

Modelling Magnetoconvection in Sunspots

Stephen David Thompson

Christ's College, Cambridge

September 2006

This dissertation is submitted for the degree of
Doctor of Philosophy

Preface

All work described in this dissertation is believed to be original except where explicit reference has been made to other authors. This dissertation is the result of my own work and includes nothing which is the outcome of work done in collaboration except where specifically indicated in the text. No part of this thesis has been submitted for any qualification other than the degree of Doctor of Philosophy at the University of Cambridge.

Acknowledgements

I would first of all like to thank the people that I have shared an office with during my PhD, whose company has definitely made my time at Cambridge more enjoyable: Paul Bushby, Steve Houghton, Claire Postlethwaite, Jennifer Siggers and Mike Tildesley. I am particularly grateful to Paul and Steve for their assistance with matters of numerical methods and Fortran and IDL programming.

I am grateful to Professor Mike Proctor, my supervisor, and Professor Nigel Weiss, for their all of their help and guidance throughout this work. I am also grateful to Paul Bushby, Steve Houghton and Neil Roxburgh for their assistance in proof-reading parts of this dissertation.

I would also like to thank my family for all of their support and encouragement during my PhD. Finally, I would like to acknowledge PPARC for funding this work.

Summary

Although sunspots have been observed for centuries, we do not yet have a detailed understanding of their structure and properties. This thesis considers a number of approaches to modelling sunspots, and in particular the magnetoconvective processes that take place within them. Particular emphasis is given to the penumbra, since little is currently known theoretically about this region.

Our models will fall into two categories. The first category includes simplified, reduced models, covering the linear and weakly nonlinear regimes, while the second category consists of full three-dimensional numerical simulations of compressible magnetoconvection.

We begin by constructing a simple linearized model of magnetoconvection in uniform inclined fields. We then move on to the weakly nonlinear regime, where competition between different patterns of convection can be investigated. Finally we extend the model by allowing the angle of inclination of the field to vary with position. This allows us to build up a reasonable reduced model of convection in a sunspot, showing a transition between ‘umbra’ and ‘penumbra’; although this model has its limitations, it serves to show how much can be achieved using a minimal approach, without needing a detailed knowledge of the physics of sunspots.

The remainder of the thesis presents a number of three-dimensional simulations of magnetoconvection. Our simulations were designed to illustrate how the pattern of convection changes as the angle of inclination of the field is varied, with the aim of investigating the patterns that might be expected within the penumbra. We begin with simple simulations with uniform fields, and then move on to a model in which the field ‘fans out’ with position, mimicking the situation in a sunspot. The results show an interesting transition from hexagonal to more roll-like patterns, reminiscent of the transition between umbra and penumbra found in real sunspots.

Contents

1	Introduction	9
1.1	Internal structure of the Sun	9
1.2	Sunspot observations: past and present	12
1.3	Overall structure of sunspots	17
1.3.1	The model of Jahn and Schmidt (1994)	18
1.3.2	Subsurface structure: ‘cluster’ and ‘monolith’ pictures	21
1.4	Fine structure in sunspots	22
1.4.1	Umbral fine structure	22
1.4.2	Penumbral fine structure	23
1.5	Magnetoconvection	26
1.5.1	Simulations with vertical magnetic fields	27
1.5.2	Magnetoconvection in inclined magnetic fields	28
1.5.3	Flux pumping	29
1.5.4	Simulations in cylindrical geometry	29
1.6	Formation and decay of sunspots	30
1.6.1	Emergence of magnetic fields at the solar surface	30
1.6.2	Formation of the penumbra	32
1.6.3	Decay of sunspots	33
1.6.4	The 11-year solar cycle	34
1.7	Outline of the thesis	35
2	Linear theory	37
2.1	Introduction	37
2.2	Problem setup	38
2.2.1	Boundary conditions	39
2.2.2	Basic state and linearization	40

2.3	The vertical field case	42
2.4	Symmetry considerations for inclined fields	44
2.4.1	General considerations in the two-dimensional case ($k_y = 0$)	44
2.4.2	The oscillatory case (small ζ)	46
2.4.3	The steady case	50
2.5	Results	51
2.5.1	Small ζ (oscillatory case)	51
2.5.2	Steady case	62
2.6	Conclusions	65
3	Weakly nonlinear models	67
3.1	Introduction	67
3.2	General notes on pattern formation	68
3.3	Steady bifurcation on a hexagonal lattice	72
3.3.1	Weakly breaking the isotropy	75
3.3.2	Changes to the bifurcation diagrams	77
3.4	Oscillatory bifurcation on a hexagonal lattice	86
3.4.1	An additional solution branch	90
3.4.2	Weakly broken symmetry	92
3.5	Oscillatory bifurcation on a rhombic lattice	99
3.5.1	The symmetric case	99
3.5.2	Weakly broken symmetry	101
3.5.3	Results	102
3.5.4	Comparison with results below the Lifshitz point	107
3.5.5	A calculation of the coefficients a and b	109
3.6	Conclusions	111
4	Swift-Hohenberg model for tilted magnetic fields	115
4.1	Introduction	115
4.2	The model	116
4.3	Landau-Ginzburg equations	117
4.3.1	Setting up the problem	121
4.3.2	Theoretical ideas	122
4.3.3	Numerical results	126
4.4	Numerical simulations of the full Swift-Hohenberg equation	137

4.4.1	Cartesian geometry	139
4.4.2	Cylindrical geometry	140
4.5	Conclusions	147
5	Numerical simulations	152
5.1	Introduction	152
5.2	Definition of the problem	153
5.3	Numerical methods	156
5.4	Runs with uniform fields	159
5.5	Runs with non-uniform fields: ‘Arch’ initial condition	164
5.6	Runs with non-uniform fields: ‘Flux sheet’ initial condition	171
5.6.1	Initial results	176
5.6.2	Preventing the flux from escaping	182
5.7	Conclusions	193
5.7.1	Runs with uniform fields	193
5.7.2	Runs with non-uniform fields	195
6	Conclusions and future work	200
A	Magnetic boundary conditions	204
A.1	Perfectly conducting boundaries (tied field lines)	204
A.2	Perfectly insulating boundaries (matching to a potential field)	205
B	Representations for B	208
B.1	The vector potential	208
B.1.1	Boundary conditions	209
B.2	Poloidal/toroidal decomposition	211
B.2.1	A note about periodicity	212
B.2.2	Boundary conditions	213
C	Weakly nonlinear derivations	215
C.1	The formula for A_c in the steady hexagonal model	215
C.2	Solving the weakly nonlinear amplitude equations: Some predefined equation systems	216
C.2.1	System 1	217
C.2.2	System 2	217

C.2.3	System 3	218
C.2.4	System 4	218
C.2.5	System 5	218
C.2.6	System 6	219
C.3	Rhombic model	220
C.3.1	One non-zero amplitude	220
C.3.2	Two non-zero amplitudes	221
C.3.3	Three nonzero amplitudes	223
C.3.4	Four nonzero amplitudes	223
C.4	Oscillatory hexagonal model	224
C.4.1	One non-zero amplitude	225
C.4.2	Two non-zero amplitudes	226
C.4.3	Three non-zero amplitudes	229
C.4.4	Four non-zero amplitudes	231
C.4.5	Five non-zero amplitudes	232
C.4.6	Six non-zero amplitudes	233

Chapter 1

Introduction

The Sun is familiar to all of us, but what are perhaps less familiar are the dark spots that sometimes appear on its surface (Figure 1.1). These ‘sunspots’ were first observed telescopically by Galileo and others in the seventeenth century, but even today, many of their properties remain unexplained.

Sunspots are created when strong magnetic fields, generated deep within the Sun, rise up to the solar surface. The magnetic field partially inhibits the convection that is normally found at the surface, and this weakened convection is less able to transport heat into the sunspot, making it cooler, and hence darker, than its surroundings.

What is less well understood is the detailed nature of this ‘magnetoconvection’ that is taking place within a sunspot. As we shall see in the following sections, sunspots contain detailed fine structure which is ultimately the result of magnetoconvective processes. The aim of this thesis is to try to better understand these processes by studying magnetoconvection in various situations, beginning with simple models and working up to more complex numerical simulations.

We will start by giving, in this chapter, a broad introduction to the subject of sunspots, covering both observational and theoretical topics. We will then give a brief summary of the remainder of the thesis, indicating how it relates to sunspots and to some of the outstanding theoretical questions about them.

1.1 Internal structure of the Sun

Before describing sunspots, it may be helpful to give a quick outline of the structure of the Sun itself. The Sun is a (more or less) spherical ball of gas, 696 000 km in radius,



Figure 1.1: Full disc image of the Sun taken by the Swedish 1-m Solar Telescope on 15 July 2002. Courtesy Royal Swedish Academy of Sciences.

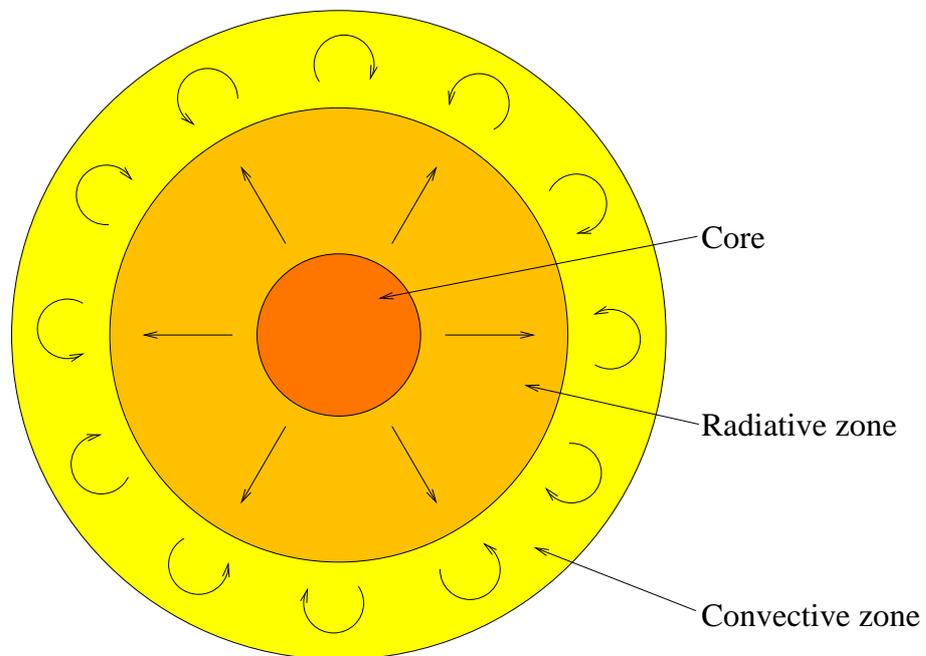


Figure 1.2: Sketch of the internal structure of the Sun.

composed mostly of hydrogen and helium. A cross-section is shown in Figure 1.2. At the centre is the core, containing approximately the inner 25% (by radius), where the density reaches values up to 150 times that of water, and the temperature reaches some 15 million degrees Kelvin. Under these extreme conditions, the nuclear fusion reactions that generate the Sun's energy can take place. Above the core, there is the radiative zone (which technically also includes the core), where heat energy is transported towards the surface by radiation, and the convective zone (including the outer 30% by radius), where convection becomes the principal means of energy transport. This convection is indeed observed at the solar surface, where it is known as 'granulation'.

The interior of the Sun cannot be observed directly, of course (although helioseismology can provide indirect measurements of the interior), and so the internal structure is calculated using models. Indeed, models of stellar structure have been quite successful in explaining the observed properties of the Sun and other stars. The basic idea in these models is to assume that the star is in a static equilibrium state. There are then five equations to be solved: one for mass continuity, one for hydrostatic force balance (the pressure gradient must balance gravity), one for energy conservation (the net outward heat flux must balance the energy generation from nuclear reactions), one for heat transport (this gives the temperature gradient required to transport a given heat flux), and finally an equation of state (relating the temperature, pressure and density of the gas). These can be solved (together with appropriate boundary conditions) to determine the internal structure of stars of various masses and compositions.

For the heat transport equation, one must consider the three different methods of heat transfer: conduction, radiation and convection. For stars like the Sun, conduction turns out to be negligible, so only the last two need to be considered. The equations describing radiative heat transfer are well understood; convection, on the other hand, is a highly nonlinear process and there is no simple formula relating the convective heat flux to the temperature gradient in a compressible fluid. Therefore, convection is usually approximated by a 'mixing length' formalism, where a fluid element is assumed to rise adiabatically by a given length (the mixing length) before giving up its excess heat to the surroundings. The mixing length itself is left as a free parameter, usually expressed as a proportion of the local pressure scale height.

The solution of such models yields (for the Sun) the structure shown in Figure 1.2. One can also use a similar approach to model subsurface structure in sunspots; the mixing length theory for convection can be adapted (in a simple way) for magnetocon-

vection, by taking a reduced mixing length parameter in order to model the reduced convective efficiency. This will be discussed further in section 1.3.

1.2 Sunspot observations: past and present

The earliest recorded observations of sunspots go back to Chinese astrologers, who were apparently observing them at least as early as the 11th century BC. The largest spots would have been visible to them with the naked eye at sunrise or sunset, or reflected on the surface of still waters. There are also various records of sunspot observations in the Western world, apparently going back as far as the ancient Greeks. However, Western religious and philosophical thinking dictated that the Sun was a celestial body, and therefore perfect in every way, and spotless. Thus, the sunspot observations were usually ignored or forgotten, and it was not until the invention of the telescope, in the seventeenth century, that the existence of sunspots became widely acknowledged in the West.

These first telescopic observations were made by Galileo, Christoph Scheiner, and David and Johannes Fabricius, around 1610. (There is still apparently some controversy as to who of these made the first observation.) One of Galileo's sunspot drawings is shown in Figure 1.3. Even from these early observations, it was clear that sunspots are composed of two distinct regions: a dark central area and a lighter outer part. Today, the inner region is known as the *umbra* (from the Latin for 'shade' or 'shadow'), and the outer area is called the *penumbra* (from Latin *paene*, 'nearly' or 'almost', + *umbra*).

Further progress was slow, in part because there followed, between the years of 1645 and 1715, a period of extremely low sunspot activity, which later became known as the Maunder Minimum. The main discovery in the eighteenth century was that of the Wilson depression, which refers to the fact that the visible surface of a sunspot is located at a deeper vertical level in the Sun than the normal photosphere. It was detected by Alexander Wilson in 1769 by careful observations of sunspots near the edge (limb) of the solar disc.

When larger telescopes became available, in the nineteenth century, it became possible to observe fine details of umbral and penumbral structure. Telescope technology has been improving continually since then and images of stunning detail are now being obtained by instruments such as the Swedish 1-metre Solar Telescope, which came into service in May 2002. Images obtained from this telescope are displayed in Figures 1.4

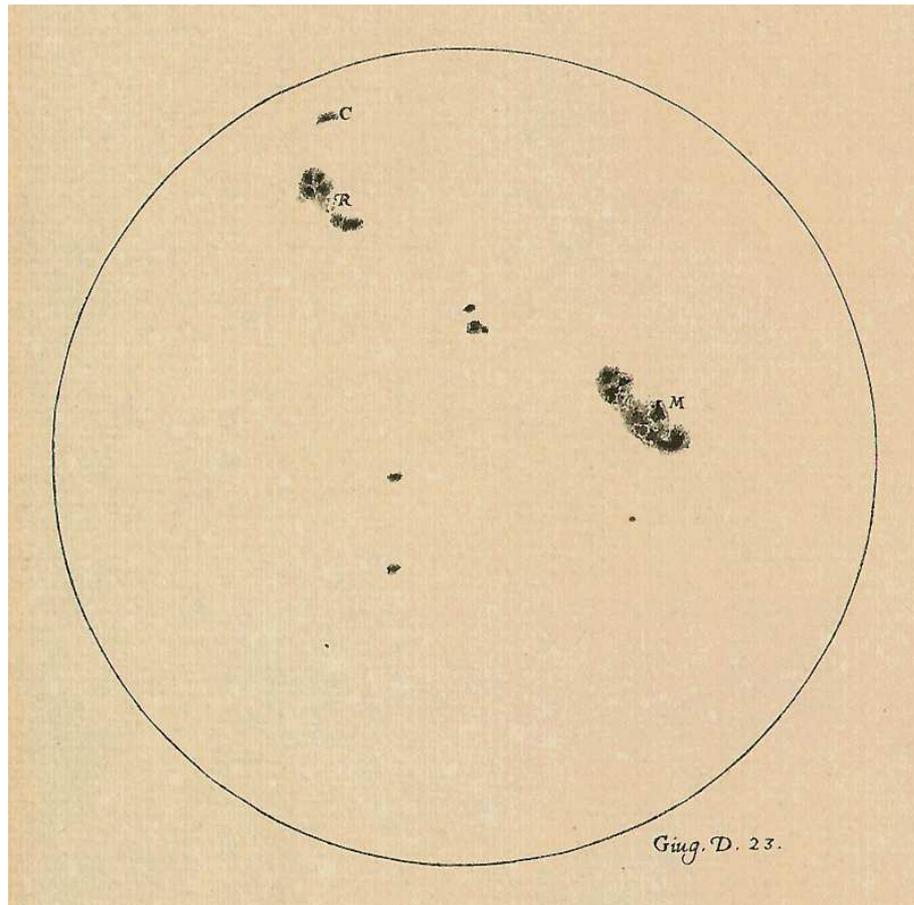


Figure 1.3: One of Galileo's sunspot drawings, from 23 June 1613.

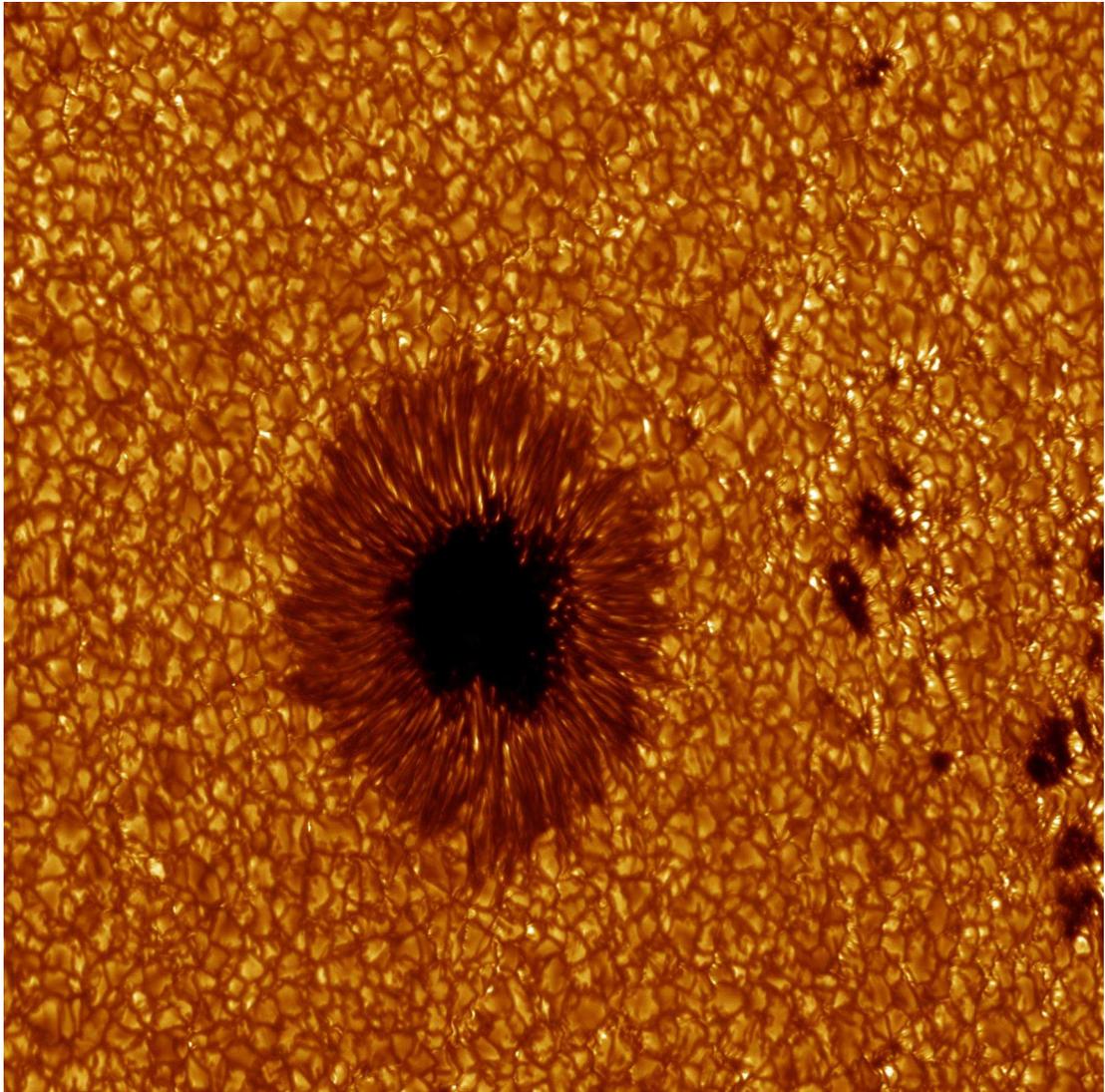


Figure 1.4: *An image taken by the Swedish 1-m Solar Telescope showing a small regular sunspot. Note also the pores visible at the right-hand side. Courtesy Royal Swedish Academy of Sciences.*

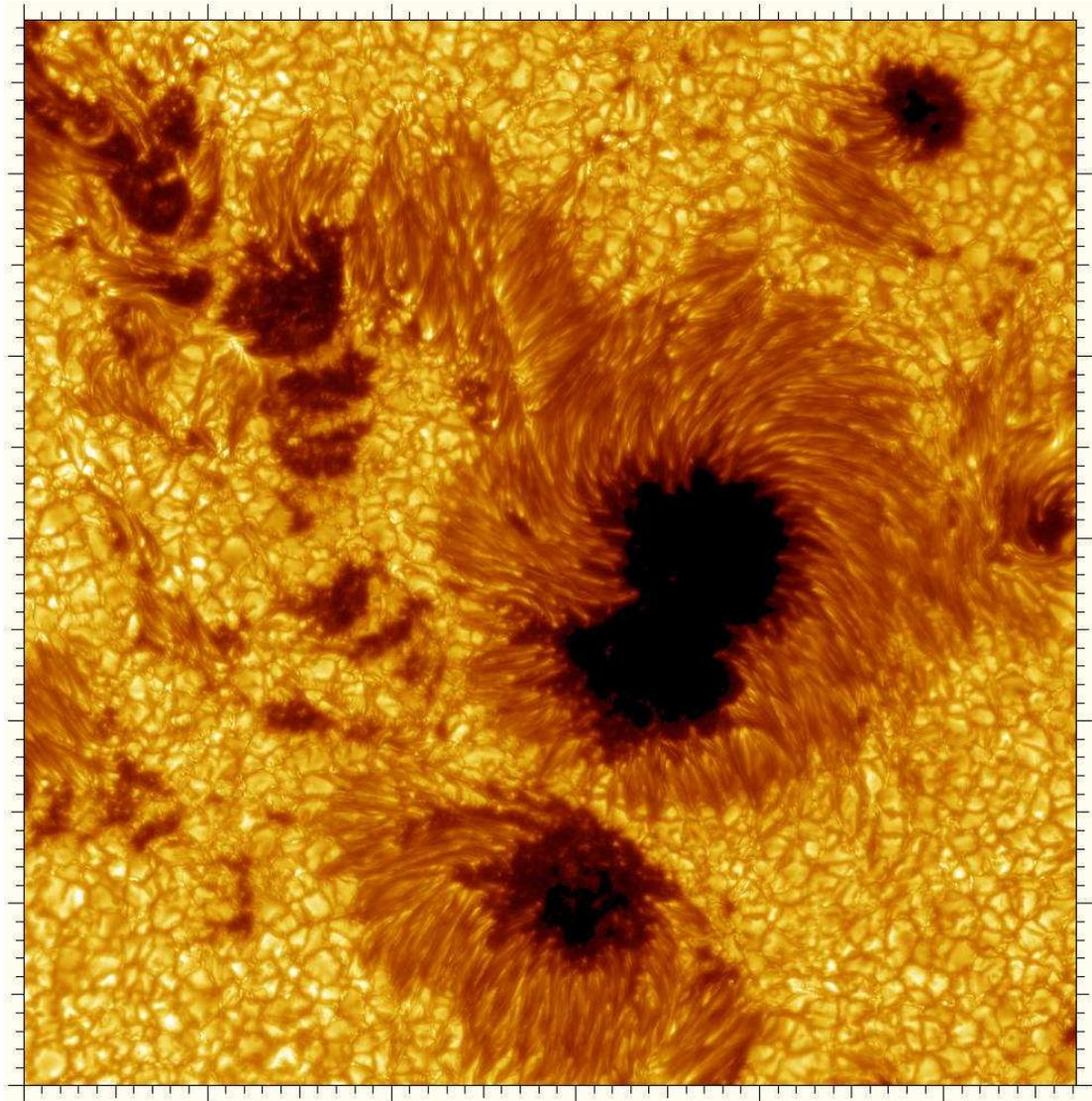


Figure 1.5: A close-up view of the sunspot group visible in Figure 1.1. These sunspots have a more irregular structure.

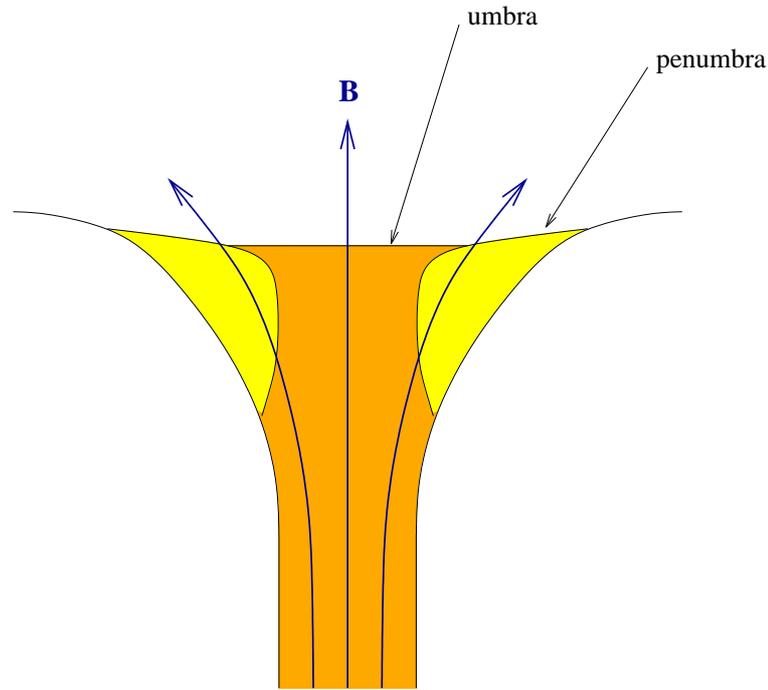


Figure 1.6: A sketch of a cross-section through a typical sunspot.

and 1.5. Sunspots, with many small-scale structures and fine details, can be seen; we will return to these fine structures in more detail in section 1.4. Note also the presence of smaller dark patches, known as pores; these are similar phenomena to sunspots, but they are smaller and they do not have penumbrae.

Our modern understanding of sunspots begins with George Ellery Hale's discovery, in 1908, that sunspots are associated with strong magnetic fields. He found this by measuring the so-called Zeeman effect, the splitting of spectral lines by a magnetic field, using an instrument that he himself invented, the spectroheliograph. In due course it was realized that this magnetic field would inhibit convection, explaining why sunspots are cooler and darker than the normal solar surface.

The magnetic field strength at the centre of a large sunspot might reach 0.3 tesla (T).¹ For comparison, the Earth's magnetic field at the surface has a strength of about 5×10^{-5} T; the field strength of a typical bar magnet might be 0.01 T; a medical MRI unit produces fields of about 1.5 T.

1.3 Overall structure of sunspots

Sunspots of many different sizes are observed. The smallest have radii of 1800 km or so, while very large spots can sometimes attain radii of 30000 km or more (Solanki, 2003). However, all well-developed regular sunspots (such as the one shown in Figure 1.4) have a similar structure, and so, at least in the case of the more regular, circular spots, we can talk about the structure of a typical sunspot.

Figure 1.6 shows a rough sketch of a vertical slice through a sunspot. The basic physical picture is that there is a vertical column of magnetic flux (indicated by the orange and yellow regions) surrounded by the ordinary field-free solar plasma. The orange region marks areas where the field remains near-vertical, which gives rise to the umbra at the surface. In the yellow regions, the field has become significantly inclined to the vertical, and it is these inclined fields that (somehow) allow the penumbra to form.

Over the years, a number of models have been put forward to try to account for the large-scale structure of sunspots.² One of the first of these models was due to Schlüter and Temesváry (1958), and a great many have been developed since; an extensive review is given in sections 4.4 to 4.8 of Solanki (2003).

To make the problem mathematically tractable, these models typically assume that the spot is circular and axisymmetric; this is not true for real sunspots, at least on small scales (e.g. the penumbra is considerably non-axisymmetric on fine scales), but it is a reasonable assumption for modelling the large-scale overall structure. The other main assumption made is that the sunspot is in equilibrium with its surroundings (both mechanically and thermally). Again, this can be justified in an average sense, since while real sunspots do have dynamic features, occurring on timescales of perhaps an hour or so, the total lifetime of a spot is much longer than this (large spots can persist for several months).

To describe the equilibrium state we basically need to ensure both a hydrostatic and a thermal equilibrium. The equation for the hydrostatic force balance is

$$-\nabla p + \rho \mathbf{g} + \mathbf{j} \wedge \mathbf{B} = 0. \quad (1.1)$$

The horizontal force balance in a sunspot is therefore between the gas pressure gradient, which exerts an inward force (because the cool interior is at a lower pressure than the

¹Magnetic field strengths are also commonly quoted in gauss (G); 1 G = 10⁻⁴ T.

²These models can also be used for pores, if the penumbra is omitted.

hot exterior), and the Lorentz force ($\mathbf{j} \wedge \mathbf{B}$), which is in an outward direction (this can be thought of as a gradient of magnetic pressure $B^2/2\mu_0$, with a higher magnetic pressure inside the spot than outside). In the vertical direction, one must balance the gas pressure gradient against gravity (as well as, possibly, Lorentz forces).

Equation 1.1, together with the equation of state, give three equations in the four unknowns p , ρ , T and \mathbf{B} . (Equation 1.1 counts as two equations because it has both vertical and horizontal components; also \mathbf{B} only really counts as one scalar unknown because we have the additional constraint $\nabla \cdot \mathbf{B} = 0$.) To close the system, an additional condition must be added. The ‘proper’ way to do this is to include an energy equation, describing the thermal equilibrium of the sunspot. However, in some models, a simpler approach is taken, in which one simply prescribes the pressure (or some other variable) as a function of depth. This produces a model which is not strictly in thermal equilibrium, but this is not as bad as it sounds, since one can vary the assumed profile until a reasonable match with observations is found.

Even from fairly simple models, it was realized early on (Schlüter and Temesváry, 1958) that the energy flux through the umbra is too great to be provided by radiation alone, and therefore convection must be occurring below the visible surface of the umbra. In other words, the magnetic field of the umbra must inhibit convection only partially, not completely.

1.3.1 The model of Jahn and Schmidt (1994)

We will now describe a recent model, that of Jahn and Schmidt (1994), as we can use it to illustrate one or two points about sunspot structure. The model is shown in Figure 1.7.

The first thing we need to do is make some assumptions about the structure of the magnetic field. Earlier models approached this by using a self-similarity method, but this does not give a particularly good fit to observations. An alternative approach, and the one used by Jahn and Schmidt (1994), is to assume that $\mathbf{j} = 0$ everywhere except on isolated current sheets. It is found that one current sheet enclosing the entire sunspot is insufficient to model the penumbra; Jahn and Schmidt (1994) chose to take two current sheets, an inner one separating the umbra from the penumbra, and an outer one enclosing the entire structure. This approach produces a sharp distinction between umbra and penumbra (as is observed) and allows the different energy transport

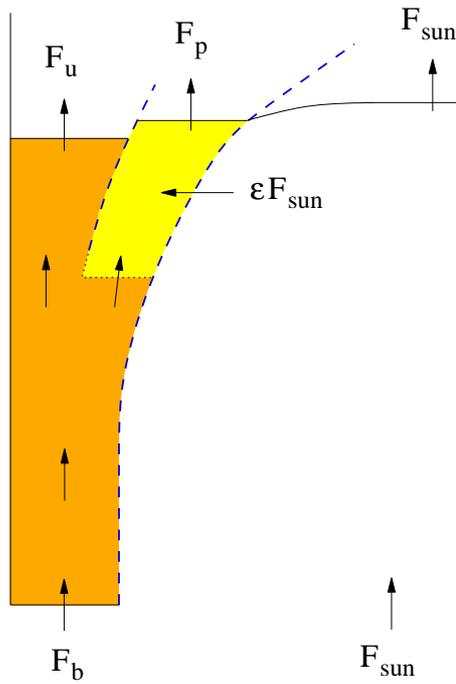


Figure 1.7: The axisymmetric sunspot model proposed by Jahn and Schmidt (1994). The small black arrows indicate heat fluxes (per unit area). The blue dashed lines indicate the position of current sheets within the model. The yellow region indicates the extent of the penumbra, defined in this model as the region where interchange convection, with associated lateral heat flux of ϵF_{sun} , takes place.

and thermodynamic properties between the two to be described. It also avoids the complexity of dealing with volume currents in the penumbra which would otherwise be needed (e.g., Jahn 1989). (One criticism, however, is that it produces a sharp jump in the magnetic field strength at the umbra-penumbra boundary, which is not observed.)

An energy equation is included (with convection being described by a mixing-length theory, using a reduced mixing length in the umbra and penumbra in order to simulate the reduced convective efficiency there). When solving this equation, boundary conditions for the energy flux are needed. At the top we can simply match to observed umbral and penumbral energy fluxes; there are also (more indirect) observations that can help with the lateral boundary conditions. In particular, it is observed that the brightness and surface temperature of the umbra both vary during the 11-year solar cycle.³ This can only be explained if the umbra is thermally well insulated from its surroundings. The penumbra shows a similar intensity variation, although with a much lower amplitude, implying that significant heat flux is being transported from the quiet Sun into the penumbra. This is usually interpreted as a convective process (sometimes called ‘interchange convection’), with material (and thermal energy) being exchanged between the penumbra and quiet Sun.

Therefore, Jahn and Schmidt (1994) assume that the flux tube comprising the sunspot is completely thermally insulated from its surroundings, except for the interchange convection that is taking place in the penumbra. This penumbral convection is modelled as a heat flux of ϵ times the normal solar heat flux, which is carried into the penumbra from the external plasma. (It is found that values of ϵ between about 0.6 and 0.7 seem to be required in order to give plausible results.) This interchange convection (and associated heat flux) extends down from the surface to a depth z_{bp} ; this depth is computed as part of the model (so that the stratifications of umbra and penumbra match), and it can be thought of as the bottom of the penumbra. Its value is typically found to be about 4000 km in this model.

We see that any model for convection in sunspots would ultimately have to explain how this behaviour for the heat flux arises. In particular, it should explain the differences between convection in the umbra, in which the heat flux is channelled along the field lines, with almost no mixing across the boundary of the flux tube, and in the penumbra, where convective interchanges apparently do occur between the penumbra and the external plasma, with a corresponding heat flux into the spot. (Notice also how

³The solar cycle is covered in section 1.6.4.

these convective processes, which occur at small scales, have an impact on the overall large-scale structure of the spot.)

1.3.2 Subsurface structure: ‘cluster’ and ‘monolith’ pictures

So far we have considered a sunspot to be a single homogeneous flux tube extending downwards beneath the surface. In fact, this is a simplification, and the flux tube will have more detailed structure than this. There are two competing theories as to what the subsurface structure looks like: either a monolithic column of flux, or a cluster of separate ‘mini’ flux tubes. (A discussion of the differences between the two models can be found in section 4.3 of Solanki 2003.)

In the monolithic model, the sunspot consists of a single flux tube that remains a coherent structure even down to great depths. The magnetic field within the tube is strong enough to reduce the efficiency of convection, but not so strong as to suppress it completely. In other words, magnetoconvection (as opposed to ordinary field-free convection) occurs inside the flux tube, and this different type of convection is why the heat flux is lower inside the flux tube than outside.

The alternative, cluster, model was first proposed by Parker (1979), who suggested that the flux tube would be subject to the so-called fluting instability, which would cause it to break up into a tight cluster of separate ‘mini flux tubes’ just below the solar surface. The spaces in between these individual tubes would contain field-free plasma. The field inside the individual tubes would be strong enough to completely suppress convection, but within the field-free regions, convection would be able to take place as normal. Thus, in this model the reduced convective efficiency is caused not by a different form of convection, but simply by a reduced spatial filling factor.

The actual conditions for the fluting instability were set out by an earlier calculation, due to Meyer et al. (1977); it was found that near the surface, the instability would be suppressed by magnetic buoyancy processes, but deeper down, the fluting instability could indeed take place. However, this calculation assumed a static equilibrium configuration, while in reality, convection takes place both inside and outside the sunspot, and this must be taken into account. The central question is therefore whether convection would be able to stabilize the flux tube against fluting. This question remains to be answered, but it has been suggested that the necessary effect could be provided by a so-called ‘collar flow’, a supergranular-scale convection flow outside the sunspot. (See

also section 1.5.4, below).

1.4 Fine structure in sunspots

We now move on from overall properties of sunspots, and turn to their fine structure. Sunspots are now known to have considerable small-scale structure, in both the umbra and the penumbra (cf. Figures 1.4 and 1.5). In addition, the magnetic field has been found to have an extraordinary non-axisymmetric structure in the penumbra.

Further information about sunspots and their fine structure can be found in the recent review articles by Thomas and Weiss (2004) and Solanki (2003).

1.4.1 Umbral fine structure

The main feature of note in the umbra is the presence of so-called umbral dots, the first recorded observation of which was by Thiessen (1950). Umbral dots are small bright features visible against the dark background, with temperatures typically about 1000 K hotter than the coolest part of the umbra. The latest observations (Sobotka and Hanslmeier, 2005) appear to have spatially resolved most of the dots; their typical diameter is found to be about 100 km. Umbral dots are also often observed to travel radially towards the centre of the spot; this is particularly true of dots near the edge of the umbra. The dots are generally interpreted to be convective features, since they are warm compared to the umbral background and generally show upward motions.

Convection is not observed directly in the umbra, since there is a ‘radiative blanket’, a stably stratified region near the surface in which heat is transported purely by radiation; however, we know that convection must be taking place below this layer. The umbral dots presumably represent particularly vigorous rising convective plumes, which are able to penetrate through the radiative blanket and reach the surface. (The theory of umbral magnetoconvection is discussed further in section 1.5.1.)

The pattern of umbral dots also seems to be better explained by the monolithic, as opposed to the cluster, model (section 1.3.2). This is because a cluster of isolated flux tubes, with convection in between, would be expected to produce a network of bright lines rather than isolated bright points; on the other hand, magnetoconvection (in a monolithic flux tube) seems to be able to explain the umbral dots as a pattern of spatially modulated oscillations (see section 1.5.1).

Another feature of umbrae worth mentioning is that they sometimes contain lanes of bright material, crossing from one side to the other, which are known as light bridges. These mark out ‘fissures’ in the umbra and generally contain weaker, more horizontal magnetic fields than the rest of the umbra. In addition, there are also sometimes ‘dark nuclei’, dark regions containing few if any umbral dots; these probably correspond to regions of stronger magnetic field (cf. section 1.5.1).

1.4.2 Penumbral fine structure

The penumbra shows a very rich structure. The most prominent feature is the presence of alternating light and dark filaments which extend (approximately) radially outwards from the centre. Modern high resolution images show that the bright filaments are themselves composed of separate ‘grains’, which are typically about 350 km wide (or less), with lengths ranging from about 350 to 2500 km. The grains often show internal structure with a few dark bands crossing them (Roupe van der Voort et al., 2004). The grains are also observed to move radially, with the motion being in an inward direction within the inner 60% or so of the penumbra (by radius), and an outward direction in the outer 40%, with typical speeds of about 500 m s^{-1} . The inward-moving grains sometimes penetrate into the umbra, where they become umbral dots.

The magnetic field in the penumbra has a curious ‘interlocking-comb’ structure, in which the inclination of the field to the vertical varies significantly in the azimuthal direction (see Figure 1.8). This variation appears also to be correlated with the intensity variations between the bright and the dark penumbral filaments. The picture that we have is as follows: in the bright filaments, the inclination to the vertical varies from about 30° in the inner penumbra to about 60° at the outer edge; in the dark filaments, the field is inclined at about 65° to the vertical in the inner penumbra, becoming nearly horizontal at the outer edge of the spot. Indeed some of these field lines even reverse direction and plunge back down beneath the surface.

The two families of field lines also differ in their larger-scale connectivity. It appears that the more nearly horizontal fields (from the dark filaments) remain close to the surface, whereas the more vertical component of the field (from the bright filaments) extends high into the atmosphere, and can even extend many thousands of kilometres across the Sun (Sams et al., 1992; Winebarger et al., 2001).

