorithm described in Chapter 4 was used to investigate the magnetic buoyancy instability. In magnetic buoyancy the algorithm finds the maximum growth rate over $k_x$ and $k_y$ space. For historical reasons the diagnostic for magnetic buoyancy problems is the growth rate of the instability, rather than a critical parameter, and this is the diagnostic which I will use in the following discussion (Acheson, 1979).

For magnetic buoyancy, the polytropic index is fixed to subadiabatic values, so that the layer of fluid is weakly stable to thermal convection. However, in a stratified atmosphere

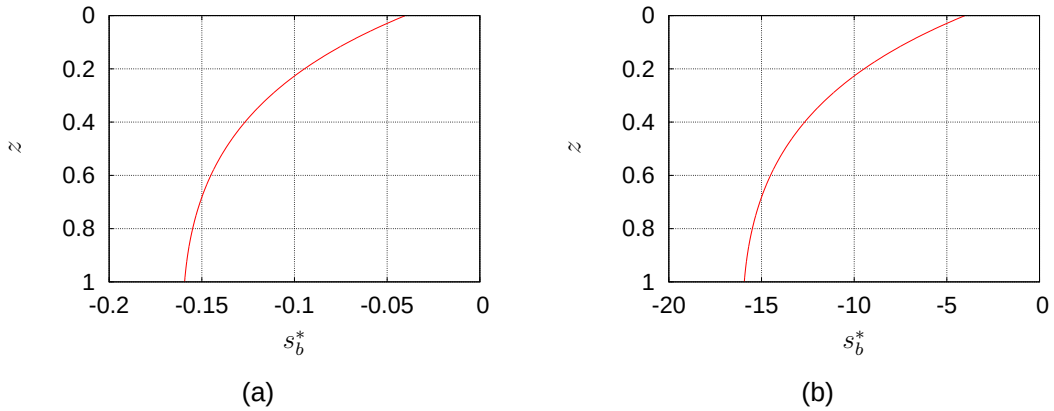



Figure 6.3: As for 6.1 but showing the anelastic entropy basic state profile for (a) $\mathcal{F} = 10^{-5}$ and (b) $\mathcal{F} = 0.001$.

the basic state magnetic field, horizontal and increasing with depth, can be unstable to the magnetic buoyancy instability. The growth rate is scaled using sound crossing time-scale, $d/\sqrt{p_0/\rho_0}$. I selected this because using a dynamical time means that the scaling applied to time varies when $\theta$ or $m$ are altered. To convert the growth-rate from sound crossing to dynamical time-scale multiply the growth-rate by $C_k\sqrt{\theta(m+1)}$. For all plots, unless otherwise stated, the following parameters are fixed as

$$\gamma = 5/3, \quad m = 1.505, \quad \theta = 0.5, \quad C_k = 0.01, \quad \mathrm{Pr} = 0.5,$$

$$\zeta = 5 \times 10^{-4}, \quad H_b = 10.$$

### 6.3.1 The effect of altering $\mathcal{F}$

The presence of a magnetic field in the $x$-direction differentiates between the two horizontal directions; a distinction between the two-dimensional interchange modes, with $k_x = 0$, and three-dimensional modes can be made (see e.g. Hughes, 2007). For a given profile of field, and with all the other parameters fixed, I therefore calculate the maximum value of $\Re\{\sigma\}$ when optimised over $k_x$ and $k_y$ and compare both the value of $\Re\{\sigma\}$ and the wavenumbers at which this growth rate is achieved, for the compressible and anelastic cases.

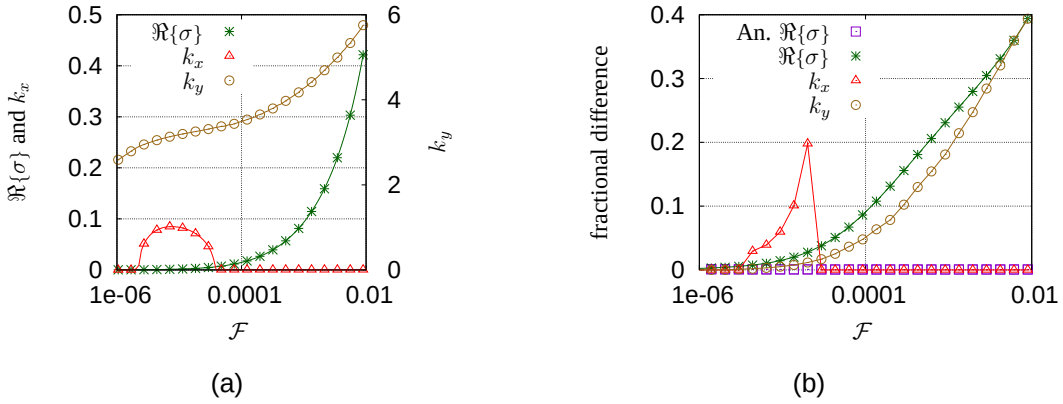



Figure 6.4: Variations with $\mathcal{F}$ of (a) the growth rate, $\Re\{\sigma\}$, and corresponding wavenumber, $k_x$ and $k_y$, of the most unstable mode for the compressible model; (b) the fractional differences in $\Re\{\sigma\}$, $k_x$ and $k_y$ between compressible and anelastic models, and the fractional difference in $\Re\{\sigma\}$ between anelastic models with and without the Lantz-Braginsky simplification, noted An. $\tilde{R}_c^m$. The value of $\epsilon = 10^{-3}$.

Figure 6.4 shows the effect of the magnetic field strength on the instability. Increasing the magnetic field increases the growth rate of the instability as it is magnetically driven. For the largest and weakest field strengths interchange modes are preferred, whereas in the intermediate regime the preferred mode becomes three-dimensional for both the anelastic and fully compressible systems. As the magnetic field is increased the fractional difference in all the parameters also increases. The fractional difference in $k_x$ can only have a non-zero value for the region where three-dimensional modes are preferred. It is interesting that, even for these cases where the system is close to adiabatic, the anelastic approximation performs badly when the field is strong and the instability has a large growth rate.

Next, I examine the difference between the eigenfunctions for compressible and the full anelastic equations. The eigenfunctions shown in figure 6.5 correspond to the most unstable modes (interchange) found for the same parameter values as figure 6.4 with a fixed $\mathcal{F} = 0.001$, where the anelastic approximation gives a 20% difference in the

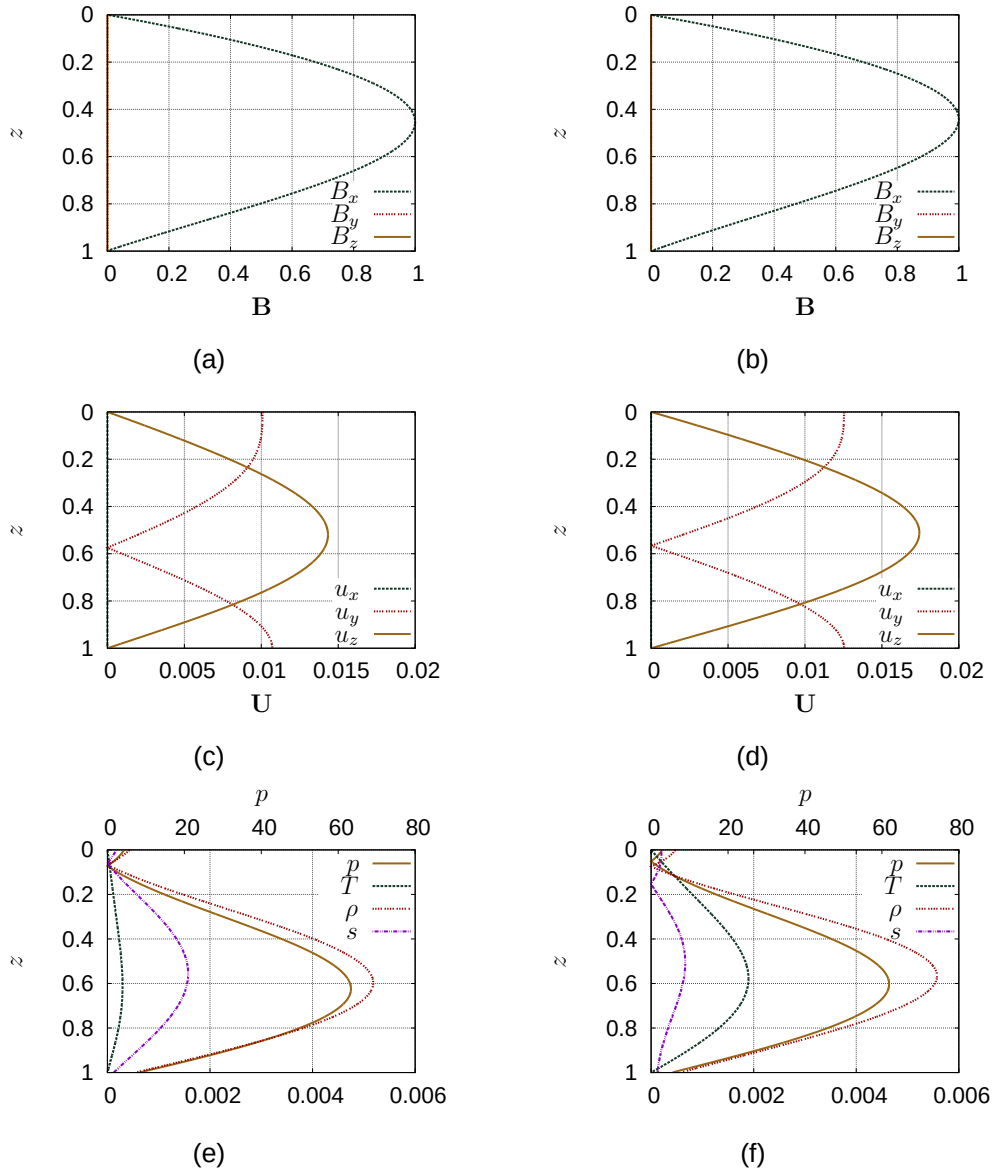



Figure 6.5: The absolute value of the eigenfunctions corresponding to the mode of maximum growth rate in figure 6.4 ($k_x = 0$ and $k_y \simeq 4.276$) for $\mathcal{F} = 0.001$; left and right panels correspond to the compressible and anelastic models respectively. (a)-(b) Magnetic field $B_x$ ($B_y$ and $B_z$ are zero for an interchange mode); (c)-(d) components of the fluid velocity $u_x$, $u_y$ and $u_z$; (e)-(f) thermodynamic variables $p$, $\rho$, $T$ and $s$.

real eigenvalue compared with the compressible equations. For both the compressible and anelastic models, I use the same normalisation, $\max(B_x) = 1$, to facilitate the comparison of the eigenfunctions. The significant difference in figure 6.5 (c) and (d) is the amplitude of the flows. The differences in the thermodynamic variables are, however, more significant. The relative amplitude of the thermodynamic variables has changed place and the profiles have been altered with the eigenfunctions passing through zero at a different depth. In the anelastic model the density does not have a hyperbolic evolution equation and I expect the thermodynamic variables to be the most affected by the approximation. The Lantz-Braginsky approximation was not used in the plot as the atmosphere was very close to being adiabatic it showed very little difference to the full anelastic equations.

### 6.3.2 The effect of altering $m$

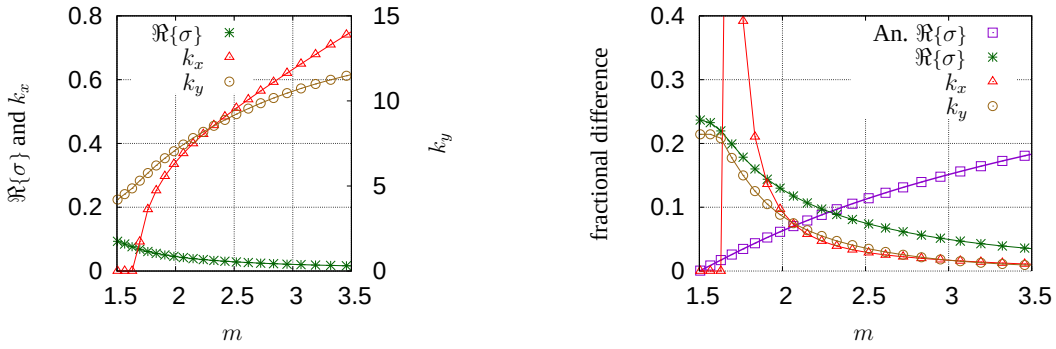



Figure 6.6: As for figure 6.4 but for variations vs. $m$. $\theta = 0.5$ is held fixed and $\mathcal{F} = 10^{-3}$. Here $\epsilon$ lies in the range $10^{-3} - 4 \times 10^{-1}$.

In increasing the polytropic index $m$ the atmosphere is becoming more stable and the parameter $\epsilon$ is becoming larger. The small parameter was used in the asymptotic expansion to derive the anelastic approximation in Chapter 3 and so when $\epsilon$ becomes large the assumptions are violated. In this subsection I assess when the assumptions underpinning the anelastic approximation are no longer valid.