As far as the vertical structure goes, we know that the more horizontally-oriented

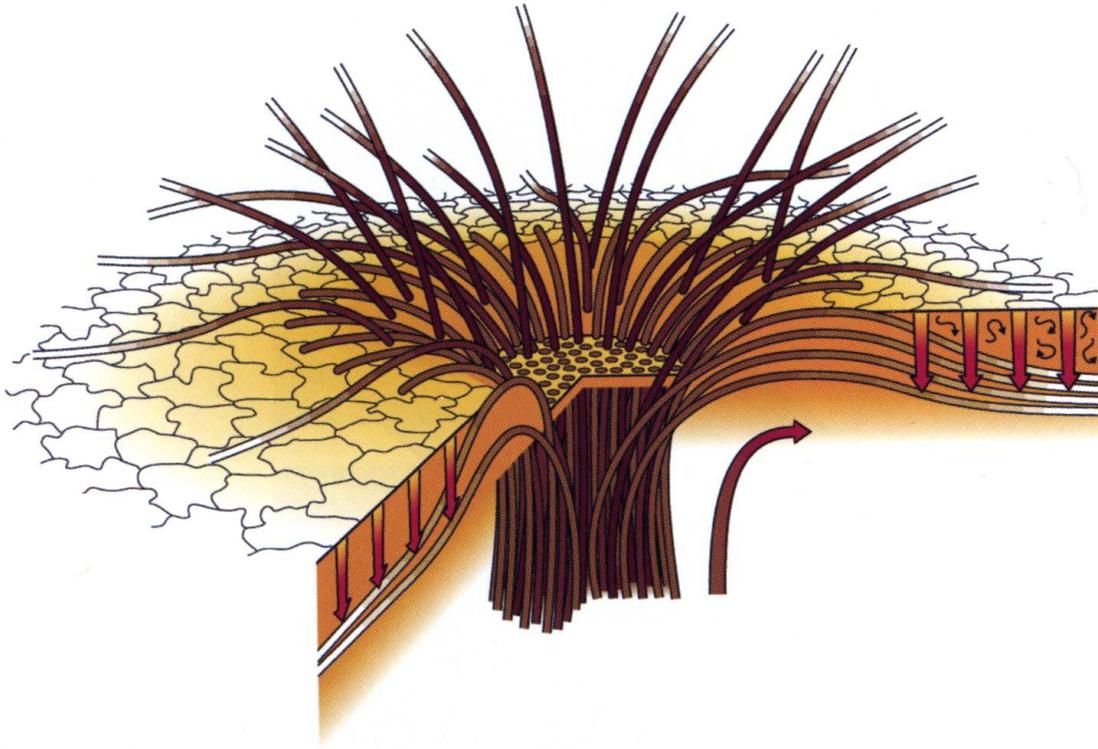


Figure 1.8: Sketch showing the interlocking-comb structure of the magnetic field of a sunspot penumbra (from Thomas and Weiss 2004). The brown tubes represent the magnetic field lines, of which there are two families, which coexist side by side: the field associated with the bright penumbral filaments rises up into the solar atmosphere, and the field associated with the dark penumbral filaments remains close to the solar surface or sometimes dives down beneath it. The large vertical arrows represent the effect of flux pumping (section 1.5.3) while the large curved arrow represents the sunspot's moat flow (section 1.6.3). Reprinted, with permission, from the *Annual Review of Astronomy and Astrophysics*, Volume 42 ©2004 by Annual Reviews, www.annualreviews.org.

field lines do not extend far above the surface. It is less clear how deep down below the surface these fields extend. They could be confined to isolated tubes, enclosed above and below by the more vertical field component, or they could be more like thin vertical slabs that are bounded above but extend downwards for a significant distance.

The observations of penumbral fine structure raise a number of important questions. First of all, how is the extraordinary interlocking-comb magnetic structure formed, and how is it maintained after its formation? Secondly, what is the origin of the observed fine structures – both the bright and dark filaments themselves, and also the smaller-scale details within them? Finally, why is the penumbra so different in appearance to the umbra?

It is likely that the observed penumbral features are ultimately of a convective origin, and therefore we should be able to answer these questions through studies of magnetoconvection. However, at present, the physical processes involved are far from well understood. (We will discuss some of the current theories below – see in particular sections 1.5.2 and 1.6.2.) In fact, the main aim of this thesis is to try to better understand some of these convective processes (and, in the process, to begin to answer some of the above questions), by starting from simple models and working our way up to more complicated (and more sunspot-like) situations. (See section 1.7 below for more details.)

A final noteworthy feature of penumbral structure is the so-called Evershed flow, discovered by John Evershed in 1909. This is a radial, near-horizontal outflow across the penumbra, and appears to be confined to the dark penumbral filaments, containing the near-horizontal fields. (This is consistent with our expectation from magnetohydrodynamics that the flow and magnetic field would tend to be aligned.) The flow is generally interpreted to be a ‘siphon flow’ along the ‘returning’ magnetic flux tubes, which dive back down below the solar surface near the outer boundary of the sunspot. What is happening here is that there is a different magnetic field strength at each end of the tube. At the inner footpoint, the field strength may be around 0.1 T (a typical penumbral value). The outer footpoints of these tubes typically correspond to magnetic features in the photosphere, with field strengths typically around 0.15 T. Therefore, there is a higher magnetic pressure, and hence lower gas pressure, at the outer edge of the tube, and this gas pressure difference drives the outflow along the tube.

1.5 Magnetoconvection

The models of overall sunspot structure show that energy is transported through the spot predominantly by convection, and the form taken by this magnetoconvection will influence the appearance and structure of the sunspot. Indeed, the differences in appearance between the umbra and the penumbra must ultimately be caused by changes in the pattern of magnetoconvection as the magnetic field becomes progressively more inclined to the vertical. In order to better understand these changes it is necessary to develop a theoretical understanding of magnetoconvection.

The study of magnetoconvection is an old subject; the linear theory for incompressible magnetoconvection was described by Chandrasekhar (1961), and a review of linear and weakly nonlinear work is given by Proctor and Weiss (1982). With the availability of modern computers, more recent work has turned to three-dimensional simulations of compressible convection, which can be used to investigate the nonlinear behaviour in a variety of regimes.

These simulations are of two kinds. The first kind attempts to include details of all relevant physics (e.g. radiative transfer and partial ionization effects) and to produce results which are directly comparable with observations. These simulations require large computing resources, and so the number of different cases that can be investigated is limited, but such simulations have been quite successful in modelling certain solar convective processes. For example, the MURAM code (Vögler et al., 2005) has been used recently to produce simulations of umbral dots which compare favourably with observations (Schüssler and Vögler, 2006).

A second approach is to simplify the physics and study more idealized models. This allows individual physical processes to be separated and studied in isolation. This kind of simulation also tends to require less computing power, so that many runs can be performed, and the different types of behaviour (occurring for different parameter values) can be catalogued and studied. A disadvantage is that quantitative comparisons with observations are not possible; one can only gain a qualitative understanding.

In general, one wants to run the more idealized type of simulation first, in order to gain a broad understanding of the physical processes involved, and then to follow up with more realistic simulations in order to make comparisons with observations. In the case of sunspot umbrae, realistic simulations are starting to be carried out (as mentioned above), but for the penumbra, we are still at the stage of trying to gain a qualitative

understanding of the physics. Therefore, in this thesis we will focus on the more idealized kind of simulations (leaving out radiative transfer and so forth).

We now summarize some of the calculations of magnetoconvection that have been carried out (of the more idealized kind), and explain their relevance to sunspots.

1.5.1 Simulations with vertical magnetic fields

The simplest form of magnetoconvection calculation considers a computational box containing compressible fluid with an imposed uniform vertical magnetic field. The upper and lower surfaces of the box are taken to be impermeable stress-free boundaries, with an imposed temperature difference between the two (in order to drive convection). In the horizontal directions, periodic boundary conditions are usually used (for reasons of computational convenience).

Examples of this type of calculation include Weiss et al. (1990); Matthews et al. (1995); Weiss et al. (1996); Tao et al. (1998); Rucklidge et al. (2000); Weiss et al. (2002). The results vary depending on the imposed magnetic field strength. For very strong fields, the Lorentz forces are strong enough to completely inhibit convection, and the fluid remains motionless. As the field strength is reduced, we come first into the *strong field* regime, in which convection takes the form of small-scale, steady, hexagonal convection cells. As the field strength is reduced further, the convection becomes weakly time-dependent, taking the form of spatially modulated oscillations, where adjacent plumes alternately wax and wane in amplitude. For weaker field strengths still, we enter a *flux separation* regime, where the magnetic field is separated from the convection. There are regions with large-scale, field-free convection, from which the magnetic flux has been expelled, and there are regions with strong fields and weak, small-scale convection. Finally there is a *weak field* regime where the magnetic flux is confined to intermittent regions of intense fields, with the rest of the domain being essentially field-free.

Calculations with a vertical magnetic field are applicable to the umbra of a sunspot (particularly in the central region). The umbra appears to be in the regime where the time-dependent spatially modulated oscillations are found; the umbral dots are then interpreted as regions where a particularly vigorous oscillation has protruded through the radiative blanket. In addition, the dark nuclei observed in some umbrae may be examples of flux separation (Weiss et al., 2002).

1.5.2 Magnetoconvection in inclined magnetic fields

Less attention has been given to the case of an inclined magnetic field (which would be appropriate for the penumbra, or the outer regions of the umbra). Tilting the field has two main effects. Firstly, the problem is no longer rotationally symmetric about a vertical axis, and this can lead to travelling wave phenomena. Secondly, the actual pattern of convection changes when the field is tilted. In a vertical field, hexagons are found near the onset of convection (as mentioned above); as the field becomes progressively more inclined, the hexagons become elongated along the direction of tilt, and eventually give way to field-aligned rolls. This latter solution is the preferred one in a horizontal magnetic field (Danielson, 1961).

Matthews et al. (1992) have looked at the linear theory for compressible convection in tilted fields, as well as some simple nonlinear models, and they were the first to point out that a tilted field would lead to travelling wave solutions. Hurlburt et al. (1996) have carried out two-dimensional simulations with inclined fields, showing the nonlinear development of these travelling rolls. Hurlburt et al. (2000) report some three-dimensional simulation results, which appear to show a transition between cellular and more roll-like patterns as the tilt of the field is increased; however, they describe only a limited number of results, in small computational boxes, so it is perhaps difficult to interpret what is happening.

In addition, Julien et al. (2000, 2003) have performed an asymptotic calculation of this problem in the limit of very strong magnetic fields. They find a transition to a new form of convection for nearly-horizontal fields, which may explain the differences between the bright and dark filaments in sunspots. However, their work makes certain assumptions and approximations and it is difficult to know if the result will still hold in a more general calculation.

We also point out a calculation by Busse and Clever (1990), which studies the linear instabilities of finite-amplitude convection rolls within tilted magnetic fields. Although the paper is more focused on the laboratory rather than the astrophysical case (for example, low magnetic Reynolds numbers are assumed throughout), some of the results may still be applicable to astrophysical situations.

1.5.3 Flux pumping

A recent calculation (Weiss et al., 2004; Thomas et al., 2002b) has demonstrated another phenomenon that can occur in compressible magnetoconvection: flux pumping. In this calculation, a layer of compressible fluid, with no magnetic field, is simulated until a steady convecting state is reached. A thin layer of magnetic field is then inserted into the centre of the box. This layer does not stay where it is, but is moved about by the turbulent convection; the calculation shows that the magnetic field lines are transported preferentially downwards. This process, known as flux pumping, is essentially caused by the strong asymmetry in compressible convection: there are weak broad upflows but very strong downflows.

The pumping is able to hold down the field lines even against their natural tendency to rise via magnetic buoyancy. The calculations show that flux pumping is a robust process, and it provides a possible explanation for how the returning flux tubes, observed in the penumbra (see above), can remain submerged below the solar surface.

1.5.4 Simulations in cylindrical geometry

An alternative approach is to run simulations in cylindrical geometry. For modelling the small-scale features, the geometry is probably not all that important, but to look at the larger-scale structure of pores and sunspots, the cylindrical geometry may need to be included.

The problem is simplified if one assumes axisymmetry, as was done by Hurlburt and Rucklidge (2000) and Botha et al. (2006) (for example), so that only a two-dimensional computation is required. They used a cylinder containing an initially vertical magnetic field. Their results show that this field is quickly swept into a concentrated flux tube at the centre of the cylinder – a form of flux separation. Convection is suppressed inside the flux tube, while convective cells form towards the edge of the domain. The authors suggest that their solution is a good representation of a pore, but the solution cannot really be applied to sunspots, since their penumbrae are highly non-axisymmetric.

An interesting feature of the simulations is that they tend to produce inflows along the surface just outside the flux tube. These flows could represent the ‘collar’ flow that would be needed to stabilize the sunspot against the magnetic fluting instability. Because the flow observed at the surface outside of sunspots is always *outwards* (the so-called moat flow), the authors suggest that in sunspots, these collar flows would be

hidden from view underneath the inclined outer edge of the flux tube.

The next step would be to extend these calculations to non-axisymmetric cylindrical geometry. Preliminary non-axisymmetric calculations have been presented by Hurlburt et al. (2000), which show that the axisymmetric ‘pore’ results can be unstable to non-axisymmetric perturbations. This may be related to the formation of sunspots from pores (which is discussed below).

1.6 Formation and decay of sunspots

So far we have considered sunspots as essentially static structures. We now move on to describe the processes leading to their formation and eventual disappearance.

1.6.1 Emergence of magnetic fields at the solar surface

Sunspots are magnetic features, and so before we can explain the processes that lead to the formation of a sunspot, we must first of all explain how magnetic fields reach the solar surface at all.

The study of how astrophysical bodies generate magnetic fields is known as *dynamo theory*. In the case of the Sun, magnetic fields are believed to be generated near the base of the convection zone, where a combination of differential rotation and turbulent convection allows large toroidal flux tubes to be generated.

These flux tubes may subsequently be subjected to instabilities driven by magnetic buoyancy, which will cause part of the tube to start rising up through the convective zone. The tube then becomes shredded and frayed by the turbulent convection, before arriving at the surface. This explains the observation of *bipolar active regions*, which are areas where small patches of magnetic field, of both polarities, start to emerge at the solar surface, as illustrated in Figure 1.9. A large active region will typically be composed of several of these rising flux tubes.

The emergence of the magnetic field is accompanied by the appearance of small pores, located at the points where the flux tubes (or flux tube fragments) intersect the solar surface. These pores then start to move towards each other and coalesce, forming larger and larger pores. This coalescence is driven by magnetic buoyancy: as the individual flux tube fragments continue to rise, they draw closer together, like the strings of a bunch of balloons held in one hand.

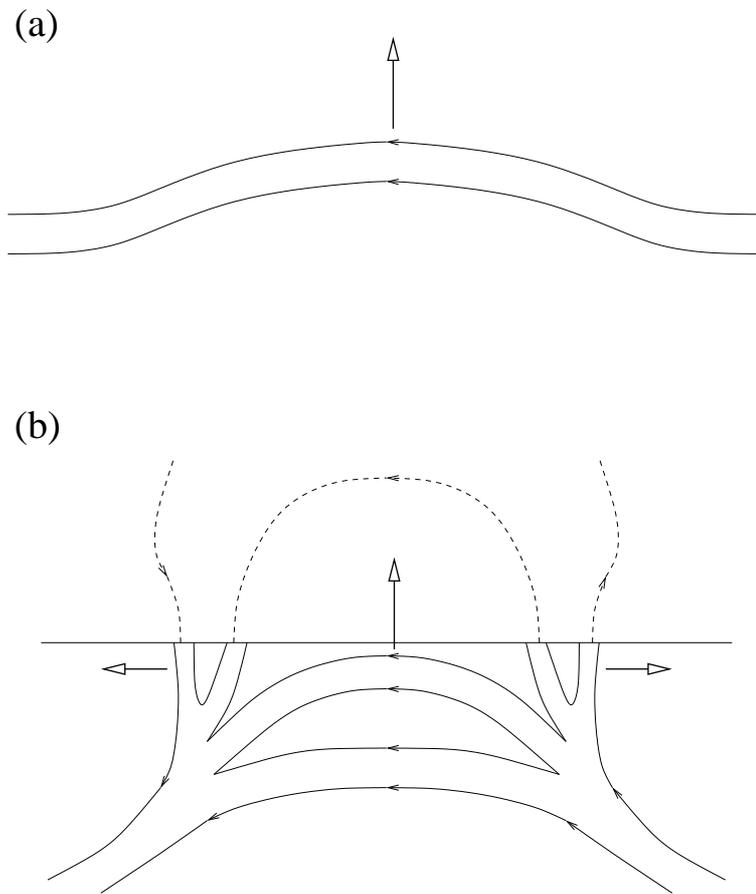


Figure 1.9: Sketches of a rising flux tube (after Figure 4 of Zwaan 1992). (a) A magnetic flux tube near the base of the convection zone. Part of the tube has started to rise towards the surface. The small arrows indicate the magnetic field direction, while the larger open arrow indicates the motion of the flux tube. (b) The flux tube has reached the surface. On the way up, it has fragmented into several parts. Pores or sunspots will be observed at the points where the fragments intersect the solar surface (the horizontal line). The dashed lines show the magnetic fields that would be seen above the surface. Note how this model naturally produces bipolar active regions containing adjacent groups of sunspots with opposite magnetic polarities.

1.6.2 Formation of the penumbra

Once the radius of a pore exceeds a certain critical value (approximately 2000 km, according to Rucklidge et al. 1995), a penumbra appears, and the pore develops into a sunspot. The formation of the penumbra is a very rapid event. The penumbra forms sector by sector, with each sector being completed in less than twenty minutes (Leka and Skumanich, 1998), and the whole process is over in under half a day (Zwaan, 1992).

Although the theoretical processes underlying penumbral formation remain largely unknown at present, the observed formation of penumbral sectors, within a short dynamical timescale of 20 minutes or so, does suggest that the transition from a pore to a sunspot is triggered by an instability of some sort. Models of pore structure indicate that the corresponding flux tubes must fan out with height, and as the magnetic flux contained in the pore increases, the radius of the flux tube correspondingly increases, as does the angle of tilt of the magnetic field at the edge of the pore. It is conjectured (Rucklidge et al., 1995) that once this angle exceeds a critical value, an instability sets in which leads to the formation of a penumbra.

One possible candidate is the fluting instability, which we have already discussed earlier in this chapter. The calculation of Meyer et al. (1977) showed that pores and sunspots would be stabilized (at least near the surface) by magnetic buoyancy effects. However, their calculation did not take account of convection, which in reality occurs both inside and outside the pore; therefore, the possibility remains that pores may be susceptible to a modified form of the fluting instability, driven by convection. The conjecture is that such an instability does indeed exist – but only if the field at the outer edge of the pore is sufficiently inclined to the vertical. This condition would be met only once the pore exceeded a certain size, and the resulting instability would drive the transition from a pore to a sunspot.

There are one or two illustrative calculations that demonstrate the possibility of such an instability. In Cartesian geometry, Tildesley (2003b) has found a convectively driven filamentary instability of a highly idealized configuration representing a pore. The nonlinear development of this instability was investigated by Tildesley and Weiss (2004) (see also Tildesley 2003a); it is found that a pattern of bright and dark filaments develops, although these are much wider than the filaments observed in real sunspots. In addition, Hurlburt et al. (2000) have performed a calculation in cylindrical geometry showing that an axisymmetric pore-like configuration is unstable to non-axisymmetric

perturbations, again leading to filamentary structure (although again the filaments are rather wide compared with observations).

These calculations suggest that pores could develop convectively driven instabilities at their outer edges. The nonlinear development of this instability has not yet been calculated within a realistic pore model, but it is conjectured (Weiss et al., 2004) that the instability would lead to fluting at the outer boundary, with some families of field lines being displaced upwards, and some downwards. (Hence the term ‘convectively driven fluting instability’.) Clearly, more realistic simulations, starting from more realistic pore configurations, are needed to test this hypothesis.

It is further conjectured that once this mildly fluted structure develops, the depressed field lines are then ‘grabbed’ by the turbulent convection outside the spot, and pulled down even further, via the flux pumping process described above. This would help the initial filamentary structure to develop into a full-blown penumbra, with its interlocking-comb magnetic field, and would explain how the ‘returning’ flux in the penumbra can form. It would also provide an explanation of why there is hysteresis between pores and sunspots, since once the field lines in the dark filaments are submerged, it would become more difficult to return them to the surface again.

1.6.3 Decay of sunspots

Sunspots have finite lifetimes, varying from hours to months depending on the size of the spot (the larger spots tend to live longer). The decay of a sunspot begins as soon as, or even before, the sunspot is fully formed. There are two main processes by which sunspots decay: gradual decay, and fragmentation.

The gradual decay is marked by a continual reduction in the size of the sunspot and the associated magnetic flux. Observationally, we see that a well-developed sunspot is always surrounded by a so-called *moat flow*, which is essentially a supergranular convection cell centred on the sunspot’s position. This leads to a radial outflow around the periphery of the sunspot. Within this outflow are embedded so-called *moving magnetic features*, small flux elements that move radially outwards from the spot. These appear to represent portions of the sunspot’s magnetic flux that are slowly chipped away by the convection currents in the moat, leading to the gradual decay of the spot.

Sometimes sunspots can also decay by fragmentation, whereby the spot suddenly splits into several smaller components, the components then disappearing shortly after-

DAILY SUNSPOT AREA AVERAGED OVER INDIVIDUAL SOLAR ROTATIONS

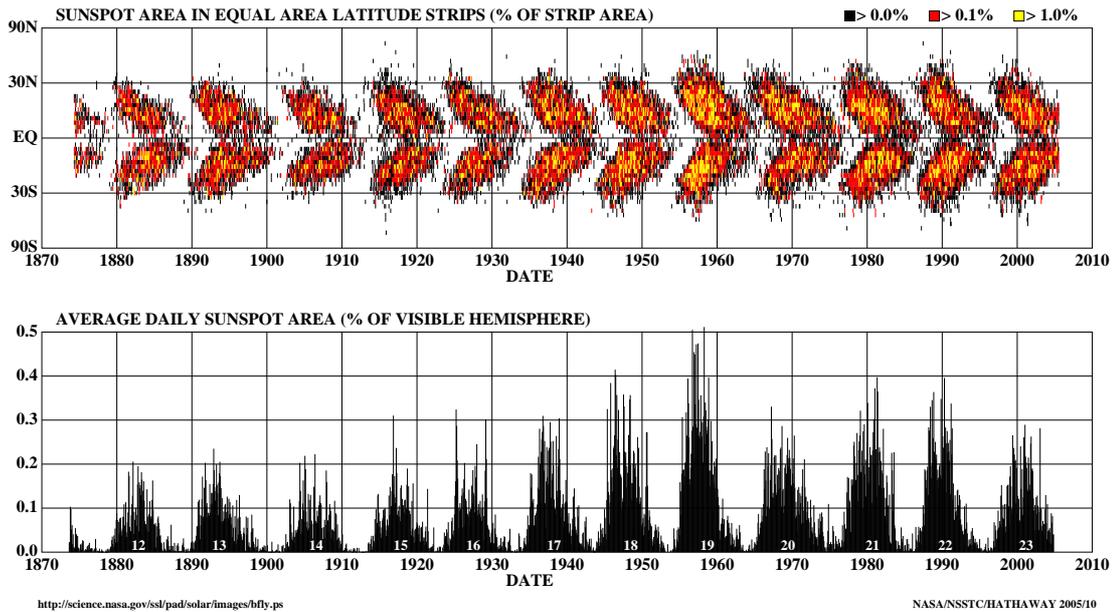


Figure 1.10: Butterfly diagrams showing (a) sunspot latitudes and (b) total area of the Sun covered by sunspots as a function of time. Courtesy NASA / David Hathaway.

wards. Often, the lines along which a sunspot splits are the same as the outlines of the pores that came together to form the sunspot in the first place. These lines are also often associated with light bridges.

1.6.4 The 11-year solar cycle

No description of sunspots would be complete without a mention of the 11-year solar cycle. This refers to a cyclic variation both in the number of sunspots observed and in the latitudes at which they appear. This is illustrated in Figure 1.10 which shows the area of the solar surface covered by sunspots at various latitudes for the last century or so. This is usually known as the ‘butterfly diagram’ (for obvious reasons). Note that the cycle is not completely periodic; the amplitude is modulated from cycle to cycle. There are also periods known as ‘grand minima’ where sunspot activity becomes weak or non-existent; the most recent of these was the Maunder Minimum of c.1645–1715.

Because sunspots are a surface manifestation of magnetic fields that are actually produced much deeper in the Sun, the sunspot cycle is an indication that the dynamo processes that create these fields operate in a cyclic manner. The solar dynamo is not yet fully understood, but it is widely believed that the Sun produces large amounts

of toroidal magnetic flux near the base of the convective zone, where helioseismology indicates that a strong shear layer exists. (A portion of this flux then rises to the surface to create sunspots.) This toroidal magnetic flux is thought to migrate towards the equator as a ‘dynamo wave’, resulting in the characteristic shape of the butterfly diagram.

1.7 Outline of the thesis

So far, we have given a broad outline of the subject of sunspots, as well as trying to point out some of the unanswered questions in the field. The area that remains most puzzling is the penumbra, with its complex filamentary structure. We currently do not understand the detailed nature of convection within the penumbra, nor do we know how this convection gives rise to the pattern of bright and dark filaments observed at the surface, together with the finer-scale structures found within these filaments. Related to this is the question of why the convection within the penumbra takes on such a radically different form from that within the umbra (as indicated by the difference in appearance of the two regions and the sharp transition between them). In addition, we cannot yet explain how the intricate penumbral structure comes about to begin with, nor how it is maintained in the presence of magnetic buoyancy and other effects.

Clearly, in order to answer these questions, we will need to gain a better understanding of magnetoconvection as it applies to sunspot-like magnetic field configurations. We begin by looking at a simplified problem, that of a uniform magnetic field, inclined at a fixed angle to the vertical. This should shed some light on the form taken by convection in the different parts of a sunspot, where the angle of the field to the vertical varies from nearly zero in the umbra, to $30\text{--}60^\circ$ in the bright penumbral filaments, to $60\text{--}90^\circ$ in the dark penumbral filaments. In Chapter 2 we will investigate the linear stability theory for a simplified model, which will clarify some of the symmetry effects and other aspects of the problem. In Chapter 3 we develop these ideas into weakly nonlinear models, and in the first part of Chapter 5, fully nonlinear numerical simulation results (using the full compressible MHD equations) are presented.⁴

The second part of our work will be to look beyond uniform fields and to produce models in which different parts of the sunspot (e.g. umbra and penumbra) are present within the same model. We do this by setting up a simplified ‘sunspot-like’ magnetic

⁴A condensed version of Chapters 2 and 3 is given in Thompson (2005).

field at the beginning of the calculation, and then investigating the forms of convection that arise within that field structure.⁵ The ultimate aim would be to produce a model showing umbral and penumbral convection, intricate filamentary structure, and all of the other details that are observed on the Sun.

Sadly, we have not been able to achieve this aim. However, our models do at least show a clear difference between the umbral and penumbral forms of convection, as well as a noticeable sharp transition between these two patterns. Thus, they do begin to answer some of the questions posed above; they should also provide a useful starting point for any future research into these problems. (The main feature missing from our results is the interlocking-comb magnetic structure in the penumbra; it seems that more detailed modelling techniques will be needed before this can be reproduced.) The models themselves will be described in Chapter 4 (where a simplified approach based on the Swift-Hohenberg equation is followed) and the second half of Chapter 5 (where the full compressible MHD equations are again used).

⁵This pre-supposes the existence of such a field structure. In other words, we are not attempting to address the question of how a sunspot's field structure is created; we are instead merely trying to understand the convective patterns that arise once such a structure has been formed.

Chapter 2

Linear theory

2.1 Introduction

In this chapter, we will begin our investigations of magnetoconvection in tilted fields by studying the linear stability theory for a simple model problem. Our philosophy will be to create a simplified, idealized model rather than trying to include all the details of the physics. This will allow us to learn something about the problem without over-complicating the calculations or having to do time-consuming numerical simulations. Once we understand the basics, we will then be in a position, in later chapters, to move on to more complicated (and more realistic) models.

With this in mind, we will look at Boussinesq convection in a layer containing a uniform inclined magnetic field. Unlike previous work (Matthews et al., 1992) we do not consider fully compressible convection, although it will turn out that many of the features of the compressible problem are captured in our model. We do, however, consider fully three-dimensional solutions; the work of Matthews et al. (1992) was restricted to two dimensions. This work will give us a rough idea what happens near the onset of convection, and will provide a good starting point for the weakly nonlinear models of the next chapter.

We will first of all, in section 2.2, describe in detail our problem and our method of solving it. In section 2.3 we review what happens when the field is vertical, and in section 2.4 we show how symmetry arguments may be used to ascertain what happens when the field is tilted. Finally, we present our results in section 2.5, and give conclusions in section 2.6.

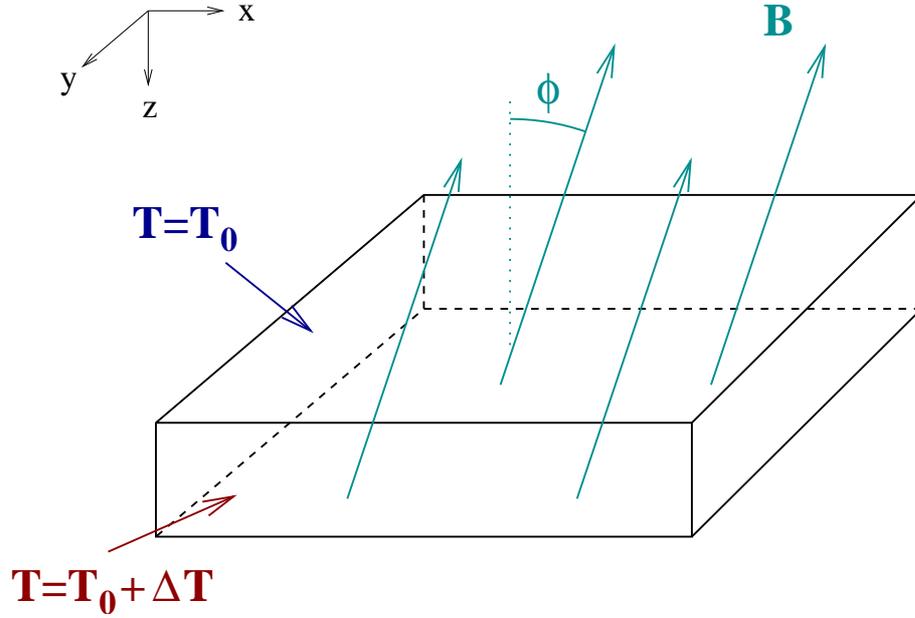


Figure 2.1: Diagram showing the system under investigation.

2.2 Problem setup

The problem that we will investigate consists of an infinite horizontal plane layer of Boussinesq fluid, heated from below, and permeated by an inclined magnetic field tilted at an angle ϕ to the vertical. Refer to Figure 2.1.

The equations of motion for the layer, in non-dimensionalized form, are as follows: the Navier–Stokes equation

$$\frac{1}{\sigma} \left(\frac{\partial \mathbf{u}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{u} \right) = -\nabla \Pi + \frac{d^3}{\nu \bar{\kappa}} \mathbf{g} - R T \mathbf{e}_z + \zeta Q \mathbf{B} \cdot \nabla \mathbf{B} + \nabla^2 \mathbf{u}, \quad (2.1)$$

the temperature equation (where we assume that radiative transfer can be treated in the diffusive approximation)

$$\frac{\partial T}{\partial t} + \mathbf{u} \cdot \nabla T = \nabla \cdot (\hat{\kappa} \nabla T), \quad (2.2)$$

and the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} + \mathbf{u} \cdot \nabla \mathbf{B} = \mathbf{B} \cdot \nabla \mathbf{u} + \zeta \nabla^2 \mathbf{B}, \quad (2.3)$$

together with

$$\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{B} = 0, \quad (2.4)$$

where \mathbf{u} is the fluid velocity, \mathbf{B} is the magnetic field, T is the temperature, and $\Pi = p + \mathbf{B}^2/2\mu_0$ is the total pressure (gas pressure plus magnetic pressure). In these equations

we have scaled all lengths by the layer depth d , times by the thermal diffusion timescale $d^2/\bar{\kappa}$, magnetic fields by the imposed field strength B_0 , and temperatures by the imposed temperature difference ΔT . This leaves the following dimensionless parameters: the Rayleigh number R , the Chandrasekhar number Q , the Prandtl number σ and the ratio ζ of magnetic to thermal diffusivity, given by

$$R = \frac{\alpha_T \Delta T d^3 |\mathbf{g}|}{\nu \bar{\kappa}}, \quad Q = \frac{B_0^2 d^2}{\mu_0 \rho \nu \eta}, \quad \sigma = \frac{\nu}{\bar{\kappa}}, \quad \zeta = \frac{\eta}{\bar{\kappa}}. \quad (2.5)$$

Here α_T is the coefficient of thermal expansion, \mathbf{g} is the acceleration due to gravity, ν is the fluid viscosity, η the magnetic diffusivity, μ_0 the permeability, and ρ the density. The thermal diffusivity of the fluid is denoted by κ ; the quantity $\bar{\kappa}$ in the above equations represents the value of κ at mid-layer ($z = 1/2$). Initially we will assume that κ is a constant (so that $\kappa = \bar{\kappa}$ everywhere), but later on we will allow for the possibility of a varying thermal diffusivity with depth. In this case we write κ as

$$\kappa = \hat{\kappa}(z) \bar{\kappa}, \quad (2.6)$$

where $\hat{\kappa}$ is a dimensionless function of depth. (For consistency, $\hat{\kappa}$ must satisfy $\hat{\kappa}(1/2) = 1$.)

2.2.1 Boundary conditions

We must also specify boundary conditions for our layer. We will assume that our layer is of infinite extent in the horizontal, so we only need to specify boundary conditions for the top and bottom surfaces. We take these surfaces to be impenetrable and stress free, so that

$$u_z = \frac{\partial u_x}{\partial z} = \frac{\partial u_y}{\partial z} = 0 \quad (2.7)$$

at $z = 0$ and at $z = 1$. We also take the temperature to be fixed at top and bottom, which in the non-dimensional units adopted, reduces to

$$T = 0 \text{ at } z = 0; \quad T = 1 \text{ at } z = 1. \quad (2.8)$$

This leaves the question of what to do with the magnetic boundary conditions. There are actually several plausible choices. The simplest would be to fix the angle of the magnetic field at the top and bottom surfaces, so that

$$B_x = -B_z \tan \phi; \quad B_y = 0 \quad (2.9)$$

on the boundary. However this is not physically realistic and can in some circumstances lead to energy being injected into the system from the boundaries (N.Roxburgh, private communication).

Instead, there are two physically motivated boundary conditions that we could use: we could take the boundary to be either a perfect electrical insulator or a perfect electrical conductor. The former is equivalent to matching to a potential field outside the layer, while the latter is equivalent to ‘tying’ the field lines (and in 2D it is also equivalent to holding the vector potential fixed at its initial value). We derive equations for these conditions in Appendix A. Although still somewhat idealized, these conditions are slightly more realistic than simply fixing the tilt angle.

For example, in the Sun we can to a first approximation assume that there is a vacuum (or at least, a current-free region) above the surface, so it is reasonable to match our solution onto a potential field above the top of the layer. We take the potential field to be inclined at an angle ϕ as $z \rightarrow -\infty$ (which is how the inclination of the field is controlled).

Below the layer, matching to a potential field is inappropriate. Instead, we take a perfectly conducting lower boundary, which ‘ties’ the footpoints of the field lines in place. (This seems reasonable if we think of the field in a sunspot as being anchored to a flux tube which is deeply rooted near the base of the convection zone.)

2.2.2 Basic state and linearization

The equations given above have a trivial solution in which the fluid velocity is zero and the magnetic field is uniform: $\mathbf{B} = \mathbf{B}_0 \equiv (-\sin \phi, 0, \cos \phi)$ in non-dimensional units. The temperature profile is linear if κ is uniform; if κ varies with depth, it is instead given by

$$T = T_0(z) \equiv C \int \frac{1}{\hat{\kappa}} dz + D \quad (2.10)$$

where C and D are constants of integration, chosen to ensure that the boundary conditions ($T = 0$ at $z = 0$, $T = 1$ at $z = 1$) are satisfied.

We now wish to analyse the stability of this basic state. The problem is separable

in x , y and t , so we introduce perturbations of the form

$$\mathbf{u} = \mathbf{u}'(z)e^{i\mathbf{k}\cdot\mathbf{x}+st} + \text{c.c.} \quad (2.11)$$

$$\mathbf{B} = \mathbf{B}_0 + \mathbf{B}'(z)e^{i\mathbf{k}\cdot\mathbf{x}+st} + \text{c.c.} \quad (2.12)$$

$$T = T_0(z) + T'(z)e^{i\mathbf{k}\cdot\mathbf{x}+st} + \text{c.c.} \quad (2.13)$$

$$\Pi = \Pi_0 + \Pi'(z)e^{i\mathbf{k}\cdot\mathbf{x}+st} + \text{c.c.} \quad (2.14)$$

(where ‘c.c.’ stands for ‘complex conjugate’). The vector \mathbf{k} is the wavevector of the disturbance, which we also write in components as $\mathbf{k} = (k_x, k_y, 0)$.

Choosing the perturbations in this form essentially corresponds to considering one Fourier mode at a time. (Obviously in linear theory different Fourier modes will just superpose, so it is valid to consider them each individually.)

After linearizing in the perturbations, and writing the equations out in components, the following eigenvalue problem for the growth rate s is obtained:

$$\frac{s}{\sigma}u'_x = -ik_x\Pi' + \zeta Q \left(-ik_x \sin \phi B'_x + \cos \phi \frac{dB'_x}{dz} \right) - (k_x^2 + k_y^2)u'_x + \frac{d^2u'_x}{dz^2} \quad (2.15)$$

$$\frac{s}{\sigma}u'_y = -ik_y\Pi' + \zeta Q \left(-ik_x \sin \phi B'_y + \cos \phi \frac{dB'_y}{dz} \right) - (k_x^2 + k_y^2)u'_y + \frac{d^2u'_y}{dz^2} \quad (2.16)$$

$$\frac{s}{\sigma}u'_z = -\frac{d\Pi'}{dz} + \zeta Q \left(-ik_x \sin \phi B'_z + \cos \phi \frac{dB'_z}{dz} \right) - (k_x^2 + k_y^2)u'_z + \frac{d^2u'_z}{dz^2} - RT' \quad (2.17)$$

$$sT' + \frac{C}{\hat{\kappa}}u'_z = -\hat{\kappa}(k_x^2 + k_y^2)T' + \hat{\kappa} \frac{d^2T'}{dz^2} + \frac{d\hat{\kappa}}{dz} \frac{dT'}{dz} \quad (2.18)$$

$$sB'_x = -ik_x \sin \phi u'_x + \cos \phi \frac{du'_x}{dz} - \zeta(k_x^2 + k_y^2)B'_x + \zeta \frac{d^2B'_x}{dz^2} \quad (2.19)$$

$$sB'_y = -ik_x \sin \phi u'_y + \cos \phi \frac{du'_y}{dz} - \zeta(k_x^2 + k_y^2)B'_y + \zeta \frac{d^2B'_y}{dz^2} \quad (2.20)$$

$$sB'_z = -ik_x \sin \phi u'_z + \cos \phi \frac{du'_z}{dz} - \zeta(k_x^2 + k_y^2)B'_z + \zeta \frac{d^2B'_z}{dz^2} \quad (2.21)$$

$$ik_x u'_x + ik_y u'_y + \frac{du'_z}{dz} = 0 \quad (2.22)$$

$$ik_x B'_x + ik_y B'_y + \frac{dB'_z}{dz} = 0 \quad (2.23)$$

subject to

$$u'_z = \frac{du'_x}{dz} = \frac{du'_y}{dz} = T' = 0 \quad (2.24)$$

at both top and bottom.

If the fixed-angle boundary condition is being used for the magnetic field, then the relevant linearized boundary condition is

$$B'_x = -B'_z \tan \phi \quad (2.25)$$

$$B'_y = 0, \quad (2.26)$$

to be applied at $z = 0$ and $z = 1$.

The potential field boundary condition (at $z = 0$) is

$$B'_x = \frac{ik_x}{\sqrt{k_x^2 + k_y^2}} B'_z \quad (2.27)$$

$$B'_y = \frac{ik_y}{\sqrt{k_x^2 + k_y^2}} B'_z. \quad (2.28)$$

The (linearized) tied field boundary condition (at $z = 1$) is

$$\zeta \frac{dB'_x}{dz} + u'_x \cos \phi - \zeta ik_x B'_z = 0 \quad (2.29)$$

$$\zeta \frac{dB'_y}{dz} + u'_y \cos \phi - \zeta ik_y B'_z = 0 \quad (2.30)$$

These equations were solved numerically using the program NRK, an implementation of the Newton-Raphson-Kantorovich method for solving boundary value problems (Cash and Moore, 1980). The eigenvalue s can be found in terms of the various parameters. Alternatively the critical Rayleigh number R_c for the onset of instability (at which $\text{Re } s = 0$) can be calculated.

2.3 The vertical field case

Before moving on to inclined fields, it is important to understand the case of a vertical field. Here, the layer is isotropic (that is to say, there is rotational symmetry about the z -axis), so without loss of generality we can consider only the two-dimensional problem, taking $k_y = 0$. (We also take k_x to be positive, again without loss of generality.) The isotropy implies a symmetry between leftward- and rightward-travelling waves, which means that the eigenvalues s will always be either purely real (representing steady convection), or come in complex conjugate pairs (representing waves travelling in either direction).

In the Boussinesq case, this problem has been investigated by Chandrasekhar (1961) and Proctor and Weiss (1982). There are two qualitatively different types of behaviour,

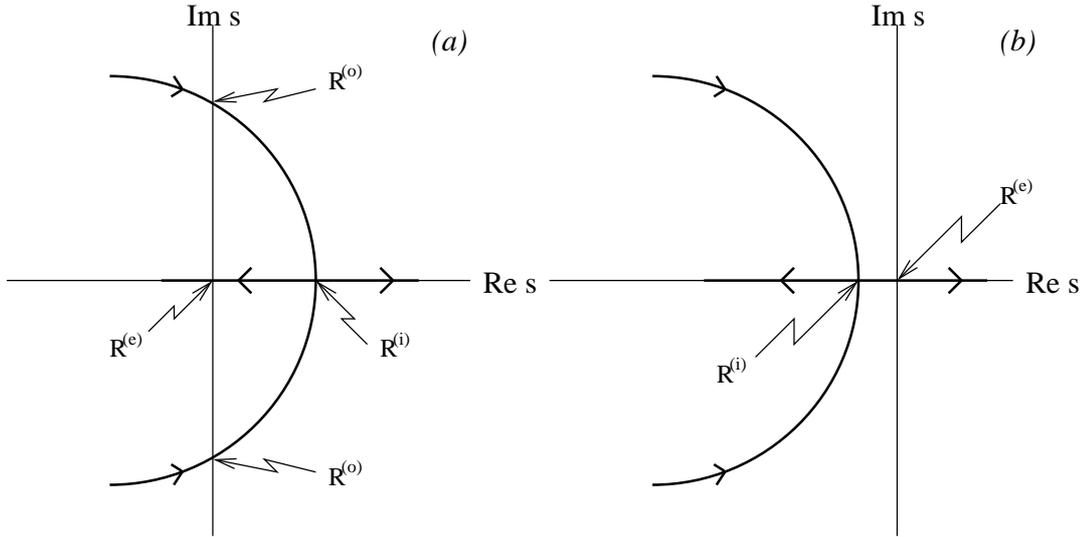


Figure 2.2: The paths of the eigenvalues through the complex plane as R is varied, when the field is vertical. (a) Oscillatory case (small ζ); (b) steady case (large ζ). (After Proctor and Weiss 1982, figure 6.)

depending on ζ . The first is the ‘oscillatory’ case which occurs for $\zeta < 1$ and Q sufficiently large. The second type of behaviour is the ‘steady’ case and this occurs when $\zeta \geq 1$ or when Q is small enough.

Note that these conditions apply for uniform κ ; when κ is non-uniform, there is a slight complication, since the ratio η/κ is now effectively a function of depth. The cut-off between the two types of behaviour will no longer be exactly at $\zeta = 1$; however, we can still say that the oscillatory case occurs for ζ small enough, and the steady case occurs for ζ large enough.

Figure 2.2 shows how the eigenvalues s move through the complex plane as R is increased, in both the oscillatory and steady cases. Note that $\text{Re } s$ corresponds to the growth rate (positive indicating instability), while $\text{Im } s$ gives the oscillation frequency of the mode. Since we are taking k_x to be positive, we have that $\text{Im } s < 0$ represents a rightward-travelling wave, $\text{Im } s > 0$ represents a leftward-travelling wave, and $\text{Im } s = 0$ indicates a steady mode.

Figure 2.2(a) shows the oscillatory case; various special Rayleigh numbers are marked. $R = R^{(o)}$ corresponds to a Hopf bifurcation, where the static state loses stability to oscillatory convection. $R = R^{(i)}$ represents a transition from oscillatory to steady convection. $R = R^{(e)}$ is where one of the eigenvalues becomes negative again, but this has no physical significance since we are still unstable to the other (positive) eigenvalue.

The steady case is illustrated in Figure 2.2(b). There is now a bifurcation to steady convection at $R = R^{(e)}$, and this time it is $R^{(i)}$ that has no physical significance. Note that the eigenvalues in the right half-plane never leave the real axis, so there is no instability to travelling waves this time.

2.4 Symmetry considerations for inclined fields

We can gain some insight into the inclined field case ($\phi > 0$) by considering the symmetries involved. One important symmetry has already been mentioned, and that is the rotational symmetry present when the field is vertical. When the field is inclined, this symmetry will be broken, and rolls of different orientations will have different properties. We can no longer take $k_y = 0$ and must instead consider the wavenumber to be a two-dimensional vector. Nevertheless, it will be helpful to first consider the ‘two-dimensional’ case (with $k_y = 0$), in order to explain the various ideas in a simpler setting, before moving on to fully three-dimensional solutions.

2.4.1 General considerations in the two-dimensional case ($k_y = 0$)

Here, we have seen that the isotropy (present for vertical fields) manifests itself as a symmetry between left- and right-going waves, meaning that the eigenvalues come in complex conjugate pairs. When the field is tilted, one would expect this left-right symmetry to be broken. However, for Boussinesq convection this is not the case; this is because there is actually an extra symmetry in the problem.

This extra symmetry is an *up-down* reflection symmetry, inherent in the Boussinesq equations. In our problem, we need to clarify what this means; we are not talking about a simple reflection $z \rightarrow -z$, since this would reverse the tilt of the field lines. Instead there is a symmetry π consisting of a reflection in the mid-plane ($z \rightarrow -z$) followed by another reflection in the y - z plane ($x \rightarrow -x$). (π is equivalent to a 180° rotation, in effect turning the layer ‘upside-down’.)

This symmetry transforms a left-going wave into a right-going wave, and vice versa. In other words, the presence of an up-down symmetry also implies that there is a symmetry between left-going and right-going waves, even when the field is tilted.

Contrast this to compressible convection, in which there is no up-down symmetry,

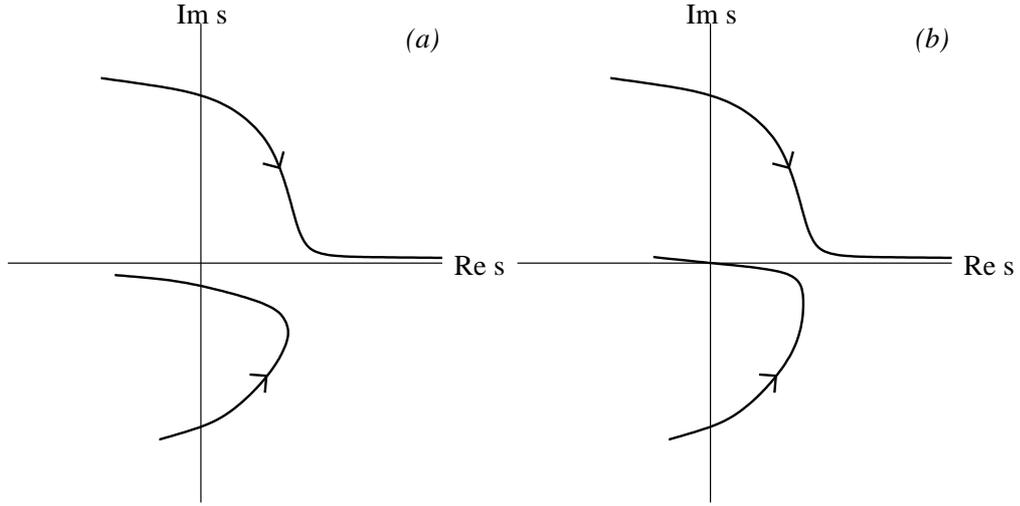


Figure 2.3: *Perturbations to Figure 2.2(a) after breaking the left-right symmetry. (a) The general case. (b) The case where the up-down symmetry is broken only by a change in the magnetic boundary conditions between top and bottom; in this case, the curve must pass through the origin.*

and therefore tilting the field *would* break the symmetry between left-going and right-going waves. Since we are trying to model sunspots, where compressibility is important, it is clearly unsatisfactory that the left-right symmetry is *not* broken in our model.

We are therefore motivated to look for ways in which this symmetry can be broken within our Boussinesq model. First of all, note that one way of achieving this would be to choose different boundary conditions at the top and bottom boundaries. In fact, we have already suggested that an appropriate choice of magnetic boundary conditions would be to use an insulating upper boundary, but a conducting lower boundary. This in itself should be sufficient to break the up-down symmetry.

We will see later on, however, that the approach of breaking the symmetry purely by changing the magnetic boundary conditions is not entirely satisfactory. An alternative method is to allow the thermal diffusivity κ to vary with depth – a possibility that we allowed for when we formulated the equations. This approach allows us to break the symmetry in a somewhat more fundamental way than simply by tweaking the boundary conditions.

We now turn to the changes that we would expect to see in Figure 2.2 when the field is tilted (assuming that the up-down symmetry in the problem has indeed been broken). The main change is that we would expect either left-going waves to be preferred over right-going ones, or vice versa (we cannot say in advance which way around it will be).

We now consider how this works in more detail, starting with the oscillatory case (low ζ) and then moving on to the steady case (high ζ).

2.4.2 The oscillatory case (small ζ)

Here, we wish to describe how Figure 2.2(a), which applies to oscillatory convection when $\phi = 0$, will change when ϕ is increased above zero. Here we are considering the case where the left-right symmetry has been broken (as discussed above); suppose that it is the left-going waves that are preferred. This would shift the top half of the diagram (representing left-going waves) to the right (i.e. becoming more unstable), while the bottom half would be shifted to the left. Thus, we might expect to see a picture something like Figure 2.3(a). (The other possibility, in which right-going waves are preferred, would correspond to a reflection of this figure about the real axis.)

For this result, however, it is important to consider *how* the symmetry π was broken. If this was done solely by changing the magnetic boundary conditions at top and bottom, then we must consider a result of Roberts (1967, p.200), who has shown that when $s = 0$, the magnetic boundary conditions actually have no effect on the problem. That is, if $s = 0$ at $R = R^{(e)}$ for the symmetric problem, then s will still be zero, at the same Rayleigh number, when the magnetic boundary conditions are changed. This is equivalent to saying that the locus of our eigenvalues must still pass through the origin, at the same value of R – although there is nothing to stop it from being perturbed from its original path elsewhere (where $s \neq 0$). Figure 2.3(b) shows the revised picture.

Note that Roberts (1967) only derives this result for the two-dimensional case, but it is a simple matter to extend it to the three-dimensional case (i.e. when $k_y \neq 0$). Following Proctor and Weiss (1982), section 5, we can write the linearized equations in a poloidal-toroidal form. First of all, take the curl of the momentum and induction equations to obtain:

$$\frac{1}{\sigma} \frac{\partial \boldsymbol{\omega}}{\partial t} = -R(\nabla T') \wedge \mathbf{e}_z + \zeta Q \mathbf{B}_0 \cdot \nabla \mathbf{j}' + \nabla^2 \boldsymbol{\omega} \quad (2.31)$$

$$\frac{\partial \mathbf{j}'}{\partial t} = \mathbf{B}_0 \cdot \nabla \boldsymbol{\omega} + \zeta \nabla^2 \mathbf{j}' \quad (2.32)$$

where $\boldsymbol{\omega} = \nabla \wedge \mathbf{u}$ and $\mathbf{j}' = \nabla \wedge \mathbf{B}'$.

The z -components of (2.31) and (2.32) decouple, giving a pair of equations for ω_z

and j'_z , as follows:

$$\frac{1}{\sigma} \frac{\partial \omega_z}{\partial t} = \zeta Q \mathbf{B}_0 \cdot \nabla j'_z + \nabla^2 \omega_z \quad (2.33)$$

$$\frac{\partial j'_z}{\partial t} = \mathbf{B}_0 \cdot \nabla \omega_z + \zeta \nabla^2 j'_z. \quad (2.34)$$

(2.33) and (2.34) are the equations for damped Alfvén waves travelling along the \mathbf{B}_0 direction. Since there is no driving term, the solutions will decay exponentially and j'_z and ω_z may be set to zero.

Because $\nabla \cdot \mathbf{u} = \nabla \cdot \mathbf{B}' = 0$, we can now introduce a poloidal-toroidal representation for both \mathbf{u} and \mathbf{B}' . In fact, since ω_z and j'_z are both zero, it follows that the toroidal component will be zero and that \mathbf{u} and \mathbf{B}' will both be purely poloidal. We write them as

$$\mathbf{u} = \nabla \wedge (\nabla \wedge (F \mathbf{e}_z)) \quad (2.35)$$

$$\mathbf{B}' = \nabla \wedge (\nabla \wedge (H \mathbf{e}_z)). \quad (2.36)$$

We can then rewrite the linearized momentum, temperature and induction equations in terms of F and H :

$$\frac{1}{\sigma} \frac{\partial}{\partial t} (\nabla^2 F) = RT' + \zeta Q (\mathbf{B}_0 \cdot \nabla) \nabla^2 H + \nabla^4 F \quad (2.37)$$

$$\frac{\partial T'}{\partial t} - (\nabla_H^2 F) \left(\frac{\partial T'}{\partial z} \right) = \frac{d\hat{\kappa}}{dz} \frac{dT_0}{dz} + \hat{\kappa} \nabla^2 T' \quad (2.38)$$

$$\frac{\partial H}{\partial t} = \mathbf{B}_0 \cdot \nabla F + \zeta \nabla^2 H. \quad (2.39)$$

Note that (2.37) couples to (2.39) only through $\nabla^2 H$, while (2.38) does not contain H at all. If we look for steady solutions (in which $\partial/\partial t \equiv 0$), we have from (2.39) that

$$\nabla^2 H = -\frac{1}{\zeta} \mathbf{B}_0 \cdot \nabla F \quad (2.40)$$

and we can substitute this into (2.37). Note that (2.37) and (2.38) now decouple completely from (2.39). We can solve the former for the temperature and velocity eigenfunctions, *independently* of the magnetic boundary conditions, thus obtaining our critical Rayleigh number R (when $s = 0$); we can then substitute these solutions into (2.39) and find the magnetic field separately. Therefore, we have proved the above claim that when $s = 0$, the magnetic boundary conditions do not influence the value of the critical Rayleigh number.

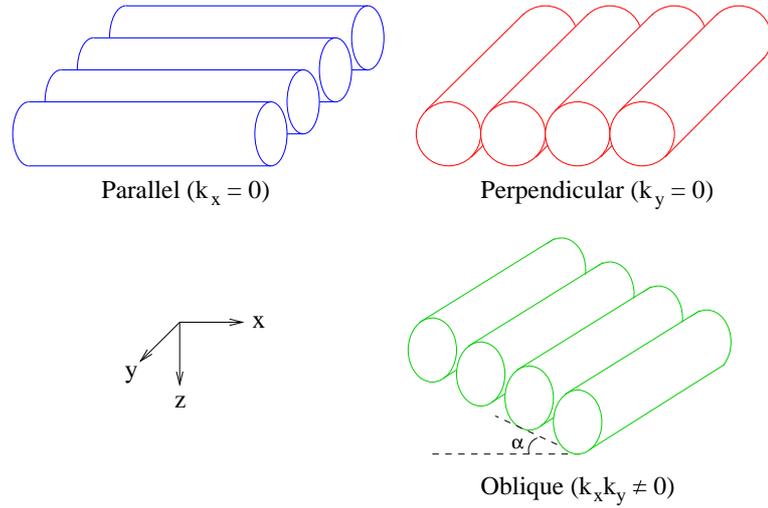


Figure 2.4: Definitions of ‘parallel’, ‘perpendicular’ and ‘oblique’ rolls. The cylinders in this picture represent the convection rolls. The magnetic field (which is not drawn) lies in the x - z plane.

The three-dimensional case ($k_y \neq 0$)

We now move on to the general case where k_y can be non-zero. To make things clearer we will introduce an angle α describing the orientation of the wavevector \mathbf{k} with respect to the x -axis, defined as follows:

$$k_x = k \cos \alpha, \quad k_y = k \sin \alpha. \quad (2.41)$$

The case $\alpha = 0$ (corresponding to the two-dimensional problem) will be called the ‘perpendicular’ case (since the roll axes are perpendicular to the plane of the tilt), and $\alpha = 90^\circ$ will be referred to as the ‘parallel’ case. There are also ‘oblique’ rolls (with $0^\circ < \alpha < 90^\circ$). These possibilities are illustrated in Figure 2.4.

Rolls of each of these types will have different properties, because of the anisotropy (if $\phi > 0$). As was shown by Matthews et al. (1992), when ϕ is small these modes can be analysed using simple symmetry arguments. We imagine that, for given parameters, we have calculated the critical Rayleigh numbers (minimized over wavenumber) for a vertical field, and we then consider what happens as ϕ is perturbed slightly from zero.

In the case of parallel rolls, the resulting perturbation to R_c is constrained by symmetry. There are two travelling parallel roll modes – one travelling in the positive y direction, the other in the negative y direction – but both of these are equivalent, in the sense that they are related by the $y \rightarrow -y$ reflection symmetry, and therefore they both have the same critical Rayleigh number. Moreover, a parallel roll mode is unchanged by

a reflection $x \rightarrow -x$ (which changes the sign of ϕ). This implies that ΔR_c for parallel rolls is an *even* function of ϕ ; and for small ϕ , it will (generically) be proportional to ϕ^2 at leading order.

Now consider the perpendicular rolls, which are essentially the same as the two-dimensional solutions discussed above. Assuming that we have found a way to break the up-down symmetry in our layer, then, as explained above, there is asymmetry between left- and right-going waves. Therefore, the perturbation to R_c will *not* be an even function of ϕ ; it will instead be proportional to ϕ at leading order.

Therefore, we expect the two-dimensional perpendicular mode to be preferred for small ϕ , but for larger ϕ the parallel rolls ‘overtake’ and become the preferred mode.

Turning to the oblique rolls, we can examine these by using the following simple model equation (from Matthews et al., 1992):

$$\Delta R_c = \pm\phi B \cos \alpha - \phi^2 C \sin^2 \alpha, \quad (2.42)$$

where ΔR_c represents the perturbation to the critical Rayleigh number (for small ϕ). This is effectively a way of combining the effects on the parallel and perpendicular rolls into a single equation. We choose $B > 0$ (without loss of generality) and $C > 0$ (because when $\alpha = 90^\circ$, we expect R_c to fall as the tilt increases¹), and minimize ΔR_c over α to find the preferred mode. We find that $\alpha = 0$ is preferred for $\phi < B/2C \equiv \phi_c$, and $\alpha = \cos^{-1}(\phi_c/\phi)$ for $\phi > \phi_c$.

In other words, perpendicular rolls are preferred for small ϕ , and oblique rolls are preferred for larger ϕ . The transition between the two occurs at $\phi = \phi_c = B/2C$, and we will refer to this point as the ‘Lifshitz point’ after a similar phenomenon in nematic liquid crystals (see e.g. Silber et al., 1992).

For this model to be useful we require ϕ_c to be small; if not, then the assumption of small ϕ would not hold, and the model would break down. However, in our actual numerical results we find that $B/2C$ is typically less than 0.1, so this is not really a problem.

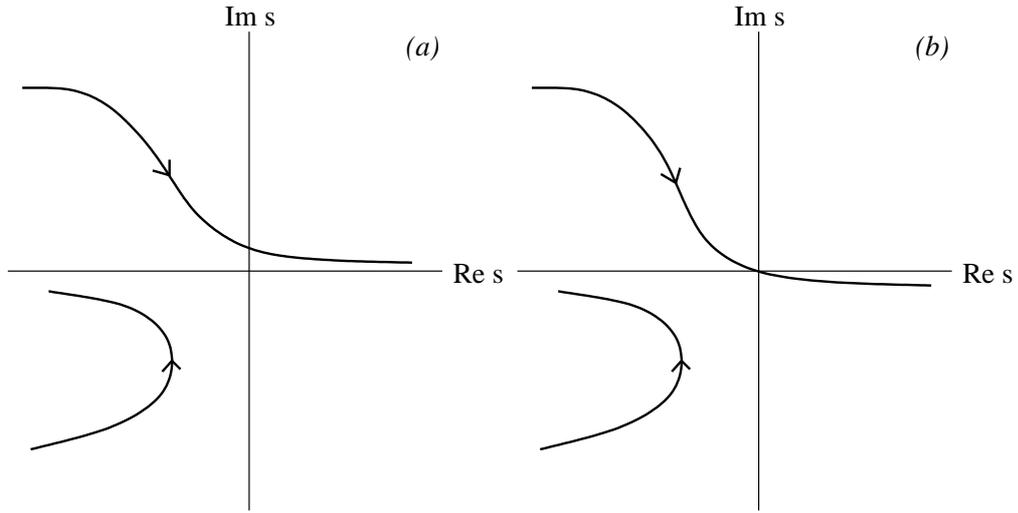


Figure 2.5: The equivalent of Figure 2.3 for the steady case.

2.4.3 The steady case

The above results have concentrated on the oscillatory case; the steady case (large ζ) is a little different. Note that the name ‘steady’ is slightly misleading, since the solutions are only truly steady if the field is vertical; this is illustrated in Figure 2.2(b), where the eigenvalues pass through the origin at $R^{(e)}$, indicating a steady-state bifurcation. As we shall see shortly, when the field is non-vertical the solutions do start to travel; the speed of travel is proportional to ϕ for small ϕ .

The perturbations to Figure 2.2, that would be expected in the steady case, are shown in Figure 2.5. This shows the path that the eigenvalues would take as R is increased, for a small non-zero ϕ . This is obtained by starting from Figure 2.2(b) and imagining a small perturbation that breaks the left-right symmetry – we are again assuming that we have found a suitable way of breaking the up-down symmetry in the layer. If the symmetry is only broken by changing the magnetic boundary conditions, then as discussed above, the curve must pass through the origin, so Figure 2.5(b) results; otherwise, we get Figure 2.5(a).

Notice that the eigenvalues in the right half-plane no longer remain purely real. In Figure 2.5(a), the curve crosses the imaginary axis at a non-zero value of $\text{Im } s$, meaning that the bifurcation to convection is no longer a steady-state one, but an ‘imperfect’

¹When $\alpha = 90^\circ$ (or equivalently, $k_x = 0$), we see from equations (2.15)–(2.23) that ϕ now only enters the problem in the combination $Q \cos \phi$. This indicates that increasing ϕ is equivalent to reducing Q , which in turn increases instability to convection and reduces R_c . We conclude that (when $\alpha = 90^\circ$ at least) R_c must fall as the tilt increases.

bifurcation in which a small non-zero oscillation frequency (or equivalently, speed of travel) is present. Compare this to the oscillatory case in which the speed of travel is non-zero even when the field is vertical. In Figure 2.5(b), we do have a steady-state bifurcation at onset (the curve goes through the origin), but $\text{Im } s$ still becomes non-zero for $R > R_c$.

Another difference between the steady and oscillatory cases is that in the former, the graph only crosses the imaginary axis once, indicating that waves can only travel in one direction (either leftwards or rightwards). This indicates that we could transform to a moving frame in which the pattern appeared to be steady. (Contrast this to the oscillatory case, where the graph crossed the imaginary axis at points both above and below the origin, indicating that both left-going and right-going waves could be present simultaneously, and no such Galilean transformation could be made.)

We now turn to the three-dimensional problem. In the oscillatory case, we noted above that one expects perpendicular rolls for small ϕ , and oblique rolls for large ϕ . This arose by considering the separate left-going and right-going perpendicular roll modes, and the symmetry breaking between them; this meant that ΔR_c for these modes was proportional to ϕ .

In the steady case, there are not separate left-going and right-going modes, and so the symmetry breaking does not apply; instead, ΔR_c for the perpendicular rolls is proportional to ϕ^2 (not ϕ as before). Also, the parallel rolls do not travel; since there is no reason for them to travel in either the positive or negative y direction, in the steady case they do not travel at all. Thus ΔR_c is proportional to ϕ^2 for the parallel rolls as well. So, without knowing the respective constants of proportionality (which cannot be obtained through simple symmetry arguments alone), we cannot say anything about which orientation of roll is going to be preferred in the steady case.

2.5 Results

2.5.1 Small ζ (oscillatory case)

Two-dimensional results ($k_y = 0$)

In this section, we give the results of our numerical solution of the eigenvalue problem described in section 2.2. The problem solved is exactly the one described in that section, with the following choice of parameters: $Q = 200$, $\zeta = 0.1$, $\sigma = 1$. (However, we do

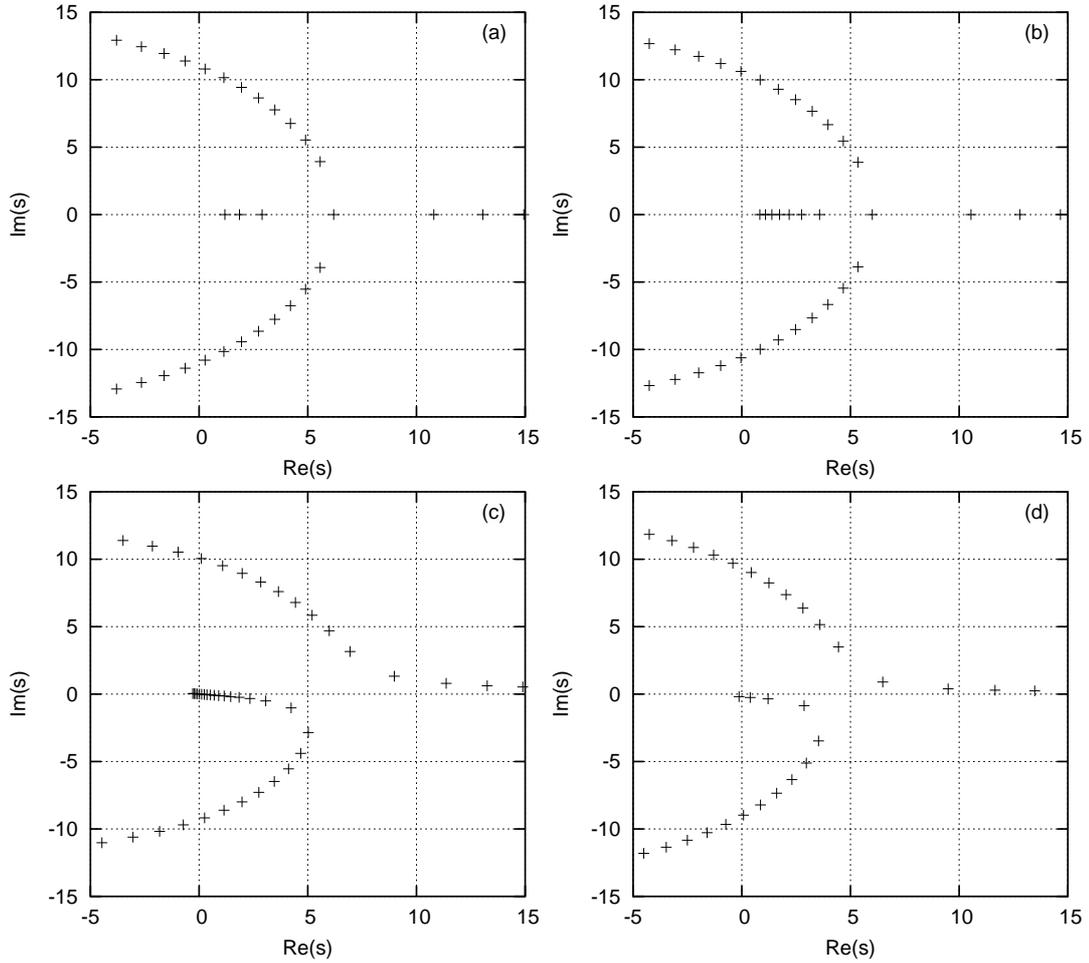


Figure 2.6: Plots of eigenvalues s in the complex plane as R is varied. The crosses are plotted at intervals of 200 in R . In each case \mathbf{k} is fixed and equal to $(k_c, 0)$, where k_c is the critical wavenumber (i.e. the one that minimizes R_c). (a) Vertical field ($\phi = 0^\circ$), with $\hat{\kappa} = 1$. (b) $\phi = 15^\circ$ and $\hat{\kappa} = 1$. (c) As (b), but with a fixed angle condition at the bottom of the layer (the other three cases use a tied field condition at the bottom.) (d) $\phi = 15^\circ$ and $\hat{\kappa} = (z + 1/2)^{-3}$. Note: in case (c) the locus of eigenvalues passes through the origin, whereas in case (d) it merely passes very close to the origin – it in fact crosses the imaginary axis slightly below the origin.

not find any significant qualitative difference for different choices of parameters, except for the change between oscillatory and steady behaviour as ζ varies.) The magnetic boundary conditions are to match onto a potential field above the layer, and to have a perfectly conducting lower boundary (as discussed in section 2.2), except where stated otherwise below.

Figure 2.6(a) shows a plot of the eigenvalues in the complex plane for a typical case with a vertical magnetic field. As expected on symmetry grounds, the eigenvalues come in complex conjugate pairs, showing that there is no preference for either left-going or right-going waves (cf. Figure 2.2a). Figure 2.6(b) shows the same situation but with ϕ increased to 15° . In this case a diagram similar to Figure 2.3(b) is expected, as explained above. Surprisingly, however, the eigenvalues still come in conjugate pairs, and the left-going and right-going waves still have identical growth rates (at any given Rayleigh number)!

This unexpected result is apparently a consequence of our particular choice of boundary conditions. For example, when using a fixed-angle condition at the bottom (instead of the ‘tied field’ condition), but keeping the potential field condition at the top, Figure 2.6(c) results. (A similar graph is also obtained if one uses a tied field at one surface with a fixed angle condition at the other.) Therefore, it appears that the left-right symmetry is not *always* broken by choosing different magnetic boundary conditions at top and bottom; apparently it sometimes is and sometimes isn’t, depending on precisely which boundary conditions are chosen.

This is unfortunate since, as discussed above, the whole point of introducing the different boundary conditions at top and bottom was to try to break this symmetry. Of course, we could abandon our preferred choice of boundary conditions (Figure 2.6(b)), and use a combination that does break the symmetry (e.g. Figure 2.6(c)). However, we have argued that the fixed-angle condition is unphysical, and we prefer not to use it.

At this point, therefore, we bring in our alternative method of breaking the symmetry: a non-uniform thermal diffusivity κ . This is not unreasonable since in the Sun, the thermal diffusivity decreases as a function of depth, due to hydrogen ionization effects (Meyer et al., 1974). The profile $\hat{\kappa} = (0.5 + z)^{-3}$ was chosen. (In fact, the precise choice of this function does not make much qualitative difference; any monotonically decreasing function seems to give similar results.) Figure 2.6(d) shows the path of the eigenvalues that is now obtained. As expected, the left-right symmetry is broken, and the left-going waves are preferred (in this case). (If we ‘reverse’ the choice of κ , so that

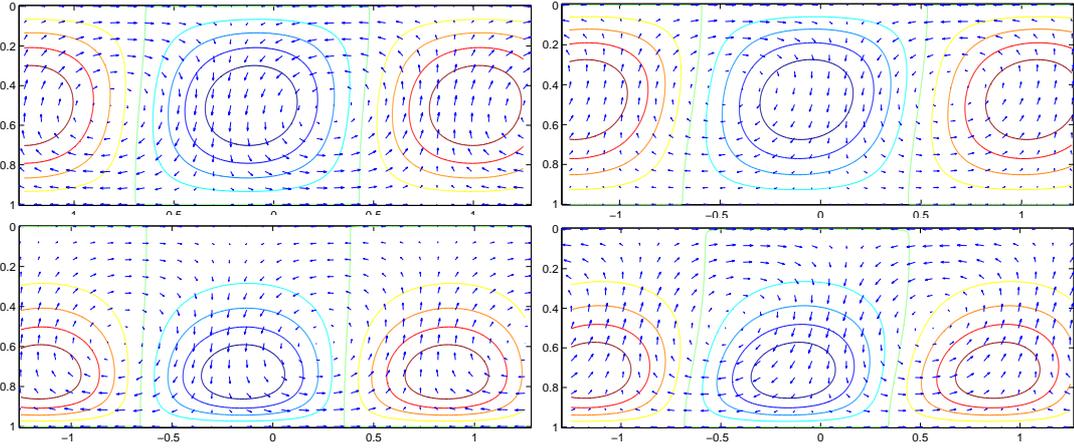


Figure 2.7: Plots of the solutions obtained, with $\phi = 45^\circ$. The contours show the temperature perturbation, and the arrows show the fluid velocity. The top two plots have $\hat{\kappa} = 1$ while the bottom two have $\hat{\kappa} = (1/2 + z)^{-3}$. The left panels show the left-travelling waves, while the right panels show the right-going solutions.

it increases rather than decreases with depth, then the preferred direction of travel is to the right – as might be expected.)

Figure 2.7 shows some of the solutions themselves. These are produced by multiplying the eigenfunctions by $\exp(ikx)$ and plotting the real part of the result (as explained earlier). The upper two panels show the solutions when κ was chosen to be uniform; note that these are not simply mirror images of each other, so the symmetry π has indeed been broken here, even though the growth rates (and critical Rayleigh numbers) for these two solutions are, for whatever reason, identical. In the bottom two pictures, in which κ varies with depth, this is no longer the case. Notice that the temperature perturbation is now concentrated towards the bottom of the layer; this is because heat is being transported more by conduction near the surface, where the heat conductivity is greater, and more by convection at the deeper levels, where the conductivity is lower.

Figure 2.8 shows the variation of the critical Rayleigh number with wavenumber. The first plot shows the vertical field case, in which the oscillatory mode is preferred when it exists (for k_x sufficiently small). The second plot shows what happens when the field is inclined; the preference for left-going waves is clearly visible. Note that the right-going wave has two critical Rayleigh numbers for each wavenumber; this is because the eigenvalue crosses the imaginary axis twice as R changes (see Figure 2.6d). The lower of these two values is the physically relevant one.

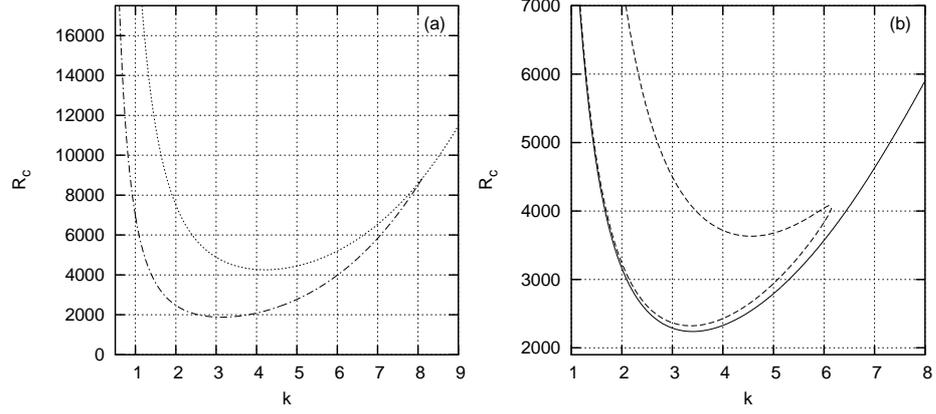


Figure 2.8: R_c as a function of k_x . (a) Vertical field. We obtain both steady solutions (dotted curve) and oscillatory solutions (dash-dotted curve). (b) Tilted field ($\phi = 10^\circ$). Here all solutions travel, either leftwards (solid curve) or rightwards (dashed curve).

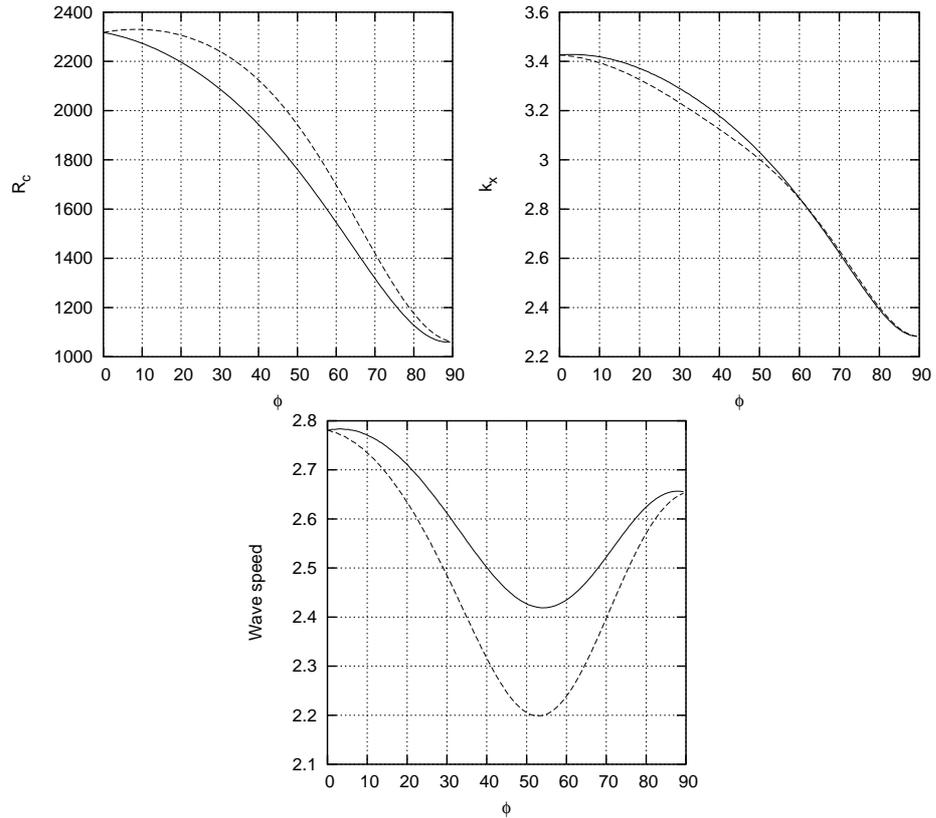


Figure 2.9: Graphs of the critical Rayleigh number (R_c), the wavenumber (k_x) and the wave speed, as functions of ϕ .

Figure 2.9 shows how R_c , k_x and the wave speed vary as a function of the tilt angle ϕ . These graphs confirm that the perturbations to each quantity are proportional to ϕ for small ϕ , as expected on symmetry grounds (see previous section). This applies for ϕ up to about $5\text{--}10^\circ$. After this point, the critical Rayleigh number decreases for both left- and right-going waves. Note also that the wavenumber decreases (equivalently, the wavelength increases) with increasing tilt, indicating that the solution is being stretched out along the tilt direction. This is understandable in that one would expect the convection cells to try to align themselves with the magnetic field lines, and this would cause them to appear to stretch out along the tilt direction.

Three-dimensional results

So far our results have been restricted to the two-dimensional case with $k_y = 0$; we now move on to three-dimensional solutions. To clarify, we are at this point continuing to use our depth-dependent profile for $\hat{\kappa}$, which, as mentioned above, is sufficient to break the symmetry between the leftward- and rightward-travelling modes.

Figure 2.10 shows contour plots of R_c as a function of both k_x and k_y for four different tilt angles. (Note that the convention that we have used is that left-going waves are plotted for k_x negative, and right-going waves are plotted for k_x positive.²) The preferred mode corresponds to the point in \mathbf{k} -space that minimizes R_c ; this point has been marked by an asterisk on each diagram. Figure 2.11 shows the angle α (as defined in equation 2.41) corresponding to the preferred mode. This shows clearly which type of roll (from Figure 2.4) is preferred at each ϕ ; $\alpha = 0^\circ$ corresponds to perpendicular rolls, $\alpha = 90^\circ$ represents parallel rolls, while an α in between these two values means that oblique rolls are preferred. Finally, Figure 2.12 shows the critical Rayleigh number for the different types of rolls as a function of ϕ .

Figures 2.10 and 2.11 are related, in that the position of the minimum on Figure 2.10, for a given ϕ , determines the value of α corresponding to that ϕ on Figure 2.11. Hence, we see that the minimum is located on the negative k_x axis for the smallest ϕ values (Figure 2.10a), which means that α is zero for small ϕ . When ϕ reaches a critical value (just above 10° in this case), the minimum leaves the k_x axis and moves

²Effectively what we are doing here is demanding that $\text{Im } s$ be negative, and then using the sign of k_x to indicate the direction of travel. This is in contrast to our previous figures where k_x is assumed positive and the sign of $\text{Im } s$ was used to indicate the direction of travel. The two are of course equivalent; this can be seen from equations (2.11)–(2.14), which remain invariant if one changes \mathbf{k} to $-\mathbf{k}$ and s to $-s$.

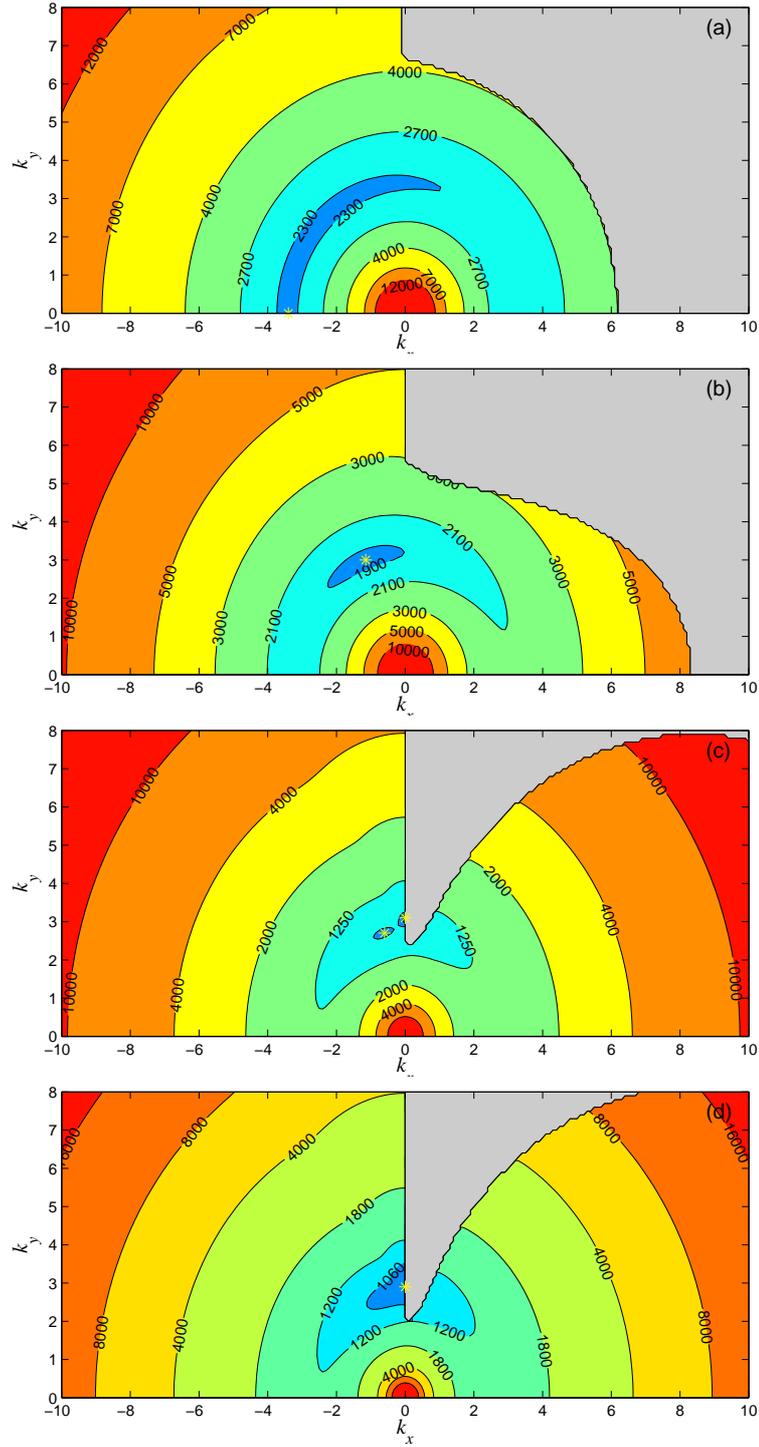


Figure 2.10: Contour plots showing critical Rayleigh number as a function of wavevector, with positive (negative) k_x representing right-going (left-going) waves. The shaded areas represent regions where no right-going solution exists. The preferred mode (corresponding to the minimum R_c) is indicated by a small asterisk in each plot. The four cases are (a) $\phi = 10^\circ$, (b) $\phi = 40^\circ$, (c) $\phi = 72.5^\circ$, (d) $\phi = 75^\circ$.