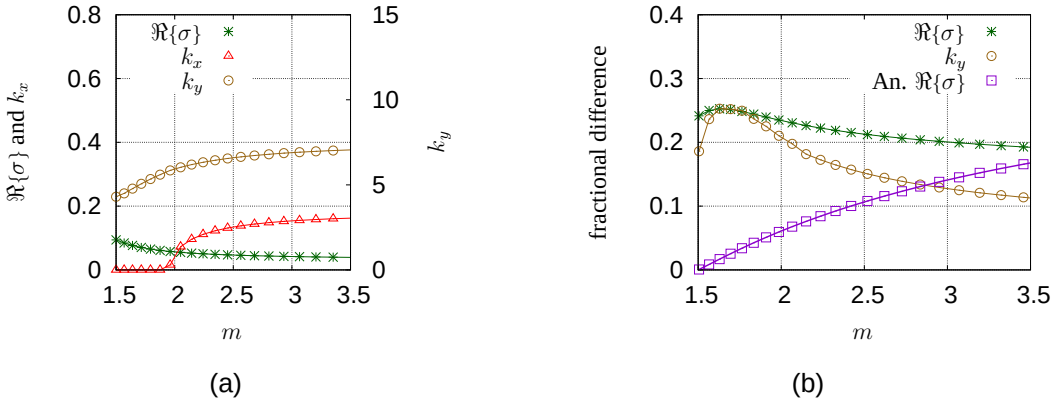



Figure 6.7: As for figure 6.6 but for $\chi$ held constant, again with $\mathcal{F} = 10^{-3}$.

Figure 6.6(a) shows the maximum growth rate and wavenumbers $k_x$ and $k_y$ for the compressible case when $m$ is altered with $\theta$ fixed, so the density stratification $\chi$ changes. $\chi$ is defined in equation (5.8). As $m$ is increased, so that the layer becomes more stably stratified, the maximum growth rate decreases, as might be expected. Also, there is a transition from two-dimensional modes to three-dimensional modes being preferred, although for all cases $k_y$ remains much larger than $k_x$. Figure 6.6(b) shows how the anelastic system compares with the fully compressible system as the system moves away from being adiabatic. Counter-intuitively, the anelastic system starts off as a poor approximation to the fully compressible system but the fractional difference decreases as $m$ is increased and the growth rate decreases. The anelastic approximation also captures the transition from two-dimensional to three-dimensional modes, though this occurs at a different value of $m$.

For figure 6.7 $m$ is altered with a fixed $\chi = 1.84$ (so that $\theta$ decreases as $m$ increases). Figure 6.7 shows a similar trend to figure 6.6, but where $\chi$ is held constant. The instability is an interchange mode when the atmosphere is close to being adiabatic. The anelastic approximation has a large error of greater than 20% for most values of $m$. As the polytropic index increases the gravity decreases (as $\chi$ is held constant) and the instability is weakened so that the growth rate decreases. The decreasing growth rate coincides with a reduced error in the anelastic approximation. The Lantz-Braginsky

approximation diverges from the anelastic equations (3.20) as $m$ is increased.

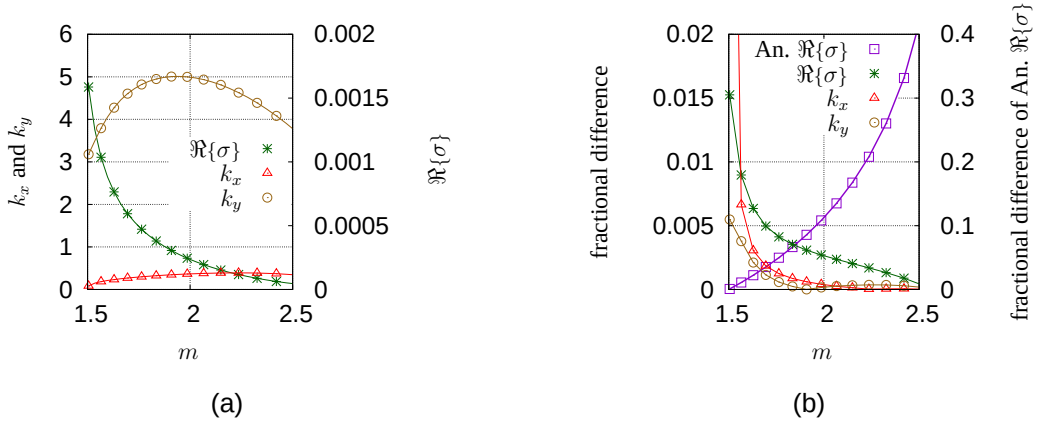



Figure 6.8: As for figure 6.6 for $\mathcal{F} = 10^{-5}$ where $\epsilon$ lies in the range $10^{-3} - 2\times10^{-1}$.

Figures 6.8(a) and (b) show similar behaviour as in figures 6.6(a) and (b) but with a weaker field. Note that the anelastic approximation is more accurate at these low magnetic field values. Figure 6.8(b) shows that the Lantz-Braginsky simplification, which requires the layer to be nearly adiabatic, agrees exactly with the anelastic approximation at $m = 1.5$ and has a fractional difference of $40\%$ compared to the full anelastic approximation by $m = 2.5$.

### 6.3.3 The effect of altering $C_k$

The thermal conduction and viscosity are proportional to $C_k$ and, from the derivation in Chapter 3, altering $C_k$ should not effect the accuracy of the anelastic approximation.

Figure 6.9 shows how the instability depends upon $C_k$. At low values of $C_k$ the instability grows more rapidly than at higher $C_k$ values as diffusion of heat (which leads to a loss of buoyancy) takes a long time. For an intermediate range of values of $C_k$ the mode of maximum growth rate is three-dimensional, but for larger $C_k$ an interchange mode becomes dominant. The anelastic approximation captures these transitions in wavenumber; however it does better as $C_k$ is increased when the growth rate of

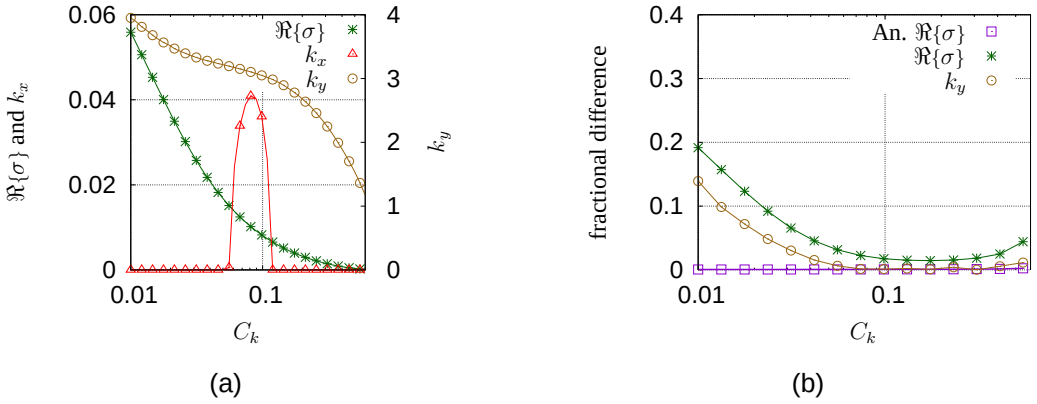



Figure 6.9: As for figure 6.4 but variations vs. $C_k$ with $\epsilon = 10^{-3}$ and $m = 1.505$.

the instability decreases. Because the model is nearly adiabatic the Lantz-Braginsky approximation reproduces the anelastic equations (3.20) results very well.

### 6.3.4 The effect of altering $\theta$

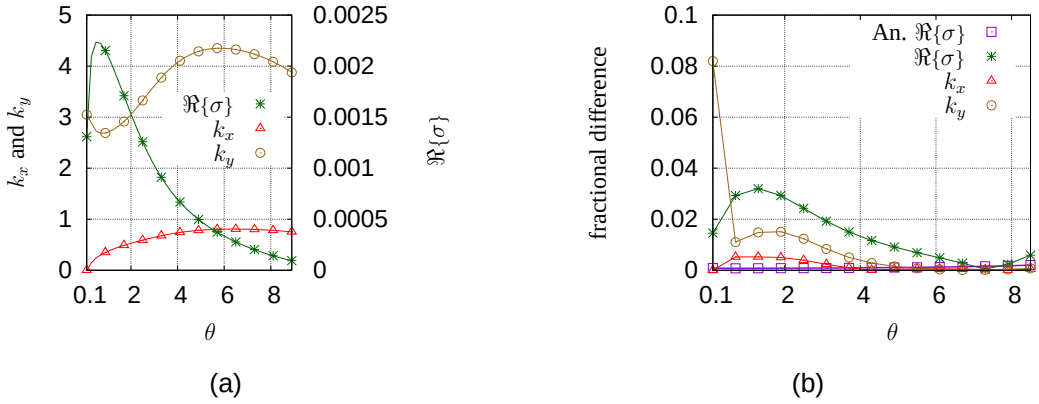



Figure 6.10: As for figure 6.4 but variations vs. $\theta$, for $\mathcal{F} = 10^{-3}$ and $H_b = 1$. Here $\epsilon$ lies in the range $10^{-4} - 1.2 \times 10^{-2}$.

The temperature flux at the bottom of the layer is controlled by the dimensionless number $\theta$. So far in this section it has appeared that, when the growth rate of the instability increases, the anelastic approximation performs worse in reproducing the compressible results. This impression is confirmed by figure 6.10 which compares the results for a

series of calculations where the thermal gradient, and thus the stratification, is changed whilst $m$ is held fixed.

On the one hand, large thermal gradients produce large density gradients which suppress the instability; on the other hand, for very low values of $\theta$, the gravity is weak (see equation (3.22)) and so the destabilising effect of magnetic fields is also reduced. Consequently, as demonstrated in figure 6.10(a), the growth rate first increases and then decreases with $\theta$ — and, as shown in figure 6.10(b), so does the accuracy of the anelastic approximation. It therefore appears as though the relative accuracy is controlled by the growth rate of the instability. This growth rate can be thought of as providing a time-scale for the evolution of the instability and so, if this is long, then the anelastic approximation performs well; if, on the contrary, the instability develops rapidly then the anelastic approximation is less accurate.

### 6.3.5 The effect of altering $\zeta$

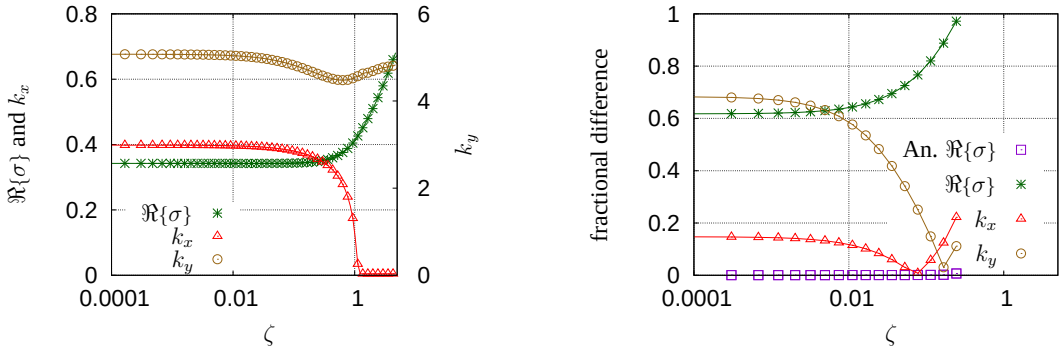



Figure 6.11: As for figure 6.4 but variations vs. $\zeta$, for $\mathcal{F} = 0.01$ and $\epsilon = 10^{-3}$.

The magnetic diffusivity is proportional to the parameter $\zeta$ but altering $\zeta$ should not effect the accuracy of the anelastic approximation.

For low $\zeta$ values figure 6.11(a) shows the instability is three dimensional and changes to an interchange mode as $\zeta$ passes through unity whereupon the growth rate of the

instability increases rapidly for larger $\zeta$ values. Figure 6.11(b) shows that the fractional differences in this case are very large and it would not be valid to use the anelastic approximation. This is an extreme case where the magnetic field is large and so is the growth rate of the instability. As $\zeta$ increases the growth rate in the compressible approximation increases, as shown in figure 6.11(a), but in the anelastic case the growth rate decreases. This explains the large and increasing fractional difference as $\zeta$ is increased. The wavenumbers coincide at $\zeta \sim 0.1$ but this is not significant as they diverge again at $\zeta$ values around this point.