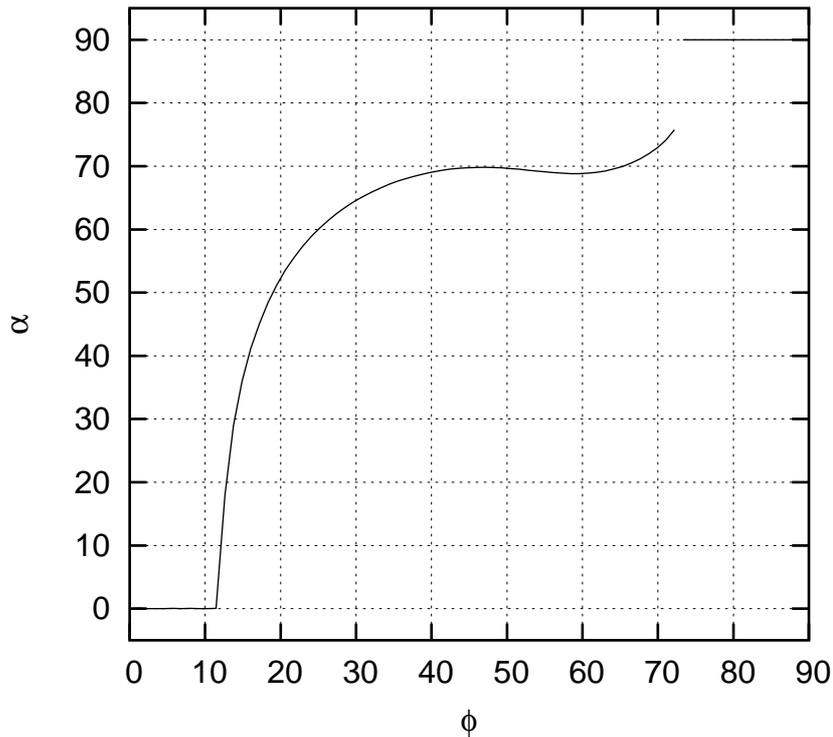


Figure 2.11: *The preferred orientation α of the convection rolls as a function of ϕ .*

up and to the right. This is quite rapid at first and is visible as a sharp increase in α on Figure 2.11. Physically, this means that the preferred mode has changed from being perpendicular rolls to being oblique rolls. Indeed this transition is the same as the one predicted by symmetry arguments (section 2.4.2). By $\phi = 40^\circ$ (Figure 2.10b) the minimum has settled to a position corresponding to $\alpha \approx 70^\circ$. For larger ϕ , a second local minimum appears on the positive k_y axis; this is not initially the lowest of the two minima, but it rapidly decreases in value and eventually ‘overtakes’ the original minimum to become the new global minimum. This results in a discontinuous change in α and is the reason for the discontinuity in Figure 2.11. The point where the minima ‘cross’ can be seen in Figure 2.10(c). For higher values of ϕ the parallel rolls are now dominant (Figure 2.10d).

This behaviour can be compared to the predictions of the simple model from section 2.4.2, and in particular, equation (2.42). As already stated, the main prediction was that there would be a transition from perpendicular to oblique rolls, and this transition was expected to occur at $\phi = B/2C$ where B and C are as defined in (2.42). The values of B and C can be estimated by inspecting the numerical results near $\phi = 0$; we find $B = 163.6$ and $C = 923.2$, which gives $B/2C = 0.089$ radians, or 5.1 degrees, which is

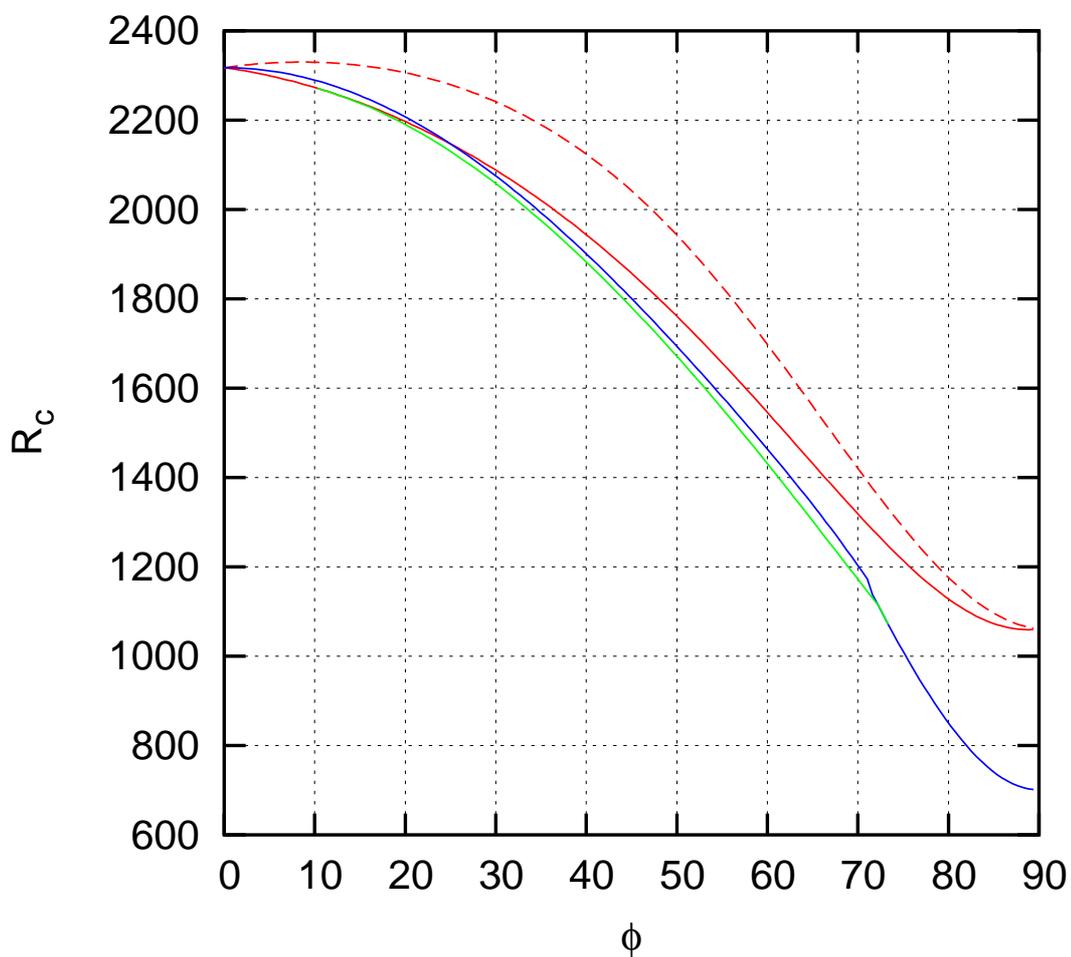


Figure 2.12: The critical Rayleigh numbers R_c for various modes as a function of ϕ . The colour coding is the same as in Figure 2.4. The red curves show critical Rayleigh numbers for perpendicular rolls, corresponding to our two-dimensional solutions; the dashed curve is for the right-going rolls, and the solid curve is for the leftward-travelling solutions. The blue curve is for parallel rolls, and the green for oblique (more precisely, the green curve shows the critical Rayleigh number for the most unstable oblique mode, when such a mode is more unstable than either the perpendicular or parallel rolls).

about a factor of two too small (compared to the actual location of the Lifshitz point). Therefore, the model equation (2.42) is accurate qualitatively but not quantitatively, although the latter is hardly surprising given the ad hoc nature of the model (e.g. the factors of $\cos \alpha$ and $\sin \alpha$ were invented in order to satisfy the right properties, rather than to be an exact quantitative description).

One thing that the model equation (2.42) does get wrong, even qualitatively speaking, is that it does not predict the discontinuous jump in α , and the associated appearance of a second local minimum in R_c , which occurs for larger values of ϕ . Therefore, we need to explain why this second minimum appears.

We can do this by considering the governing equations when $\alpha = 90^\circ$. In this case the horizontal component of the imposed field disappears from the equations, and this mode sees only the vertical component. As the tilt of the field increases, the vertical component of the field decreases and so the effective Chandrasekhar number decreases. Beyond a certain tilt the effective Q becomes too small for the oscillatory instability, and the mode becomes steady. This is the fact that was missing from the analysis of section 2.4. The critical Rayleigh number falls more quickly after this point (this is visible on Figure 2.12 as a ‘kink’ in the curve for parallel rolls, near $\phi = 70^\circ$) and the steady parallel mode soon becomes the dominant one.

The preference for parallel rolls for large ϕ is consistent with the work of Danielson (1961), who showed that in a *horizontal* magnetic field ($\phi = 90^\circ$), convection would consist of parallel rolls. Essentially, this is because this form of motion can just bodily displace the field lines, without distorting them; perpendicular or oblique rolls would have to twist the field lines, and this would be resisted by the Lorentz force.

One final point concerning Figure 2.10 is the meaning of the grey, shaded areas on the contour plots. (Note that the jagged appearance of these regions is not real, but is due to limited resolution.) These regions arise because, for certain wavevectors, no right-going wave exists, only a left-going one. This can be seen clearly on Figure 2.8(b), for example, where the right-going branch (the dashed curve) does not extend to values of k_x greater than about 6.

It is also of interest to consider what happens to the position of the Lifshitz point as the amount of up-down asymmetry (as determined by the choice of κ) is varied. To investigate this, we set $\kappa = (0.5 + z)^{-\kappa_0}$, and then plotted the value of ϕ corresponding to the Lifshitz point as a function of κ_0 ; see Figure 2.13. It can be seen that as the amount of up-down asymmetry is increased (corresponding to larger values of κ_0), the

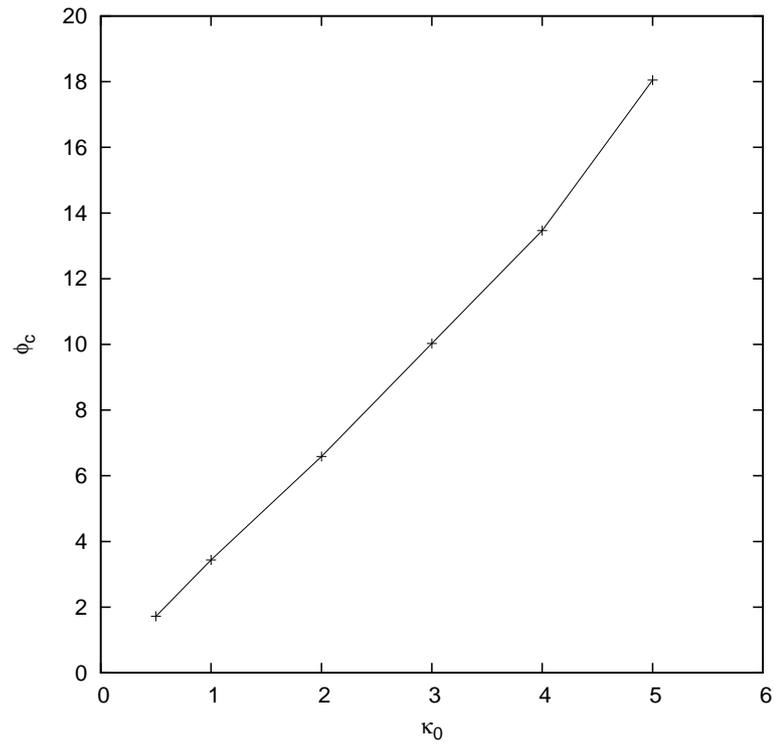


Figure 2.13: The variation of ϕ_c (the value of ϕ corresponding to the Lifshitz point, measured in degrees) with κ . Here $\kappa = (0.5 + z)^{-\kappa_0}$ and κ_0 has been allowed to vary. (Other parameters: $Q = 300$, $\zeta = 0.1$, $\sigma = 1$.)

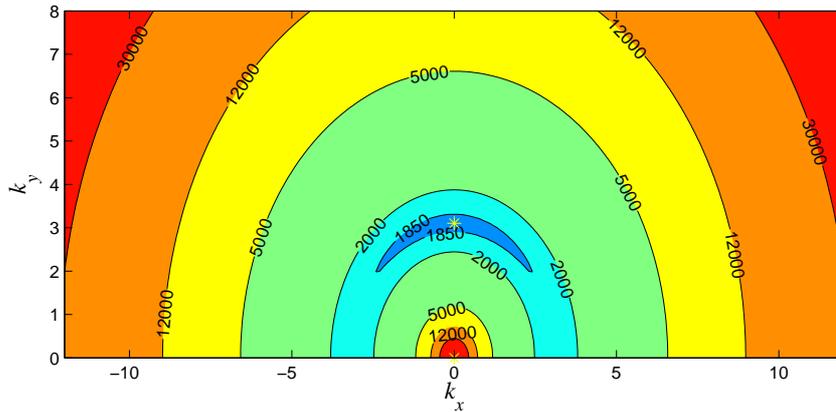


Figure 2.14: Contour plots of R_c as a function of wavevector in the case where $\hat{\kappa}$ does not vary with depth. The tilt angle $\phi = 15^\circ$. Notice that although these results are anisotropic, they are still left-right symmetric. This shows that the left-right symmetry, found in the two-dimensional results above, continues even when $k_y \neq 0$.

Lifshitz point occurs at larger and larger values of ϕ .

For completeness, we also present the three-dimensional results for the case where $\hat{\kappa}$ does *not* vary with depth. (Recall from the previous section that in this case, the left-right symmetry is *not* broken, despite expectations to the contrary.) The contour plot of R_c against k_x and k_y is shown in Figure 2.14. Notice how this unexpected left-right symmetry persists even when $k_y \neq 0$.

2.5.2 Steady case

The results here are actually a little simpler than the oscillatory case, and so less space will be devoted to these. As noted above, the solutions are expected to start travelling as soon as ϕ increases above zero, but we cannot say from symmetry arguments which direction the solutions will travel in. Nor can we say in advance which of parallel, perpendicular, or oblique rolls will be preferred.

In order to obtain a steady, rather than oscillatory, bifurcation (when $\phi = 0$), we have taken $\zeta = 1.1$ instead of 0.1 in this section (leaving the other parameters unchanged). We are also using the depth-dependent profile for $\hat{\kappa}$ in this section.

A contour plot of R_c against k_x and k_y , for the case $\phi = 45^\circ$, is shown in Figure 2.15. Note that in this case, all solutions travel to the left (i.e. against the tilt), and no solutions travel to the right; consequently, the figure only shows results for $k_x < 0$.

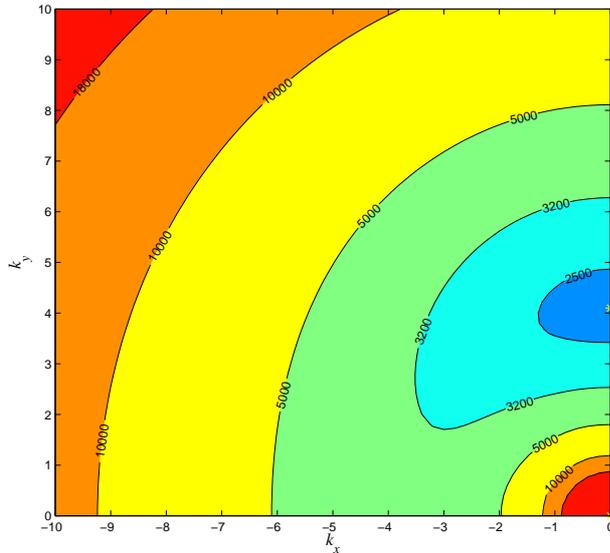


Figure 2.15: Contour plot of R_c against k_x and k_y in the steady case. Note that all solutions travel leftwards (or else are steady), hence only the left-hand half of the diagram is shown.

Notice how the anisotropy introduced by the tilt is clearly visible. The minimum R_c occurs on the k_y axis, indicating that parallel rolls are preferred. In fact, we find that parallel rolls are the preferred mode for all ϕ in this case.

In Figure 2.16 we plot R_c , the wavenumber and the wave speed as functions of ϕ , for both parallel and perpendicular rolls. As regards the wave speed, this is zero for the parallel rolls (this is just because of the $y \rightarrow -y$ reflection symmetry), but it is non-zero for perpendicular rolls (and in fact oblique rolls as well). Note how the speed of travel is proportional to ϕ for small ϕ , as expected (indeed, this is true up to quite large ϕ , about 60° or so).

We can also compare our results to those of Chandrasekhar (1961) who briefly considered convection in inclined fields in the steady (large ζ) case. He also found that parallel rolls (with $k_x = 0$) would be the preferred mode for all ϕ . However, his model was purely Boussinesq, with uniform κ and identical boundary conditions at top and bottom. Therefore, his model contained the up-down symmetry that we talked about in the previous section, and so his solutions did not travel at all. By contrast, all our solutions (except the parallel rolls) travel to the left. This may seem irrelevant, since the preferred mode is the parallel rolls, which do not travel. However, once we move on to nonlinear solutions, all modes will be present in some sense, and we might expect to start to see the effects of the travelling modes.

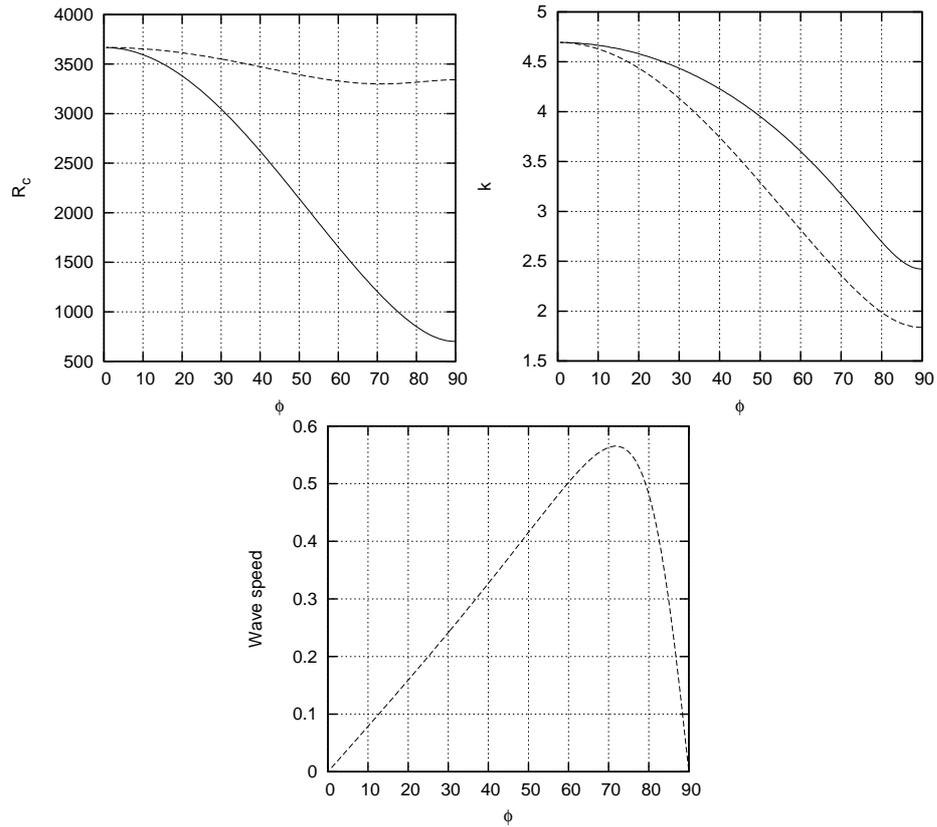


Figure 2.16: Graphs of the critical Rayleigh number (R_c), the wavenumber ($|\mathbf{k}|$) and the wave speed, as functions of ϕ . The solid line is for parallel rolls, while the dashed line is for perpendicular rolls.

2.6 Conclusions

In this chapter we have investigated the linear stability theory for magnetoconvection in a Boussinesq fluid layer with an inclined magnetic field and a depth-dependent thermal diffusivity κ .

We included the depth-dependent κ as a way of breaking the up-down symmetry of the fluid layer, a symmetry that is present in Boussinesq but not in compressible convection. Since compressibility is important in sunspots, it was important to find a way to break this symmetry in our model. We at first tried to do this simply by modifying the boundary conditions at the top and bottom of the layer, but that was found to be unsatisfactory (for reasons explained above), and so we included the depth-dependent κ instead.

The most important effect of tilting the field is that the layer becomes anisotropic, so that rolls of different orientations have different stability properties, and modes can now exhibit a preferred direction of travel (either towards or away from the tilt direction).

When ζ is large (the ‘steady’ case) the preferred mode is parallel rolls for all ϕ . These are steady, but rolls in any other orientation travel (with a speed proportional to ϕ for small ϕ). These travelling modes are not preferred near onset, but their presence suggests that patterns in the nonlinear regime might travel.

For small ζ (the ‘oscillatory’ case), the results are more complicated. In the two-dimensional problem ($k_y = 0$), we find broadly similar results to those of Matthews et al. (1992), with symmetry-breaking between left- and right-going modes. In the three-dimensional problem ($k_y \neq 0$), we find that travelling perpendicular rolls are preferred for small ϕ , travelling oblique rolls for intermediate ϕ and steady parallel rolls for large ϕ . Note that Matthews et al. (1992) predicted that the transition between perpendicular and oblique rolls would be found, but they did not predict the transition between oblique and parallel rolls for larger ϕ . This comes about because the vertical component of the magnetic field drops as ϕ increases, which triggers a change from oscillatory to steady convection.

Perhaps the main difference between our model and that of Matthews et al. (1992) is in the direction of travel of the preferred mode. We find waves travelling to the left (against the tilt) for all ϕ ; they find waves travelling to the right for most ϕ , but for a small range of ϕ near 90° , this reverses and the waves travel to the left. They can also produce left-going waves for all ϕ by changing their boundary conditions. Clearly, the

direction, and speed, of travel are highly model-dependent. Nevertheless, in terms of application to sunspots, it is encouraging that some models, at least, show a reversal of direction of travel as ϕ increases – even if it is the wrong way around (in a sunspot, features are observed to move inwards, i.e. ‘left’, for small ϕ and right for large ϕ)!

Chapter 3

Weakly nonlinear models

3.1 Introduction

In this chapter we will move beyond simple linear models and start to cover weakly nonlinear effects. We will use ideas from bifurcation theory and pattern formation to construct fairly general models, which will allow us to learn something about how convective patterns change as the inclination of the field is varied. In the next chapter, we will also use this work as a starting point for building more detailed models.

Consider first of all the case of a vertical magnetic field. In the previous chapter, we investigated the linear stability of the trivial solution, finding that for large enough Rayleigh number, the layer became unstable to roll-like (plane wave) perturbations of the form $\exp i\mathbf{k} \cdot \mathbf{x}$. At the critical Rayleigh number, we found that there was a circle of critical modes, with $|\mathbf{k}| = k_c$, which were neutrally stable; all other modes decayed exponentially. For R slightly above critical, there was a narrow annulus of modes in Fourier space, with wavenumbers close to k_c , which could grow exponentially.

We can use this linear analysis to try to predict what we might see when R is just above critical. We might expect the convection to take the form of rolls, with wavenumber $|\mathbf{k}|$ equal to (or very close to) k_c . However, there are an infinite number of such modes, because of the rotational symmetry of the problem; any wavevector lying on the critical circle $|\mathbf{k}| = k_c$ will do. Moreover, we can superpose two or more of these roll modes to obtain a new pattern. Examples of this are shown in Figure 3.1, which shows how hexagons can be obtained by superposing three sets of rolls at 120° to each other, and how squares can be produced by superposing two sets of rolls at right angles.

Linear theory alone cannot distinguish between a single roll mode, or a superposi-

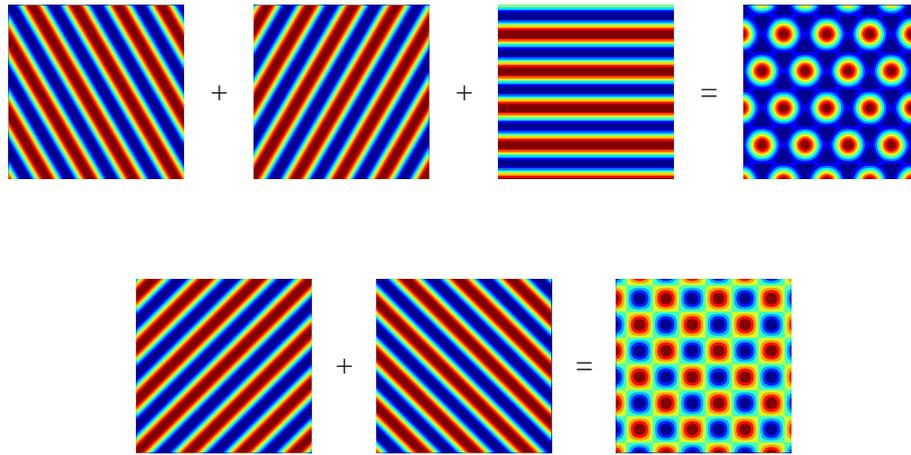


Figure 3.1: Hexagon and square patterns. The first line shows a hexagonal pattern, created by adding together three sets of rolls at an angle of 120° to each other. The second line shows a square pattern, created by superposing two sets of rolls at right angles to each other.

tion such as those shown in Figure 3.1, because the different modes would all evolve completely independently of each other. However, once we add nonlinearities into the problem, interactions between the different Fourier modes become possible. This means that the theory can now distinguish between the different patterns, and we will be able to predict which of the patterns we would see near onset. If the nonlinearities are assumed to be weak, it is possible to model the nonlinear interactions based on minimal assumptions about the underlying physics, which is what we will do in the rest of this chapter.

3.2 General notes on pattern formation

In this section we will explain some general methods that are used for studying pattern formation in weakly nonlinear systems. (In the following sections, we will go on to show how these methods can be applied in particular cases.)

A standard method of dealing with systems close to a bifurcation point (in our case, with R only just above R_c) is to apply the centre manifold theorem. This allows us to simplify the problem by considering only the dynamics on the (extended) centre manifold. Basically, this can be thought of as eliminating all modes of wavenumber not equal to k_c ; these are ‘fast’ modes which decay exponentially with time. The critical modes, with wavenumber equal to k_c , constitute the slow dynamics. Unfortunately, with

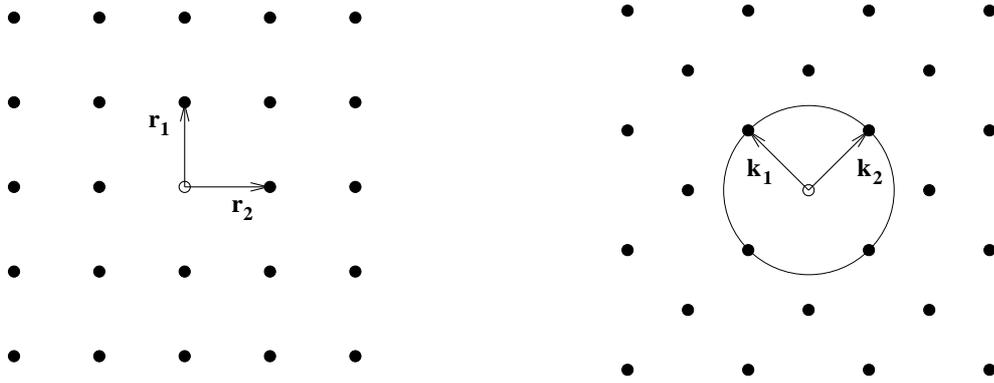


Figure 3.2: A periodic square lattice. The left-hand picture shows the lattice in real space, and the right-hand picture shows the ‘dual’ lattice in Fourier space.

the problem in its present form, the conditions of the centre manifold theorem are not met, for two reasons: firstly, there are an infinite number of critical modes, occupying the circle $|\mathbf{k}| = k_c$; secondly, the non-critical modes can be arbitrarily close to this circle and have growth rates which can be arbitrarily close to zero.

One way to avoid both of these problems is to restrict ourselves to solutions which are doubly periodic on some lattice; in other words, if we denote the solution by $\Psi(\mathbf{x})$, then there should exist two ‘lattice vectors’, \mathbf{r}_1 and \mathbf{r}_2 , such that for all integers n_1 and n_2 and for all \mathbf{x} , Ψ satisfies $\Psi(\mathbf{x} + n_1\mathbf{r}_1 + n_2\mathbf{r}_2) = \Psi(\mathbf{x})$.

This restriction to a doubly periodic lattice can also be thought of as a restriction on the Fourier transform of the solution. If we consider possible Fourier modes $\exp i\mathbf{k} \cdot \mathbf{x}$, then clearly only certain wavevectors \mathbf{k} will produce a function which is doubly periodic on the given lattice. The set of such wavevectors \mathbf{k} itself forms a lattice, in Fourier space, which is known as the *dual lattice*.

These concepts are best explained by giving examples. Two lattices which are commonly used in these sorts of problems are the square and hexagonal lattices (so-called because they admit the square and the hexagonal patterns from Figure 3.1).

The square lattice is illustrated in Figure 3.2. The left-hand picture shows the two lattice vectors \mathbf{r}_1 and \mathbf{r}_2 , together with the lattice points themselves (i.e. points of the form $n_1\mathbf{r}_1 + n_2\mathbf{r}_2$ for n_1 and n_2 integer). For the square pattern, these points would correspond to the centres of the squares. The right-hand picture shows the dual lattice vectors \mathbf{k}_1 and \mathbf{k}_2 , together with the dual lattice points (points of the form $n_1\mathbf{k}_1 + n_2\mathbf{k}_2$ for n_1 and n_2 integer). The points of the dual lattice do not correspond to physical

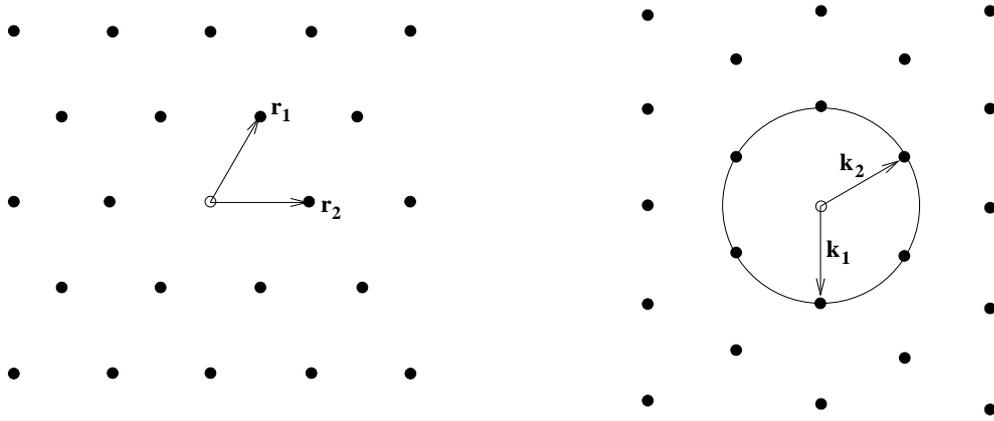


Figure 3.3: The equivalent of Figure 3.2 for the hexagonal lattice.

points but rather to wavevectors of ‘admissible’ Fourier modes. The equivalent diagram for the hexagonal lattice is shown in Figure 3.3. (The other common type of lattice is the ‘rhombic’ lattice, where the angle between the two wavevectors \mathbf{k}_1 and \mathbf{k}_2 takes a general value, equal to neither 90° nor 120° .)

If we scale our lattice such that the dual vectors \mathbf{k}_1 and \mathbf{k}_2 are of length k_c , then the critical circle will intersect four wavevectors (in the square or rhombic case) or six wavevectors (in the hexagonal case), as illustrated in the diagram. In other words, whilst in the full problem there was an entire circle of critical modes, once we restrict to a periodic lattice, there are a finite number of permissible critical modes. Moreover, all other modes on the dual lattice are a finite distance away from the critical circle $|\mathbf{k}| = k_c$, and therefore their growth rates are bounded away from zero. Hence, both of the difficulties mentioned above are resolved, and we can reduce the problem to the evolution of a small number of modes on the extended centre manifold.

Thus, the restriction to a periodic lattice greatly simplifies the mathematics, but it also unfortunately restricts the set of possible patterns that can be investigated. For example, if we chose the square lattice, we could look at rolls and square patterns. The hexagonal lattice allows rolls and hexagons as well as some other patterns (e.g. triangles and rectangles), but it leaves out the squares. (There is no lattice on which both squares and hexagons can exist simultaneously.) The results must be interpreted with this in mind, although usually we have some physical intuition about the problem, which can inform the decision of which lattice is most appropriate to a particular situation.

Therefore we are led to consider a superposition of N different roll modes. In the

case of a steady bifurcation, the solutions will take the following form:

$$\Psi = \sum_{j=1}^N z_j(t) e^{i\mathbf{k}_j \cdot \mathbf{x}} \Psi_0^{(j)} + \text{c.c.} \quad (3.1)$$

Here Ψ stands for (one or more of) the physical variables from the original problem (\mathbf{u} , \mathbf{B} etc.). The \mathbf{k}_j are the wavevectors of each mode intersecting the critical circle $|\mathbf{k}| = k_c$. Any vertical dependence has been factored out into $\Psi_0^{(j)}$ (which essentially represents the vertical eigenfunction of the mode). The z_j are complex functions of time representing the amplitudes of the different modes.

In the case of the square lattice, it appears that $N = 4$, since four points on the dual lattice intersect the critical circle; see Figure 3.2. However, in fact N is only 2; this is because wavevectors \mathbf{k} and $-\mathbf{k}$ are in fact equivalent (the *direction* of the wavevector determines the orientation of a mode, but its *sense*, either $+$ or $-$, is irrelevant). Similarly, for steady bifurcation on a hexagonal lattice, $N = 3$ and not 6.

For an oscillatory bifurcation, the form is slightly different:

$$\Psi = \sum_{j=1}^N (z_j(t) e^{i\mathbf{k}_j \cdot \mathbf{x}} + w_j(t) e^{-i\mathbf{k}_j \cdot \mathbf{x}}) e^{-i\omega t} \Psi_0^{(j)} + \text{c.c.} \quad (3.2)$$

Note the addition of the $\exp(-i\omega t)$ term, which represents the oscillation frequency of each mode. This time there *is* a difference between wavevectors \mathbf{k} and $-\mathbf{k}$; the sense of the wavevector indicates the direction of travel of each mode (each mode can travel in one of two directions).

Our objective now is to come up with a set of ordinary differential equations that govern the time evolution of the z_j . These are known as the *amplitude equations*. In linear theory they would be very simple: $\dot{z}_j = s z_j$, where s is the corresponding linear growth rate. In weakly nonlinear theory we have to add further terms representing nonlinear interactions between the different modes. These would generally be determined from a perturbation analysis of the original partial differential equations; the required amplitude equations usually come out as a solvability condition.

However, we do not need to go to such lengths to gain some useful insight into the problem. Rather, we can make use of symmetry. We know that the problem possesses certain symmetries; in our case, the relevant symmetries are rotations, reflections and translations of the plane, which together form the Euclidean symmetry group $E(2)$. (More specifically, we must consider the subset of $E(2)$ that leaves our periodic lattice invariant.) The amplitude equations must be consistent with these symmetries; not just

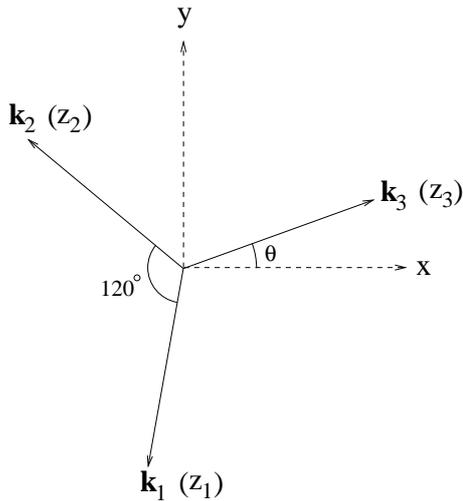


Figure 3.4: Diagram showing the three wavevectors \mathbf{k}_j and their corresponding amplitudes z_j ($j = 1, 2, 3$), for the steady hexagonal model. Note the orientation (represented by the angle θ); the direction of tilt of the magnetic field is along the x -axis.

any amplitude equations are allowed. Together with the assumption that the amplitudes are small, this usually provides enough information to constrain the amplitude equations quite strongly; typically the form of the equations is known, but there are one or more undetermined coefficients.

We will now consider how these principles apply to particular choices of periodic lattice.

3.3 Steady bifurcation on a hexagonal lattice

We start with the hexagonal lattice, since this is a natural choice for magnetoconvection, for the following reason. From studies of magnetoconvection in vertical fields, we expect to see a pattern of hexagons near onset. When the field is tilted we would presumably still see a (slightly distorted) hexagonal pattern, but when the field is nearly horizontal, the preferred pattern will be field-aligned rolls (Danielson, 1961). The interesting question is what happens for field inclinations in between these two extremes.

To investigate this competition between hexagons and rolls, we consider the hexagonal lattice, with three wavevectors \mathbf{k}_1 , \mathbf{k}_2 and \mathbf{k}_3 , as illustrated in Figure 3.4. The angle θ represents the orientation of the lattice; this is unimportant when the field is vertical (because of the rotational symmetry), although it will become important later on, when

we introduce non-vertical fields.

Staying with a vertical magnetic field for the moment, we can determine the most general amplitude equations satisfying the symmetry requirements. (The relevant symmetry group is the subset of $E(2)$ that leaves the hexagonal lattice invariant, which is the group $T^2 \rtimes D_6$; T^2 represents x and y translations, while D_6 represents rotational and reflectional symmetries of the hexagonal lattice.) This is actually a well-known, standard problem (e.g. Golubitsky et al. 1984). The equations are:

$$\dot{z}_1 = \mu z_1 + \alpha \bar{z}_2 \bar{z}_3 - \gamma z_1 (|z_1|^2 + \beta |z_2|^2 + \beta |z_3|^2) \quad (3.3)$$

$$\dot{z}_2 = \mu z_2 + \alpha \bar{z}_3 \bar{z}_1 - \gamma z_2 (|z_2|^2 + \beta |z_3|^2 + \beta |z_1|^2) \quad (3.4)$$

$$\dot{z}_3 = \mu z_3 + \alpha \bar{z}_1 \bar{z}_2 - \gamma z_3 (|z_3|^2 + \beta |z_1|^2 + \beta |z_2|^2) \quad (3.5)$$

The coefficients α , β , γ and μ are all real. These will depend on the various parameters in the original problem (Q , ζ , etc.) and we cannot determine their values without more detailed calculations. Clearly, however, we can see that μ will be an increasing function of R , with $\mu = 0$ when $R = R_c$. The coefficient α represents the amount of up-down symmetry breaking; it is zero in an up-down symmetric case (such as Boussinesq convection).

We have dropped terms higher than third order, which may be justified by assuming the following scalings. The small parameter in the problem is the amount by which the Rayleigh number exceeds critical (i.e., $R - R_c$). If we define ϵ such that $R - R_c \propto \epsilon^2$, then the amplitudes $|z_j|$ are of order ϵ . We must also assume for consistency that $\alpha = O(\epsilon)$, i.e. that the amount of up-down symmetry breaking is small. Each term is then of the same order (ϵ^3), and all the neglected terms are of higher order.

By rescaling time and the amplitudes, we can assume without loss of generality that $\alpha = \gamma = 1$. We cannot scale out β , but we do make the assumption that $\beta > 1$, which ensures that rolls are stable in the absence of the quadratic term (which is the case for convection).

We may also take the z_j to be real (after a shift of origin if necessary). This can be shown quite easily by writing out the equations in terms of amplitudes and phases of the z_j ; see e.g. Malomed et al. (1994).