### 6.3.6 The effect of altering $H_b$

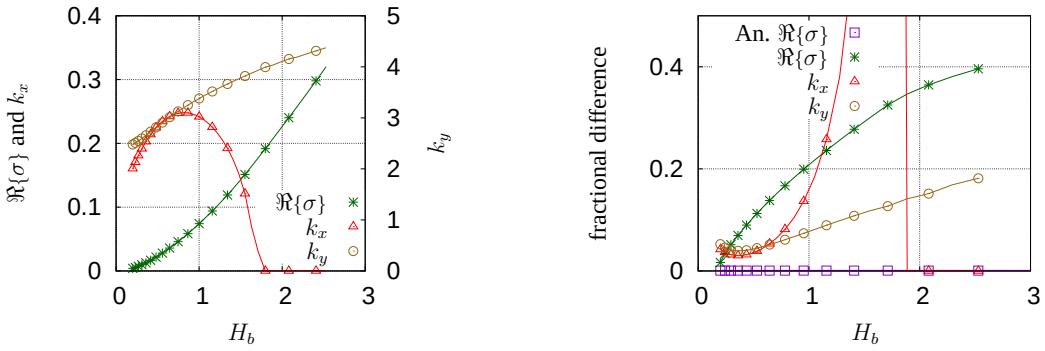



Figure 6.12: As for figure 6.4 but variations vs. $H_b$, for $\mathcal{F} = 0.1$, $C_k = 0.025$ and $\epsilon = 10^{-3}$.

The basic state magnetic field is $B_{xb} = 1 + H_b z$. Altering $H_b$ will effect the strength of the magnetic field and so from the assumptions used in Chapter 3 it is likely to effect the accuracy of the anelastic approximation.

Figure 6.12 shows the effects of varying the magnetic field gradient $H_b$. For large field gradients, interchange modes are preferred and the instability grows faster. The anelastic approximation again fails when the growth rate becomes large, but interestingly, whereas the growth rate increases in a concave manner, the fractional difference increases in a

convex manner. This demonstrates that there is not a simple direct relationship between the growth rate of the instability and the accuracy of the anelastic approximation.

## 6.4 Summary

The derivation of the anelastic approximation depends on a parameter $\epsilon$, the departure from an adiabatic atmosphere, being small. It also depends on the time-scale being slow. When the atmosphere is nearly adiabatic then the magnetic buoyancy instability grows quickly, or there are waves travelling quickly, so the time-scale used in the anelastic approximation is inadequate. When the atmosphere departs from being adiabatic then the parameter $\epsilon$ increases, violating one assumption in the anelastic approximation, but the growth-rate of the instability slows, satisfying one of the assumptions in the anelastic approximation. From the results it is clear that as $m$ increases from its adiabatic value of $1.5$ the anelastic approximation performs better, suggesting that violating the time-scale assumption has a larger effect than violating the assumption of a nearly adiabatic atmosphere.

# Chapter 7

# Non-linear Anelastic Codes

To investigate fully how the anelastic approximation differs from the fully compressible equations requires a non-linear comparison of the two sets of equations. Studies from a meteorological perspective list some of the limitations of the anelastic approximation (see e.g. Nance & Durran, 1994). The conclusion Nance and Durran drew was that using a version of the anelastic approximation similar to equations (3.20), but without magnetic field, produces results with errors, "significantly less than the errors generated in real world models". Meteorological studies do not include magnetic field and are focused on different types of instability to those that occur in the Sun. Therefore it would still be useful to have non-linear models and work out a range of parameters where the anelastic approximation performs well in the context of the Sun's convection zone, radiative interior, and tachocline. It is also important to characterise any artefacts introduced by the anelastic approximation.

## 7.1 Anelastic Time-Stepping

Even using a simple model, the range of spatial scales in modelling the Sun is astronomical and added complexity comes from time-stepping. With a finite-difference time-stepping numerical code the domain is discretised using a distance $\Delta x$ and the time

is discretised using step $\Delta t$. A spectral code does not use the spatial discretization $\Delta x$ but for the purpose of this argument it can be considered that this is true. An explicit finite-difference method will not capture correctly a wave of speed $c$ if, in one time step, the wave crosses more than the distance $\Delta x$. This argument is also true for advection as well as for waves. This leads to the Courant-Friedrichs-Lewy (CFL) condition for stability

$$\mathcal{C} \equiv \left| \frac{c\Delta t}{\Delta x} \right| \leq 1, \tag{7.1}$$

This is a necessary but not sufficient condition for stability. The ratio $\mathcal{C}$ is termed the Courant number (Courant et al., 1967). It is very clear from (7.1) that reducing the fastest wave speed in the system by a factor of ten will lead to a factor of ten improvement in the size of time-step that can be taken. This is important as some instabilities develop over a long time so there can be a sacrifice in resolution and accuracy to stop the simulation becoming too computationally expensive. It is often the case that, to counter the small time-step, the instability will be driven harder with less realistic parameters so that it develops on a faster time-scale. Driving a system harder may mean an increase in the Reynolds number and turbulence which requires a smaller $\Delta x$ if the system is to be accurately modelled.

An explicit time-stepping method is where the variable at time $n + 1$ can be written in terms of the variable at time $n$ whereas in an implicit method the variable at $n + 1$ is written in terms of the variable at time $n$, $n + 1$, and maybe other time-steps. The CFL condition is a limiting factor when the equations are time-stepped explicitly but if the time-stepping is done in an implicit manner then the CFL condition is less limiting, in the sense that (7.1) is no longer a strict inequality. In astrophysical fluid dynamics turbulence is thought to play a crucial role and the advection term, which generates much of the turbulence, is not dealt with as satisfactorily in the implicit case. It is also hard to evolve the non-linear terms implicitly. This means that the turbulence would not be correctly modelled and so the wrong dynamics and turbulent transport coefficients would be obtained. Taking large time-steps in an implicit scheme could lead to a stable

but inaccurate solution.

## 7.2 Non-linear Formulation

A full three-dimensional set of equations (3.20) is expensive to run in terms of computer resources. For reasons of simplicity and for the purpose of illustration of methods I will outline a two-dimensional reduction of the full three-dimensional set of equations (3.20). The two-dimensions are $(x, z)$ where $z$ is the direction of gravity and $x$ is homogeneous and periodic.

The velocity components of any two-dimensional divergence-free flow can be represented by a stream-function, which reduces the number of equations that need to be solved. Each variable is decomposed using $\xi(\mathbf{x}, t) = \overline{\xi}(z) + \xi^*(x, z, t)$, where $\xi$ is any variable, $\overline{\xi}$ is the variable in a reference state, and $\xi^*$ is the fluctuation variable. I will drop the superscript $^*$ on the fluctuation variables. A well constructed stream-function will also ensure that the flow evolves whilst obeying the anelastic continuity equation (3.20a). A stream-function $\psi$ satisfying this is

$$\overline{\rho}\mathbf{u} = \nabla \times (\psi\hat{\mathbf{y}}) + \overline{\rho}v\hat{\mathbf{y}} = \overline{\rho}(u, v, w). \tag{7.2}$$

Since $\nabla \cdot \mathbf{B} = 0$ a similar technique can be used for the magnetic field to reduce the number of unknowns and also to ensure the field evolution obeys the solenoidal condition (3.20e). The field potential $A$ is thus

$$\mathbf{B} = \nabla \times (A\hat{\mathbf{y}}) + \beta\hat{\mathbf{y}}, \tag{7.3}$$

where $\beta$ is the $\hat{\mathbf{y}}$ component of the magnetic field. It is also useful to define the vorticity $\omega$ as

$$\boldsymbol{\omega} \cdot \hat{\mathbf{y}} = \omega = -\left[\frac{\nabla^2}{\overline{\rho}} + \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\overline{\rho}}\right)\frac{\mathrm{d}}{\mathrm{d}z}\right]\psi, \tag{7.4}$$

and the Laplacian of the field potential $H$, which is the $\hat{\mathbf{y}}$ component of the current density $\boldsymbol{J}$, defined as

$$H = \nabla^2 A = (\mathbf{B} \cdot \hat{\mathbf{x}})_z - (\mathbf{B} \cdot \hat{\mathbf{z}})_x, \tag{7.5}$$

where $(\mathbf{B}\cdot\hat{\mathbf{z}})_x$ is the $x$ partial derivative of the $z$ component of the magnetic field. With these defined, I can go on to develop equations for a two-dimensional Cartesian non-linear anelastic code using the Lantz-Braginsky approximation.

### 7.2.1 Non-linear Anelastic Equations

Taking the curl of the simplified Lantz-Braginsky momentum equation (3.35), with entropy diffusion and no temperature diffusion, gives

$$\frac{\partial\boldsymbol{\omega}}{\partial t} = -\nabla\times(\mathbf{u}\cdot\nabla\mathbf{u}) - \nabla\times(s\hat{\mathbf{z}}) + \tilde{\mathcal{F}}\nabla\times\left[\frac{1}{\overline{\rho}}(\nabla\times\mathbf{B})\times\mathbf{B}\right] + \left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\nabla\times\left(\frac{1}{\overline{\rho}}\nabla\cdot\boldsymbol{\tau}\right). \quad (7.6)$$

This is in vector form but for solving the system numerically only the $\hat{\mathbf{y}}$ component is needed, which is given by

$$\frac{\partial\omega}{\partial t} = -(w_z+u_x)\,\omega - w\omega_z - u\omega_x + s_x + \tilde{\mathcal{F}}\left[\frac{H_zA_x - H_xA_z}{\overline{\rho}} + \frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\overline{\rho}}\right)(A_xH - \beta\beta_x)\right] + \left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\left[\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\overline{\rho}}\right)\left(\omega_z + \frac{4}{3}\left[w_{xz}+u_{xx}\right]\right) + \frac{\omega_{xx}+\omega_{zz}}{\overline{\rho}}\right]. \quad (7.7a)$$

This is the evolution equation for the vorticity. The stream function $\psi$ can be obtained by solving equation (7.4) and the velocity components $u$ and $w$ can be obtained from $\psi$ using equation (7.2). The $\hat{\mathbf{y}}$ component of the velocity field $v$ requires another evolution which is the $\hat{\mathbf{y}}$ component of equation (3.35), i.e.

$$\frac{\partial v}{\partial t} = \frac{v_x\psi_z}{\overline{\rho}} - \frac{v_z\psi_x}{\overline{\rho}} + \frac{\tilde{\mathcal{F}}}{\overline{\rho}}\left[\beta_zA_x - \beta_xA_z\right] + \frac{1}{\overline{\rho}}\left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}(v_{xx}+v_{zz}). \quad (7.7b)$$

The induction equation

$$\frac{\partial\mathbf{B}}{\partial t} = \nabla\times(\mathbf{u}\times\mathbf{B}) + C_k\zeta\nabla^2\mathbf{B},$$

is solved in a similar manner to the momentum equation; only the $\hat{\mathbf{y}}$ component of the curl of the induction equation is required. The $\hat{\mathbf{y}}$ component of the curl of the induction is

$$\frac{\partial H}{\partial t} = -uH_x - wH_z + A_x\left(\omega_z + w_{xz} + u_{xx}\right) + 2w_z A_{zz} + 2\left(u_z + w_x\right)A_{xz} + 2u_x A_{xx} + A_z\left(w_{zz} + w_{xx}\right) + C_k\zeta\left(H_{xx} + H_{zz}\right), \quad (7.7c)$$

which is the evolution equation for $H$. The potential $A$ can then be found using equation (7.5). To find $\beta$, the $\hat{\mathbf{y}}$ component of the induction equation is required,

$$\frac{\partial \beta}{\partial t} = \frac{\beta\psi_x\overline{\rho}' + \overline{\rho}\left(\psi_z\beta_x - \beta_z\psi_x\right)}{\overline{\rho}^2} + v_z A_x - A_z v_x + C_k\zeta\left(\beta_{xx} + \beta_{zz}\right). \quad (7.7d)$$

The energy equation written in terms of entropy is

$$\overline{\rho}\overline{T}\left[\frac{\partial s}{\partial t} + \left(\mathbf{u}\cdot\nabla\right)\left(\overline{s} + s\right)\right] = \frac{1}{\left(\mathrm{Pr}\,\tilde{R}\right)^{1/2}}\nabla\cdot\overline{T}\nabla s + \frac{\gamma-1}{\gamma}\tilde{C}_k^2\left[\left(\mathrm{Pr}^3\tilde{R}\right)^{1/2}\frac{\partial u_i}{\partial x_j}\tau_{ij} + \left(\mathrm{Pr}\tilde{R}\right)^{1/2}\tilde{\mathcal{F}}\zeta\left(\nabla\times\mathbf{B}\right)^2\right],$$

which can be written in terms of $u$, $v$, $w$, $H$ and $\beta$ as

$$\frac{\partial s}{\partial t} = -us_x - ws_z - w\frac{\mathrm{d}\overline{s}}{\mathrm{d}z} + \frac{1}{\sqrt{\mathrm{Pr}\,\tilde{R}}}\frac{1}{\overline{\rho}}\left(s_{xx} + s_{zz} + \frac{\mathrm{d}\overline{T}}{\mathrm{d}z}\frac{s_z}{\overline{T}}\right) + \frac{\gamma-1}{\gamma}\tilde{C}_k^2\left(\mathrm{Pr}\tilde{R}\right)^{1/2}\tilde{\mathcal{F}}\zeta\frac{1}{\overline{\rho}\overline{T}}\left[H^2 + \beta_x^2 + \beta_z^2\right] + \frac{\gamma-1}{\gamma}\tilde{C}_k^2\sqrt{\mathrm{Pr}^3\tilde{R}}\frac{1}{\overline{T}\overline{\rho}}\left[\frac{4}{3}\left(w_z\left[w_z - u_x\right] + u_x^2\right) + \left(u_z + w_x\right)^2 + v_x^2 + v_z^2\right]. \quad (7.7e)$$

**Means**

In this chapter all terms are assumed to have periodic boundary conditions in the $\hat{\mathbf{x}}$ direction. From the definition of the velocity component $w = \partial_x\psi$ there are no terms in $w$ that are a functions of $z$ only as terms such as $\psi = xf(z)$, which corresponds to $w$ being a function of $z$ only and are not periodic in $\hat{\mathbf{x}}$. This turns out to not be a problem

as the taking the horizontal average of the anelastic continuity equation (3.20a) shows that this average will equal zero, in that

$$\int_0^L \frac{\partial \psi}{\partial x} dx = 0,$$

as $\psi$ is periodic with period $L$. The same argument can be given for $\mathbf{B} \cdot \hat{\mathbf{z}}$.

### 7.2.2 Hybrid Spectral Methods

In this outline of a possible non-linear code the $\hat{\mathbf{x}}$ direction is periodic and solved in Fourier space, allowing the use of spectral methods, and so the error decays exponentially as the resolution increases (Canuto, 2006). Spectral methods, or hybrid spectral methods where only some of the terms are solved in spectral space, are used in many non-linear anelastic codes (see e.g. Clune et al., 1999; Jones & Kuzanyan, 2009; Glatzmaier, 1984). In this outline I will leave the $z$ direction in real space where $z \in [0, 1]$ and only the $x$ direction will be in Fourier space with $x \in (0, L]$, where $L$ defines the aspect ratio of the $\hat{\mathbf{x}}$ to $\hat{\mathbf{z}}$ dimension. It is a discrete Fourier transform which is performed in the $\hat{\mathbf{x}}$ direction which can be done using e.g. the FFTw library (Frigo & Johnson, 2005). In a discrete Fourier transform all the variables are represented in arrays and if the Fourier transform is taken in the $\hat{\mathbf{x}}$ direction from the real array $X$ of size $N$ the result is a Hermitian array $Y$ via

$$Y_i = \sum_{j=0}^{N-1} X_j e^{2\pi i j/N\sqrt{-1}}.$$

The periodic boundary condition in the $\hat{\mathbf{x}}$ direction is implemented naturally in Fourier space. The $\hat{\mathbf{z}}$ direction is solved in real space using finite difference methods as each variable has different boundary conditions and implementing these in Fourier space would not be spectrally accurate and removes the banded structure of a matrix representation of a derivative, see §4.1 for more detailed explanation of representation of finite difference differential operators. To make the code stable the linear terms are

solved using an exponential integrating factor and so are solved exactly. The equations in the $\hat{\mathbf{z}}$ direction are solved using Crank-Nicolson method (Crank & Nicolson, 1947). The non-linear terms are calculated using an Adams-Bashforth time-stepping method.

To calculate the right-hand-sides of equations (7.7) requires calculating convolutions of some terms in Fourier space. Working out the convolutions is more computationally expensive than converting the terms that are in Fourier space into real space, multiplying them, and then converting back into Fourier space. A convolution requires two matricies to be multiplied so will require $\sim N^2$ operations where as converting into real space and back requires $\sim N \ln N$ operations.

Aliasing errors in numerical models occur when the high frequencies become too small to be resolved allowing them to interact with low frequency modes. This high to low frequency transfer of energy causes the model to blow-up. Orszag (1971) showed that filtering out the highest $\frac{1}{3}$ of the wavenumbers is sufficient to stop aliasing errors associated with quadratic non-linearities in Fourier space.

## 7.3 Boundary Conditions

For the magnetic buoyancy case and magnetoconvection cases to implement stress free boundaries requires

$$\frac{\partial u}{\partial z} = \frac{\partial v}{\partial z} = w = 0, \tag{7.8}$$

and using (7.2) this translates to

$$\psi = 0 \text{ and } -\frac{\psi_{zz}}{\overline{\rho}} - \psi_z \frac{\mathrm{d} \ln \overline{\rho}}{\mathrm{d}z} = 0 \qquad \text{at } z = 0, 1. \tag{7.9}$$

The boundary conditions on $\omega$ are therefore

$$\omega = \hat{\mathbf{y}} \cdot \nabla \times \mathbf{u} = u_z - w_x = 0 \tag{7.10}$$

For the magnetic buoyancy (or horizontal field) case

$$\frac{\partial B_x}{\partial z} = \frac{\partial B_y}{\partial z} = B_z = c, \tag{7.11}$$

where the $c$ is a constant due to the imposed field of $B_x = H_b z + 1$. Similarly to $\psi$, the boundary condition is

$$A(z{=}0) = 0,\ A(z{=}1) = -\left(2 + H_b\right)/2, \quad \text{and } H = 0.$$

In the magnetoconvection (or a vertical field) case

$$B_x = B_y = \frac{\partial B_z}{\partial z} = 0 \qquad \text{so} \qquad \frac{\partial A}{\partial z} = \frac{\partial H}{\partial z} = 0. \tag{7.12}$$

The flow is isentropic so the boundary condition is $s = 0$ on $z = 0, 1$.

### 7.3.1 Pressure Term

Without using the Lantz-Braginsky approximation the anelastic equations (3.20) contain a pressure term in the momentum equation (3.20b). In the compressible equations there is an evolution equation for the pressure but in the anelastic case the pressure responds instantaneously, as shown in equation (4.13). The divergence of the momentum equation is required to advance the pressure fluctuations via the equation

$$\begin{aligned}
&\left(\nabla^2 + \frac{\overline{p}' - \overline{p}\,\overline{\rho}'}{\overline{p}^2} - \frac{1}{\overline{p}}\frac{\mathrm{d}}{\mathrm{d}z}\right) p = \\
&-\overline{\rho}\left[\left(\langle w\rangle_z + w_z\right)^2 + u_x^2 + 2u_z w_x + \left(\langle w\rangle + w\right)\left(\langle w\rangle_{zz} + w_{zz} + u_{xz}\right) + u\left(w_{xz} + u_{xx}\right)\right] \\
&\qquad - \frac{T}{\overline{T}}\overline{\rho}' - \frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\left[\left(\langle w\rangle + w\right)\left(\langle w\rangle_z + w_z\right) + u w_x\right] + \overline{\rho}\left(\frac{T\overline{T}'}{\overline{T}^2} - \frac{T_z}{\overline{T}}\right) \\
&\qquad + \tilde{\mathcal{F}}\Big(-H^2 + A_z H_z + H_x A_x - \beta_z^2 - \beta\beta_{zz} - \beta_x^2 - \beta\beta_{xx} + \Big) \\
&+\left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\left[-\frac{7}{3}\frac{\mathrm{d}}{\mathrm{d}z}\left(\frac{1}{\overline{\rho}}\right)\omega_x + \frac{8}{3\overline{\rho}}\frac{\mathrm{d}^2}{\mathrm{d}z^2}\left(\frac{1}{\overline{\rho}}\right)\psi_{xz} - \frac{4\left(8\left(\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\right)^3 - 7\rho\frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\frac{\mathrm{d}^2\overline{\rho}}{\mathrm{d}z^2} + \overline{\rho}^2\frac{\mathrm{d}^3\overline{\rho}}{\mathrm{d}z^3}\right)}{3\overline{\rho}^5}\right].
\end{aligned} \tag{7.13}$$

This is an elliptic equation where the right-hand-side is known from the hyperbolic evolution equations. A typical way to solve the equation (7.13) would be to use an LU-decomposition of the operator acting on $p$ on the left-hand-side e.g. using the LAPACK

(Anderson et al., 1999) algorithm GBTRF for the banded matrix operator. The top and bottom boundary conditions are applied by replacing the representation of the operator with the appropriate boundary condition at the first and last row of the matrix, e.g. for a Dirichlet boundary condition $p = a, b$ on $z = 0, 1$ then the matrix would be

$$\begin{pmatrix} 1 & 0 & \cdots \\ & \nabla^2 + \frac{\overline{p}'-\overline{p}\,\overline{\rho}'}{\overline{p}^2} - \frac{1}{\overline{p}}\frac{\mathrm{d}}{\mathrm{d}z} & \\ \cdots & 0 & 1 \end{pmatrix} p = \begin{pmatrix} a \\ \text{R.H.S} \\ b \end{pmatrix},$$

where R.H.S. is the right-hand-side of (7.13), and the operator would fill the internal points in the matrix. The LU-decomposition only needs to be done once at the start of the simulation since the linear differential operator is constant.

The problem in solving equation (7.13) comes from the boundary conditions. The boundary conditions involve the velocity which comes from the momentum equation and will be from the previous time-step and it is non-trivial to apply a boundary condition that is based on the current time-step. In the Anelastic Spherical Harmonic (ASH) code Clune et al. (1999) take the horizontal divergence of the momentum equation so that the pressure is calculated based on the previous time-step, as seen in Clune et al. (1999) equation (A.4). As the equation solved to calculate the pressure is elliptic then the boundaries effect every point in the domain. Although so far the difficulties have been framed with regards to the anelastic equations (3.20), a similar problem occurs in the simplified Lantz-Braginsky equations if the temperature diffusion is used rather than the entropy diffusion. This gives strong arguments to use the Lantz-Braginsky simplification and entropy diffusion together where only the entropy is required.

## 7.4 The Non-linear Algorithm

This is an outline of the operations a two-dimensional Cartesian non-linear code undertakes in one time-step. The initial condition is equivalent to the basic state used in

the linear code but now the code is advancing in time. To cause the simulation to grow from a stable basic state then a small amount of noise may need to be added.

- A Fourier transform is taken in the $\hat{\mathbf{x}}$ direction.
- The derived variables ($\psi, A, u, w$) are calculated, these are used in equations (7.7a - d). In a similar method as used to solve the elliptic equation discussed in § 7.3.1, $\psi$ can be calculated from equation (7.2) and $A$ can be calculated from (7.5). Then from equation (7.4) $u$ and $w$ can be calculated from $\psi$.
- The $x$ derivatives of all the variables are calculated, along with higher $x$ derivatives where required. The variables are in Fourier space so calculating the $N^{\text{th}}$ derivative is equivalent to multiplying the variable by $(2\pi ik/L)^N$.
- In addition the $z$ derivatives are calculated for all the variables. The variables in the $\hat{\mathbf{z}}$ direction are in real space and so $z$ derivatives are calculated using a finite difference representation of a derivative, see § 4.1 for more on finite difference schemes and § 7.3 for the boundary conditions to be applied.
- An inverse Fourier transform is taken in the $\hat{\mathbf{x}}$ direction so that all the variables are in real space. This means that no convolutions are required to calculate the right-hand-side of equations (7.7) as these are more computationally expensive than inverse Fourier transforms.
- The equations (7.7) are in the form $\partial_t \xi = \mathcal{L}(\xi) + \mathcal{N}$, where $\mathcal{L}$ is a function of the linear terms and $\mathcal{N}$ represents all the inhomogeneous and non-linear terms. $\mathcal{N}$ can be calculated by multiplying the correct terms together from the previous time-step.
- The non-linear terms and inhomogeneous terms on the right-hand-side of equations (7.7), denoted $\mathcal{N}$, can be calculated for any given time-step using the Euler method, i.e. the value of the right-hand-sides of equations (7.7) is found and then multiplied by the time-step size. A more accurate way to deal with $\mathcal{N}$ is

through an explicit multi-step method such as the Adams-Bashforth method (see e.g. Hairer et al., 1993). The Adams-Bashforth method is an extension of the Euler method and uses information about previous steps so the next time-step can be calculated more accurately, e.g the fourth-order Adams-Bashforth method is

$$\xi^{n+1} = \xi^n + \frac{\Delta t}{24}\left(55\mathcal{N}^n - 59\mathcal{N}^{n-1} + 37\mathcal{N}^{n-2} - 9\mathcal{N}^{n-3}\right),$$

where $\mathcal{N}(z, n\Delta t) = \mathcal{N}^n$. In the non-linear codes mentioned in this section it much more common for an Adams-Bashforth method to be used. This is used in preference to an Euler method as it increases the stability and is used in preference over a Runge-Kutta as calculating the non-linear terms is numerically expensive and a Runge-Kutta method requires multiple calculations at mid-steps.