We now turn to solutions of these equations. The bifurcation diagram as μ is varied is shown in Figure 3.5. The trivial solution, $z_1 = z_2 = z_3 = 0$, is stable for μ negative and unstable for μ positive. There is a branch of rolls (e.g. $z_1 = \sqrt{\mu}$, $z_2 = z_3 = 0$), which bifurcates supercritically from $\mu = 0$, and is stable for μ sufficiently large. The

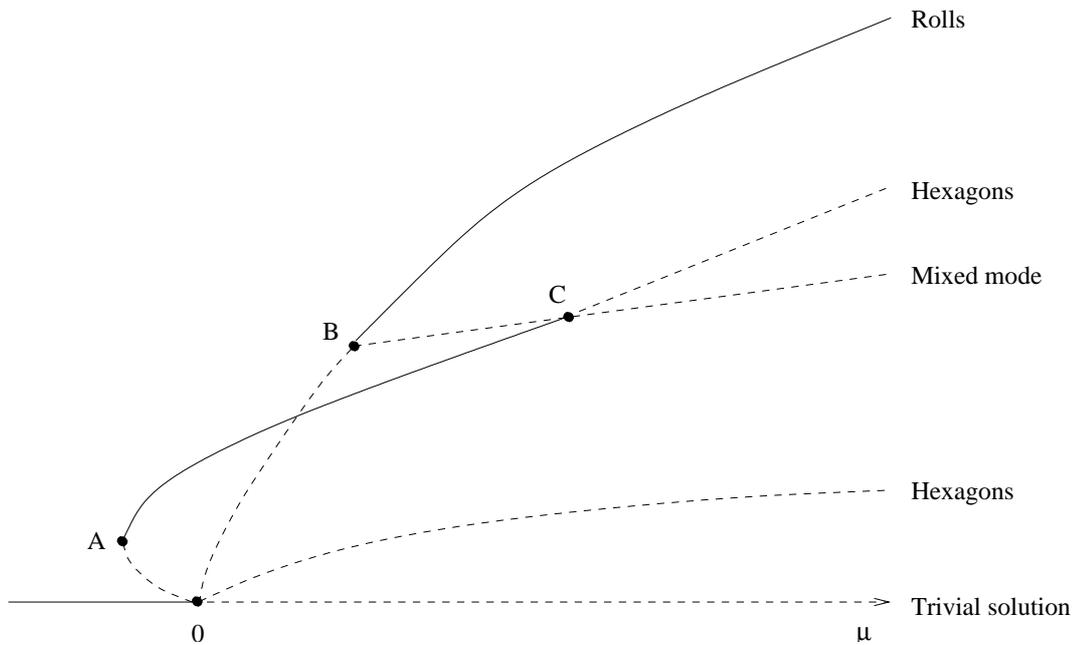


Figure 3.5: Bifurcation diagram for the symmetric case (when all three linear growth rates are equal). The bifurcations A, B and C occur at the following values of μ : A at $\mu = -1/(4 + 8\beta)$; B at $\mu = 1/(\beta - 1)^2$; C at $\mu = (\beta + 2)/(\beta - 1)^2$.

bifurcation at which the rolls gain stability also creates a branch of mixed modes, which are rectangular in appearance, and have (e.g.) $z_1 = z_2 \neq z_3$; this solution is always unstable.

Finally, there are also two branches of hexagons. These are related by a sign change (in the three z_j). The uppermost branch can be stable (for some μ values at least) while the lower branch is always unstable (and hence unobservable). In convection, one branch would represent a solution with upflows at the centre of each convection cell (so-called ‘up’ hexagons), while the other would have downflows there (‘down’ hexagons). In general, we cannot say which is which, i.e. we cannot say whether the uppermost, stable branch corresponds to the ‘up’ or to the ‘down’ hexagons. For compressible convection, however, we know that the ‘up’ hexagons are the stable solution (as is confirmed by numerical simulations, for example; see also section 1.5.1).

The problem also contains hysteresis. This occurs in two separate ranges of μ values. The first is between points A and O on Figure 3.5, where hexagons and the trivial solution are simultaneously stable. This indicates that the onset of convection itself is associated with hysteresis. The second region of hysteresis occurs between points B and C, where there is bistability between rolls and hexagons. This indicates that the

transition between hexagons (for smaller μ) and rolls (for larger μ) is hysteretic.

3.3.1 Weakly breaking the isotropy

So far, we have considered only vertical fields. When the field is tilted, anisotropy is introduced and the symmetry between the three different modes is broken. The translational symmetry remains, as does the reflection symmetry $y \rightarrow -y$, but the other reflection symmetries, as well as the isotropy, are lost. We could repeat the analysis using the new (smaller) symmetry group, and come up with a different set of ordinary differential equations for the z_j . However, it is more enlightening to consider a situation where the symmetry is only *weakly broken*. For example, if we consider a small tilt angle ϕ , then the problem is strictly speaking anisotropic, but we are still very close to the isotropic situation, and this can be exploited.

We will model this by breaking the symmetry in the linear terms only, replacing the single value μ with different values μ_1 , μ_2 and μ_3 in each equation. Strictly speaking the other coefficients (α , β and γ) should be changed as well, but this is not actually necessary to break the symmetry, and we will just leave them unchanged. (We will justify this further below.)

The new equations, with weakly broken symmetry, are:

$$\dot{z}_1 = \mu_1 z_1 + \alpha \bar{z}_2 \bar{z}_3 - \gamma z_1 (|z_1|^2 + \beta |z_2|^2 + \beta |z_3|^2) \quad (3.6)$$

$$\dot{z}_2 = \mu_2 z_2 + \alpha \bar{z}_3 \bar{z}_1 - \gamma z_2 (|z_2|^2 + \beta |z_3|^2 + \beta |z_1|^2) \quad (3.7)$$

$$\dot{z}_3 = \mu_3 z_3 + \alpha \bar{z}_1 \bar{z}_2 - \gamma z_3 (|z_3|^2 + \beta |z_1|^2 + \beta |z_2|^2) \quad (3.8)$$

Since we now have two small quantities, the tilt angle ϕ and the ‘effective’ Rayleigh number $r \equiv (R - R_c)/R_c$, it is important to clarify the scalings involved. We introduce a small parameter ϵ and scale ϕ as $O(\epsilon)$, while r is (as before) taken to be $O(\epsilon^2)$. We can now write the μ_j as functions of both r and ϕ . The correct form for small ϵ is

$$\mu_j = A_j \phi^2 + B_j i \phi + C r + O(\epsilon^3), \quad (3.9)$$

where A_j , B_j and C are real constants, with $A_1 \geq A_2 \geq A_3$. (The argument that the growth rates must take this form for small ϕ was given in the previous chapter.) Ignoring the imaginary parts (for the moment), each μ_j is of order ϵ^2 , and the amplitudes $|z_j|$ may be scaled to be $O(\epsilon)$. We also assume once again that $\alpha = O(\epsilon)$; each term on the right-hand side is then of the same order (ϵ^3).

Unfortunately, the imaginary parts of the μ_j appear at order ϵ and thus do not satisfy this balance. However, we can deal with this by making a Galilean transformation. To show this we first of all write $z_j = R_j e^{i\psi_j}$ (with R_j and ψ_j real) and decompose (3.6)–(3.8) into amplitude and phase equations:

$$\dot{R}_1 = \mu_{1r} R_1 + \alpha R_2 R_3 \cos \Psi - \gamma(R_1^2 + \beta R_2^2 + \beta R_3^2) R_1 \quad (3.10)$$

$$\dot{R}_2 = \mu_{2r} R_2 + \alpha R_3 R_1 \cos \Psi - \gamma(R_2^2 + \beta R_3^2 + \beta R_1^2) R_2 \quad (3.11)$$

$$\dot{R}_3 = \mu_{3r} R_3 + \alpha R_1 R_2 \cos \Psi - \gamma(R_3^2 + \beta R_1^2 + \beta R_2^2) R_3 \quad (3.12)$$

$$\dot{\Psi} = \Omega - \alpha \left(\frac{R_2 R_3}{R_1} + \frac{R_3 R_1}{R_2} + \frac{R_1 R_2}{R_3} \right) \sin \Psi. \quad (3.13)$$

Here $\Psi = \psi_1 + \psi_2 + \psi_3$ and $\Omega = \mu_{1i} + \mu_{2i} + \mu_{3i}$. (Note that the ‘r’ and ‘i’ subscripts refer to real and imaginary parts; e.g. $\mu_1 = \mu_{1r} + i\mu_{1i}$.)

When $\phi = 0$, $\Omega = 0$ and (3.13) shows that Ψ tends to zero. This means that we can take all three z_j to be real. To see this, note that we have two translational degrees of freedom, so we can set two of the ψ_j to zero by shifting the origin. The condition $\Psi = 0$ then implies that all three ψ_j are zero, i.e. all z_j are real.

When $\phi \neq 0$ it appears that Ω will become non-zero (see equation 3.9), invalidating the argument. However in our problem it can be shown that $B_1 + B_2 + B_3 = 0$ (see below), so that Ω is in fact $O(\epsilon^3)$; to balance (3.13), Ψ must then remain of order ϵ . Therefore in (3.10)–(3.12), we should consistently set $\cos \Psi = 1$ at this order; Ψ then decouples and the net result is that we are left with (3.6)–(3.8) again, except that now the imaginary parts of the μ_j can be ignored and one can assume that all the z_j (and μ_j) are real. (Therefore, we will drop the suffix ‘r’ from now on, writing μ_j in place of μ_{jr} .) This is equivalent to transforming to a moving frame in which the patterns appear steady.

The condition $B_1 + B_2 + B_3 = 0$ is in fact clear from a small- ϕ expansion (up to first order) of the linearized equations of the previous chapter (equations 2.15–2.23). Since ϕ now appears only in the combination $k_x \phi$, it can be seen that the equations are now invariant under rotations of the wavevector \mathbf{k} as long as one adjusts ϕ to ensure that $k_x \phi$ remains unchanged. Therefore, since we know that the oscillation frequency is proportional to ϕ at this order, this argument shows that it is also proportional to k_x (at fixed $|\mathbf{k}|$). Hence, since $k_{1x} + k_{2x} + k_{3x} = 0$, it follows that $B_1 + B_2 + B_3 = 0$.

We can now explain why we do not have to break the symmetry in the coefficients α , β and γ , i.e. why we do not have to write out separate values α_1 , α_2 and α_3 in the three equations rather than just the single value α (and similarly for β and γ). Clearly, all

three α_j must be equal when $\phi = 0$, and so we have $\alpha_j = \alpha + O(\phi)$; since $\phi = O(\epsilon)$, the corrections that this would produce would all appear at fourth order in ϵ (at least). Since we are dropping all terms of order ϵ^4 and higher, these corrections should be consistently neglected, and the symmetry should be maintained in α ; the same applies to β and γ . (Another way of putting this is that there exists a near-identity transformation that reduces the system to the above form.) Conversely, note that with these scalings the asymmetry in μ appears at order ϵ^3 , and the broken symmetry in μ must be included.

As before, we assume $\beta > 1$, and rescale time and the amplitudes to ensure that $\alpha = \gamma = 1$. By rescaling ϕ and r , we can also choose $A_3 = 1$ and $C = 1$.

We remark that once the isotropy is broken (for $\phi \neq 0$), the orientation of the hexagonal lattice, represented by the angle θ on Figure 3.4, is now important. Note that without loss of generality we can choose $0^\circ \leq \theta \leq 30^\circ$. This is for two reasons: firstly, the lattice itself is invariant under rotations by 60° (with a corresponding cyclic permutation of the \mathbf{k}_j and z_j); secondly, the $y \rightarrow -y$ reflection symmetry means that θ and $-\theta$ will be equivalent, which halves the range of angles that must be considered.

The angle $\theta = 30^\circ$ would probably be the most relevant choice, since the linear theory showed that parallel rolls (with the wavevector pointing in the y -direction) were the most unstable mode, and this mode is included (as \mathbf{k}_1) when $\theta = 30^\circ$. However, it is interesting to look at other values of θ as well, because in a sense, all modes would be present in the full problem (for example, if numerical simulations were to be performed).

This problem, in the form (3.10)–(3.12), has been studied before, although in a different context, by previous authors. Malomed et al. (1994) give bifurcation diagrams which use (μ_1, μ_2, μ_3) directly as bifurcation parameters; these are not directly related to our parameters r and ϕ . Nevertheless they provide several useful analytic results which we have made use of in this chapter. Matthews (1998) has produced bifurcation diagrams showing how the standard picture of Figure 3.5 changes once the symmetry is broken – we have reproduced these below, since they are directly equivalent to our problem in the case when ϕ is held fixed at a small non-zero value, and r is changed.

3.3.2 Changes to the bifurcation diagrams

Once the symmetry is broken, the bifurcation diagrams of Figure 3.5 undergo several changes. The first relates to the roll solutions. Previously, the three possible orientations of rolls all had the same properties, because they were related by symmetry; therefore

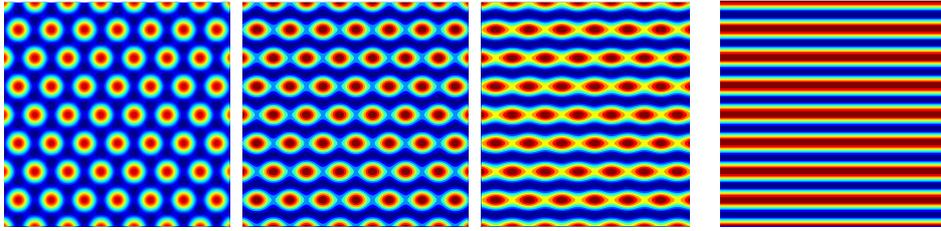


Figure 3.6: Distortion from a pure hexagon solution as the asymmetry is increased.

there was effectively only one branch of rolls. Now that the symmetry is broken, this splits into three separate branches, one for each possible orientation, and each having slightly different properties. We will denote the roll branch with z_j non-zero by R_j . The solution R_1 is stable if

$$2\beta\mu_1 - \mu_2 - \mu_3 > 0 \quad \text{and} \quad (\beta\mu_1 - \mu_2)(\beta\mu_1 - \mu_3) > \mu_1 \quad (3.14)$$

(Malomed et al., 1994).

The second change relates to the hexagon and rectangle (or mixed mode) branches. Previously we could make a clear distinction between the hexagons and the rectangles, in the sense that hexagons have $z_1 = z_2 = z_3$, while rectangles have e.g. $z_1 = z_2 \neq z_3$. Once the symmetry is broken, this distinction does not apply, and we can only really describe these branches as ‘mixed modes with all three z_j different’. However, since the symmetry breaking is weak, we will still be able to recognize the mixed mode branches as slightly distorted versions of the original patterns (either hexagons or rectangles). Therefore, we will continue using the name ‘hexagons’ (and the label H) for the stable ‘mixed mode’ branch, although it should be understood that these are not perfectly regular hexagons, but are instead slightly distorted in appearance.

This effect is illustrated in Figure 3.6, where we have plotted a pure hexagon solution ($|z_1| = |z_2| = |z_3|$) at the far left, a pure roll solution ($z_2 = z_3 = 0$) at the far right, and ‘distorted hexagon’ solutions (with $|z_1| > |z_2| = |z_3|$) in between. For very weakly broken symmetry (second picture), the solutions are only slightly distorted from the original hexagon pattern, but for more strongly broken symmetry (third picture) the pattern becomes much closer to the rolls in appearance.

We will now plot bifurcation diagrams showing how Figure 3.5 changes once the symmetry is broken. We do this in three different cases. The first (corresponding to the work of Matthews 1998) has ϕ fixed and r varying. The second is the reverse, with fixed r and variable ϕ . Finally, we plot diagrams for the general case where both r and ϕ

Solution	Condition for stability	ϕ_{pf}^2
R_1	(always)	$\frac{A_1}{(\beta A_1 - A_2)(\beta A_1 - 1)}$
R_2	$A_1 < \min(2\beta A_2 - 1, \beta A_2)$	$\frac{A_2}{(\beta A_2 - A_1)(\beta A_2 - 1)}$
R_3	$\beta > \max(A_1, A_2)$	$\frac{1}{(\beta - A_1)(\beta - A_2)}$

Table 3.1: Stability properties for rolls. There are two possibilities. If the condition in the second column is met then the rolls are stable for $\phi^2 > \phi_{\text{pf}}^2$ where ϕ_{pf}^2 is given in the third column. If the condition in the second column is not met then the rolls are never stable.

are allowed to vary, although for simplicity we have restricted this to the special values $\theta = 0^\circ$ and $\theta = 30^\circ$.

Bifurcation diagrams for fixed ϕ

The bifurcation diagrams given by Matthews (1998), reprinted in Figure 3.7, correspond to taking a fixed small value of ϕ and allowing r to vary.¹ Apart from the roll branch splitting into two or three separate branches, as mentioned above, the changes from the isotropic case are minimal: both regions of hysteresis still exist, and we still have the stable trivial solution for low r and stable hexagons for intermediate r . There is one other change, relating to the rolls: they are stable for large r as before, but they can also be stable in a (very narrow) range of r values near the initial bifurcation from the trivial solution. (This occurs for $\theta \neq 0^\circ$.)

Bifurcation diagrams for fixed r

We consider the case $r = 0$, which corresponds to setting the Rayleigh number equal to the critical Rayleigh number for a vertical field (so that the trivial solution is neutrally stable when $\phi = 0$). We then allow the tilt angle to increase above zero. As this is done, the critical Rayleigh number falls (because of the decreasing vertical component of the magnetic field, as was discussed in the previous chapter), and so the growth rates for the three modes become positive. (In other words, there should be a bifurcation at $\phi = 0$ at which the three roll modes appear.)

¹These figures are from Matthews (1998), ©1998 Elsevier, reproduced with permission.

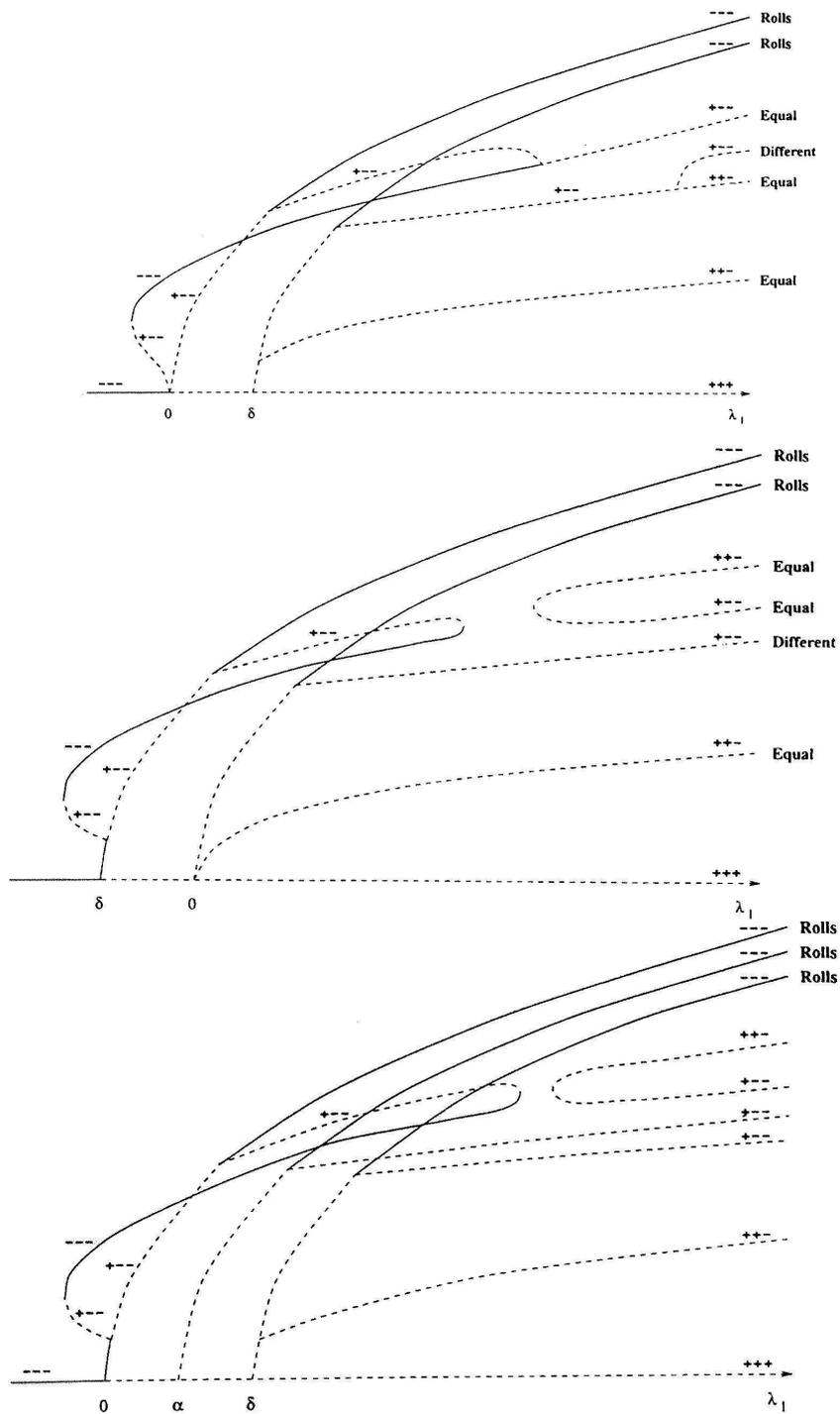


Figure 3.7: Bifurcation diagrams (from Matthews 1998) showing the changes to Figure 3.5 once the symmetry is broken. Note that Matthews' λ_1 corresponds to our r (with ϕ fixed at a small non-zero value). The three pictures correspond to different values of θ , as follows: Top picture: $\theta = 0^\circ$; middle picture: $\theta = 30^\circ$; bottom picture: $0^\circ < \theta < 30^\circ$.

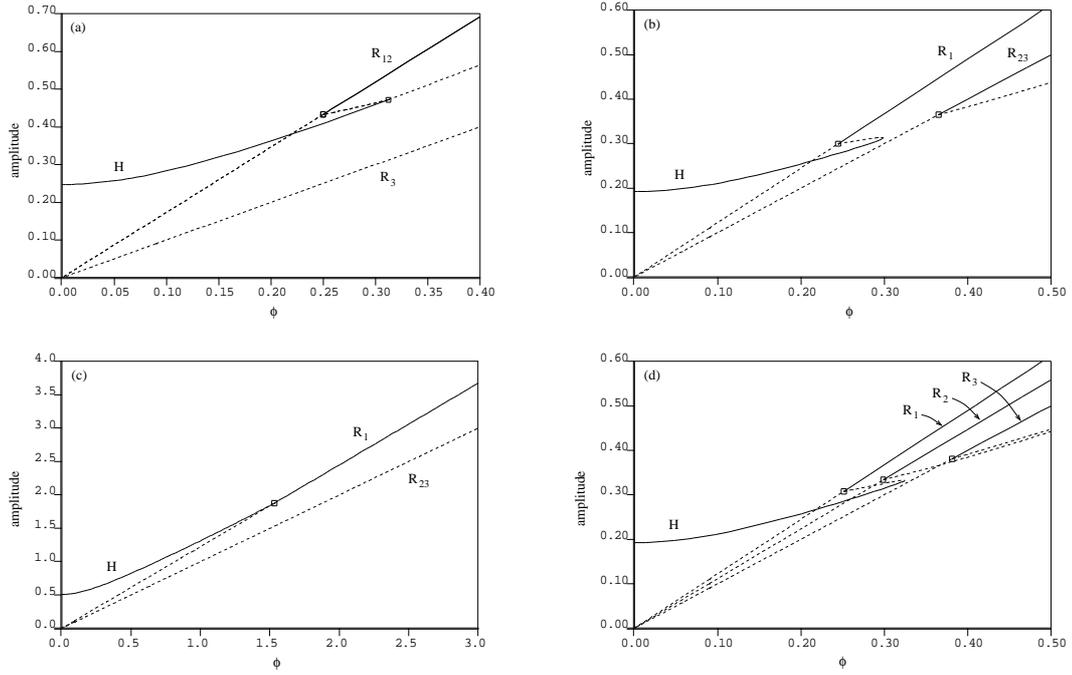


Figure 3.8: Bifurcation diagrams for $r = 0$ and variable ϕ . The amplitude $(|z_1|^2 + |z_2|^2 + |z_3|^2)^{1/2}$ is plotted against ϕ for (a) $\beta = A_1 = A_2 = 3$ (corresponding to $\theta = 0^\circ$); (b) $\beta = 4$, $A_1 = 1.5$, $A_2 = 1$ (corresponding to $\theta = 30^\circ$); (c) $\beta = 1.2$, $A_1 = 1.5$, $A_2 = 1$ (also corresponding to $\theta = 30^\circ$); (d) $\beta = 4$, $A_1 = 1.5$, $A_2 = 1.25$ (corresponding to $0^\circ < \theta < 30^\circ$). (Note that for clarity, some of the unstable solution branches have been omitted.)

Bifurcation diagrams for a number of values of β , A_1 and A_2 , produced with the program AUTO (Doedel, 1981), are shown in Figure 3.8. We see that the three roll modes do indeed bifurcate from $\phi = 0$, and that they are all unstable initially. It is possible for them to become stable for larger ϕ (at pitchfork bifurcations). The conditions for stable rolls are given in Table 3.1; these can be obtained from (3.14) using the equation $\mu_j = A_j\phi^2$. Note that R_1 rolls are always stable for ϕ large enough, while R_2 and R_3 may or may not become stable.

For the hexagons, there are two qualitatively different cases, depending on θ . The case $\theta = 0^\circ$ (corresponding to $A_1 = A_2$), illustrated in Figure 3.8(a), is special. In this case the hexagon branch joins onto an unstable branch of ‘mixed modes’, which then connects back to the R_{12} branch; there is always hysteresis between rolls and hexagons for this value of θ . The general case is $\theta \neq 0$ (or $A_1 \neq A_2$), corresponding to Figures 3.8(b)–3.8(d). Here the hexagons always connect to the R_1 branch at the point where the latter becomes stable (this is the pitchfork bifurcation $\phi = \phi_{\text{pf}}$ in the notation of Table 3.1). This pitchfork can be either supercritical (Figure 3.8c) or subcritical (Figures 3.8b and 3.8d). In the supercritical case, there is no possibility of hysteresis, but in the subcritical case, the hexagon branch turns around at a saddle-node bifurcation, and there is hysteresis between rolls and hexagons. We have not been able to find analytically the position of this saddle-node bifurcation, and so we cannot give a formula for the ‘amount’ of hysteresis, i.e. the width of the interval in ϕ for which there is bistability, although we find that this interval tends to be widest when β is small (close to 1), or when the three A_j are close together in value.

In the specific case of $\theta = 30^\circ$ (where $A_2 = 1$), we can determine analytically whether the bifurcation is supercritical or subcritical. First compute A_c by the following formula:

$$\frac{1}{A_c} = \beta + \frac{2(1 + \beta - 2\beta^2)}{2\beta + \sqrt{2\beta + 2}}. \quad (3.15)$$

If A_c is negative, then the bifurcation is subcritical. If A_c is positive, then the bifurcation is supercritical for $A_1 > A_c$, or subcritical for $A_1 < A_c$. (The derivation of this condition is given in Appendix C.)

Bifurcation diagrams varying both r and ϕ

Finally we consider the general case where both r and ϕ are allowed to vary simultaneously. For simplicity we restrict our results to the cases $\theta = 0^\circ$ and $\theta = 30^\circ$. In such cases two of the μ_j will be equal; for example, we might have $\mu_1 = \mu_2 = \mu$, so that the

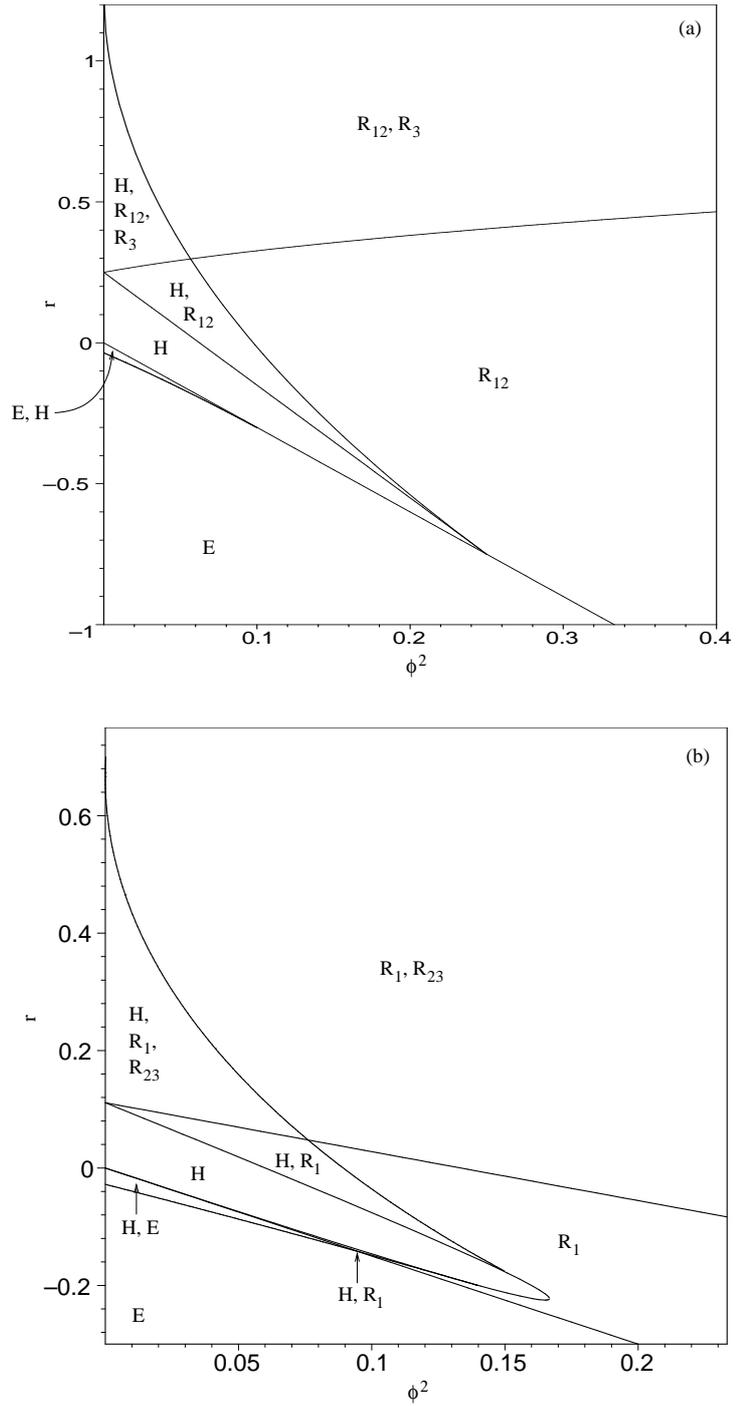


Figure 3.9: Diagrams showing the stability regions of rolls (R), hexagons (H) and the trivial (equilibrium) solution (E) as functions of r and ϕ , for (a) $\beta = A_1 = A_2 = 3$, corresponding to $\theta = 0^\circ$; (b) $\beta = 4$, $A_1 = 1.5$, $A_2 = 1$, corresponding to $\theta = 30^\circ$.

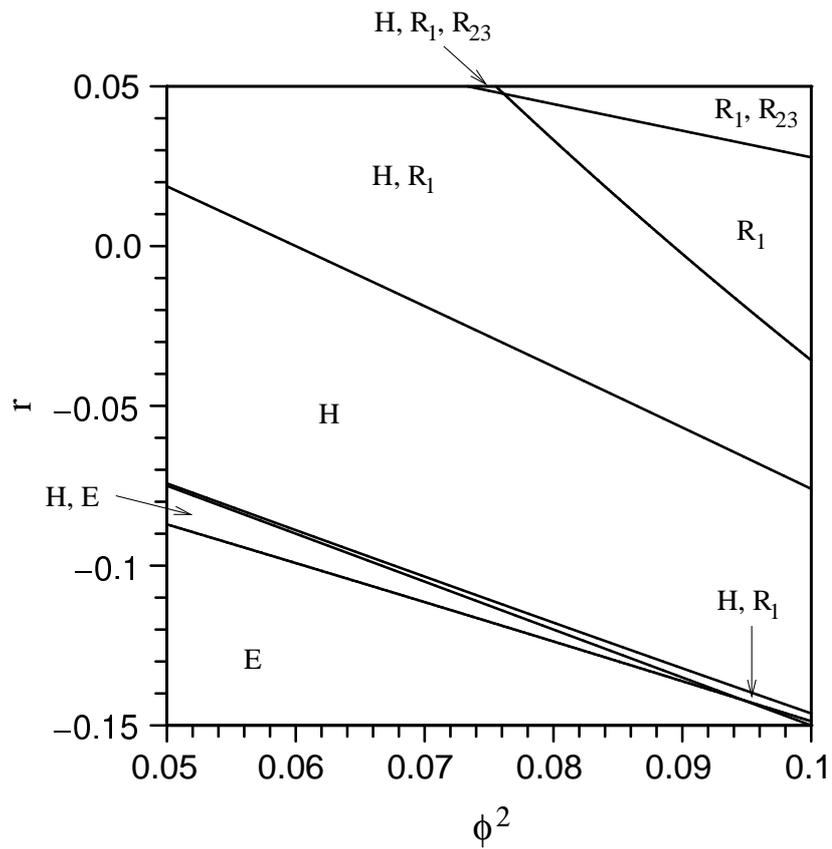


Figure 3.10: Close-up of part of Figure 3.9(b). Note the small region ‘H, R₁’ near the bottom-right; this in fact extends upwards and leftwards all the way to the r -axis, although it does become extremely narrow as $\phi \rightarrow 0$.

two bifurcation parameters are μ and μ_3 . These can be related to our parameters r and ϕ by a simple change of variables. In such cases, Malomed et al. (1994) give plots of the stability regions in the μ - μ_3 plane. We can easily convert their results to the new coordinate system defined by r and ϕ and thus produce diagrams of the stability regions of the various solutions in terms of r and ϕ . See Figure 3.9, and also Figure 3.10, which is an enlargement of part of Figure 3.9(b).

The basic features of the diagrams are essentially independent of the parameters; although the sizes and positions of the stability regions can change, the topology seems to remain unaltered. We see rolls for large r or large ϕ , with a region of stable hexagons for low ϕ and intermediate r . In addition there is typically an area somewhere in parameter space of bistability between rolls and hexagons, indicating that hysteresis is a generic possibility, at least for some paths through parameter space.

Note that the bifurcation diagrams shown previously can be related to the graphs of Figure 3.9. For example, the diagrams of section 3.3.2, corresponding to the work of Matthews (1998), correspond on Figure 3.9 to moving upwards along the line $\phi = \eta$ (where η is a small positive constant), a vertical line slightly to the right of the r -axis. Moving along this path on Figure 3.9(a), which corresponds to $\theta = 0^\circ$, would give the bifurcation diagram shown in Figure 3.7 (top panel). For Figure 3.9(b), corresponding to $\theta = 30^\circ$, the middle panel of Figure 3.7 would result. Note that in both cases the sequences of stable and unstable patterns shown in our diagram (Figure 3.9) agree with the work of Matthews (1998) (Figure 3.7). In the latter case, this is perhaps more difficult to see because of the small region near onset where rolls can be stable, visible on Figure 3.7 (middle panel) near $\lambda_1 = \delta$. On our diagram (Figure 3.9b) this in fact corresponds to the narrow wedge marked ‘ H, R_1 ’ which actually extends leftwards all the way to the ϕ -axis; this can hopefully be seen more clearly on the enlargement shown in Figure 3.10.

The bifurcation diagrams of section 3.3.2, in which $r = 0$ and ϕ was varied, of course correspond to moving along the horizontal line $r = 0$ on Figure 3.9. On Figure 3.9(a), this line moves from the region ‘ H ’, to ‘ H, R_{12} ’, and finally to ‘ R_{12} ’; therefore, there is hysteresis between rolls and hexagons. Indeed there will always be such hysteresis, no matter what path one takes through parameter space, because the only way to move from region ‘ H ’ to region ‘ R_{12} ’ is via the region of bistability (except for the non-generic case in which one moves through the codimension-2 point at the very bottom right of the ‘ H ’ region).

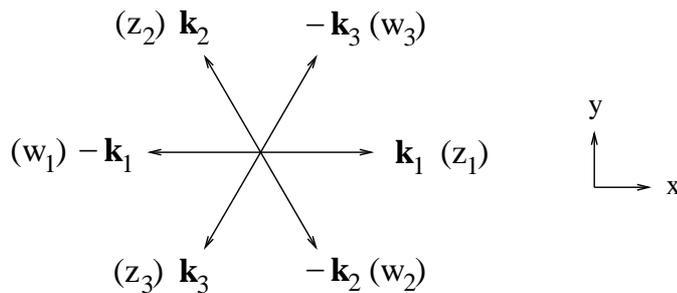


Figure 3.11: The wavevectors (\mathbf{k}_j) and corresponding complex amplitudes (z_j and w_j) used for the oscillatory hexagonal model. The arrows represent the direction of travel of each mode. As before, the direction of tilt of the magnetic field will be in the x -direction.

By contrast, on Figure 3.9(b), there are two distinct routes from the tongue-shaped ‘ H ’ region to the ‘ R_1 ’ region. One involves going through the ‘side’ of the tongue, up and into the ‘ H, R_1 ’ region, and then into the ‘ R_1 ’ region. The other route goes through the ‘tip’ of the tongue at the far right-hand side, and thus goes directly into the ‘ R_1 ’ region without any hysteresis or bistability in between. Now, with the parameter values as chosen in Figure 3.9(b), the line $r = 0$ takes the former route, going via a region of bistability before finally reaching the region where only rolls are stable. However, if different parameters are chosen, then the regions can shift around in such a way that the line $r = 0$ moves through the ‘tip’ of the tongue rather than its ‘side’; this would correspond to the supercritical case mentioned in section 3.3.2, and there would be no hysteresis in this case. However, note that the region of bistability still exists – it has just been moved up to a higher value of r . Therefore, hysteresis could still be seen for these values of A_1 and β , but one would have to move on a different path through parameter space, e.g. by varying ϕ and fixing r , not at zero, but at some appropriate positive value; or indeed by fixing ϕ , at an appropriate value, and varying r .

3.4 Oscillatory bifurcation on a hexagonal lattice

For small ζ (and sufficiently large Q), the initial bifurcation to convection is oscillatory rather than steady. This means that the representation (3.2) must be used instead of (3.1) for our solutions (with $N = 3$ for the hexagonal lattice). The six modes, with complex amplitudes z_j and w_j ($j = 1, 2, 3$), are shown in Figure 3.11.

Using the same approach as before, we can determine the most general set of am-

plitude equations for these six modes that are consistent with the symmetries of the problem. In the case of a vertical field, the relevant symmetry group is $(T^2 \times D_6) \times S^1$. (T^2 and D_6 have the same meanings as before; the extra symmetry S^1 corresponds to a time shift symmetry, which essentially comes about because t now explicitly appears in the representation (3.2)). The amplitude equations consistent with this symmetry group are (Roberts et al., 1986):

$$\dot{z}_1 = [\mu + a|z_1|^2 + b|w_1|^2 + c(|z_2|^2 + |z_3|^2) + d(|w_2|^2 + |w_3|^2)]z_1 + f(z_2w_2 + z_3w_3)\bar{w}_1 \quad (3.16)$$

$$\dot{z}_2 = [\mu + a|z_2|^2 + b|w_2|^2 + c(|z_3|^2 + |z_1|^2) + d(|w_3|^2 + |w_1|^2)]z_2 + f(z_3w_3 + z_1w_1)\bar{w}_2 \quad (3.17)$$

$$\dot{z}_3 = [\mu + a|z_3|^2 + b|w_3|^2 + c(|z_1|^2 + |z_2|^2) + d(|w_1|^2 + |w_2|^2)]z_3 + f(z_1w_1 + z_2w_2)\bar{w}_3 \quad (3.18)$$

$$\dot{w}_1 = [\mu + a|w_1|^2 + b|z_1|^2 + c(|w_2|^2 + |w_3|^2) + d(|z_2|^2 + |z_3|^2)]w_1 + f(z_2w_2 + z_3w_3)\bar{z}_1 \quad (3.19)$$

$$\dot{w}_2 = [\mu + a|w_2|^2 + b|z_2|^2 + c(|w_3|^2 + |w_1|^2) + d(|z_3|^2 + |z_1|^2)]w_2 + f(z_3w_3 + z_1w_1)\bar{z}_2 \quad (3.20)$$

$$\dot{w}_3 = [\mu + a|w_3|^2 + b|z_3|^2 + c(|w_1|^2 + |w_2|^2) + d(|z_1|^2 + |z_2|^2)]w_3 + f(z_1w_1 + z_2w_2)\bar{z}_3 \quad (3.21)$$

Here, μ is the (complex) linear growth rate, and a , b , c , d and f are complex constants. Note that there are no quadratic terms in the oscillatory case, even when the up-down symmetry is broken. (This is a consequence of the extra time shift symmetry.)

Roberts et al. (1986) found eleven branches of solutions to these equations, which are listed in Table 3.2. Note that not all of these solutions will be found to be stable in our model (more on this below).

Figures 3.12 and 3.13 show graphical depictions of some of the solutions. These plots are obtained from the six amplitudes by using equation (3.2). Note that the solutions are functions of time: for the solutions shown in Figure 3.12, the only time-dependence is that the pattern travels with a constant speed, while for the patterns in Figure 3.13, there is a more complicated time-dependence and we have plotted snapshots at various time points. The time is expressed in units such that the period of the solution is 1. Only the first half of this period is plotted; the second half (from $t = 1/2$ to $t = 1$) is the same as the first half (from $t = 0$ to $t = 1/2$), but with a sign change.

<i>Solution</i>	<i>Label</i>	<i>Amplitudes</i>
oscillating triangles	OT	$(r, r, r, 0, 0, 0)$
standing hexagons	SHe	(r, r, r, r, r, r)
standing rectangles	SRe	$(0, r, r, 0, r, r)$
standing regular triangles	SRT	$(r, r, r, -r, -r, -r)$
standing rolls	SRo	$(r, 0, 0, r, 0, 0)$
travelling rectangles (1)	TRe1	$(r, 0, r, 0, 0, 0)$
travelling rectangles (2)	TRe2	$(r, 0, 0, 0, 0, r)$
travelling rolls	TRo	$(r, 0, 0, 0, 0, 0)$
twisted rectangles	TwRe	$(r, re^{2\pi i/3}, re^{4\pi i/3}, r, re^{2\pi i/3}, re^{4\pi i/3})$
wavy rolls (1)	WR1	$(r, 0, r, r, 0, -r)$
wavy rolls (2)	WR2	$(r, re^{2\pi i/3}, re^{4\pi i/3}, -r, -re^{2\pi i/3}, -re^{4\pi i/3})$

Table 3.2: The solutions found by Roberts et al. (1986) to equations (3.16)–(3.21), together with typical values for the amplitudes $(z_1, z_2, z_3, w_1, w_2, w_3)$ (r represents a real number).