- Using an Euler or multi-step time-stepping method for the homogeneous linear terms on the right-hand-sides of equations (7.7a - d) can be unstable (see e.g. Burden & Faires, 2005). Terms dependent on the $z$ derivative can be dealt with using the Crank-Nicolson time-step algorithm (Crank & Nicolson, 1947). For a partial differential equation such as

$$\frac{\partial \xi}{\partial t} = \mathcal{L}(\xi), \tag{7.14}$$

where $\xi$ is the variable to be advanced in time then the Crank-Nicolson algorithm is

$$\frac{\xi^{n+1} - \xi^n}{\Delta t} = \frac{1}{2}\left[\mathcal{L}^{n+1}\left(\xi, z, t, \frac{\partial \xi}{\partial z}, \frac{\partial^2 \xi}{\partial z^2}\right) + \mathcal{L}^n\left(\xi, z, t, \frac{\partial \xi}{\partial z}, \frac{\partial^2 \xi}{\partial z^2}\right)\right], \tag{7.15}$$

where $\Delta z$ is the spatial discretization width, and $\xi(z, n\Delta t) = \xi^n$. Given the equation

$$\frac{\partial \xi}{\partial t} = \frac{\partial^N \xi}{\partial z^N} + \mathcal{N}, \tag{7.16}$$

the Crank-Nicolson algorithm would be

$$(2I - \Delta t D_N)\,\xi^{n+1} = (2I + \Delta t D_N)\,\xi^n + 2\Delta t \mathcal{N}, \tag{7.17}$$

where the $I$ is the identity matrix and $D_N$ the finite difference matrix representation of $\partial^N/\partial z^N$. This can be solved using LU-decomposition described in § 7.3.1.

- An exponential integrating factor can solve exactly the homogeneous linear terms involving $x$ derivatives on the right-hand-side equations (7.7a - d). For example, the equation

$$\frac{\partial \xi}{\partial t} = -a\frac{\partial^2 \xi}{\partial x^2} + \mathcal{N}, \tag{7.18}$$

can be advanced in time exactly using

$$\xi_k^{n+1} = e^{-\Delta t(2a\pi ik/L)^2}\left(\xi_k^n + \mathcal{N}_k^n\right), \tag{7.19}$$

where $\xi(k, n\Delta t) = \xi_k^n$.

- Combining the methods above for a hybrid-spectral time-step using the example equation

$$\frac{\partial \xi}{\partial t} = \frac{\partial^N \xi}{\partial z^N} - a\frac{\partial^2 \xi}{\partial x^2} + \mathcal{N}, \tag{7.20}$$

then the time-step, using Euler's method for $\mathcal{N}$, would be

$$\left(2I - \Delta t D_N\right)\xi_k^{n+1} = e^{-\Delta t(2a\pi ik/L)^2}\left[\left(2I + \Delta t D_N\right)\xi_k^n + 2\Delta t\mathcal{N}_k^n\right] \tag{7.21}$$

## 7.5 The Current Non-Linear Codes

There is a an attempt under-way to create a set of anelastic benchmarks by Jones et al. (2011). The aim is to compare the currently available magnetic spherical anelastic codes by developing a standard benchmark against which the codes can be validated. The codes included so far are the Anelastic Spherical Harmonic (ASH) code (Clune et al., 1999), the Leeds code (Jones & Kuzanyan, 2009), Gary Glatzmaier's code (Glatzmaier, 1984), and a code in development from Johannes Wicht and Thomas Gastine.

Glatzmaier's code is based on a spherical harmonic expansion of the variables in the anelastic system equivalent to equations (3.20). The Lantz-Braginsky approximation

has not been made and so to update the pressure Glatzmaier uses an additional boundary conditions on $\psi$ from the stress free condition. The stress free boundary condition in equation (7.9) gives two boundary conditions on $\psi$, namely

$$\psi = 0 \qquad \text{for } z = 0, 1$$

which is the boundary condition applied to $\psi$, and

$$-\frac{\psi_z z}{\overline{\rho}} + \frac{\mathrm{d}\overline{\rho}}{\mathrm{d}z}\frac{\psi_z}{\overline{\rho}} = 0 \qquad \text{for } z = 0, 1$$

which can be substituted into the vorticity equation (7.7a) to give a boundary condition for the pressure. This means that the pressure boundary is calculated based on thestream-function from the previous time-step. In Glatzmaier's code the reference state is time-independent, which is compatible with the anelastic approximation.

The Anelastic Spherical Harmonic (ASH) code was developed from Glatzmaier's code and the equations (3.20). Unlike in Glatzmaier's code the ASH code updates the reference state and this allows the reference state to have large departures from an adiabatic atmosphere. The reasoning is that so long as the Mach number of the flow is small then the results are valid and updating the reference state ensures that the fluctuation terms will be small in comparison to the reference state. Updating the reference does indeed make

$$\frac{\rho}{\overline{\rho}} \ll 1$$

but it also makes the small parameter $\epsilon$ used in the asymptotic expansion large. The approximation is only valid when the $\partial\rho/\partial t$ is much smaller than the other terms and can be neglected. When the reference state is far from adiabatic then a source of energy is introduced into the temperature equation, this can be seen clearly from the derivation of the temperature equation in §3.7.2. This spurious energy source is due to inconsistencies with the model when the atmosphere is far from adiabatic. The ASH code also does not use the Lantz-Braginsky simplification and to update the pressure the velocity at the previous time-step is used as a boundary condition.

Rogers & Glatzmaier (2005) developed a non-linear code based on the anelastic equations but where the temperature, rather than entropy, equation was used, see §3.7.2 for more on the temperature equation. Here the reference state is updated periodically by adding the horizontal mean. In the paper the reference state is not shown separated out from the fluctuation state, both appear in Rogers & Glatzmaier (2005) equation (3).

The code in Lantz & Fan (1999) is in a Cartesian geometry and has a time-independent reference state. It also uses the Lantz-Braginsky approximation, explained in §3.7, so the pressure is de-coupled from the system and entropy is the only remaining thermodynamic variable. All the equations are solved in real space using a finite-difference representation of the differential operators, except for the elliptic equation (7.4) for $\psi$ where the horizontal direction is solved in spectral space.

The Leeds code, used in Jones & Kuzanyan (2009) is similar to the code developed by Lantz. It too uses the Lantz-Braginsky approximation. It solves the equations in a spherical geometry and is pseudo-spectral using Legendre polynomials and Fourier modes for the non-radial derivative and a non-uniform finite-difference mesh in the radial direction.

# Chapter 8

# Conclusions

## 8.1 Discussion of New Results

There are many sets of equations that are in the anelastic 'family', such as the anelastic equations (3.20) derived in Gilman & Glatzmaier (1981), the set of anelastic equations derived in Ogura & Phillips (1962), the Lantz-Braginsky simplification as derived in §3.7, and the anelastic equations derived in Gough (1969). I have used an asymptotic expansion to develop the anelastic temperature equation given in §3.7.2. All of these sets of equations are equivalent, if the same type of diffusion is used and magnetic fields neglected, when the atmosphere is perfectly adiabatic and differ at higher orders of $\epsilon$ when the atmosphere is not adiabatic. The definition of $\epsilon$, which is a measure of the departure from the atmosphere being adiabatic, is given in (3.4). There are situations where one set of equations will give results that are closer to the fully compressible than some of the other sets, as demonstrated in Chapters 5 and 6, but this is also when the validity of making the anelastic approximation itself is in doubt, such as when the Alfvénic time-scale becomes small or the growth rate of the instability is large.

In Chapter 4 I developed the linear code I used to test the anelastic approximation and in Chapters 5 and 6 I have investigated the range of validity of the anelastic

approximation by solving linear stability problems in both the fully compressible and anelastic models. I calculated either the critical Rayleigh number for magnetoconvection or the largest growth rate of a magnetic buoyancy instability; I then determined the fractional difference between the fully compressible results and those calculated using the anelastic approximation.

For the problem of magnetoconvection in Chapter 5, the results are not surprising. The anelastic approximation performed well where expected and not so well (though still reasonably well) when pushed beyond its range of formal validity. The only slightly surprising result is that the inclusion of a stronger field or larger thermal flux gradient can lead to a decrease in the accuracy of the anelastic approximation. With a strong magnetic field the Alfvénic time becomes small; this is inconsistent with the time-scale assumption made in the anelastic approximation. In the studies presented in this thesis the magnetic field is imposed but in the Sun the field is likely to be the result of dynamo action and so the magnetic energy would be comparable to the kinetic energy. This would mean in low Mach number flows the field would obey the weak field assumptions. In Chapter 5 the Takens-Bogdanov point in the magnetoconvection instability was captured by the anelastic approximation with no substantial difficulties. When altering the boundary condition from isentropic to isothermal then the growth rate for the magnetoconvection instabilities reduced, with the reduction more pronounced at large $\theta$ values and low $\zeta$ values. When studying magnetoconvection in a tilted field then, in agreement with Matthews et al. (1992), I found that the stability reversal, described in §5.3.8, is very sensitive. The results for stability reversal in this work differ from the results in Roxburgh (2007), which is presumed due to the sensitivity. These results should be contrasted with those of the magnetic buoyancy instability.

For magnetic buoyancy, even when the expansion parameter $\epsilon$ is small (i.e. the atmosphere is nearly adiabatic), it is possible to break the anelastic approximation; on the contrary, slowly growing magnetic buoyancy modes appear to be captured accurately even for large $\epsilon$. It would seem that in the case when the atmosphere is subadiabatic

the accuracy of the approximation is controlled by the growth rate of the magnetically driven instability. It has previously been assumed that the approximation is valid so long as the Mach number is sufficiently small (see e.g. Gough, 1969; Miesch, 2005) and the Alfvénic time sufficiently large (see e.g. Glatzmaier, 1984; Lantz & Fan, 1999). There are a number of possible reasons why the approximation may break down. It may be that the Mach number is too large but this is not applicable to the linear case. Another possible reason is that hydrostatic balance is assumed at leading order and when the magnetic fields become large this balance may be upset. This seems unlikely for the cases considered here as the magnetic pressure is only $1\%$ of the terms in the hydrostatic balance. Here the strong magnetic field, and the correspondingly fast Alfvén waves, or a fast growth rate leads to a violation of the assumption about the scalings used on the unit of time. This leads to the situation where $\epsilon$ can be increased but, as the growth rate decreases, the approximation becomes more accurate. In the magnetic buoyancy case, the full anelastic approximation appears to be more sensitive to the time-scales than to the departure from an adiabatic state. As this instability is magnetically driven, when the magnetic field becomes large it is hard to distinguish if the growth of the instability or the increased magnetic field strength causes the anelastic approximation to become inaccurate.

I also consider the accuracy of the Lantz-Braginsky simplification and demonstrate that, as expected, the Lantz-Braginsky simplification appears accurate (i.e. close to the results of the full anelastic approximations) when the atmosphere is nearly adiabatic. This remains true even in cases with large magnetic fields. To reiterate the Lantz-Braginsky approximation only differs from the full anelastic approximation by order $\epsilon$, i.e. the difference is of the same order as the $\partial\rho/\partial t$ term which is neglected in the anelastic continuity equation. This would suggest that both the Lantz-Braginsky simplification and the anelastic equations (3.20) are equally valid. The linear results show that the terms that the Lantz-Braginsky approximation is neglecting make a significant contribution when the atmosphere departs from adiabatic stratification. There are also arguments given in §7.3.1 in favour of using the Lantz-Braginsky simplification with

the entropy diffusion in time-stepping non-linear codes as solving the elliptic equation for the pressure requires boundary conditions which are based on the previous time-step. Chapter 7 also briefly gives an argument that evolving the reference state is not compatible with the anelastic approximation.

I conclude by reiterating that care must be taken when performing calculations based on the anelastic approximation in stably stratified regime. It remains to be seen how well the anelastic approximation performs in fully nonlinear simulations of these instabilities but the validity of updating the reference state is called into question.

## 8.2 Where the Anelastic Approximation is used

It is important to our understanding of dynamo theory to be able to model the interior of the Sun. Understanding how magnetic fields are generated and rise to the surface is a problem that requires large-scale and high resolution simulations.

The initial rise of magnetic field through the convection zone is due to the magnetic buoyancy instability. Understanding how magnetic fields rise in a stably stratified atmosphere is important as the likely generation site for large-scale magnetic field is in the stably stratified Tachocline. Above the Tachocline is the weakly superadiabatic convection zone and so magnetic buoyancy simulations which are capable of modelling both regions accurately are important.

Sunspots can be seen at the surface of the Sun. Sunspots are sites of localised strong magnetic field. The mechanism which generates such a strong field is still debated. To model the flux tubes rise through the convection zone an understanding of magnetoconvection is important. Studies of magnetoconvection can help to explain the structure of sunspots and the interaction between a strongly magnetized rising element in a convecting plasma.

The anelastic approximation is used in many nonlinear codes that model the Sun and

other stars. Anelastic codes are also used in atmospheric modelling (see e.g. Barranco & Marcus, 2006; Delden, 1992; Ashworth et al., 1997). The anelastic approximation is used as it greatly reduces the required time-step and simplifies interpreting the results compared with solving a fully compressible system. The anelastic approximation is also used in theoretical work as it can simplify the analytics of the problem. The anelastic approximation has advantages over the Boussinesq approximation in that the anelastic equations allow density fluctuations and a stratified background state.

Which version of the anelastic equations used is important as it effects the types of equation that any computer program will have to solve. It is known in the case of a perfectly adiabatic atmosphere all of the versions of the anelastic approximation will be equivalent. The results presented in this thesis may be of help in understanding when one version is more appropriate than another; i.e. in the linear case when the atmosphere departs from adiabatic then the Lantz-Braginsky approximations produces less accurate results whereas in a nearly perfectly adiabatic simulation in a nonlinear code then the process for solving for the pressure could create inaccuracies and so the Lantz-Braginsky approximations maybe more appropriate.

## 8.3 Extensions

The work presented in this thesis can be extended by building the two-dimensional nonlinear code outlined in Chapter 7; then simulations with the fully compressible, the full anelastic, and the Lantz-Braginsky approximation could be compared over a range of parameters. Particular attention should be given to when the magnetic field is strong or when the atmosphere departures from being adiabatic as these are the cases where the differences were most marked in the linear results of Chapters 5 and 6. Furthermore the potential problem with how the pressure term is solved, expanded on in §7.3.1, could be further explored by having a code using the Lantz-Braginsky approximation which solves for the pressure and uses this in the next time-step against one that does use the

pressure explicitly in the next time-step. A nonlinear investigation into the different diffusivities used in the energy or entropy equation could also be interesting as entropy diffusion has becoming increasingly common since proposed in Gilman & Glatzmaier (1981).

It would also be useful to include other aspects such as rotation. Rotation breaks some of the symmetries in the magnetoconvection problem and so would make the simulations more realistic. Although the Sun is not a rapid rotator this does not preclude rotation from playing a crucial role. It may have an effect on how well the anelastic approximation performs and it will certainly effect the results for the magnetoconvection and magnetic buoyancy instabilities. Extending the code from two-dimensional to three-dimensional would help in modelling the magnetic buoyancy instability. Magnetic buoyancy tends to favour three-dimensional modes and so restricting it to two-dimensions means that these modes cannot be investigated.

There are other approximations which are of interest in solar modelling such as the sound-proof approximations which have been compared against the anelastic approximation, albeit without a magnetic field, in Nance & Durran (1994). It would also be possible to develop implicit solvers for the fully compressible equations and compare these against the anelastic approximation. A full non-dimensional parameter study would allow the most suitable approximation for each type of problem to be known before the precise problem of interest was tackled.

# Appendices

## A Compressible Linear Equations

The compressible equation may be linearised. On assuming a normal mode solution for the perturbations about a basic state $\xi_b$, in which all variables take the form $\xi(x,y,z,t) = \xi_b(z) + \hat{\xi}(z)\exp(\sigma t + \mathrm{i}k_x x + \mathrm{i}k_y y)$ where $\xi$ is the full variable and $\hat{\xi}$ is the perturbation. The linear equations in the compressible model (equations (2.6a-d)) may be expressed as

$$\sigma\hat{\rho} = -\mathrm{i}k_x\rho_b\hat{u}_x - \mathrm{i}k_y\rho_b\hat{u}_y - {\rho_b}'\hat{u}_z - \rho_b{\hat{u}_z}', \tag{A.1}$$

$$\begin{aligned}\sigma\hat{u}_x = {} & -\mathrm{i}k_x\hat{T} - \left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\left(\frac{4k_x^2+3k_y^2}{3\rho_b}\right)\hat{u}_x - k_xk_y\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}\hat{u}_y + \frac{\mathcal{F}B_{zb}}{\rho_b}{\hat{b}_x}' \\ & + \frac{\mathcal{F}}{\rho_b}\left(B'_{xb} - \mathrm{i}k_xB_{zb}\right)\hat{b}_z - \mathrm{i}k_x\frac{T_b}{\rho_b}\rho + \mathrm{i}k_x\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}{\hat{u}_z}' + \left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{\rho_b}{\hat{u}_x}'',\end{aligned} \tag{A.2}$$

$$\begin{aligned}\sigma\hat{u}_y = {} & -\mathrm{i}k_y\hat{T} - k_xk_y\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}\hat{u}_x - \left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\left(\frac{3k_x^2+4k_y^2}{3\rho_b}\right)\hat{u}_y \\ & + \mathrm{i}k_y\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}{\hat{u}_z}' + \left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{\rho_b}{\hat{u}_y}'' - \mathrm{i}k_y\frac{\mathcal{F}B_{xb}}{\rho_b}\hat{b}_x + \mathrm{i}k_x\frac{\mathcal{F}B_{xb}}{\rho_b}\hat{b}_y \\ & - \mathrm{i}k_y\frac{\mathcal{F}B_{zb}}{\rho_b}\hat{b}_z + \frac{\mathcal{F}B_{zb}}{\rho_b}{\hat{b}_y}' - \mathrm{i}k_y\frac{T_b}{b_b}\hat{\rho},\end{aligned} \tag{A.3}$$

$$\sigma\hat{u}_z = -\frac{\rho_b{}'}{\rho_b}\hat{T} - \left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{\rho_b}\left(k_x^2+k_y^2\right)\hat{u}_z - \frac{\mathcal{F}B'_{xb}}{\rho_b}\hat{b}_x + \mathrm{i}k_x\frac{\mathcal{F}B_{xb}}{\rho_b}\hat{b}_z + \frac{1-T_b{}'}{\rho_b}\hat{\rho}$$
$$- \hat{T}' + \mathrm{i}k_x\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}\hat{u_x}' + \mathrm{i}k_y\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}\hat{u_y}' - \frac{\mathcal{F}B_{xb}}{\rho_b}\hat{b_x}' - \frac{T_b}{\rho_b}\hat{\rho}'$$
$$+ 4\left(\frac{\mathrm{Pr}}{R}\right)^{1/2}\frac{1}{3\rho_b}\hat{u_z}'', \tag{A.4}$$

$$\sigma\hat{T} = -\frac{\gamma}{(\mathrm{Pr}\,R)^{1/2}\,\rho_b}(k_x^2+k_y^2)\hat{T} + \frac{\gamma}{(\mathrm{Pr}\,R)^{1/2}\,\rho_b}\hat{T}'' - \mathrm{i}k_x(\gamma-1)T_b\hat{u}_x$$
$$- \mathrm{i}k_y(\gamma-1)T_b\hat{u}_y - T_b{}'\hat{u}_z - (\gamma-1)T_b\hat{u_z}' - 2\mathrm{i}k_x(\gamma-1)C_k^2\,(\mathrm{Pr}\,R)^{1/2}\,\zeta\mathcal{F}\frac{B'_{xb}}{\rho_b}\hat{b}_z$$
$$+ 2(\gamma-1)C_k^2\,(\mathrm{Pr}\,R)^{1/2}\,\mathrm{Pr}\,\frac{B'_{xb}}{\rho_b}\hat{b_x}', \tag{A.5}$$

$$\sigma\hat{b}_x = B_{zb}\hat{u_x}' - \mathrm{i}k_yB_{xb}\hat{u_y} - B'_{xb}\hat{u_z} - B_{xb}\hat{u_z}' - \left(k_x^2+k_y^2\right)\zeta C_k\hat{b}_x + \zeta C_k\hat{b_x}'', \tag{A.6}$$

$$\sigma\hat{b}_y = \mathrm{i}k_xB_{xb}\hat{u_y} + B_{zb}\hat{u_y}' - \left(k_x^2+k_y^2\right)\zeta C_k\hat{b_y} + \zeta C_k\hat{b_y}'', \tag{A.7}$$

$$\sigma\hat{b}_z = -\mathrm{i}B_{zb}k_x\hat{u_x} - \mathrm{i}B_{zb}k_y\hat{u_y} + \mathrm{i}B_{xb}k_x\hat{u_z} - \left(k_x^2+k_y^2\right)\zeta C_k\hat{b_z} + \zeta C_k\hat{b_z}'', \tag{A.8}$$

# B Anelastic Linear Equations

The linear perturbation equations of the anelastic approximation are similar to the linear compressible equations, except there is a reference state $\overline{\xi}$ about which the anelastic approximation is derived. I again assume a normal mode solution of the form $\xi(x,y,z,t) = \xi_b^*(z) + \hat{\xi}(z)\exp(\sigma t + \mathrm{i}k_x x + \mathrm{i}k_y y)$ where $\xi$ is the fluctuation variable and $\hat{\xi}$ is the perturbation about the basic state $\xi_b^*(z)$. Using this normal mode form the linear equations for $s, \mathbf{u}, \mathbf{B}$ and $p$ in the anelastic model (equations (3.20a-f) but where temperature diffusion is dominant so entropy diffusion is ignored) may be expressed as

$$\sigma s = \left(\frac{\overline{\rho}''\overline{\rho} - 2\overline{\rho}'^2}{\overline{\rho}^2} + k_x^2 + k_y^2\right)\frac{(1-\gamma)}{\gamma\left(\Pr\tilde{R}\right)^{1/2}\overline{p}}p + \frac{2(1-\gamma)\overline{\rho}'}{\gamma\left(\Pr\tilde{R}\right)^{1/2}\overline{\rho p}}p' - \frac{1-\gamma)}{\gamma\left(\Pr\tilde{R}\right)^{1/2}\overline{p}}p'' - (\overline{s}+s_b^*)'u_z + \left(\frac{\overline{T}''}{\overline{T}} - k_x^2 - k_y^2\right)\frac{1}{\overline{\rho}}s + \frac{2\overline{T}'}{\overline{T}\overline{\rho}}s' + \frac{1}{\overline{\rho}}s'' + 2(\gamma-1)\tilde{C_k}^2\left(\Pr\tilde{R}\right)^{1/2}\tilde{\mathcal{F}}\zeta\frac{B'_{xb}}{\overline{T}\overline{\rho}}b_x{}' - 2\mathrm{i}(\gamma-1)\tilde{C_k}^2\left(\Pr\tilde{R}\right)^{1/2}\tilde{\mathcal{F}}\zeta\frac{B'_{xb}}{\overline{T}\overline{\rho}}k_x b_z, \tag{B.9}$$

$$\sigma u_x = -\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\left(\frac{4}{3}k_x^2 + k_y^2\right)u_x + \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}u_x'' - \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_x k_y}{3}u_y + \mathrm{i}\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_x}{3}u_z{}' + \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B_{zb}b_x{}' + \frac{\tilde{\mathcal{F}}}{\overline{\rho}}\left(B'_{xb} - \mathrm{i}k_x B_{zb}\right)b_z - \mathrm{i}\frac{k_x}{\overline{\rho}}p, \tag{B.10}$$

$$\sigma u_y = -\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_x k_y}{3}u_x - \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\left(k_x^2 + \frac{4}{3}k_y^2\right)u_y + \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}u_y'' + \mathrm{i}\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_y}{3}u_z{}' - \mathrm{i}\frac{\tilde{\mathcal{F}}}{\overline{\rho}}k_y B_{xb}b_x + \mathrm{i}\frac{\tilde{\mathcal{F}}}{\overline{\rho}}k_x B_{xb}b_y + \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B_{zb}b_y{}' - \mathrm{i}\frac{\tilde{\mathcal{F}}}{\overline{\rho}}k_y B_{zb}b_z - \mathrm{i}\frac{k_y}{\overline{\rho}}p, \tag{B.11}$$

$$\sigma u_z = \mathrm{i}\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_x}{3}u_x{}' + \mathrm{i}\left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_y}{3}u_y{}' - \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\left(k_x^2 + k_y^2\right)u_z + \left(\frac{\Pr}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{4}{3}u_z'' - \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B'_{xb}b_x - \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B_{xb}b_x{}' + \mathrm{i}\frac{\tilde{\mathcal{F}}}{\overline{\rho}}k_x B_{xb}b_z - s + \frac{1}{\gamma\overline{p}}p - \frac{1}{\overline{\rho}}p', \tag{B.12}$$

$$\sigma b_x = B_{zb}u_x{}' - \mathrm{i}k_y B_{xb}u_y - B'_{xb}u_z - B_{xb}u_z{}' - \left(k_x^2 + k_y^2\right)\zeta\tilde{C_k}b_x + \zeta\tilde{C_k}b_x'', \tag{B.13}$$

$$\sigma b_y = \mathrm{i}k_x B_{xb}u_y + B_{zb}u_y{}' - \left(k_x^2 + k_y^2\right)\zeta\tilde{C_k}b_y + \zeta\tilde{C_k}b_y'', \tag{B.14}$$

$$\sigma b_z = -\mathrm{i}B_{zb}k_x u_x - \mathrm{i}B_{zb}k_y u_y + \mathrm{i}B_{xb}k_x u_z - \left(k_x^2 + k_y^2\right)\zeta\tilde{C_k}b_z + \zeta\tilde{C_k}b_z'', \tag{B.15}$$

$$0 = \mathrm{i}k_x\overline{\rho}u_x + \mathrm{i}k_y\overline{\rho}u_y + \overline{\rho}'u_z + \overline{\rho}u_z{}'. \tag{B.16}$$