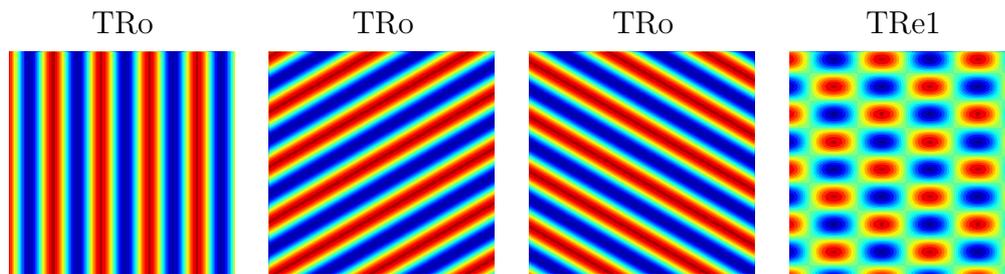


Figure 3.12: Solutions to equations (3.16)–(3.21). The TRo travel in a direction perpendicular to the roll axes; the TRe1 branch shown travels in the x -direction.

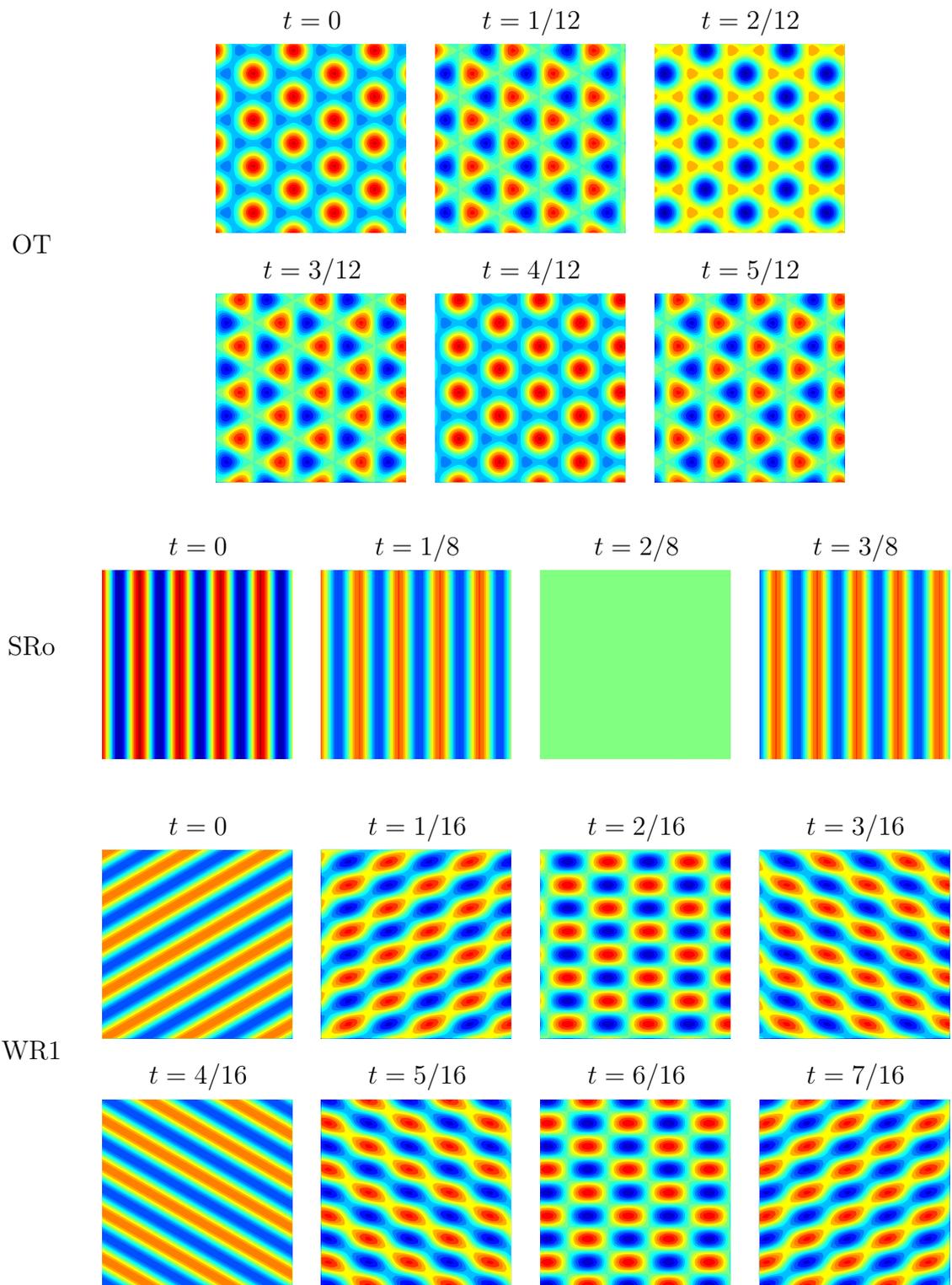


Figure 3.13: Solutions to equations (3.16)–(3.21).

3.4.1 An additional solution branch

In finding the above solutions, Roberts et al. (1986) used the equivariant Hopf theorem, which is an effective method, although it has the drawback that it is not guaranteed to find all possible solutions. We have taken a more elementary approach, simply by inspecting the equations in various cases (the full details are given in Appendix C). Our method still does not find all possible solutions, but we do find the eleven solution branches found by Roberts et al. (1986), plus one additional solution. This new solution does not have a one-complex-dimensional fixed point subspace, which is why it was not encountered by Roberts et al. (1986).

This new solution has the following form:

$$z_1 = z_2 \neq 0; \quad w_3 \neq 0; \quad z_3 = w_1 = w_2 = 0 \quad (3.22)$$

$$|z_1|^2 = \frac{\mu_r(d_r - a_r)}{a_r^2 - 2d_r^2 + a_r c_r} \quad (3.23)$$

$$|w_3|^2 = \frac{\mu_r(2d_r - a_r - c_r)}{a_r^2 - 2d_r^2 + a_r c_r} \quad (3.24)$$

$$\frac{d}{dt}(\arg z_1) = \mu_i + (a_i + c_i)|z_1|^2 + d_i|w_3|^2 \quad (3.25)$$

$$\frac{d}{dt}(\arg w_3) = \mu_i + a_i|w_3|^2 + 2d_i|z_1|^2. \quad (3.26)$$

These equations, together with stability criteria for the solution, are derived in Appendix C. Unfortunately it is not possible to write down the stability conditions in a simple form; however, it is a straightforward matter to compute the stability eigenvalues numerically. We have found that the solution is often unstable, but it is stable for (at least) the following choice of parameters: $a = -1$, $b = -1.5$, $c = -0.2$, $d = -0.3$ and $f = -0.1$ (for which $|z_1| = |z_2| = 0.828\sqrt{\mu_r}$ and $|w_3| = 0.767\sqrt{\mu_r}$). We have confirmed this result by numerically solving (3.16)–(3.21) as an initial value problem; one such run is shown in Figure 3.14, which demonstrates the existence and stability of the solution in this case.

The solution in this case is shown in Figure 3.15. Note that it resembles the oscillating triangle solution, albeit with a slight asymmetry (since $|z_1|$ and $|w_3|$ take slightly different values). The difference is that as well as oscillating between regular hexagons and regular triangles, the new solution also appears to drift (upwards and rightwards in the case shown in Figure 3.15). The other main difference is that the OT solution oscillates between hexagons and triangles much more quickly (in fact, three times more quickly)

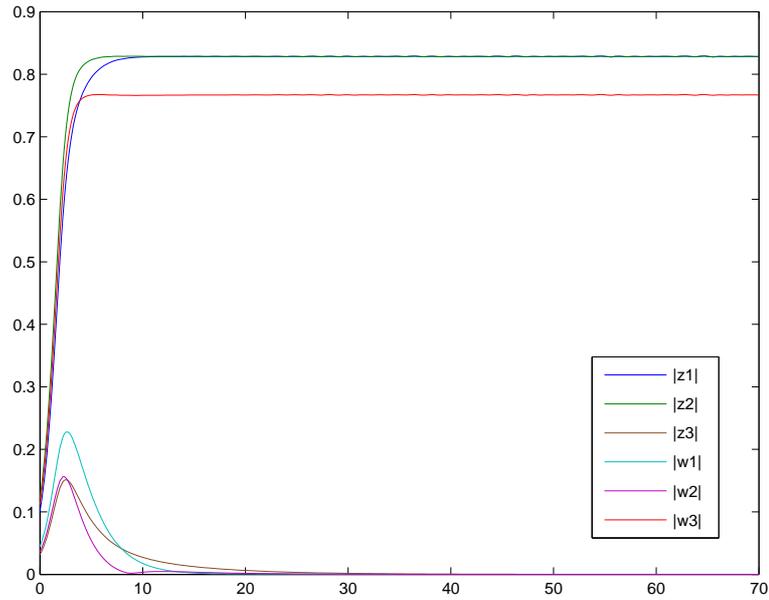


Figure 3.14: Numerical solution of (3.16)–(3.21), illustrating the existence and stability of the new solution branch. The six amplitudes $|z_1|$, $|z_2|$, $|z_3|$, $|w_1|$, $|w_2|$ and $|w_3|$ are plotted against time. (The parameters were: $a = -1$, $b = -1.5$, $c = -0.2$, $d = -0.3$ and $f = -0.1$, with $\mu = 1$.)

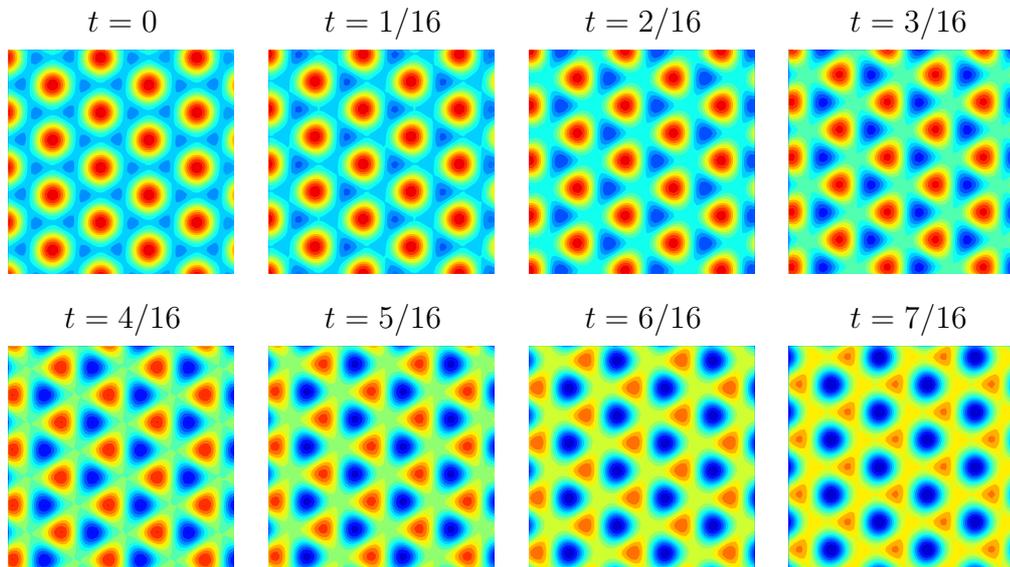


Figure 3.15: The new ‘drifting oscillating triangles’ solution.

than this new solution does, as can be seen by comparing the relevant figures. (In terms of symmetries, the oscillating triangles have a spatiotemporal symmetry consisting of a time shift of $1/3$ of the total period together with a translation; the new ‘drifting oscillating triangles’ actually have no spatiotemporal symmetries.)

We have not considered what happens to this solution when the equations are truncated at higher order than third order. A version of equations (3.27)–(3.32) truncated at fifth order, instead of third, is given by Roberts et al. (1986) (see their Proposition 2.2). These equations contain non-zero terms at fifth order in the equations for \dot{z}_3 , \dot{w}_1 and \dot{w}_2 (these are the three variables that are zero in our third-order solution). Therefore, the solution at higher order will have small (rather than identically zero) values for z_3 , w_1 and w_2 ; this might cause the solution to have slightly different properties (although we have not investigated this in detail).

The reason that Roberts et al. (1986) did not find this new solution branch is that it does not have maximal isotropy. In fact, the only symmetry possessed by the new solution is the reflection symmetry that exchanges z_1 and z_2 , and w_1 and w_2 . None of the other elements of D_6 are symmetries, nor are there any spatiotemporal symmetries. Therefore the isotropy subgroup of this solution is just Z_2 . The corresponding fixed point set is $(z_1, z_2, z_3, w_1, w_2, w_3) = (p, p, q, r, r, s)$ (where p , q , r and s are arbitrary complex numbers) which is 4-D complex; a 1-D complex fixed point set would have been needed for the solution to be detectable via the usual methods (the equivariant Hopf theorem).

3.4.2 Weakly broken symmetry

We would now like to investigate non-vertical magnetic fields, and we (once again) do this by weakly breaking the isotropy, introducing a small tilt angle ϕ . As before, we

break the symmetry in the linear terms only, leading to:

$$\dot{z}_1 = [\mu_1 + a|z_1|^2 + b|w_1|^2 + c(|z_2|^2 + |z_3|^2) + d(|w_2|^2 + |w_3|^2)]z_1 + f(z_2w_2 + z_3w_3)\bar{w}_1 \quad (3.27)$$

$$\dot{z}_2 = [\mu_2 + a|z_2|^2 + b|w_2|^2 + c(|z_3|^2 + |z_1|^2) + d(|w_3|^2 + |w_1|^2)]z_2 + f(z_3w_3 + z_1w_1)\bar{w}_2 \quad (3.28)$$

$$\dot{z}_3 = [\mu_3 + a|z_3|^2 + b|w_3|^2 + c(|z_1|^2 + |z_2|^2) + d(|w_1|^2 + |w_2|^2)]z_3 + f(z_1w_1 + z_2w_2)\bar{w}_3 \quad (3.29)$$

$$\dot{w}_1 = [\mu'_1 + a|w_1|^2 + b|z_1|^2 + c(|w_2|^2 + |w_3|^2) + d(|z_2|^2 + |z_3|^2)]w_1 + f(z_2w_2 + z_3w_3)\bar{z}_1 \quad (3.30)$$

$$\dot{w}_2 = [\mu'_2 + a|w_2|^2 + b|z_2|^2 + c(|w_3|^2 + |w_1|^2) + d(|z_3|^2 + |z_1|^2)]w_2 + f(z_3w_3 + z_1w_1)\bar{z}_2 \quad (3.31)$$

$$\dot{w}_3 = [\mu'_3 + a|w_3|^2 + b|z_3|^2 + c(|w_1|^2 + |w_2|^2) + d(|z_1|^2 + |z_2|^2)]w_3 + f(z_1w_1 + z_2w_2)\bar{z}_3 \quad (3.32)$$

Recall that once ϕ becomes non-zero, the orientation of the lattice is important. For simplicity, we have restricted ourselves to one particular orientation, as illustrated in Figure 3.11. We have picked \mathbf{k}_1 to lie on the x -axis; this corresponds to perpendicular rolls, which are the most unstable mode for small ϕ (as shown in the previous chapter), so we are including the most unstable mode plus five other modes. We then have the following form for the growth rates for small r and small ϕ :

$$\mu_1 = \gamma r - \alpha \phi \quad (3.33)$$

$$\mu_2 = \mu_3 = \gamma r + \beta \phi \quad (3.34)$$

$$\mu'_1 = \gamma r + \alpha \phi \quad (3.35)$$

$$\mu'_2 = \mu'_3 = \gamma r - \beta \phi \quad (3.36)$$

where α , β and γ are complex constants. Here we have scaled both r and ϕ as $O(\epsilon)$ and expanded the growth rates to first order in ϵ . All the terms in equations (3.27)–(3.32) are now of order ϵ^3 .

Before proceeding, we briefly discuss the effects that the symmetry-breaking has on the solutions. As in the steady case, once the symmetry is broken we find that many of the solution branches ‘split’ into two or more separate branches, that previously were related by symmetry, but now are genuinely different solutions. These are summarized in Table 3.3. For example, there are now four different branches of travelling rolls (labelled

<i>Solution</i>	<i>Label</i>	<i>Amplitudes</i>
travelling rolls	TRo(a)	$(0, 0, 0, x, 0, 0)$
	TRo(b)	$(0, x, 0, 0, 0, 0)$ or $(0, 0, x, 0, 0, 0)$
	TRo(c)	$(0, 0, 0, 0, x, 0)$ or $(0, 0, 0, 0, 0, x)$
	TRo(d)	$(x, 0, 0, 0, 0, 0)$
standing rolls	SRo(a)	$(x, 0, 0, y, 0, 0)$
	SRo(b)	$(0, x, 0, 0, y, 0)$ or $(0, 0, x, 0, 0, y)$
travelling rectangles (1)	TRe1(a)	$(0, x, x, 0, 0, 0)$
oscillating triangles	OT(a)	$(0, 0, 0, x, y, y)$
	OT(b)	$(x, y, y, 0, 0, 0)$
wavy rolls (1)	WR1(a)	$(0, x, x, 0, y, y)$

Table 3.3: *The solutions found to equations (3.27)–(3.32) and typical forms of the amplitudes $(|z_1|, |z_2|, |z_3|, |w_1|, |w_2|, |w_3|)$ in each case. (Here, x and y represent real numbers. More details on the exact form of each solution branch can be found in Appendix C.) Only solutions that are found to be stable (in the analysis below) have been shown. Note that there are multiple forms of some of the solution branches, because of the asymmetry; these different forms have been indicated by lower-case letters, e.g. TRo(a–d).*

‘a’-‘d’), instead of just one. (There are not six, one for each z_j and each w_j , because we still have the $y \rightarrow -y$ reflection symmetry; for example, ‘ z_2 -rolls’ and ‘ z_3 -rolls’ are still equivalent, but they are different to ‘ z_1 -rolls’.)

Figure 3.16 shows graphically how the symmetry breaking changes the appearance of some of the solution branches (compare this to Figure 3.13). Here we have broken the symmetry by taking $|z_j|$ and $|w_j|$ each to be slightly different, as follows: $|w_1| = 1.6$, $|z_2| = |z_3| = 1.3$, $|w_2| = |w_3| = 1.0$, and $|z_1| = 0.7$. (These values are to be substituted in place of x and y in Table 3.3.) The values have been chosen arbitrarily; they are intended purely to illustrate how the solutions change when the amplitudes are no longer all equal.

For SRO, the asymmetry means that the solution does not return to the ‘zero’ state at $t = 2/8$ as it did before. For OT(a), the solutions start to look a little more like a perpendicular roll pattern (because of the dominance of that particular roll mode). For WR1, the solution no longer reaches a ‘pure’ roll state at each end of the oscillation. Note also that the asymmetry has no effect on the TRO or TRe1 solutions.

Returning to our amplitude equations, we can rescale ϕ and r to ensure that $\alpha_r = \gamma_r = 1$, leaving β_r , α_i , β_i and γ_i as free parameters. We can also rescale time and the amplitudes to remove two degrees of freedom from the choices of a , b , c , d and f . Thus there are effectively 12 undetermined coefficients in the problem as well as the parameters r and ϕ .

The coefficients a , b , c , d and f present a problem, since they represent five complex parameters whose values cannot be determined from symmetry arguments alone. However, Clune and Knobloch (1994) have performed a perturbation analysis of the equations for Boussinesq magnetoconvection, using a vertical field and illustrative boundary conditions (magnetic field constrained to be vertical at top and bottom), and found the values of these coefficients for a wide range of Q , ζ and σ values.

Therefore, our approach will be to look at the patterns found by Clune and Knobloch (1994) to occur in magnetoconvection, and see how these change as one introduces a slightly tilted field. The most common pattern that they find is OT, and there are large regions of parameter space where this is the only stable solution. There are also fairly large regions where TRO are the only stable solution, and where OT and SRO are simultaneously stable.

Now suppose, for example, that we are in a situation where OT are the only stable pattern when $\phi = 0$, and we want to know what happens as ϕ is increased above zero. To answer this we need to choose a set of values for a , b , c etc. that would lead to

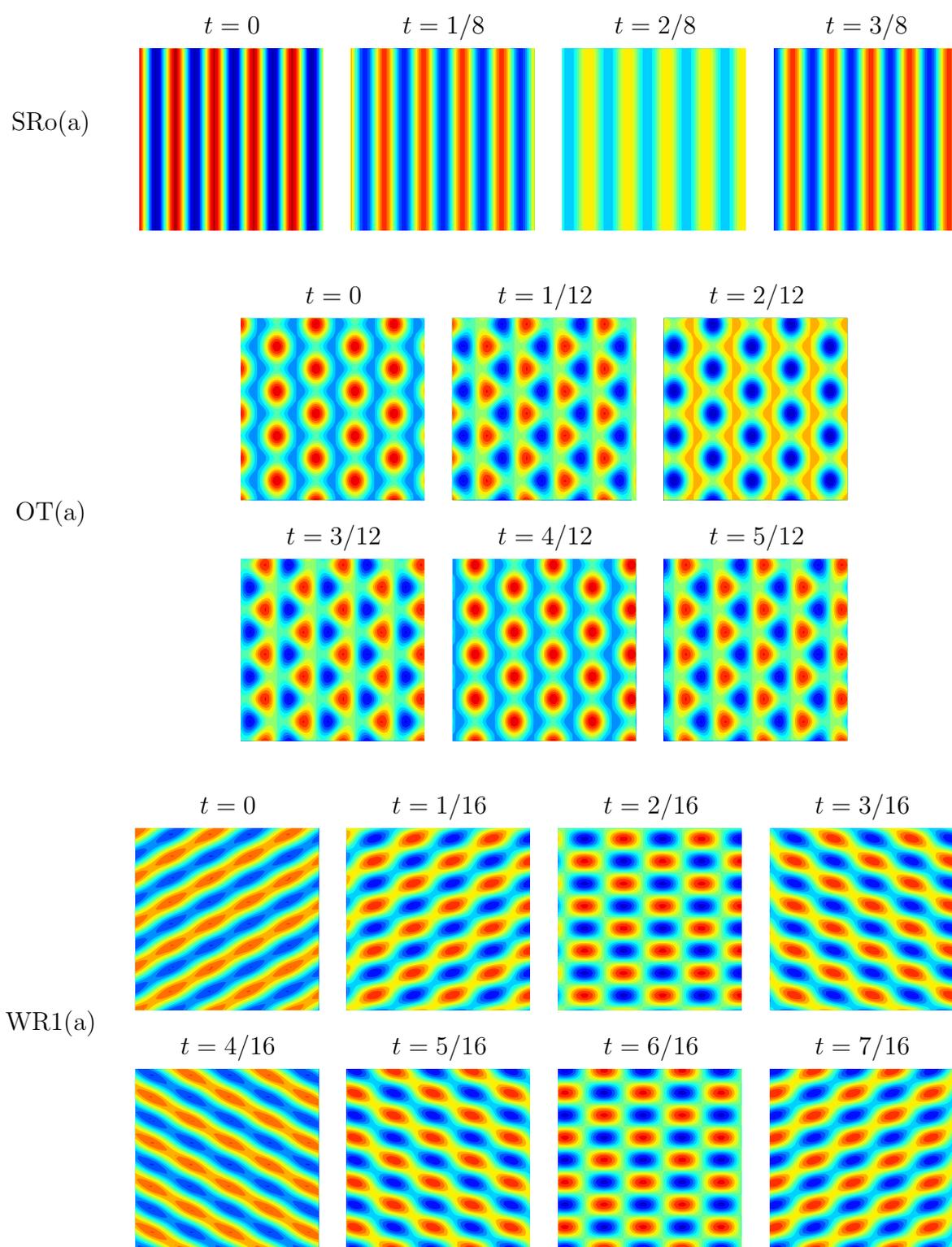


Figure 3.16: Solutions to equations (3.27)–(3.32).

<i>Case</i>	<i>a</i>	<i>b</i>	<i>c</i>	<i>d</i>	<i>f</i>	<i>Stable solutions</i>
						<i>in symmetric case</i>
(a)	$-1.5 + 2i$	$-3 + 2.2i$	$-3 + 0.5i$	$-3 + 0.8i$	$-1 - 2i$	TRo
(b)	-10	-9.6	-1	-9.5	-4.2	OT
(c)	$-3 - 1.1i$	$-1.5 + 0.3i$	$-1.7 + 0.7i$	$-4 + i$	$-1.8 + i$	SRo & OT

Table 3.4: *The values chosen for the nonlinear coefficients a , b , c , d and f .*

such a situation. Unfortunately, this choice is far from unique; there are many such sets of values that would give OT as the only stable pattern when ϕ is zero. We can only repeat our analysis for several different sets of parameter values, each having OT stable at $\phi = 0$, and try to determine which features of our results are robust. This can of course be done for other situations as well – we will also look at the cases where TRo are the only stable solution for $\phi = 0$, and where SRo and OT are the only stable solutions for $\phi = 0$.

One must also choose values for α , β and γ . We have chosen to investigate the particular values $\beta_r = 0.65$, $\alpha_i = 0.5$, $\beta_i = 0.7$, $\gamma_i = 0$ in detail. These have been chosen arbitrarily, but they illustrate the typical behaviour; the effects of changing these values will be discussed (briefly) below.

The results, showing the existence and stability regions for the various patterns, are shown in Figure 3.17. The values of the coefficients a , b , c , d and f that were used in each case are given in Table 3.4. The boundaries of the existence and stability regions are always radial lines through the origin. It can be seen that for small ϕ (or large r), the stable patterns are the same as in the symmetric problem (as required), but for larger tilt angles new patterns can be made to appear. Generally speaking, one finds that TRo(a) usually become stable for the larger values of ϕ . Note that TRo(a) corresponds to perpendicular rolls, which are the preferred mode in linear theory for ϕ below the Lifshitz point. (The model in this section is valid only for small tilt angles, as has already been discussed. Therefore we never get into the regime where either oblique or parallel rolls become preferred in linear theory.) We also find that TRe1(a) are often (but not always) stable for large ϕ , a situation illustrated by cases (b) and (c). In addition, when SRo and OT are both stable for small ϕ , illustrated by case (c), wavy rolls of the first kind (WR1) can be stable for some regions of parameter space.

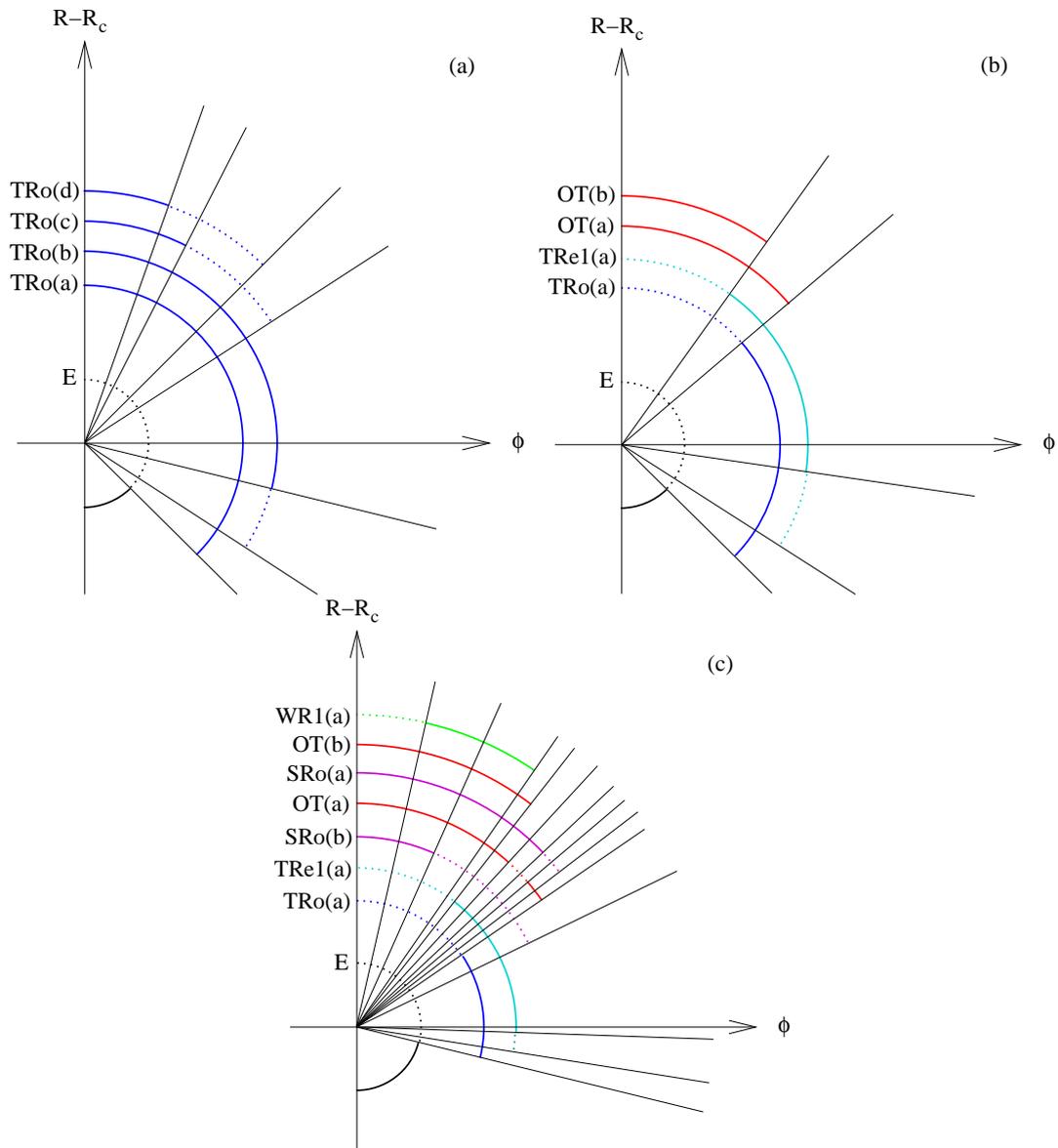


Figure 3.17: Results for the oscillatory hexagonal model. The coloured arcs represent the different solution branches; solid lines indicate a region where there is a stable solution, and dashed lines indicate an unstable solution. (The radial lines represent the bifurcations.) See text for further explanation.

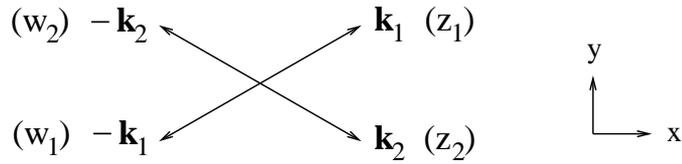


Figure 3.18: *The four modes included in the rhombic lattice (showing the wavevectors and corresponding complex amplitudes). Note that the magnetic field is tilted in the x -direction (as usual).*

Figures 3.17(a–c) were each produced with particular choices for the coefficients. We have, as explained above, also tried different values of the various parameters in order to check the robustness of the results. In cases where TRo are the only stable pattern at $\phi = 0$, it appears that the qualitative picture of Figure 3.17(a) is always found, whatever the particular parameter values chosen. In other cases, there can be some changes. The set of possible stable patterns always seems to remain the same, with the exception of TRel(a), which was found to be stable only in certain cases. However, the exact position of the bifurcation lines, and indeed the order in which those bifurcations are encountered as one moves around the origin, can vary depending on the precise parameter values chosen.

3.5 Oscillatory bifurcation on a rhombic lattice

Our final piece of work in this chapter will be to consider the oscillatory problem on a rhombic lattice instead of a hexagonal one. See Figure 3.18. The motivation for this choice of lattice comes from the linear theory, and in particular to the regime where oblique rolls are preferred. A rhombic lattice can include the two preferred modes (represented by z_1 and z_2) as well as their counterparts travelling in the opposite directions (represented by w_1 and w_2).

3.5.1 The symmetric case

The problem of oscillatory (Hopf) bifurcation on a rhombic lattice has been considered before, by Silber et al. (1992). Their work was motivated by studies of nematic liquid crystals, but it is not restricted only to liquid crystals; rather, it can be applied to any system possessing the same symmetries. (The problem has the symmetry group $T^2 \rtimes D_2$, which consists of x and y translations plus the discrete symmetries of the

rhombic lattice.)

Silber et al. (1992) obtained the following equations, truncated at third order, for the evolution of the amplitudes z_1 , z_2 , w_1 and w_2 :

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_2|^2 + c|w_1|^2 + d|z_2|^2)z_1 + fw_2z_2\bar{w}_1 \quad (3.37)$$

$$\dot{z}_2 = \mu z_2 + (a|z_2|^2 + b|w_1|^2 + c|w_2|^2 + d|z_1|^2)z_2 + fw_1z_1\bar{w}_2 \quad (3.38)$$

$$\dot{w}_1 = \mu w_1 + (a|w_1|^2 + b|z_2|^2 + c|z_1|^2 + d|w_2|^2)w_1 + fw_2z_2\bar{z}_1 \quad (3.39)$$

$$\dot{w}_2 = \mu w_2 + (a|w_2|^2 + b|z_1|^2 + c|z_2|^2 + d|w_1|^2)w_2 + fw_1z_1\bar{z}_2 \quad (3.40)$$

The coefficient μ is the (complex) linear growth rate. The coefficients a , b , c , d and f are also complex, and depend on the physical properties of the system; their values cannot be determined by symmetry arguments alone.

These five undetermined complex coefficients are problematic, since they essentially represent ten free parameters in the model. Fortunately, however, there is a limit in which some of them can be eliminated; this occurs when the angle of obliquity of the oblique rolls is small, i.e. the angle of the wavevectors \mathbf{k}_1 and \mathbf{k}_2 to the x -axis is small. This situation occurs close to the so-called Lifshitz point, the point where oblique rolls first appear (see previous chapter). In such circumstances an oblique roll can be represented as a slow modulation of a perpendicular roll.

Silber et al. (1992) performed a calculation using exactly this method, writing a general superposition of leftward- and rightward-travelling rolls as

$$\epsilon^{1/2}[A(X, Y, T)e^{i(kx-\omega t)} + B(X, Y, T)e^{i(-kx-\omega t)}] + \text{c.c.} \quad (3.41)$$

(to lowest order in ϵ), where the scalings are $X = \epsilon^{1/2}x$, $Y = \epsilon^{1/4}y$ and $T = \epsilon t$, and ϵ is a small parameter, proportional to the distance from the Hopf bifurcation. Here, solutions independent of Y indicate perpendicular rolls, while solutions depending on both X and Y indicate perpendicular rolls.

Using the symmetries of the problem, Silber et al. (1992) obtain the following evolution equations for A and B :

$$\begin{aligned} \frac{\partial A}{\partial T} = & c\frac{\partial A}{\partial X} + d\frac{\partial^2 A}{\partial X^2} + s\frac{\partial^2 A}{\partial Y^2} + g\frac{\partial^3 A}{\partial X\partial Y^2} - h\frac{\partial^4 A}{\partial Y^4} + rA \\ & + \alpha(|A|^2 + |B|^2)A + \beta|A|^2A \end{aligned} \quad (3.42)$$

$$\begin{aligned} \frac{\partial B}{\partial T} = & -c\frac{\partial B}{\partial X} + d\frac{\partial^2 B}{\partial X^2} + s\frac{\partial^2 B}{\partial Y^2} - g\frac{\partial^3 B}{\partial X\partial Y^2} - h\frac{\partial^4 B}{\partial Y^4} + rB \\ & + \alpha(|B|^2 + |A|^2)B + \beta|B|^2B, \end{aligned} \quad (3.43)$$

where c and r are real and all other coefficients are complex. The trivial solution ($A = B = 0$) is unstable to perturbations with wavevector (Q_c, P_c) if $r > r_c$, where

$$\begin{aligned} Q_c = 0, \quad P_c = 0, \quad r_c = 0 & \quad \text{if } s_r > 0 \\ Q_c = \frac{g_i P_c^2}{2d_r}, \quad P_c^2 = \frac{2d_r s_r}{g_i^2 - 4h_r d_r}, \quad r_c = \frac{s_r P_c^2}{2} < 0 & \quad \text{if } s_r < 0 \end{aligned} \quad (3.44)$$

Therefore, we see that the Lifshitz point is represented by $s_r = 0$. In other words, s_r depends on ϕ , with $s_r > 0$ if ϕ is below the Lifshitz point, where perpendicular rolls are found, while $s_r < 0$ if ϕ is above the Lifshitz point, where oblique rolls are found. (Note that we must assume $4h_r d_r > g_i^2$ in order for the model to give oblique solutions.)

In the oblique case ($s_r < 0$), we can relate the equations for A and B back to the amplitudes z_1, z_2, w_1 and w_2 by making the following substitution:

$$A = e^{iQ_c X} [z_1(T)e^{iP_c Y} + z_2(T)e^{-iP_c Y}] \quad (3.45)$$

$$B = e^{-iQ_c X} [w_2(T)e^{iP_c Y} + w_1(T)e^{-iP_c Y}] \quad (3.46)$$

If higher order harmonics are neglected, then first order ODEs can be obtained for the four complex amplitudes. These are of the form of (3.37)–(3.40) (as indeed they must be by symmetry arguments), but this time we know the values of the coefficients a, b, c, d and f (in terms of α and β), as follows:

$$d = 2a = 2(\alpha + \beta), \quad c = f = b = \alpha. \quad (3.47)$$

It may seem that we are no better off, since we still do not know the values of α and β . However, by using (3.47) we can in fact eliminate $c, d,$ and f from (3.37)–(3.40) in favour of a and b .

Therefore, by assuming that we are close to the Lifshitz point (ϕ only just above ϕ_c), the problem is greatly simplified, since in this limit there are only two undetermined coefficients instead of five.

3.5.2 Weakly broken symmetry

The problem considered so far contains a left-right reflection symmetry, which is valid for a vertical magnetic field, but not for an inclined one. As usual, we will deal with this by weakly breaking the symmetry; we write μ' instead of μ in equations (3.39) and (3.40), while leaving equations (3.37) and (3.38) unchanged. As before, the nonlinear coefficients are not modified. If we assume that we are close to the Lifshitz point, so

that (3.47) applies, then we are left with the following system of equations:

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_2|^2 + b|w_1|^2 + 2a|z_2|^2)z_1 + bw_2 z_2 \bar{w}_1 \quad (3.48)$$

$$\dot{z}_2 = \mu z_2 + (a|z_2|^2 + b|w_1|^2 + b|w_2|^2 + 2a|z_1|^2)z_2 + bw_1 z_1 \bar{w}_2 \quad (3.49)$$

$$\dot{w}_1 = \mu' w_1 + (a|w_1|^2 + b|z_2|^2 + b|z_1|^2 + 2a|w_2|^2)w_1 + bw_2 z_2 \bar{z}_1 \quad (3.50)$$

$$\dot{w}_2 = \mu' w_2 + (a|w_2|^2 + b|z_1|^2 + b|z_2|^2 + 2a|w_1|^2)w_2 + bw_1 z_1 \bar{z}_2 \quad (3.51)$$

We now consider conditions under which the assumption of weakly broken symmetry is valid. Note that we cannot use our previous tactic of assuming small ϕ ; this is because we have already assumed that we are near to the Lifshitz point, which corresponds to taking ϕ just above the value where oblique rolls first appear, a value that is not necessarily small.

Instead, we recall that with our choice of boundary conditions, there is a left-right reflection symmetry in the problem when κ does not vary with depth, even when ϕ is non-zero. Therefore, if we take δ to be some measure of $|\mathrm{d}\kappa/\mathrm{d}z|$, then the symmetry-breaking will be proportional to δ . If we also take r to be the bifurcation parameter $(R - R_c)/R_c$, then we have the following form for μ and μ' (at lowest order):

$$\mu = r - \delta; \quad \mu' = r + \delta. \quad (3.52)$$

The obvious scaling to take is $r \propto \epsilon^2$, $\delta \propto \epsilon^2$, with $|z_j|$ and $|w_j|$ of order ϵ . All terms in (3.48)–(3.51) are then of the same order, ϵ^3 . Symmetry breaking in the nonlinear terms would only appear at higher orders, as required for consistency.

Note also that we can eliminate some of the parameters in these equations by making suitable rescalings. Firstly, either μ_r or μ'_r may be scaled arbitrarily (by a positive factor), which means that only the *ratio* μ'_r/μ_r is important. Secondly, either a_r or b_r (but not both) could be scaled to ± 1 without loss of generality.

3.5.3 Results

We have looked for solutions to equations (3.48)–(3.51) in which the amplitudes remain constant (although the phases will vary with time). The results are shown in Table 3.5. Many of these branches were found analytically, with the remainder being computed numerically AUTO. Details of the derivations leading to Table 3.5 are given in Appendix C. The patterns are essentially the same as those found by Silber et al. (1992). The differences are that the solutions are slightly distorted in some cases by the presence

<i>Name of solution</i>	<i>Solutions for (z_1, z_2, w_1, w_2)</i>	<i>Stability conditions</i>
rightward-travelling rolls (TRo ^R)*	$ z_1 ^2 = -\mu_r/a_r$ $d/dt(\arg z_1) = \mu_i + a_i z_1 ^2$ $z_2 = w_1 = w_2 = 0$	$a_r < 0$ $\mu'_r/\mu_r < b_r/a_r$
perpendicular travelling rectangles (TRe [⊥])	$ z_1 ^2 = (\mu'_r b_r - \mu_r a_r)/(a_r^2 - b_r^2)$ $ w_2 ^2 = (\mu_r b_r - \mu'_r a_r)/(a_r^2 - b_r^2)$ $d/dt(\arg z_1) = \mu_i + a_i z_1 ^2 + b_i w_2 ^2$ $d/dt(\arg w_2) = \mu'_i + a_i w_2 ^2 + b_i z_1 ^2$ $z_2 = w_1 = 0$	$a_r < 0$ $ a_r > b_r $
rightward-travelling rectangles (TRe ^R)*	$ z_1 ^2 = z_2 ^2 = \mu_r / -3a_r$ $d/dt(\arg z_1) = d/dt(\arg z_2) = \mu_i + 3a_i z_1 ^2$ $w_1 = w_2 = 0$	Always unstable
standing rolls (SRo)	$ z_1 ^2 = (\mu'_r b_r - \mu_r a_r)/(a_r^2 - b_r^2)$ $ w_1 ^2 = (\mu_r b_r - \mu'_r a_r)/(a_r^2 - b_r^2)$ $d/dt(\arg z_1) = \mu_i + a_i z_1 ^2 + b_i w_1 ^2$ $d/dt(\arg w_1) = \mu'_i + a_i w_1 ^2 + b_i z_1 ^2$	$a_r < 0$ $ a_r > b_r $ $ a > b $
standing rectangles (SRe)	$ z_1 = z_2 , w_1 = w_2 ,$ $\arg z_1 - \arg z_2 + \arg w_1 - \arg w_2 = 0$	
alternating rolls (ARo)	$ z_1 = z_2 , w_1 = w_2 ,$ $\arg z_1 - \arg z_2 + \arg w_1 - \arg w_2 = 0$	

*There also exist leftward-travelling versions of these solutions (TRo^L and TRe^L), which may be obtained by exchanging $z_1 \leftrightarrow z_2, w_1 \leftrightarrow w_2$ and $\mu \leftrightarrow \mu'$.