In the Lantz-Braginsky simplification equation (B.12) is

$$\sigma u_z = \mathrm{i}\left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_x}{3}{u_x}' + \mathrm{i}\left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{k_y}{3}{u_y}' - \left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\left(k_x^2+k_y^2\right)u_z + \left(\frac{\mathrm{Pr}}{\tilde{R}}\right)^{1/2}\frac{1}{\overline{\rho}}\frac{4}{3}{u_z}'' - \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B'_{xb}b_x - \frac{\tilde{\mathcal{F}}}{\overline{\rho}}B_{xb}{b_x}' + \mathrm{i}\frac{\tilde{\mathcal{F}}}{\overline{\rho}}k_x B_{xb}b_z - s + \frac{\overline{\rho}'}{\overline{\rho}^2}p - \frac{1}{\overline{\rho}}p'. \tag{B.17}$$

# C Relations Between Dimensionless Numbers

In this thesis the following dimensionless numbers are used:

$$\tilde{C}_k = \epsilon^{-1}2C_k = \frac{k}{dc_p\rho_0\sqrt{c_p-c_v)T_0}}, \quad \tilde{\mathcal{F}} = \epsilon^{-1}\mathcal{F} = \frac{{B_0}^2}{gd\rho_0\mu_0},$$
$$\mathrm{Pr} = \frac{\mu_r}{\kappa_r\rho_0}, \quad \zeta = \frac{\eta_r}{\kappa_r}\text{and} \quad \tilde{R} = \epsilon R = \frac{gc_p d^3\rho_0^2}{k\mu}. \tag{C.18}$$

The Chandrasekhar number is

$$Q = \frac{B_0^2 d^2}{\eta_r\mu_0\mu_r}. \tag{C.19}$$

Roxburgh (2007) uses a parameter $\mathcal{P}$ defined at mid-layer

$$\tilde{\mathcal{P}} = \epsilon\mathcal{P} = \frac{\left(c_p-c_v\right)\beta z_0 d^2}{\kappa_r^2} \tag{C.20}$$

The Chandrasekhar number and $\mathcal{P}$ are related to the dimensionless numbers used in this work via

$$Q = \frac{\mathcal{F}\mathcal{P}}{\zeta\,\mathrm{Pr}\,\epsilon} \qquad \mathcal{P} = \frac{1}{C_k^2} = \frac{R\,\mathrm{Pr}}{(m+1)\theta}\left(1+\frac{\theta}{2}\right)^{-2m+1}, \tag{C.21}$$

where the last equation made use of the relation

$$R = \frac{\theta(m+1)}{C_k^2\mathrm{Pr}}.$$

The Froude number is

$$F = \kappa_0^2/(\epsilon c_p T_0 d^2), \tag{C.22}$$

and it is related to $C_k$ via

$$F = \frac{\gamma-1}{\gamma}C_k^2 \tag{C.23}$$

# Bibliography

Acheson, D. J. (1979). Instability by magnetic buoyancy. *Solar Phys.*, *62*, 23–50.

Acheson, D. J., & Gibbons, M. P. (1978). On the instability of toroidal magnetic fields and differential rotation in stars. *Philos. Trans. R. Soc. London, Ser. A*, *289*(1363), 459–500.

Alfvén, H. (1943). On the Existence of Electromagnetic-Hydrodynamic Waves. *Arkiv for Astronomi*, *29*, 1–7.

Anderson, E., Bai, Z., Bischof, C., Blackford, S., Demmel, J., Dongarra, J., Du Croz, J., Greenbaum, A., Hammarling, S., McKenney, A., & Sorensen, D. (1999). *LAPACK Users' Guide*. Philadelphia, PA: Society for Industrial and Applied Mathematics, third ed.

Anufriev, A. P., Jones, C. A., & Soward, A. M. (2005). The Boussinesq and anelastic liquid approximations for convection in the Earth's core. *Phys. Earth Planet. Inter.*, *152*, 163–190.

Ashworth, M., Foelkel, F., Glzow, V., Kleese, K., Eppel, D. P., Kapitza, H., & Unger, S. (1997). Parallelization of the gesima mesoscale atmospheric model. *Parallel Comput.*, *23*(14), 2201 – 2213.

Augustson, K. C., Brun, A. S., & Toomre, J. (2010). Convection and dynamo action in B stars. In N. Brummell, & A. Brun (Eds.) *Astrophysical Dynamics: From Galaxies to Stars*, 271, (pp. 100–101). International Astronomical Union.

Barranco, J. A., & Marcus, P. S. (2006). A 3D spectral anelastic hydrodynamic code for shearing, stratified flows. *J. Comput. Phys., 219*, 21–46.

Batchelor, G. K. (1953). The conditions for dynamical similarity of motions of a frictionless perfect-gas atmosphere. *Quart. J. Roy. Meteor. Soc., 79*, 224–235.

Beer, J. (2000). Long-term indirect indices of solar variability. *Space Sci. Rev., 94*, 53–66.

Berkoff, N. A., Kersale, E., & Tobias, S. M. (2010). Comparison of the anelastic approximation with fully compressible equations for linear magnetoconvection and magnetic buoyancy. *Geophys. Astrophys. Fluid Dyn., 104*, 545–563.

Bernstein, I. B., Frieman, E. A., Kruskal, M. D., & Kulsrud, R. M. (1958). An Energy Principle for Hydromagnetic Stability Problems. *Proc. R. Soc. A, 244*, 17–40.

Braginsky, S. I., & Roberts, P. H. (1995). Equations governing convection in earth's core and the geodynamo. *Geophys. Astrophys. Fluid Dyn., 79*, 1–97.

Brown, B. P. (2009). *Convection and dynamo action in rapidly rotating Suns*. Ph.D. thesis, University of Colorado at Boulder, Colorado, U.S.A.

Brown, B. P., Browning, M. K., Brun, A. S., Miesch, M. S., & Toomre, J. (2007). Rapid rotation, active nests of convection and global-scale flows in solar-like stars. *Astron. Nachr., 328*, 1002–1005.

Brummell, N. H., Clune, T. L., & Toomre, J. (2002). Penetration and Overshooting in Turbulent Compressible Convection. *Astrophys. J., 570*, 825–854.

Brun, A. S., & Toomre, J. (2002). Turbulent Convection under the Influence of Rotation: Sustaining a Strong Differential Rotation. *Astrophys. J., 570*, 865–885.

Burden, R., & Faires, J. (2005). *Numerical analysis*. Thomson Brooks/Cole.

Canuto, C. (2006). *Spectral methods: fundamentals in single domains*. Scientific computation. Springer.

Cattaneo, F. (1984). Oscillatory convection in sunspots. In T. D. Guyenne & J. J. Hunt (Ed.) *ESA Special Publication*, vol. 220, (pp. 47–50).

Cattaneo, F., Emonet, T., & Weiss, N. (2003). On the Interaction between Convection and Magnetic Fields. *Astrophys. J.*, *588*, 1183–1198.

Cattaneo, F., & Hughes, D. W. (1988). The nonlinear breakup of a magnetic layer - instability to interchange modes. *J. Fluid Mech.*, *196*, 323–344.

Cattaneo, F., & Hughes, D. W. (2001). Solar dynamo theory : Solar dynamo theory: a new look at the origin of small-scale magnetic fields. *Astron. Geophys.*, *42*(3), 3.18–3.22.

Chandrasekhar, S. (1961). *Hydrodynamic and Hydromagnetic Stability*. Oxford: Clarendon Press.

Charbonneau, P. (2010). Dynamo Models of the Solar Cycle. *Living Rev. Solar Phys.*, *7*, 3–94.

Charbonneau, P., Christensen-Dalsgaard, J., Henning, R., Larsen, R., Schou, J., Thompson, M., & Tomczyk, S. (1999). Helioseismic constraints on the structure of the solar tachocline. *Astrophys. J.*, *527*, 445–460.

Christensen-Dalsgaard, J. (1985). Speed of sound in the solar interior. *Nature*, *315*, 378 – 382.

Christensen-Dalsgaard, J. (1988). Study of solar structure based on p-mode helioseismology. In E. J. Rolfe (Ed.) *Seismology of the Sun and Sun-Like Stars*, vol. 286 of *ESA Special Publication*, (pp. 431–450).

Clune, T. C., Elliott, J. R., Miesch, M. S., Toomre, J., & Glatzmaier, G. A. (1999). Computational aspects of a code to study rotating turbulent convection in spherical shells. *Parallel Comput.*, *25*(4), 361–380.

Courant, R., Friedrichs, K., & Lewy, H. (1967). On the partial difference equations of mathematical physics. *IBM Journal of Research and Development*, *11*(2), 215–234.

Cowling, T. G. (1933). The magnetic field of sunspots. *Mon. Not. R. Astr. Soc.*, *94*, 39–48.

Crank, J., & Nicolson, P. (1947). A Practical Method for Numerical Evaluation of Solutions of Partial Differential Equations of the Heat Conduction Type. *Proc. Cam. phil. Soc.*, *43*, 50–64.

Davidson, P. A. (2001). *An Introduction to Magnetohydrodynamics*. Cambridge Texts in Applied Mathematics. Cambridge University Press.

Deardorff, J. W. (1964). A Numerical Study of Two-Dimensional Parallel-Plate Convection. *J. Atmos. Sci.*, *21*, 419–438.

Delden, A. V. (1992). The dynamics of meso-scale atmospheric circulations. *Phys. Rep.*, *211*(6), 251 – 374.

Di Mauro, M. P. (2008). Helioseismology. *Astrophysics and Space Sciences Transactions*, *4*(1), 13–17.
URL `http://www.astrophys-space-sci-trans.net/4/13/2008/`

Durney, B. (1970). Nonaxisymmetric Convection in a Rotating Spherical Shell. *Astrophys. J.*, *161*, 1115–1128.

Durran, D. R. (1989). Improving the Anelastic Approximation. *J. Atmos. Sci.*, *46*, 1453–1461.

Eddy, J. A. (1976). The maunder minimum. *Science*, *192*(4245), 1189–1202.

Fan, Y. (2001). Nonlinear Growth of the Three-dimensional Undular Instability of a Horizontal Magnetic Layer and the Formation of Arching Flux Tubes. *Astrophys. J.*, *546*, 509–527.

Fan, Y. (2009). Magnetic Fields in the Solar Convection Zone. *Living Rev. Solar Phys.*, *6*, 4–100.

Fearn, D. R. (1985). The method of inverse iteration.

Frigo, M., & Johnson, S. G. (2005). The design and implementation of FFTW3. *Proceedings of the IEEE, 93*(2), 216–231. Special issue on "Program Generation, Optimization, and Platform Adaptation".

Galilei, G., Welser, M., & de Filiis, A. (1613). *Istoria e dimostrazioni intorno alle macchie solari e loro accidenti comprese in tre lettere scritte allillvstrissimo signor Marco Velseri Linceo Si aggiungono nel fine le Lettere, e Disquisizioni del finto Apelle.bound with:De Maculis Solaribus Tres Epistolae. De Iisdem Et Stellis Circa Iovem Errantibus. Disquisitio Ad Marcum Velserum.*.

Galloway, D. J., Proctor, M. R. E., & Weiss, N. O. (1977). Formation of intense magnetic fields near the surface of the sun. *Nature*, *266*, 686–689.

Garaud, P. (2002). Dynamics of the solar tachocline - I. An incompressible study. *Mon. Not. R. Astron. Soc.*, *329*, 1–17.

Gilman, P. A., & Glatzmaier, G. A. (1981). Compressible convection in a rotating spherical shell. i. anelastic equations. *Astrophys. J. Suppl. S.*, *45*, 335–388.

Glatzmaier, G. A. (1984). Numerical simulations of stellar convective dynamos. I - The model and method. *J. Comput. Phys.*, *55*, 461–484.

Glatzmaier, G. A. (1985). Numerical simulations of stellar convective dynamos. II - Field propagation in the convection zone. *Astrophys. J.*, *291*, 300–307.

Glatzmaier, G. A. (2005). Planetary and stellar dynamos: challenges for next generation models. In A. M. Soward, C. A. Jones, D. W. Hughes, & N. O. Weiss (Eds.) *Fluid Dynamics and Dynamos in Astrophysics and Geophysics*, (pp. 331–357). London: CRC Press.

Glatzmaier, G. A., & Roberts, P. H. (1996). An anelastic evolutionary geodynamo simulation driven by compositional and thermal convection. *Physica D*, *97*(1-3), 81 – 94.

Gough, D. (2007). An introduction to the solar tachocline. In D. W. Hughes, R. R., & N. O. Weiss (Eds.) *The Solar Tachocline*, (pp. 3–30). Cambridge University Press.