Table 3.5: *The solutions found to equations (3.48)–(3.51), and their stability conditions (all of the listed conditions must be satisfied for the solution to be stable). Note that the SRe and ARo solutions were computed numerically with AUTO, hence the precise forms of the solutions and stability conditions are not listed in the table.*

of asymmetry, and that there are sometimes two copies of each solution, one travelling left and one travelling right (whereas previously, these would have been related by symmetry). These separate left-going and right-going branches have been indicated by adding a superscript L or R where appropriate.

Figure 3.19 shows some of these patterns in graphical form. (We have taken the wavevectors \mathbf{k}_j to be at an angle of 15° to the x -axis in this picture.) The figure shows patterns both in the symmetric case (with all amplitudes equal) and in an asymmetric case (where the $|w_j|$ have been taken to be 40% larger than the $|z_j|$). The first row shows patterns which travel uniformly with time; TRo travel either left or right, while TRe $^\perp$ travel up or down (i.e. in the y -direction). The rest of the figure shows ARo, which have a more complicated time evolution. The ARo are essentially the same as the WR1 from the oscillatory hexagonal model (the only difference is that the angle between the constituent wavevectors is different).

Figure 3.20 shows our results for the existence and stability of the various solution branches in the different regions of the μ_r - μ'_r plane. Note that only the *ratio* μ'_r/μ_r is important, so the regions are always sectors emanating from the origin, which have been represented by arcs on the diagrams. A solid (dashed) arc indicates the existence of a stable (unstable) solution. There are three cases (a)–(c) depending on a_r and b_r : case (a) applies when $|b_r| < |a_r|$, case (b) when $|a_r| < |b_r| < 3|a_r|$, and case (c) when $|b_r| > 3|a_r|$. In case (b), although the existence region of ARo can be calculated analytically, the stability region must be found numerically (it depends on all parameters including a_i and b_i). In particular we cannot say explicitly whether the stability boundary will be inside or outside the sector AOB (this is represented by diagrams (b)(i) and (b)(ii)). Finally diagram (d) is not a separate case but rather shows some additional solution branches, which exist in all three cases, but are always unstable.

The diagram can be interpreted by observing that if κ is constant (the symmetric case) then we are on the forty-five degree line $\mu_r = \mu'_r$. Increasing the Rayleigh number corresponds to moving towards the top right. If we increase the variability of κ we move towards the top left, if κ decreases with depth, or the bottom right, if κ increases downward. (Recall that in our model, κ decreased downward so $\mu'_r > \mu_r$ is the relevant part of the diagram.) For example in case (a) we would see TRe $^\perp$ and SRo for weakly varying κ , while for more strongly varying κ we would see only leftward-travelling rolls.

In addition to the solutions of Table 3.5, one or two other solution branches were found with AUTO. These had all four amplitudes ($|z_1|$, $|z_2|$, $|w_1|$ and $|w_2|$) unequal

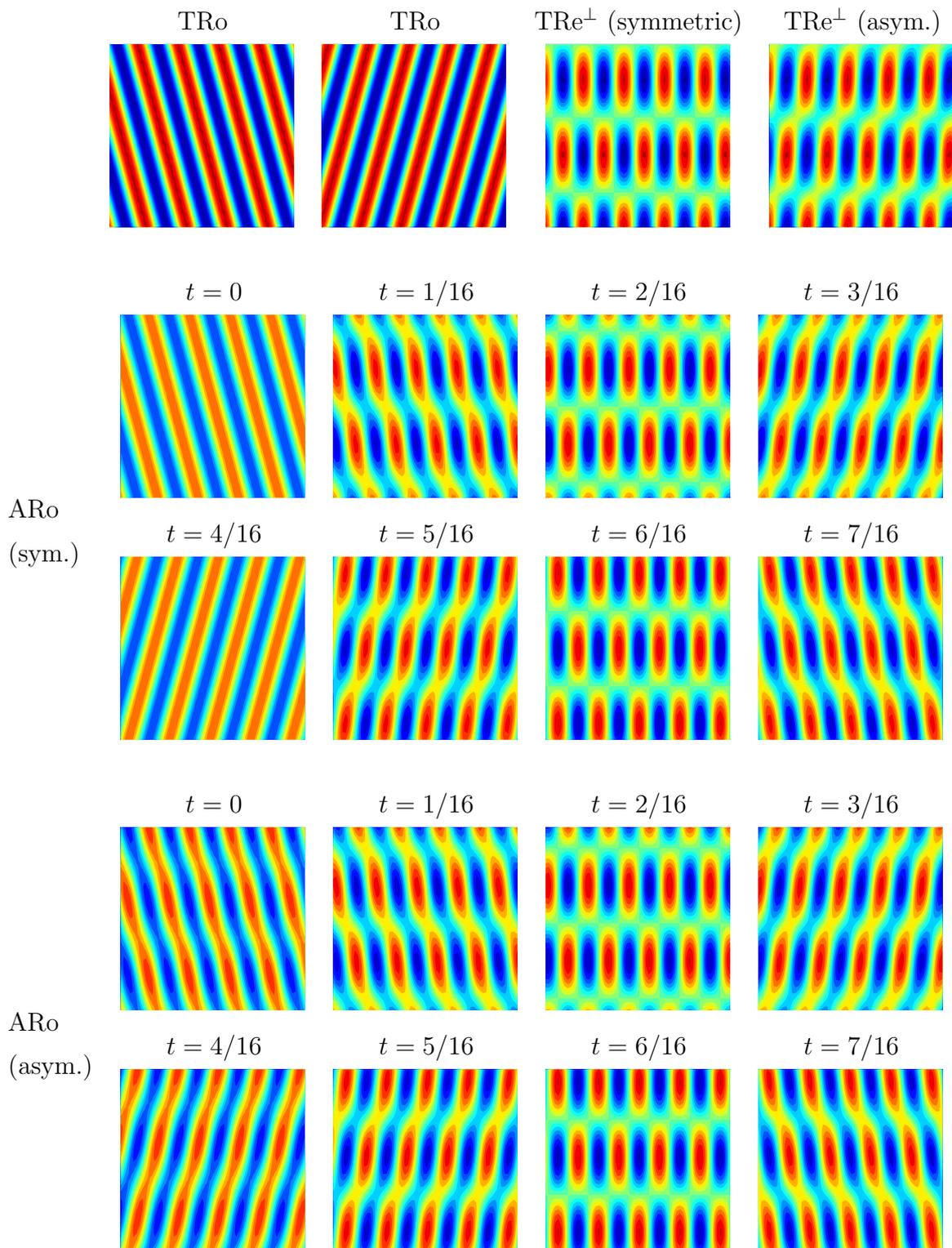


Figure 3.19: Patterns in the rhombic model.

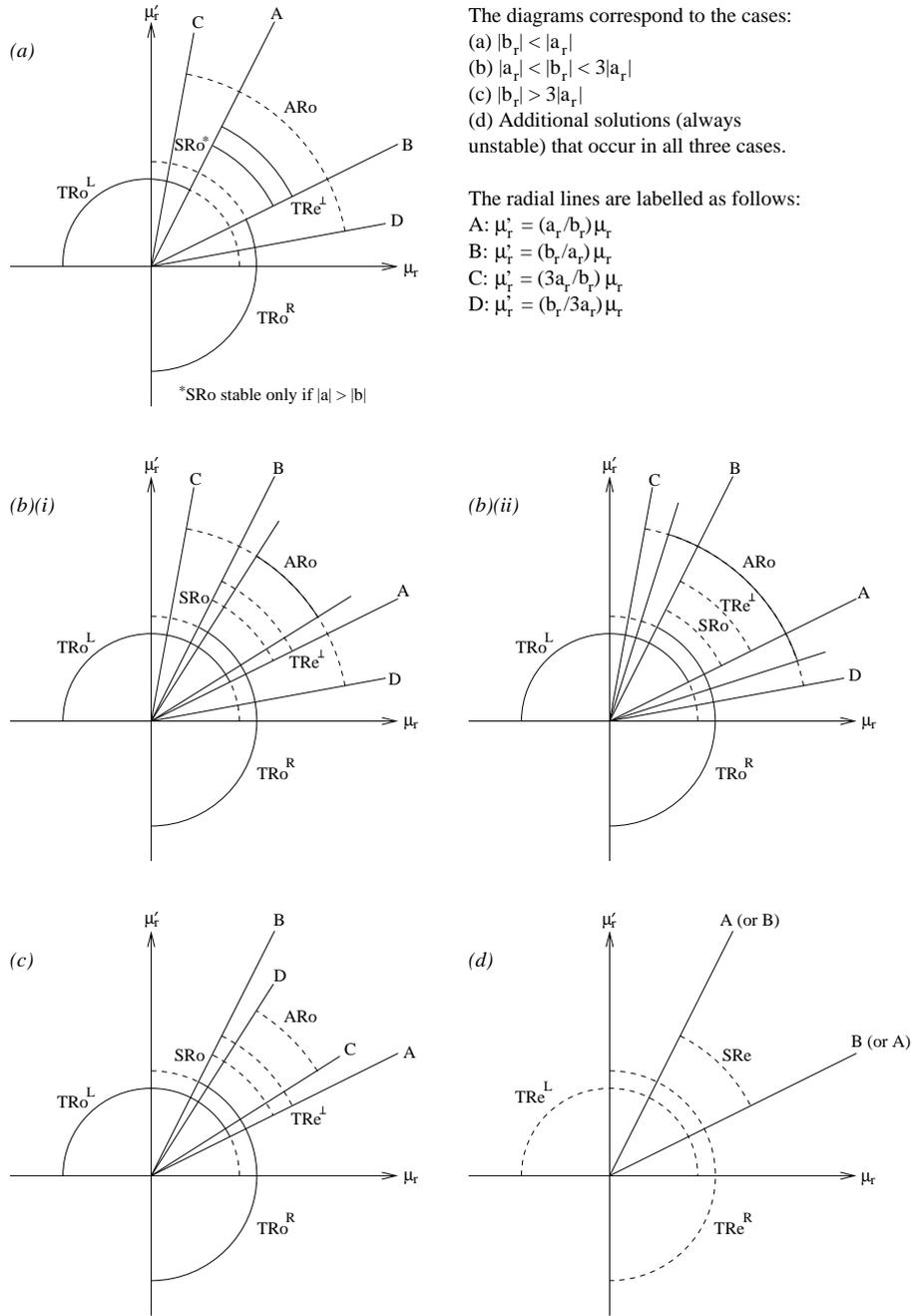


Figure 3.20: Regions of existence and stability of the various patterns in the μ_r - μ'_r plane. Solid (dashed) curves indicate stable (unstable) solutions. See text for further explanation.

and non-zero. For example, solution branches of this kind were found that connected together the (unstable) standing rectangle branch and the travelling roll solutions. In some cases there were even solutions in which the amplitudes $|z_j|$ and $|w_j|$ were periodic functions of time. However, all of these more complicated solution branches were found to be unstable (at least for the cases we considered), so they are not discussed further here.

3.5.4 Comparison with results below the Lifshitz point

The results above apply to the oblique roll regime (above the Lifshitz point). However, the model (3.42)–(3.43) also applies below the Lifshitz point, i.e. when $\phi < \phi_0$, if we choose $s_r > 0$ instead of $s_r < 0$. It is instructive to compare the results in the two different cases, since this illustrates how the solutions will change as ϕ is increased through ϕ_0 , from the perpendicular into the oblique roll regime. (See also Figure 9 of Silber et al. 1992.)

In the ‘perpendicular’ regime ($s_r > 0$) we have $Q_c = P_c = 0$ (equation 3.44), so there are only two critical modes, representing left-going and right-going perpendicular rolls. We can obtain amplitude equations for these modes simply by setting $z_1 = A$, $w_1 = B$ in (3.42)–(3.43), and dropping all spatial derivative terms, which gives:

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_1|^2)z_1 \quad (3.53)$$

$$\dot{w}_1 = \mu' w_1 + (a|w_1|^2 + b|z_1|^2)w_1 \quad (3.54)$$

(here we have also weakly broken the left-right symmetry). Note that these equations can alternatively be derived simply by setting z_2 and w_2 to zero in (3.48)–(3.51).

These equations have two solutions, corresponding to either travelling or standing rolls. In fact we can quite easily relate the existence and stability results for these solutions to the corresponding results in the oblique roll regime. We find that TRo exist and are stable below the Lifshitz point if and only if they exist and are stable above the Lifshitz point; in other words the stability conditions for TRo do not change between the perpendicular and oblique regimes. For the SRo, we find that the stability condition for SRo below the Lifshitz point corresponds to the stability condition for TRe[⊥] above the Lifshitz point; in other words, if TRe[⊥] are stable in the oblique regime then SRo will be stable in the perpendicular regime (and vice versa). This is all summarized in Figure 3.21 which shows the equivalent of Figure 3.20 for the perpendicular case.

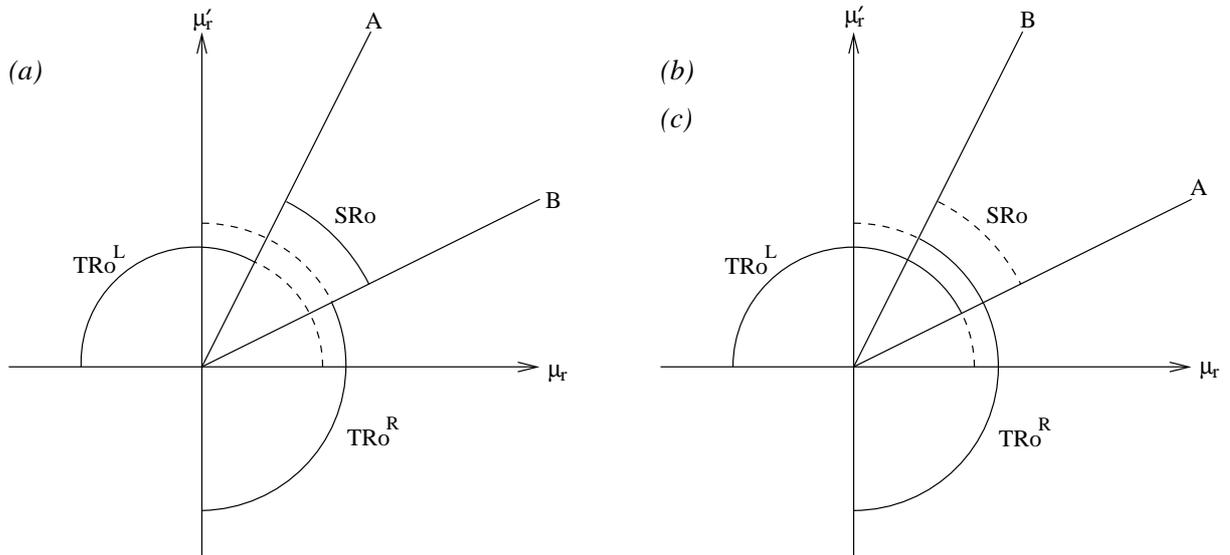


Figure 3.21: The equivalent of Figure 3.20, for the ‘perpendicular’ regime ($\phi < \phi_0$, $s_r > 0$) instead of the ‘oblique’ regime ($\phi > \phi_0$, $s_r < 0$). Note that in this case, there are only two possible solutions: perpendicular travelling rolls (TRo) or perpendicular standing rolls (SRo). The labels (a), (b), (c) and the lines A, B have the same meaning as in Figure 3.20.

We note therefore that one of the following would be observed if ϕ was set to a value just below the Lifshitz point, and then increased to a value just above it:

- If the symmetry is very strongly broken, then only travelling rolls would be stable. These will be perpendicular rolls below the Lifshitz point, or oblique rolls above it.
- If the symmetry is not too strongly broken, and $|b_r| > |a_r|$ (corresponding to cases (b) and (c) on Figures 3.20 and 3.21), then we will see again see travelling rolls on both sides of the Lifshitz point; we might also see alternating rolls above the Lifshitz point (depending on the values of μ , μ' , a and b).
- If the symmetry is not too strongly broken, and $|b_r| < |a_r|$ (corresponding to case (a) on Figures 3.20 and 3.21), then we will see perpendicular standing rolls below the Lifshitz point, and the TRe^\perp solution (see Figure 3.19) above it. If $|a| > |b|$, then we will also see oblique standing rolls above the Lifshitz point.

3.5.5 A calculation of the coefficients a and b

In order to determine which of the above cases occur in any given problem, one needs to know the values of the coefficients a and b . These can be calculated by means of a perturbation analysis of the original equations about the equilibrium solution. In fact, we can calculate these coefficients for the symmetric case ($\mu = \mu'$) if we want to, since their values do not change when the symmetry is weakly broken – the same values will apply to both the symmetric and asymmetric cases.

The symmetric case comes about when $d\kappa/dz \rightarrow 0$, i.e. when κ is uniform. Our assumptions also require us to choose ϕ at (or just above) the Lifshitz point. It turns out that as $d\kappa/dz \rightarrow 0$, the value of ϕ corresponding to the Lifshitz point tends to zero as well (see Figure 2.13 on page 61). In other words, if we let $d\kappa/dz \rightarrow 0$ with ϕ fixed just above the Lifshitz point, we will end up with $\phi \rightarrow 0$ as well; therefore, it appears that the limit we require is that of a Boussinesq fluid (uniform κ) in a vertical field.

Matthews and Rucklidge (1993) have performed a weakly nonlinear calculation in this limit. They actually use different magnetic boundary conditions to our model – they have the magnetic field constrained to be vertical at top and bottom (so $B_x = B_y = 0$) while we have used a potential field at one boundary and a tied field at the other. Therefore the results will not apply directly to our model, but should still give a useful qualitative guide to what happens.

The calculation of Matthews and Rucklidge (1993) considered only the two-dimensional case. We can simulate this by setting $z_2 = w_2 = 0$ (cf. equations 3.53–3.54), and also $\mu = \mu'$, obtaining the following equations, equivalent to theirs:

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_1|^2)z_1 \quad (3.55)$$

$$\dot{w}_1 = \mu w_1 + (a|w_1|^2 + b|z_1|^2)w_1 \quad (3.56)$$

except that they now give explicit expressions for a and b (γ and δ in their notation). Note that since, near the Lifshitz point, there are only two coefficients a and b to be determined (c , d and f having been eliminated using equation 3.47), it is enough to consider only this two-dimensional problem.

These values can then be used to determine which of cases (a), (b) or (c) occurs as a function of the parameters Q , ζ and σ of the original PDEs. This has been done in Figure 3.22. Notice that case (c) does not in fact occur for this problem. The difference between (a)* and (a) is that SRo are stable in the former ($|a|$ being greater than $|b|$ there) but unstable in the latter. The shaded region to the left of the diagram corresponds to

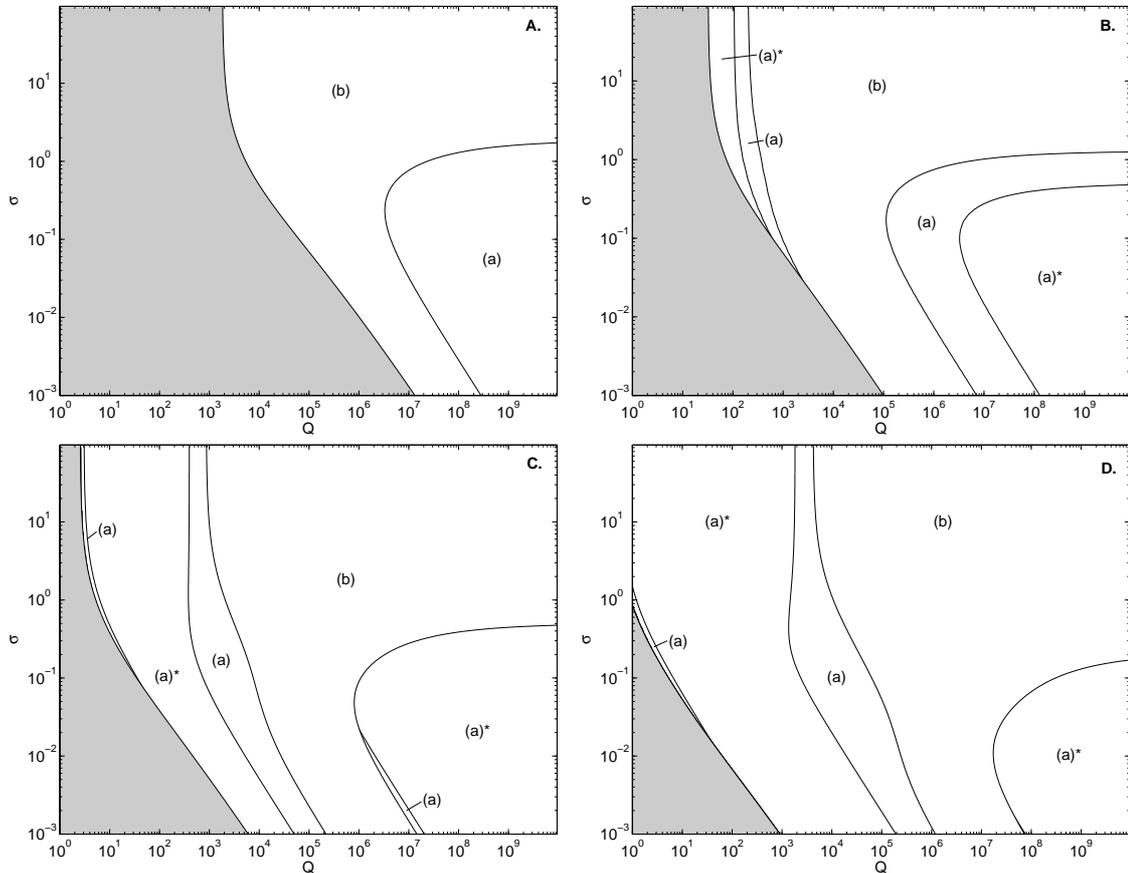


Figure 3.22: Diagrams showing which of the cases (a), (b) or (c) from Figure 3.20 occurs as a function of Q , ζ and σ (as indicated by the calculation of section 3.5.5). Note that in case (a), SRo may or may not be stable; the regions where SRo are stable have been marked with an asterisk. In the shaded areas, convection is steady at onset and the model does not apply. The four diagrams correspond to different values of ζ as follows. A: $\zeta = 0.9$; B: $\zeta = 0.5$; C: $\zeta = 0.1$; D: $\zeta = 0.02$.

parameter values in which convection is steady at onset (due to Q being too small) and hence the analysis does not apply here.

3.6 Conclusions

In this chapter, we have moved on from linearized computations and looked at some simplified nonlinear models. These can shed some light on what happens when the linear theory predicts several competing modes with similar growth rates. The great advantage of these models is that since they only use very general symmetry arguments, the results can be applied to a wide variety of situations. (For example, the results would in principle apply equally well to Boussinesq, anelastic, or compressible problems.) Their main disadvantage is that they apply only to the weakly nonlinear regime.

The method used was to select a lattice on which the solutions were to be doubly periodic. This restriction was essential for technical reasons, although it does limit the class of solutions that can be obtained. However, (approximately) doubly periodic patterns are frequently found experimentally or in numerical simulations, and therefore the method is useful in practice. Another limitation is that when we check for stability or instability of a pattern, we are in fact only checking for instability with respect to perturbations that are doubly periodic on the given lattice – for example, instabilities leading to long-wavelength modulations of the patterns will not be found by this method.

Our models were divided into two cases: the ‘steady’ case (for large ζ) and the ‘oscillatory’ case (for small ζ). We will describe each of these in turn.

The steady case is certainly the simpler of the two. Here, we investigated the problem on a hexagonal lattice (section 3.3). The hexagonal lattice is useful because it allows the competition between rolls and hexagons to be investigated; we would expect to see hexagonal patterns for small tilt angles ϕ (because of the lack of up-down symmetry), but for larger tilt angles we would expect to see parallel rolls (Danielson, 1961). The steady hexagonal model allows this transition to be investigated in more detail.

The results show that when $\phi = 0$, the convection takes the form of steady hexagons (Figure 3.6, left-hand panel). As ϕ increases, this pattern persists, although it becomes distorted in appearance (due to asymmetry between the three modes making up the hexagons; Figure 3.6, second and third panels). The hexagons also drift with a speed which is initially proportional to the tilt angle; the direction of travel was leftward in our model (see previous chapter), but this is model-dependent. When ϕ is increased

further, there is a transition to steady parallel rolls; in many cases this transition is associated with hysteresis.

The oscillatory models showed slightly more complicated results. In this case we investigated two different models: one on a rhombic lattice (section 3.5), and one on a hexagonal lattice (section 3.4).

The oscillatory rhombic model included four modes, and was for simplicity restricted to values of ϕ close to the Lifshitz point. If the amount of symmetry breaking was sufficiently weak, then there were essentially two cases. In the first case, standing rolls would be found below the Lifshitz point, while travelling rectangles (travelling in a direction perpendicular to the tilt), and possibly also standing rolls, would be found above it. In the second case, travelling rolls would be found below the Lifshitz point, while both travelling and (possibly) alternating rolls would be found above it.

The oscillatory hexagonal model was appropriate for small tilt angles, and had somewhat more complicated results (Figure 3.17). It is difficult to draw any conclusions relevant to sunspots from these results. We can however observe that there is a wide variety of interesting behaviour that can occur in this regime, even before considering fully nonlinear effects.

Note that in reality, the distinction between either a hexagonal or a rhombic lattice is somewhat artificial; modes from both lattices (plus an infinite number of other modes) would be present in the full problem. There is no real answer to this, but we can state that the hexagonal model is more appropriate to the perpendicular roll regime in the linear theory, occurring for small ϕ , while the rhombic model is more appropriate to the vicinity of the Lifshitz point, which can be found at larger values of ϕ . (Unfortunately, since we have to choose either one lattice or the other, we cannot really investigate the transition between these two regimes using these methods, except in the limited way that was discussed in section 3.5.4.)

We can now ask which of the two types of model – steady or oscillatory – is more relevant to sunspots. To answer this, we must consider the value of ζ in the solar surface layers. In fact this value varies with depth due to ionization effects; $\zeta < 1$ for depths less than 2000 km, but for depths of around 2000 to 20000 km there is a layer in which $\zeta > 1$ (Meyer et al., 1974). Therefore, if the surface convective features observed in sunspots are fairly shallow (depth less than about 2000 km), the oscillatory model is appropriate. If their depth is much greater, then the steady model would be appropriate.

There is also another possibility: a combination of both the steady and the oscillatory

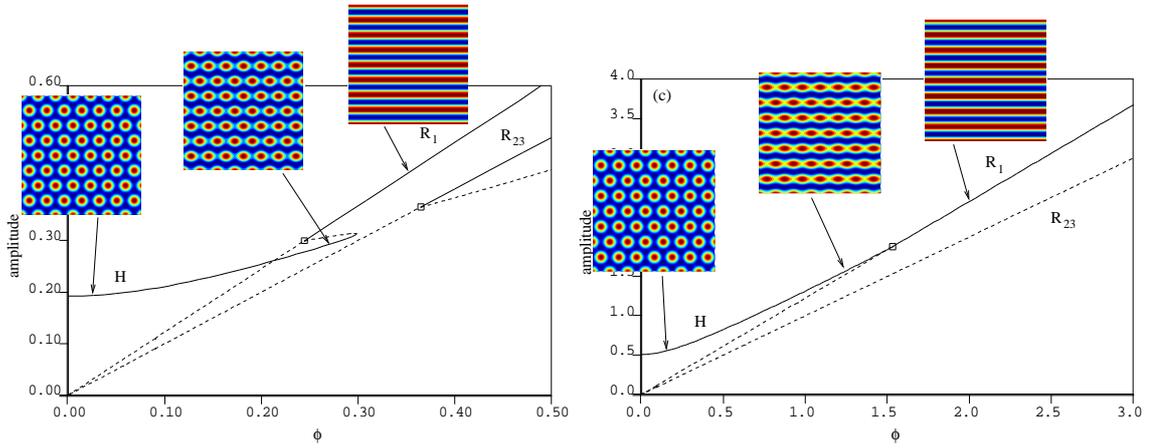


Figure 3.23: Bifurcation diagrams from the steady hexagonal model (with $\theta = 30^\circ$), showing a case with hysteresis (left-hand picture) and one without (right-hand picture).

behaviour could occur simultaneously. In our models, we get either one or the other, because we consider only the situation near onset; for higher Rayleigh numbers, more complicated behaviour can occur. For example, the simulations of Weiss et al. (1990, 1996), in which ζ is chosen to vary from 0.2 at the top of the layer to 2.2 at the base, show that convection is steady near onset, but as the Rayleigh number increases, a secondary bifurcation occurs, leading to a new solution. This solution consists of persistent overturning convection near the base of the layer, coupled to an oscillating pattern near the surface. (The models of this chapter cannot directly describe solutions of this kind; they would appear as secondary branches bifurcating from the solutions of our steady hexagonal model.)

Perhaps one of the most interesting features found in our models is the presence of hysteresis as ϕ is varied. This is seen in both the steady and oscillatory cases (although we have only investigated the former in detail). Hysteresis is usually associated with sharp transitions between different patterns, and might therefore be part of the explanation for the sharp transition between the umbra and penumbra of a sunspot.

The hysteresis is illustrated in Figure 3.23 which shows bifurcation diagrams from the steady hexagonal model in two cases, one with and one without hysteresis. Pictorial representations of the stable solutions at various points have also been added.

Consider what would happen if ϕ was allowed to vary smoothly with position, as a crude way of modelling a sunspot. In a case without hysteresis, the convection pattern would vary from regular hexagons in an area with a vertical field, to parallel rolls in

areas of large ϕ (representing near-horizontal fields), with a smoothly varying pattern in between; this is illustrated by the right-hand diagram. By contrast, in a case with hysteresis, a rather more abrupt transition would be expected between the hexagon and roll regions (left-hand picture).

In the latter case, it is tempting to relate the hexagons to the umbra, and the rolls to the penumbra.² Of course, such a model would be highly simplified and idealized, and so could not be compared directly with observations. However, the mechanism that it represents is a very general one and we can reasonably expect it to apply to more complicated models as well (in addition to whatever else might be happening in those models).

The idea of hysteresis in magnetoconvection has also been employed by previous authors. For example, Rucklidge et al. (1995) suggest the presence of hysteresis as part of their explanation of the abrupt formation of penumbrae. Thomas and Weiss (2004) conjecture that hysteresis may be caused by the phenomenon of ‘flux pumping’. Our work suggests that hysteresis may also occur naturally as an intrinsic part of magnetoconvection in inclined fields.

We also note that a similar hysteresis phenomenon was found in the asymptotic calculation of Julien et al. (2000, 2003), although their model uses somewhat different assumptions to ours. (Both our model and theirs are highly idealized, albeit in different ways.) Moreover, they find a second hysteretic transition occurring for larger ϕ , which they suggest is a possible explanation of the differing properties of bright and dark filaments. This second transition is not seen in our model, but since we assume small ϕ we would not expect to see any effects associated with more strongly inclined fields.

In the next chapter we will describe how the steady hexagonal model of section 3.3 can be modified to include a tilt angle ϕ that is a function of x . This will allow us to investigate the hysteresis phenomenon discussed above in more detail.

²The rolls do not capture the complex, filamentary structure found in real penumbrae, but we might suppose that such structure develops as a result of an instability starting from this roll state. However, such instabilities are beyond the scope of this weakly nonlinear theory.

Chapter 4

Swift-Hohenberg model for tilted magnetic fields

4.1 Introduction

The weakly nonlinear models discussed in the previous chapter used ideas from equivariant bifurcation theory to determine the patterns that might be observed in various circumstances. The main assumption made was that the patterns were to be doubly periodic with respect to some lattice (either rhombic or hexagonal), which led to a small number of ODEs describing the dynamics. We now wish to extend the model to allow ϕ (the angle of tilt of the magnetic field) to vary with position, as a simple way of modelling solar pores and sunspots, where the magnetic field fans out and becomes more inclined as one moves radially outwards.

Our approach will be to start from the Swift-Hohenberg equation (Swift and Hohenberg, 1977), a commonly used model equation for convection problems, in which the full three-dimensional governing equations are reduced to a single PDE in two spatial dimensions. This is a useful model since it exhibits many of the properties of convection, at least in the weakly nonlinear regime, but it is much simpler to work with than the full equations. We will add some extra terms to the usual Swift-Hohenberg equation, in order to represent the effects of the tilted magnetic field. These terms will be discussed in section 4.2.

Before studying this modified Swift-Hohenberg equation in its own right, we will first use it to derive a set of three Landau-Ginzburg equations. This represents a further simplification, since it converts the problem from a PDE in two spatial dimensions to a

system in one spatial dimension. This makes it easier to understand what is happening, albeit with some loss of generality. These equations will be derived in section 4.3.

We then move on, in section 4.4, to look at the properties of the modified Swift-Hohenberg equation itself. Using the Swift-Hohenberg equation allows us to look at more general situations, and one of the things we will do is to adapt the equation to a circular (rather than Cartesian) geometry, in order to represent a sunspot slightly more realistically. Finally, we present our conclusions in section 4.5.

4.2 The model

We consider the following modification of the Swift-Hohenberg equation:

$$\frac{\partial \psi}{\partial t} = r\psi - (1 + \nabla^2)^2 \psi + \alpha \psi^2 - \gamma \psi^3 + a\phi^2 \psi + b\phi^2 \frac{\partial^2 \psi}{\partial x^2} + c\phi \frac{\partial \psi}{\partial x}, \quad (4.1)$$

where $\psi = \psi(x, y, t)$ and $\phi = \phi(x)$ are real functions, and r, a, b, c, α and γ are real constants. When $\phi = 0$ we recover the original Swift-Hohenberg equation (with added quadratic term). The last three terms on the right-hand side are new and represent the effects of the tilted field (under the assumption that ϕ is small).

The effects of these new terms are best seen by looking at the linearized version of (4.1). If ϕ is constant we can look at the linear theory very easily, substituting $\psi = \exp(i\mathbf{k} \cdot \mathbf{x} + \sigma t)$ (with $\mathbf{k} = (k_x, k_y)$). The following equation for σ is obtained:

$$\sigma = (r + a\phi^2) - (1 - |\mathbf{k}|^2) - bk_x^2 \phi^2 + ik_x c\phi. \quad (4.2)$$

We can now see that the term proportional to a simply causes the growth rate to increase in proportion to ϕ^2 . This represents the fact that as the field tilt increases (at fixed total field strength), the critical Rayleigh number falls, and the convection becomes more vigorous (at fixed Rayleigh number).

The term proportional to b introduces a (linear) anisotropy, so that the linear growth rate of a mode depends on its orientation. Modes with $k_x \neq 0$ are penalized, so that the most unstable mode has $\mathbf{k} = (0, 1)$, corresponding to rolls aligned with the tilt direction of the magnetic field.

The term proportional to c models the fact that patterns travel when ϕ is non-zero. If ϕ was uniform, this term could be removed by transforming to an appropriate moving frame, but when ϕ depends on position, it must be retained.

Note that symmetry considerations imply that the travelling term is an odd function of ϕ , and therefore proportional to ϕ (for small ϕ), while the other two new terms are even in ϕ , and therefore proportional to ϕ^2 at leading order.

Modifications of the Swift-Hohenberg equation, including anisotropy, have been studied before. These can be divided into two classes: those in which the quadratic term (representing broken up-down symmetry) is included, and those in which it is not. An example of the latter is the work of Pesch and Kramer (1986), who applied a modified Swift-Hohenberg equation to anisotropic liquid crystal systems. Another example is in the field of rotating convection, where Roxin and Riecke (2002, 2004) apply an anisotropic Swift-Hohenberg model. A more detailed model of anisotropic rotating convection (with mean flows included) is given by Pollicott et al. (2003).

Examples of models including both anisotropy effects and the quadratic term are given by Walgraef and Schiller (1987) and Schmitz and Zimmermann (1997). Here, a transition from hexagons to rolls is found as the anisotropy increases (as in our model). However, the ‘travelling’ term (proportional to c) has not been included in any previous studies, and this term creates interesting new effects, as we shall see below.

4.3 Landau-Ginzburg equations

From the above model it is possible to derive a set of amplitude equations of Landau-Ginzburg type. We introduce a small parameter ϵ and make the following transformations:

$$r = \epsilon^2 r_2 \tag{4.3}$$

$$c = \epsilon c_1 \tag{4.4}$$

$$\phi = \epsilon \phi_1 \tag{4.5}$$

$$\alpha = \epsilon \alpha_1, \tag{4.6}$$

where r_2 , c_1 and α_1 are constants (having no ϵ -dependence). We assume that the other constants (b and γ) do not depend on ϵ . We also assume that ϕ_1 is a slowly varying function of x , varying on an $O(1/\epsilon)$ length scale, i.e. $\phi_1(x) = f(\epsilon x)$ where the definition of f does not itself depend on ϵ .

The scalings for r , ϕ and α represent small-amplitude convection, a small tilt angle, and a small up-down asymmetry, and all of these agree with the scalings made in the previous chapter (section 3.3). We choose $c = O(\epsilon)$ out of necessity; the calculation we

are about to perform would fail if c was taken to be of order 1. However, we can justify the choice of small c in terms of our linear model (Chapter 2). There, we adopted the Boussinesq approximation, with a depth-dependent κ being used to break the up-down symmetry. It turned out that the speed of travel of our solutions was proportional to the amount of up-down asymmetry (i.e. the relative size of $d\kappa/dz$). But here, we must assume small up-down asymmetry (equation 4.6), so it is consistent to take a slow speed of travel as well (equation 4.4).

Note also that by rescaling ψ and redefining ϵ , we can without loss of generality rescale α_1 and γ arbitrarily. We do so, setting $\alpha_1 = 1/2$, $\gamma = 1/3$ (this will simplify the equations later on).

Equation (4.1) becomes

$$\frac{\partial\psi}{\partial t} = \epsilon^2(r_2 + a\phi_1^2)\psi - (1 + \nabla^2)^2\psi + \frac{1}{2}\epsilon\psi^2 - \frac{1}{3}\psi^3 + \epsilon^2 c_1 \phi_1 \frac{\partial\psi}{\partial x} + b\epsilon^2 \phi_1^2 \frac{\partial^2\psi}{\partial x^2}. \quad (4.7)$$

We can make the separation of length and time scales explicit by introducing ‘fast’ variables $x' = x$, $y' = y$, $t' = t$ and ‘slow’ variables $X = \epsilon x/\sqrt{3}$, $Y = \epsilon y/\sqrt{3}$, $T = \epsilon^2 t$ (the factors of $\sqrt{3}$ are for later convenience). We now imagine ψ as being a function of both the fast and slow variables, writing $\psi = \psi(x', y', t', X, Y, T)$ instead of the previous $\psi = \psi(x, y, t)$. Using the chain rule, we find the following relationships between the old and new variables:

$$\frac{\partial\psi}{\partial t} = \frac{\partial\psi}{\partial t'} + \epsilon^2 \frac{\partial\psi}{\partial T} \quad (4.8)$$

$$\frac{\partial\psi}{\partial x} = \frac{\partial\psi}{\partial x'} + O(\epsilon) \quad (4.9)$$

$$\frac{\partial^2\psi}{\partial x^2} = \frac{\partial^2\psi}{\partial x'^2} + O(\epsilon) \quad (4.10)$$

$$-(1 + \nabla^2)^2\psi = L_0\psi + \epsilon L_1\psi + \epsilon^2 L_2\psi + O(\epsilon^3) \quad (4.11)$$

where

$$L_0 = -(1 + \nabla'^2)^2 \quad (4.12)$$

$$L_1 = -\frac{4}{\sqrt{3}}(1 + \nabla'^2) \left(\frac{\partial^2}{\partial x' \partial X} + \frac{\partial^2}{\partial y' \partial Y} \right) \quad (4.13)$$

$$L_2 = -\frac{2}{3}(1 + \nabla'^2) \left(\frac{\partial^2}{\partial X^2} + \frac{\partial^2}{\partial Y^2} \right) - \frac{4}{3} \left(\frac{\partial^2}{\partial x' \partial X} + \frac{\partial^2}{\partial y' \partial Y} \right)^2 \quad (4.14)$$

and $\nabla'^2 = \partial^2/\partial x'^2 + \partial^2/\partial y'^2$.

Note that because we assumed that ϕ varies on an $O(1/\epsilon)$ length scale, ϕ_1 can be treated as a function of X alone (and not x').