Gough, D. O. (1969). The anelastic approximation for thermal convection. *J. Atmos. Sci.*, *26*, 448–456.

Gough, D. O., & McIntyre, M. E. (1998). Inevitability of a magnetic field in the Sun's radiative interior. *Nature*, *394*, 755–757.

Gough, D. O., Moore, D. R., Spiegel, E. A., & Weiss, N. O. (1976). Convective Instability in a Compressible Atmosphere. II. *Astrophys. J.*, *206*, 536–542.

Hairer, E., Nørsett, S., & Wanner, G. (1993). *Solving Ordinary Differential Equations: Stiff and differential-algebraic problems*. Springer series in Computational Mathematics. Springer-Verlag.

Hale, G. E. (1908). On the Probable Existence of a Magnetic Field in Sun-Spots. *Astrophys. J.*, *28*, 315–343.

Hughes, D. W. (2007). Magnetic buoyancy instabilities in the tachocline. In D. W. Hughes, R. Rosner, & N. O. Weiss (Eds.) *The Solar Tachocline*, (pp. 275–298). Cambridge: Cambridge Univ. Press.

Hughes, D. W., & Cattaneo, F. (1987). A new look at the instability of a stratified horizontal magnetic field. *Geophys. Astrophys. Fluid Dyn.*, *39*, 65–81.

Hughes, D. W., & Proctor, M. R. E. (1988). Magnetic fields in the solar convection zone - Magnetoconvection and magnetic buoyancy. *Annu. Rev. Fluid Mech.*, *20*, 187–223.

Hurlburt, N. E., Matthews, P. C., & Proctor, M. R. E. (1996). Nonlinear compressible convection in oblique magnetic fields. *Astrophys. J.*, *457*(2, Part 1), 933.

Jones, C. A., Boronski, P., Brun, A. S., Glatzmaier, G. A., Gastine, T., & Wicht, J. (2011). Anelastic convection-driven dynamo benchmarks.

Jones, C. A., & Kuzanyan, K. M. (2009). Compressible convection in the deep atmospheres of giant planets. *Icarus*, *204*, 227–238.

Jones, C. A., Kuzanyan, K. M., & Mitchell, R. H. (2009). Linear theory of compressible convection in rapidly rotating spherical shells, using the anelastic approximation. *J. Fluid Mech.*, *634*, 291–319.

Kersalé, E., Hughes, D. W., Ogilvie, G. I., Tobias, S. M., & Weiss, N. O. (2004). Global Magnetorotational Instability with Inflow. I. Linear Theory and the Role of Boundary Conditions. *Astrophys. J.*, *602*, 892–903.

Kosovichev, A. G., Schou, J., Scherrer, P. H., Bogart, R. S., Bush, R. I., Hoeksema, J. T., Aloise, J., Bacon, L., Burnette, A., de Forest, C., Giles, P. M., Leibrand, K., Nigam, R., Rubin, M., Scott, K., Williams, S. D., Basu, S., Christensen-Dalsgaard, J., Dappen, W., Rhodes, E. J., Jr., Duvall, T. L., Jr., Howe, R., Thompson, M. J., Gough, D. O., Sekii, T., Toomre, J., Tarbell, T. D., Title, A. M., Mathur, D., Morrison, M., Saba, J. L. R., Wolfson, C. J., Zayer, I., & Milford, P. N. (1997). Structure and Rotation of the Solar Interior: Initial Results from the MDI Medium-L Program. *Solar Phys.*, *170*, 43–61.

Krause, F., & Rädler, K.-H. (1980). *Mean-Field Magnetohydrodynamics and Dynamo Theory*. Oxford; New York: Pergamon Press.

Lantz, S. R. (1992). *Dynamical Behavior of Magnetic Fields in a Stratified, Convecting Fluid Layer*. Ph.D. thesis, Cornell University, Ithaca, U.S.A.

Lantz, S. R., & Fan, Y. (1999). Anelastic magnetohydrodynamic equations for modeling solar and stellar convection zones. *Astrophys. J. Suppl. S.*, *121*, 247–264.

Leverington, D. (2003). *Babylon to Voyager and beyond: a history of planetary astronomy*. Cambridge University Press.

Lilly, D. K. (1996). A comparison of incompressible, anelastic and Boussinesq dynamics. *Atmos. Res.*, *40*, 143–151.

Matthews, P. C., Hurlburt, N. E., Proctor, M. R. E., & Brownjohn, D. P. (1992). Compressible magnetoconvection in oblique fields linearized theory and simple nonlinear models. *J. Fluid Mech.*, *240*, 559–569.

Maunder, E. W. (1904). Note on the distribution of sun-spots in heliographic latitude, 1874-1902. *Mon. Not. R. Astr. Soc.*, *64*, 747–761.

McIntyre, M. E. (2007). Magnetic confinement and the sharp tachopause. In D. W. Hughes, R. R., & N. O. Weiss (Eds.) *The Solar Tachocline*, (pp. 183–212). Cambridge University Press.

Miesch, M. S. (2005). Large-Scale Dynamics of the Convection Zone and Tachocline. *Living Reviews in Solar Physics*, *2*, 1.

Miesch, M. S., Elliott, J. R., Toomre, J., Clune, T. C., Glatzmaier, G. A., & Gilman, P. A. (2000). Three-dimensional spherical simulations of solar convection: Differential rotation and pattern evolution achieved with laminar and turbulent states. *Astrophys. J.*, *532*, 593–615.

Nance, L. B., & Durran, D. R. (1994). A Comparison of the Accuracy of Three Anelastic Systems and the Pseudo-Incompressible System. *J. Atmos. Sci.*, *51*, 3549–3565.

Newcomb, W. (1961). Convective instability induced by gravity in a plasma with a frozen-in magnetic field. *Phys. Fluids*, *4*, 391–396.

Nordlund, A. (1982). Numerical simulations of the solar granulation. I - Basic equations and methods. *Astron. Astrophys.*, *107*, 1–10.

Ogura, Y., & Phillips, N. A. (1962). Scale Analysis of Deep and Shallow Convection in the Atmosphere. *J. Atmos. Sci.*, *19*, 173–179.

Orszag, S. A. (1971). On the elimination of aliasing in finite-difference schemes by filtering high-wavenumber components. *J. Atmos. Sci.*, *28*(6), 1074–1074.

Ossendrijver, M. (2003). The solar dynamo. *Astron. Astrophys. Rev.*, *11*, 287–367.

Parker, E. N. (1955a). Hydromagnetic Dynamo Models. *Astrophys. J.*, *122*, 293–314.

Parker, E. N. (1955b). The Formation of Sunspots from the Solar Toroidal Field. *Astrophys. J.*, *121*, 491–507.

Parker, E. N. (1975). The generation of magnetic fields in astrophysical bodies. X - Magnetic buoyancy and the solar dynamo. *Astrophys. J.*, *198*, 205–209.

Pedlosky, J. (1987). *Geophysical Fluid Dynamics*. Springer-Verlag.

Priest, E., & Forbes, T. (2000). *Magnetic reconnection: MHD theory and applications*. Cambridge University Press.

Priest, E. R. (1984). *Solar Magnetohydrodynamics*. Dordrecht: Reidel.

Proctor, M. (2006). Dynamo action and the sun. In M. Rieutord & B. Dubrulle (Ed.) *EAS Publications Series*, vol. 21 of *EAS Publications Series*, (pp. 241–273).

Proctor, M. R. E. (2005). Magnetoconvection. In A. M. Soward, C. A. Jones, D. W. Hughes, & N. O. Weiss (Eds.) *Fluid Dynamics and Dynamos in Astrophysics and Geophysics*, (pp. 235–276). London: CRC Press.

Proctor, M. R. E., & Weiss, N. O. (1982). Magnetoconvection. *Rep. on Prog. in Phy.*, *45*, 1317–1379.

Rayleigh, L. (1916). On convection currents in a horizontal layer of fluid, when the higher temperature is on the under side. *Phil. Mag*, *32*(192), 529 – 546.

Reynolds, O. (1895). On the dynamical theory of incompressible viscous fluids and the determination of the criterion. *Philos. Trans. R. Soc. London, Ser. A*, (pp. 123–164).

Ribes, J. C., & Nesme-Ribes, E. (1993). The solar sunspot cycle in the Maunder minimum AD1645 to AD1715. *Astron. Astrophys.*, *276*, 549–563.

Rogers, T. M., & Glatzmaier, G. A. (2005). Penetrative convection within the anelastic approximation. *Astrophys. J.*, *620*, 432–441.

Rogers, T. M., Glatzmaier, G. A., & Jones, C. A. (2006). Numerical Simulations of Penetration and Overshoot in the Sun. *Astrophys. J.*, *653*, 765–773.

Roxburgh, N. T. (2007). *Anelastic convection in an inclined magnetic field*. Ph.D. thesis, Darwin College, Cambridge.

Schwabe, S. H. (1843). Solar observations during 1843. *Astron. Nachr.*, *21*(495), 283–288.

Silvers, L. J., Vasil, G. M., Brummell, N. H., & Proctor, M. R. E. (2011). The Evolution of a Double Diffusive Magnetic Buoyancy Instability. In *IAU Symposium*, vol. 271 of *IAU Symposium*, (pp. 218–226).

Smagorinsky, J. (1963). General circulation experiments with the primitive equations. *Mon. Weather Rev.*, *91*(3), 99–164.

Spiegel, E. A., & Veronis, G. (1960). On the Boussinesq Approximation for a Compressible Fluid. *Astrophys. J.*, *131*, 442–447.

Spiegel, E. A., & Weiss, N. O. (1982). Magnetic buoyancy and the Boussinesq approximation. *Geophys. Astrophys. Fluid Dyn.*, *22*, 219–234.

Spiegel, E. A., & Zahn, J. (1992). The solar tachocline. *Astrophys. J.*, *265*, 106–114.

Steenbeck, M., & Krause, F. (1966). Erklärung stellarer und planetarer Magnetfelder durch einen turbulenzbedingten Dynamomechanismus. *Z. Naturforsch. Teil A*, *21*, 1285–+. (English transl, in P. H. Roberts and M. Stix, NCAR-TN/IA-60, p. 49).

Thomas, J. H., & Nye, A. H. (1975). Convective instability in the presence of a nonuniform horizontal magnetic field. *Phys. Fluids*, *18*, 490–491.

Thomas, J. H., & Weiss, N. O. (2008). *Sunspots and Starspots*. Cambridge University Press.

Tobias, S., & Weiss, N. (2007). The solar dynamo and the tachocline. In D. W. Hughes, R. R., & N. O. Weiss (Eds.) *The Solar Tachocline*, (pp. 319–350). Cambridge University Press.

Tobias, S. M. (2005). The solar tachocline: Formation, stability, and its role in the solar dynamo. In A. M. Soward, C. A. Jones, D. W. Hughes, & N. O. Weiss (Eds.) *Fluid Dynamics and Dynamos in Astrophysics and Geophysics*, (pp. 193–234). London: CRC Press.

Tobias, S. M., Brummell, N. H., Clune, T. L., & Toomre, J. (1998). Pumping of magnetic fields by turbulent penetrative convection. *The Astrophysical Journal Letters*, *502*(2), L177.

Tobias, S. M., Brummell, N. H., Clune, T. L., & Toomre, J. (2001). Transport and Storage of Magnetic Field by Overshooting Turbulent Compressible Convection. *Astrophys. J.*, *549*, 1183–1203.

Tobias, S. M., & Hughes, D. W. (2004). The influence of velocity shear on magnetic buoyancy instability in the solar tachocline. *Astrophys. J.*, *603*, 785–802.

van Ballegooijen, A. A. (1982). The overshoot layer at the base of the solar convective zone and the problem of magnetic flux storage. *Astrophys. J.*, *113*, 99–112.

Vasil, G. M., & Brummell, N. H. (2008). Magnetic Buoyancy Instabilities of a Shear-generated Magnetic Layer. *Astrophys. J.*, *686*, 709–730.

Vögler, A., & Schüssler, M. (2007). A solar surface dynamo. *Astron. Astrophys.*, *465*, L43–L46.

Vonmoos, M., Beer, J., & Muscheler, R. (2006). Large variations in Holocene solar activity: Constraints from $^{10}$Be in the Greenland Ice Core Project ice core. *J. Geophys. Res. A (Space Physics)*, *111*(A10105), 1–14.

Weiss, N. O. (1966). The Expulsion of Magnetic Flux by Eddies. *Proc. R. Soc. A*, *293*, 310–328.

Wilhelmson, R., & Ogura, Y. (1972). The Pressure Perturbation and the Numerical Modeling of a Cloud. *J. Atmos. Sci.*, *29*, 1295–1307.

Wissink, J. G., Matthews, P. C., Hughes, D. W., & Proctor, M. R. E. (2000). Numerical Simulations of Buoyant Magnetic Flux Tubes. *Astrophys. J.*, *536*, 982–997.