The next step is to expand ψ as

$$\psi = \epsilon\psi_1 + \epsilon^2\psi_2 + \epsilon^3\psi_3 + \dots \quad (4.15)$$

and substitute into (4.7). Equating coefficients at $O(\epsilon)$ gives

$$\frac{\partial\psi_1}{\partial t'} = L_0\psi_1. \quad (4.16)$$

We look for solutions in which $\partial\psi/\partial t' = 0$, so (4.16) simply becomes $L_0\psi_1 = 0$, which has solutions

$$\psi_1 = z(X, Y, T) \exp(ik_x x' + ik_y y') + \text{c.c.} \quad (4.17)$$

for any k_x and k_y satisfying $k_x^2 + k_y^2 = 1$, and for an arbitrary function $z(X, Y, T)$. At this stage we are free to take any linear combination of these solutions. We choose to take three modes on a hexagonal lattice, giving us the following solution for ψ_1 :

$$\psi_1 = \sum_{j=1}^3 z_j(X, Y, T) \exp(i\mathbf{k}_j \cdot \mathbf{x}') + \text{c.c.} \quad (4.18)$$

with

$$\mathbf{k}_1 = (0, -1), \quad \mathbf{k}_{2,3} = (\mp\sqrt{3}/2, 1/2). \quad (4.19)$$

(Cf. Figure 3.4 on page 72; our choice here corresponds to $\theta = 30^\circ$.)

At $O(\epsilon^2)$ we have

$$0 = L_0\psi_2 + L_1\psi_1. \quad (4.20)$$

It happens that ψ_1 , as given by (4.18), satisfies $L_1\psi_1 = 0$, hence (4.20) \Rightarrow $L_0\psi_2 = 0$. The solutions for ψ_2 are proportional to $\exp i\mathbf{k} \cdot \mathbf{x}'$, with $|\mathbf{k}| = 1$, and also satisfy

$$L_1\psi_2 = 0. \quad (4.21)$$

At $O(\epsilon^3)$ we obtain (after some rearrangement)

$$-L_0\psi_3 = -\frac{\partial\psi_1}{\partial T} + (r_2 + a\phi_1^2)\psi_1 + L_2\psi_1 + c_1\phi_1 \frac{\partial\psi_1}{\partial x'} + b\phi_1^2 \frac{\partial^2\psi_1}{\partial x'^2} + \frac{1}{2}\psi_1^2 - \frac{1}{3}\psi_1^3. \quad (4.22)$$

Here we have used (4.21).

The operator L_0 is self-adjoint, so $\langle L_0 y_1, y_2 \rangle = \langle y_1, L_0 y_2 \rangle$ for any two functions y_1 and y_2 .¹ In particular, $\langle L_0 \psi_3, \psi_1 \rangle = \langle \psi_3, L_0 \psi_1 \rangle$; and since $L_0 \psi_1 = 0$, this implies that

¹Here $\langle f, g \rangle$ is an inner product between functions f and g , defined by $\langle f, g \rangle = \int \bar{f}(x)g(x)dx$. The integral is taken over a suitable interval in x (e.g. if the functions had periodic boundary conditions, then the integral would be over one period).

$\langle L_0\psi_3, \psi_1 \rangle = 0$. The net result of this is that $L_0\psi_3$ cannot contain any terms proportional to $\exp(\pm i\mathbf{k}_j \cdot \mathbf{x}')$.

We will now, therefore, pick out terms on the right-hand side of (4.22) that are proportional to $\exp(\pm i\mathbf{k}_1 \cdot \mathbf{x}')$. (The corresponding terms with \mathbf{k}_2 or \mathbf{k}_3 in place of \mathbf{k}_1 can be obtained by cyclic permutation of indices.) We first of all note that

$$\psi_1^2 = 2\bar{z}_2\bar{z}_3 \exp(i\mathbf{k}_1 \cdot \mathbf{x}') + \text{c.c.} + \text{other terms} \quad (4.23)$$

$$\psi_1^3 = z_1(3|z_1|^2 + 6|z_2|^2 + 6|z_3|^2) \exp(i\mathbf{k}_1 \cdot \mathbf{x}') + \text{c.c.} + \text{other terms} \quad (4.24)$$

where ‘other terms’ denotes terms that are not proportional to $\exp(\pm i\mathbf{k}_1 \cdot \mathbf{x}')$. We then obtain

$$\begin{aligned} -L_0\psi_3 = & \left[-\frac{\partial z_1}{\partial T} + (r_2 + a\phi_1^2)z_1 + \frac{4}{3} \left(k_{x1} \frac{\partial}{\partial X} + k_{y1} \frac{\partial}{\partial Y} \right)^2 z_1 + c_1\phi_1 i k_{x1} z_1 - b\phi_1^2 k_{x1}^2 z_1 \right. \\ & \left. + \bar{z}_2\bar{z}_3 - z_1(|z_1|^2 + 3|z_2|^2 + 3|z_3|^2) \right] \exp(i\mathbf{k}_1 \cdot \mathbf{x}') + \text{other terms} \end{aligned} \quad (4.25)$$

where again the ‘other terms’ are not proportional to $\exp(\pm i\mathbf{k}_1 \cdot \mathbf{x}')$. Since $L_0\psi_3$ cannot contain any terms proportional to $\exp(\pm i\mathbf{k}_1 \cdot \mathbf{x}')$, the entire right-hand side of the equation (excluding the ‘other terms’) can be set to zero as a solvability condition. Doing this for the other two equations (obtained by cyclic permutation) as well, and also substituting the values of k_{xj} and k_{yj} , yields the following equations:

$$\begin{aligned} \frac{\partial z_1}{\partial T} = & (r_2 + a\phi_1^2)z_1 + \frac{4}{3} \frac{\partial^2 z_1}{\partial Y^2} \\ & + \bar{z}_2\bar{z}_3 - z_1(|z_1|^2 + 2|z_2|^2 + 2|z_3|^2) \end{aligned} \quad (4.26)$$

$$\begin{aligned} \frac{\partial z_2}{\partial T} = & (r_2 + (a - \tilde{b})\phi_1^2)z_2 - \tilde{c}\phi_1 i z_2 + \left(-\frac{\partial}{\partial X} + \frac{1}{\sqrt{3}} \frac{\partial}{\partial Y} \right)^2 z_2 \\ & + \bar{z}_3\bar{z}_1 - z_2(|z_2|^2 + 2|z_3|^2 + 2|z_1|^2) \end{aligned} \quad (4.27)$$

$$\begin{aligned} \frac{\partial z_3}{\partial T} = & (r_2 + (a - \tilde{b})\phi_1^2)z_3 + \tilde{c}\phi_1 i z_3 + \left(\frac{\partial}{\partial X} + \frac{1}{\sqrt{3}} \frac{\partial}{\partial Y} \right)^2 z_3 \\ & + \bar{z}_1\bar{z}_2 - z_3(|z_3|^2 + 2|z_1|^2 + 2|z_2|^2) \end{aligned} \quad (4.28)$$

where $\tilde{b} = 3b/4$ and $\tilde{c} = \sqrt{3}c_1/2$ (these rescalings eliminate a couple of numeric constants that would otherwise appear in the equations). Equations (4.26)–(4.28) are the three amplitude equations (Landau-Ginzburg equations) for our model.

As a simplification, we will neglect all Y -dependence in the following sections, so that $\partial/\partial Y$ can be set to zero. (This is reasonable because we do not expect to see any significant variation along the Y direction, since the coefficients of the equations are constant along the Y direction, and vary only in the X direction.)

4.3.1 Setting up the problem

In order to specify the problem fully, we have to choose a function $\phi_1(X)$, as well as giving values for r_2 , a , \tilde{b} and \tilde{c} . Initially, we will set the constants r_2 and \tilde{c} to zero, and look at the effects of non-zero values later on.

This leaves a and \tilde{b} , which affect how the growth rates for the three modes change with ϕ . For these we can refer back to the results of our steady hexagonal model, described in section 3.3. Recall that there was a pitchfork bifurcation at $\phi = \phi_{\text{pf}}$ at which rolls became stable, and also a saddle-node bifurcation at $\phi = \phi_{\text{sn}}$ at which the hexagons became unstable. In this chapter, we will scale our parameters such that $\phi_{\text{pf}} = 1$. (This can always be done by multiplying both a and \tilde{b} by ϕ_{pf}^2 if the latter is not equal to 1.) This effectively leaves one free parameter, which we may take to be $a/(a - \tilde{b})$; this is equivalent to the parameter A_1 from section 3.3.1. (This parameter measures the amount of anisotropy in the system; $A_1 = 1$ gives an isotropic system while $A_1 \gg 1$ implies a highly anisotropic problem.)

Recall also that the pitchfork bifurcation at ϕ_{pf} could be either supercritical or subcritical, depending on A_1 . We will investigate two values of A_1 , one for each of these possibilities. Case I will have

$$A_1 = 1.11, \quad a = 0.826, \quad \tilde{b} = 0.082, \quad (4.29)$$

which produces a subcritical bifurcation (with $\phi_{\text{pf}} = 1$ and $\phi_{\text{sn}} = 1.4$). Case II will be supercritical, with

$$A_1 = 2.5, \quad a = 0.391, \quad \tilde{b} = 0.234. \quad (4.30)$$

The choice of $\phi_1(X)$ defines how much the tilt varies by, and over what length scale it varies. For computational convenience we will take ϕ_1 to be a periodic function, with period X_0 , defined as follows:

$$\phi_1 = 0.75(1 - \cos(2\pi X/X_0)). \quad (4.31)$$

This function varies sinusoidally from 0 at the edges of the domain to 1.5 in the centre. We can now impose periodic boundary conditions when solving (4.26)–(4.28), with the solution domain being $X \in [0, X_0]$, or equivalently $x \in [0, L]$, where $L = X_0\sqrt{3}/\epsilon$.

With periodic boundary conditions we can easily compute solutions to (4.26)–(4.28) using a Fourier spectral method for the horizontal differences. We couple this with an ETD (exponential time differencing) method for the time discretization (Cox and

Matthews, 2002). Given values for r_2 , a , \tilde{b} and \tilde{c} , and ϕ_0 , ϕ_1 and X_0 , we can then solve the equations numerically to find z_1 – z_3 as functions of X and T .

Then, once we have the three amplitudes z_1 – z_3 , and a value for L , we can reconstruct ψ (to lowest order) from equation (4.18). In the following, we fix L to be a constant, choosing a value that gives a reasonably large number of convection cells across the width of the domain. We allow X_0 to vary; note that this is equivalent to varying ϵ , since $\epsilon = X_0\sqrt{3}/L$.

There is one other requirement before numerical calculations can be performed, which comes about because there is no diffusion term in (4.26) (after setting $\partial/\partial Y = 0$). This causes problems numerically, and we resolve this by adding a small ‘regularizing’ term, $-\eta\partial^4 z_1/\partial X^4$, to the right-hand side of (4.26), where η is chosen to be a very small value (i.e. small enough that varying η has no qualitative effect on the results). This prevents numerical difficulties and is physically reasonable, since one of the higher order terms would be a fourth-order diffusion term anyway.

4.3.2 Theoretical ideas

In the case $\tilde{c} = 0$ (in which the ‘travelling’ term is switched off), we can make some progress theoretically.

If we momentarily neglect the diffusion terms in equations (4.26)–(4.28), then there are no X -derivatives in the equations, and we are effectively left with a set of ODEs for each point along the X axis. These ODEs are nothing other than the ‘steady hexagonal’ model from the previous chapter (section 3.3). Since we know $\phi(X)$, we can use the results of that chapter (e.g. the bifurcation diagram Figure 3.8) to plot the stable solution(s) as a function of X . This has been done in Figure 4.1.

These graphs give us an idea of what the solution to the PDEs (4.26)–(4.28) will be. In the supercritical case we would expect the solution to closely follow the graphs shown in the middle row of Figure 4.1, although perhaps modified slightly by the diffusion terms in (4.26)–(4.28), which we have so far neglected. However, in the subcritical case, pictured in the top row of the figure, there are two different solution branches, and the solution to the PDEs cannot follow both of these simultaneously. There would have to be a discontinuous change (or front) between the two patterns at at least one point.

Near such a front, the diffusion terms will clearly play an important role. In fact, they have two effects: firstly, they smooth out the front, so that instead of a sharp

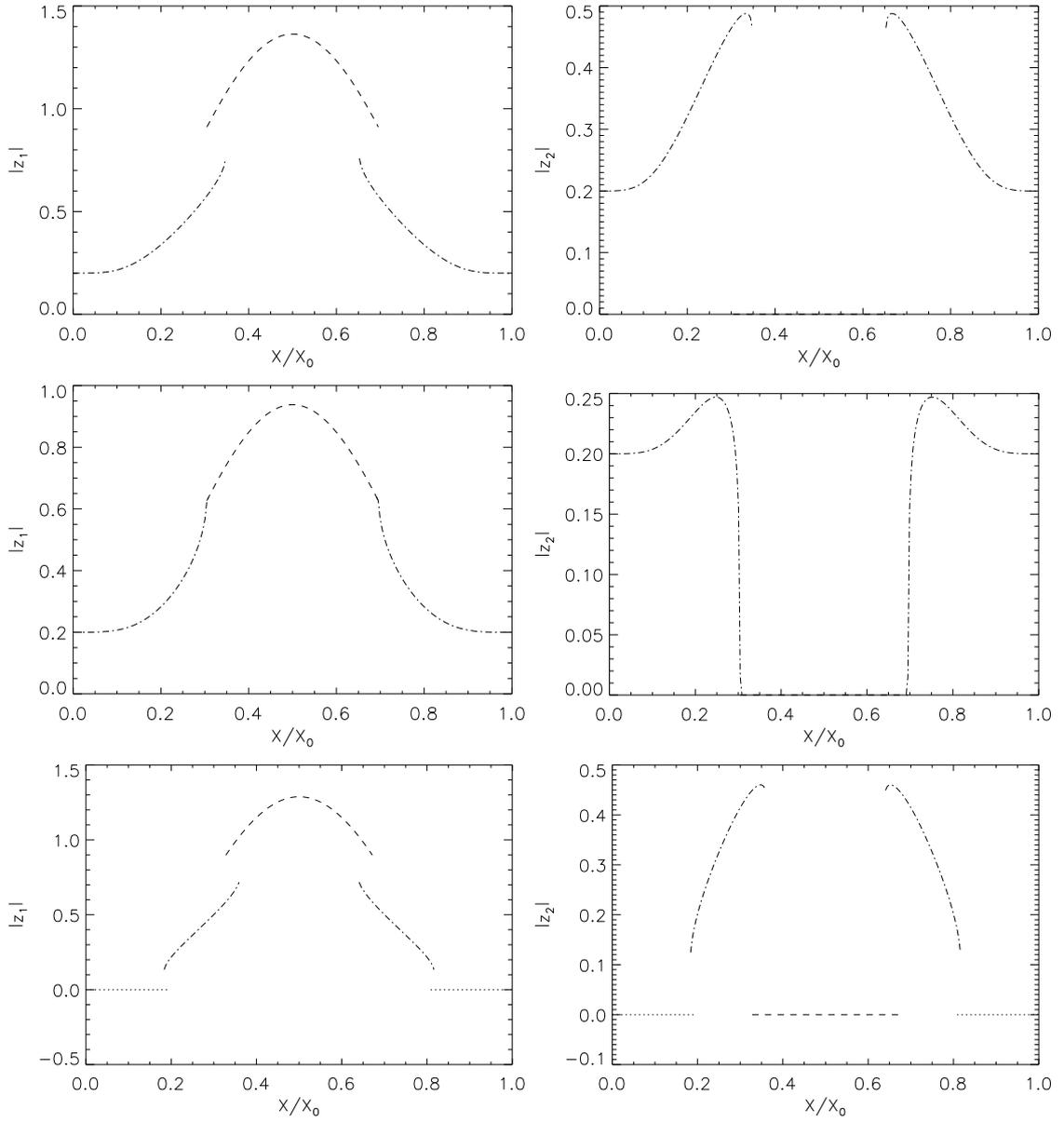


Figure 4.1: Graphs showing the stable roll (dashed line) and hexagon (dash-dotted line) branches, as well as the trivial solution (dotted line), as indicated by the analysis of section 4.3.2. The left-hand column shows $|z_1|$ against X , and the right-hand column shows $|z_2| = |z_3|$ against X . These amplitudes are only plotted at points where the respective solutions are stable. The top row is the subcritical case (see equation 4.29), the middle row is the supercritical case (equation 4.30), and the bottom row shows the subcritical case with r_2 changed to -0.2 .

discontinuity there is a continuous change over some diffusive length scale; secondly, as we will see shortly, they can actually cause the front to travel, either to the left or to the right.

We can investigate this phenomenon by considering what happens when ϕ is uniform. Although this is not the case in our model, it should nevertheless give an idea of what happens if ϕ is varying only slowly with X .

We first of all write the equations in the following form:

$$\dot{z}_1 = \frac{\partial U}{\partial z_1} \quad (4.32)$$

$$\dot{z}_2 = \frac{\partial U}{\partial z_2} + \frac{\partial^2 z_2}{\partial X^2} \quad (4.33)$$

$$\dot{z}_3 = \frac{\partial U}{\partial z_3} + \frac{\partial^2 z_3}{\partial X^2} \quad (4.34)$$

where

$$U = \frac{\mu_1}{2} z_1^2 + \frac{\mu_{23}}{2} (z_2^2 + z_3^2) - \frac{1}{4} (z_1^4 + z_2^4 + z_3^4) - (z_1^2 z_2^2 + z_2^2 z_3^2 + z_3^2 z_1^2) + z_1 z_2 z_3 \quad (4.35)$$

(here we have neglected the fourth order diffusion term in the \dot{z}_1 equation, since it is very small). U can be thought of as a ‘potential’.

We look for solutions of the form $z_j(x, t) = z_j(\xi)$, where $\xi = x - vt$, and v is a constant. These are uniformly translating solutions (i.e., they are time-independent when viewed in a frame moving with constant velocity v). The equations become:

$$-v \frac{dz_1}{d\xi} = \frac{\partial U}{\partial z_1} \quad (4.36)$$

$$-v \frac{dz_2}{d\xi} = \frac{\partial U}{\partial z_2} + \frac{d^2 z_2}{d\xi^2} \quad (4.37)$$

$$-v \frac{dz_3}{d\xi} = \frac{\partial U}{\partial z_3} + \frac{d^2 z_3}{d\xi^2}. \quad (4.38)$$

Multiplying each equation by $dz_j/d\xi$, and summing, we obtain:

$$\frac{d}{d\xi} \left[U + \frac{1}{2} \left(\frac{dz_2}{d\xi} \right)^2 + \frac{1}{2} \left(\frac{dz_3}{d\xi} \right)^2 \right] = -v \left[\left(\frac{dz_1}{d\xi} \right)^2 + \left(\frac{dz_2}{d\xi} \right)^2 + \left(\frac{dz_3}{d\xi} \right)^2 \right]. \quad (4.39)$$

If we define

$$E = U + \frac{1}{2} \left(\frac{dz_2}{d\xi} \right)^2 + \frac{1}{2} \left(\frac{dz_3}{d\xi} \right)^2, \quad (4.40)$$

then we can see that E is either an increasing function of ξ (if $v < 0$), a decreasing function (if $v > 0$), or constant (if $v = 0$).

Now consider a solution containing a front between hexagons and rolls. In particular, suppose that there are hexagons for large negative ξ , and rolls for large positive ξ , with a transition region near $\xi = 0$. In regions far from the origin, z_1 , z_2 and z_3 are constant, and so $E = U$ there. So we have $E \rightarrow U_{\text{hex}}$ for $\xi \ll 0$ and $E \rightarrow U_{\text{rolls}}$ for $\xi \gg 0$ (where U_{hex} and U_{rolls} are the potentials for the hexagon and roll solutions respectively).

Therefore, we see that if $U_{\text{hex}} > U_{\text{rolls}}$, then v will be positive, and the front will move to the right (so that the hexagons invade the rolls). If $U_{\text{hex}} < U_{\text{rolls}}$ then the opposite happens and the rolls invade the hexagons. The solution with the higher value of U ‘wins’. If $U_{\text{hex}} = U_{\text{rolls}}$, then a stationary front between the two solutions is possible.

We now consider what happens if ϕ is changing with position. This technically makes the above analysis invalid: firstly, the amplitudes in the ‘hexagon’ and ‘roll’ regions are no longer constant, but are varying with ϕ (and hence X), and secondly, the potential U now explicitly depends on X , which complicates the calculation somewhat. Nevertheless, if ϕ is varying slowly enough, then the above analysis should be at least approximately valid, and we can use it to get some ideas about what is happening. We find that U_{hex} and U_{rolls} now both depend on X (since they are functions of ϕ), and therefore v will also depend on X in general. This means that there will be some regions where the front travels to the left, some where it travels to the right, and (possibly) an ‘equilibrium’ position where the front can be stationary.

Therefore, even though there is an entire band of ϕ values in which hexagons and rolls are both (locally) stable, the front between them will be located at a well-defined position within this band; this will be the point at which the potentials U_{hex} and U_{rolls} (as defined above) are equal.

The above applies for $r \geq 0$. If $r < 0$ then there is an additional consideration: the trivial solution is now stable over part of the domain, where ϕ is sufficiently small. (This situation is illustrated in the bottom row of Figure 4.1). Note that the transition between the trivial solution and hexagons is associated with hysteresis, just as the transition between hexagons and rolls is. The arguments above can be applied equally well to a front between hexagons and the trivial solution. The trivial solution has $U = 0$, and so we would expect to see a sharp transition between the trivial solution and hexagons near the point where $U_{\text{hex}} = 0$.

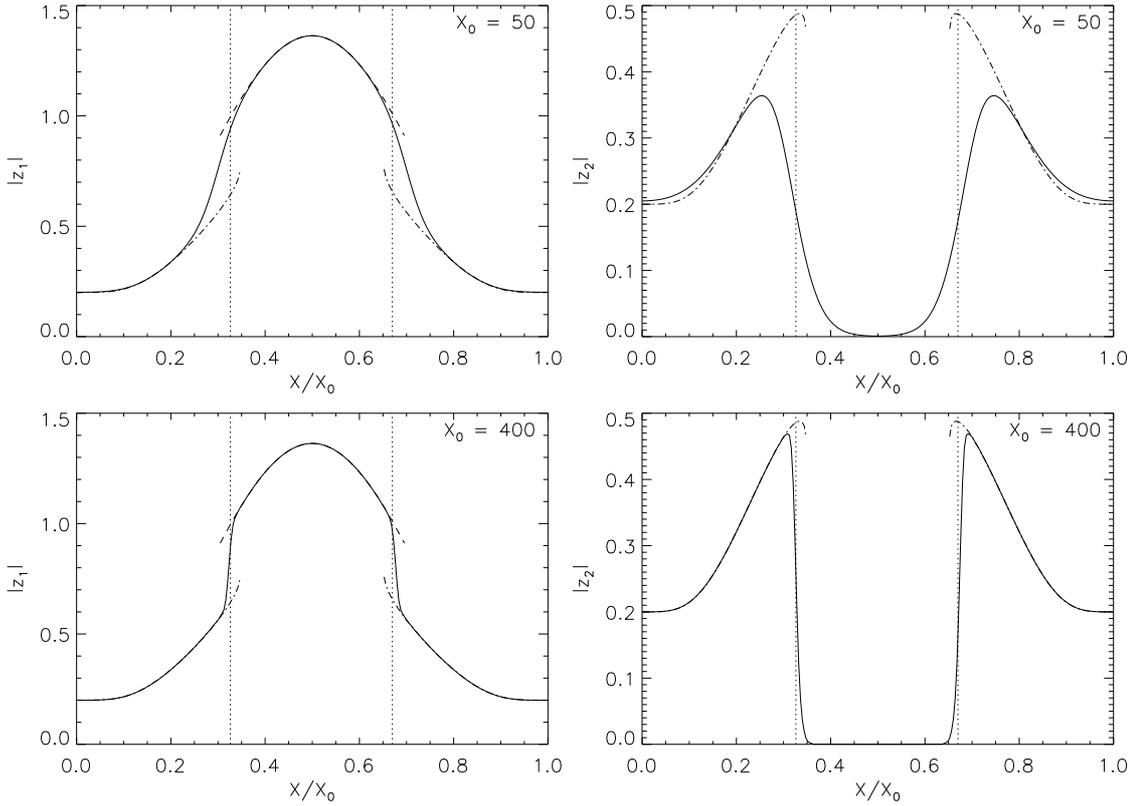


Figure 4.2: The amplitudes $|z_1|$ (left-hand column) and $|z_2| = |z_3|$ (right-hand column) corresponding to two choices for the parameter X_0 : the top row shows the case $X_0 = 50$ (corresponding to a relatively high amount of diffusion), and the bottom row shows the case $X_0 = 400$ (with much lower diffusion). The solid lines show the actual computed amplitudes. The dashed and dashed-dotted lines show the predicted stable roll and hexagon branches, respectively. The vertical dotted lines give the positions where $U_{\text{rolls}} = U_{\text{hex}}$ (see section 4.3.2).

4.3.3 Numerical results

We now discuss the results of our numerical solutions of equations (4.26)–(4.28). We will first of all look at case I (the ‘subcritical’ case, equation 4.29), and we will set $r_2 = \tilde{c} = 0$ to begin with. For our initial conditions, we set ψ to a small random value at each grid point. Initially we will take these random values to be *real*, as this simplifies the results a little; later we will move on to discuss what happens with complex initial conditions.

We begin by plotting the amplitudes $|z_1|$, $|z_2|$ and $|z_3|$ that result from such a simulation, as a function of X . This is done in Figure 4.2. We find that $|z_2|$ and $|z_3|$ end up being equal to each other, so we have plotted only two sets of graphs, one showing $|z_1|$

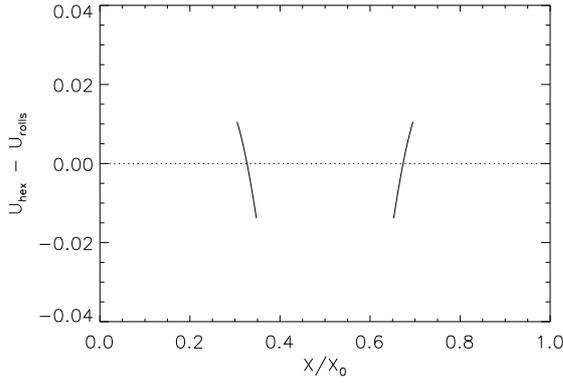


Figure 4.3: The value of $U_{\text{hex}} - U_{\text{rolls}}$ as a function of X , for the subcritical case (equation 4.29) with $r_2 = 0$. This value has been plotted only in regions where rolls and hexagons are both stable.

(left-hand column of the figure) and one showing $|z_2| = |z_3|$ (right-hand column).

The ‘ideal’ solution curves for rolls and hexagons, from Figure 4.1, are also shown, as dashed and dashed-dotted lines respectively. We find that away from the transition between rolls and hexagons, the numerical solution (solid line) follows these ideal curves quite closely. The transition itself is smoothed out over a ‘diffusion width’; this is more pronounced in the top two pictures, where there is more diffusion, than in the bottom two pictures.

We have also calculated the potentials U_{hex} and U_{rolls} (as defined in the previous section) for the parameters used in these runs. In Figure 4.3, we plot $U_{\text{hex}} - U_{\text{rolls}}$ (for those values of X where rolls and hexagons are simultaneously stable). We see that for ϕ less than a critical value, $U_{\text{hex}} > U_{\text{rolls}}$, so that the front moves towards higher ϕ ; for ϕ greater than this critical value, the opposite is true. The front therefore moves towards the critical value, which represents its stable equilibrium position. This position has been plotted on Figure 4.2 as a vertical dotted line; it can be seen that this agrees quite closely with the actual observed position of the front.

Figure 4.4 shows a representation of ψ itself for three different values of X_0 . These have been drawn at fixed $L = 40\pi\sqrt{3}$ (this value gives room for thirty hexagonal cells across the width of the box). X_0 is increasing from top to bottom; physically, increasing X_0 corresponds to reducing the amount of diffusion in the system. Unsurprisingly, the transition between rolls and hexagons is sharpest when diffusion is least (largest X_0). In fact, in this subcritical case, the sharpness of the transition is limited only by the diffusion terms.

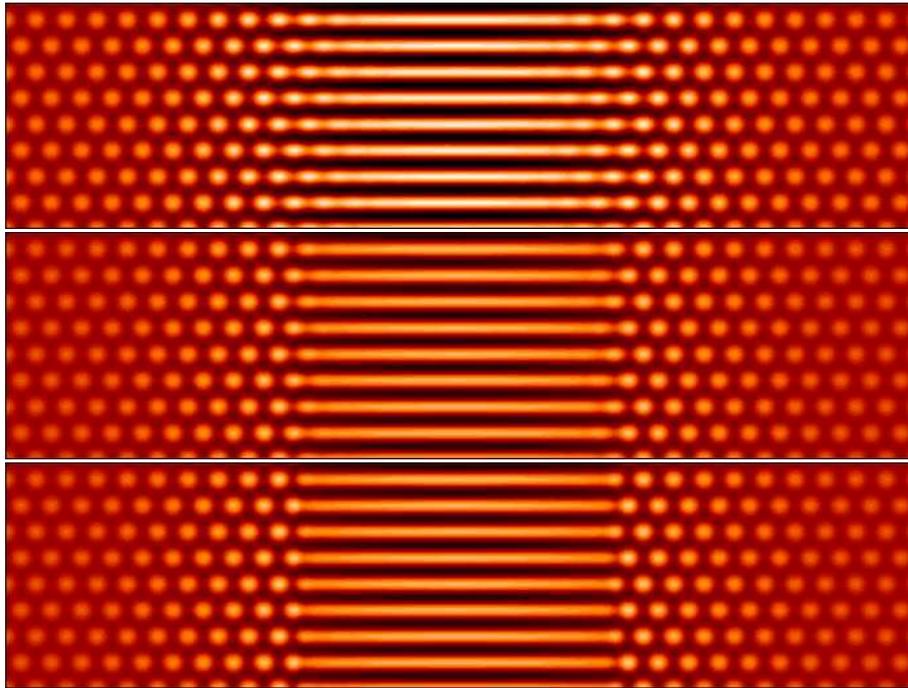


Figure 4.4: *The effect of varying X_0 (which is proportional to ϵ) in the subcritical model. The top picture shows $X_0 = 50$, the middle one $X_0 = 150$, and the bottom one $X_0 = 400$. (These are steady solutions.)*

It should also be noted that $\epsilon = X_0\sqrt{3}/L$, so increasing X_0 (at fixed L) also corresponds to increasing ϵ (our measure of the amplitude of the convection). The values of ϵ corresponding to the three panels of Figure 4.4 are 0.40, 1.19 and 3.18. These are perhaps rather large (given that ϵ is a small parameter), so the model may not strictly be valid towards the upper end of this range, but what we can conclude is that the transition becomes sharper for larger values of ϵ (which would correspond to larger Rayleigh numbers).

We also briefly mention the effect of increasing the parameter r_2 (which previously we have set to zero). Basically, this change moves the transition point between rolls and hexagons towards smaller values of ϕ (cf. Figure 3.9 on page 83). This is illustrated in Figure 4.5 where we have fixed $X_0 = 150$ and varied r_2 , taking the values 0.5, 1.2 and 1.5. As r_2 increases, the region of rolls becomes larger, until eventually rolls are seen in the entire domain.

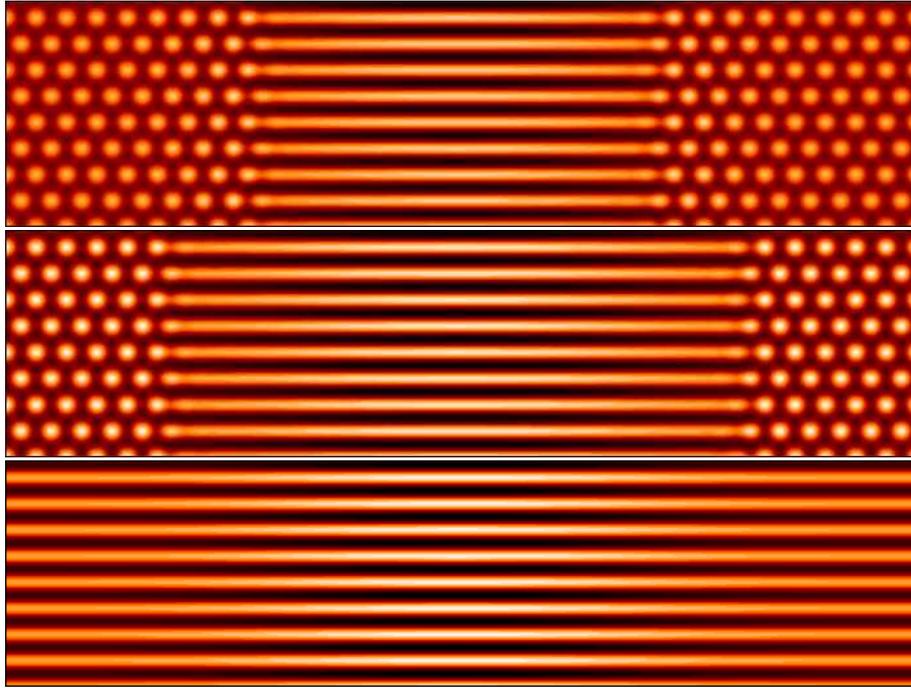


Figure 4.5: Results for (from top to bottom) $r_2 = 0.5$, $r_2 = 1.2$ and $r_2 = 1.5$. (As before, these are steady solutions.)

Results when $\tilde{c} \neq 0$

So far we have fixed $\tilde{c} = 0$; when $\tilde{c} \neq 0$ (and sufficiently large), patterns start to travel. We will choose $\tilde{c} > 0$ (without loss of generality), producing leftward-travelling patterns. This produces ‘inward’-travelling patterns on the left-hand side of the domain (corresponding to the situation in a sunspot) and ‘outward’-travelling patterns on the right-hand side.

We actually find for very small \tilde{c} that steady patterns are possible. Such a solution, for $\tilde{c} = 0.02$, is shown in Figure 4.6. For $\tilde{c} \gtrsim 0.025$, however, we find that the solutions become time-dependent, with patterns travelling to the left. Note that this travelling applies to the hexagons, but not the rolls, since \tilde{c} does not appear in the \dot{z}_1 equation.

A solution with $\tilde{c} = 0.2$ is displayed in Figure 4.7; snapshots at six different times are shown. The solution is similar in appearance to the patterns found above (with $\tilde{c} = 0$); the difference is that the hexagons now travel leftwards (although the rolls remain stationary). Moreover, the speed of this travel depends on ϕ , and hence also on X . Near the centre of the domain, the hexagons are travelling quickly, but near the left- and right-hand edges, they move more slowly. This causes the pattern to be compressed in the left-hand half of the domain, or stretched in the right-hand half, in a kind of

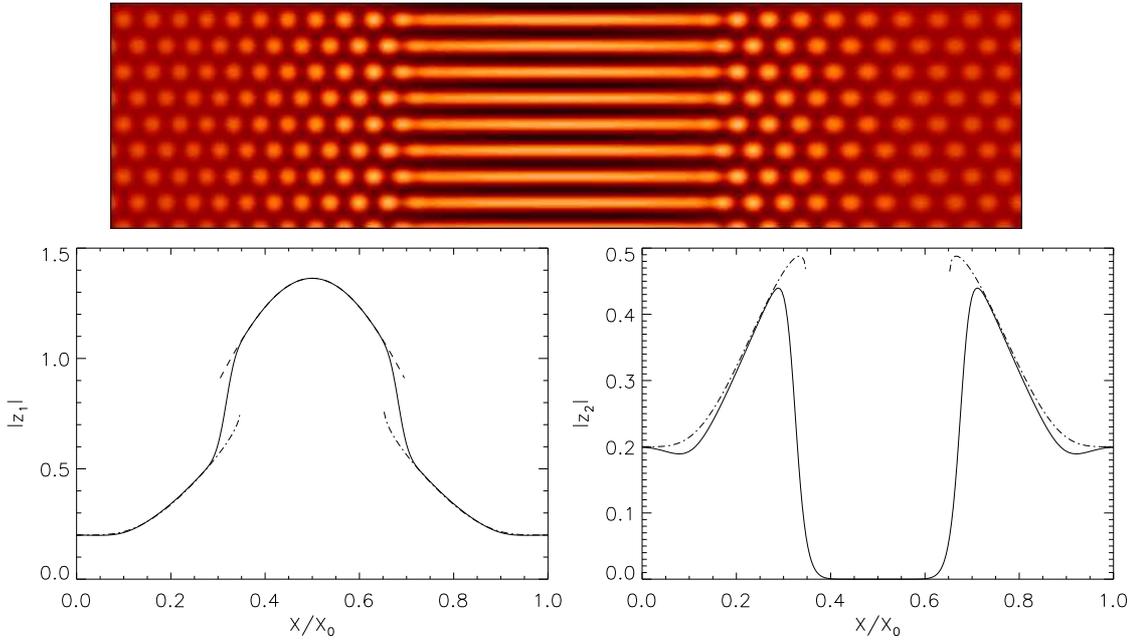


Figure 4.6: A steady solution for $\tilde{c} = 0.02$. Plots of $|z_1|$ and $|z_2| = |z_3|$ against X are also shown.

‘concertina effect’.

The small arrows on Figure 4.7 indicate points where this ‘concertina effect’ has become quite pronounced. Notice that in this situation, the wavenumber of the hexagonal pattern is now some distance away from the preferred wavenumber. The system responds to this by causing two of the ‘compressed’ hexagons join together to become one larger cell, or one ‘stretched’ hexagon to split into two smaller convection cells. After this point, the system evens itself out, with all the hexagons returning to a similar size.

This process can also be seen by looking at the amplitudes $|z_1|$ and $|z_2| = |z_3|$. These are plotted at four different times in Figures 4.8 and 4.9. To explain what’s happening here, we consider the following (linear) model equation:

$$\dot{z} = r_2 z + i\tilde{c}\phi(x)z + D \frac{\partial^2 z}{\partial x^2}. \quad (4.41)$$

This can be seen as a simplification of either (4.27) or (4.28), where only three of the right-hand-side terms have been kept. (The coefficient D is assumed to be real and positive.)

Substituting $z = R(x) \exp(i\theta(x))$ (where R and θ are real) we obtain

$$\dot{R} = (r_2 - D\theta'^2)R + DR'' \quad (4.42)$$

$$\dot{\theta} = \tilde{c}\phi(x) + 2D \frac{R'\theta'}{R} + D\theta''. \quad (4.43)$$

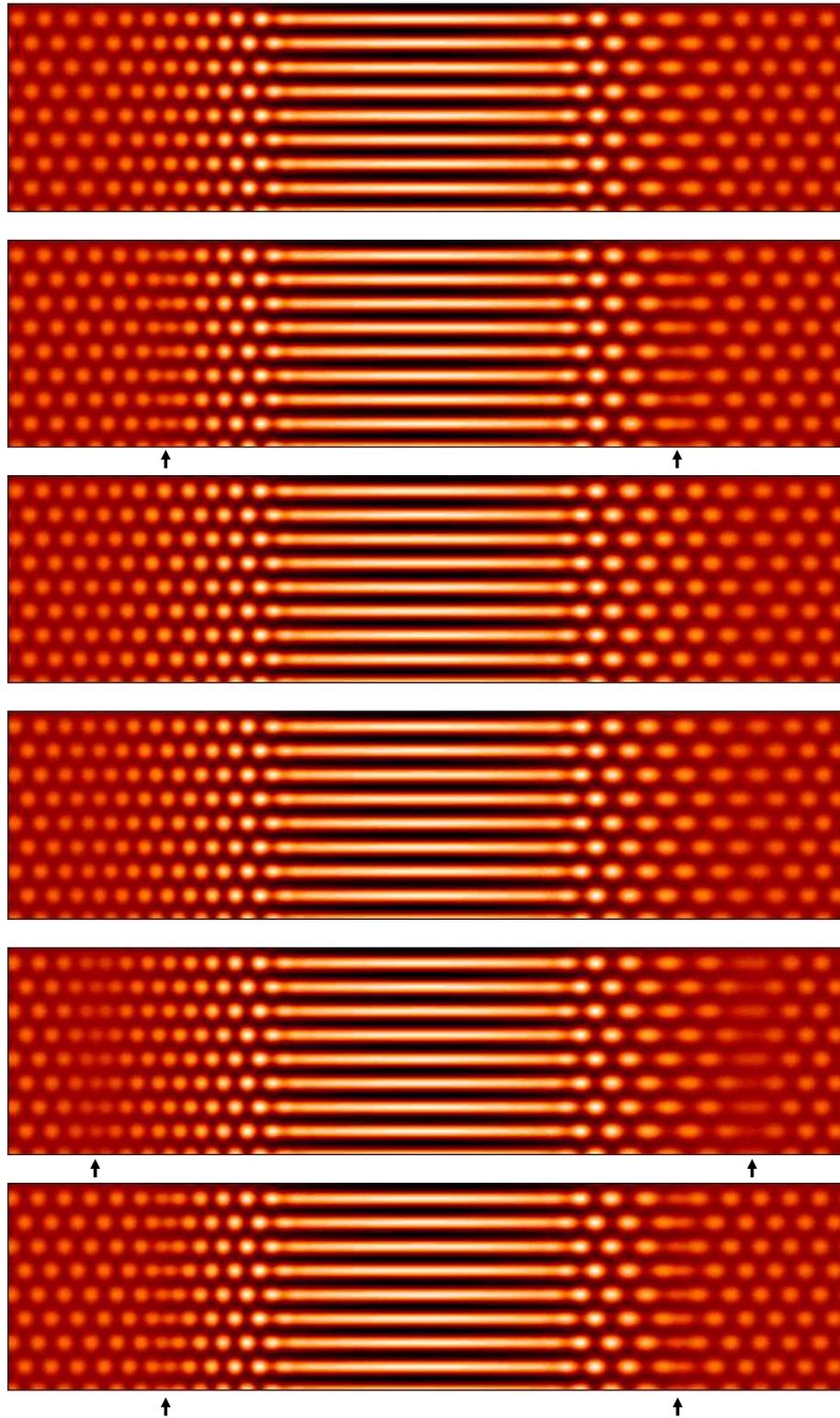


Figure 4.7: Solution for $\tilde{c} = 0.2$. Snapshots are shown every 20 time units. The arrows indicate points where the hexagons become distorted and the amplitude falls.

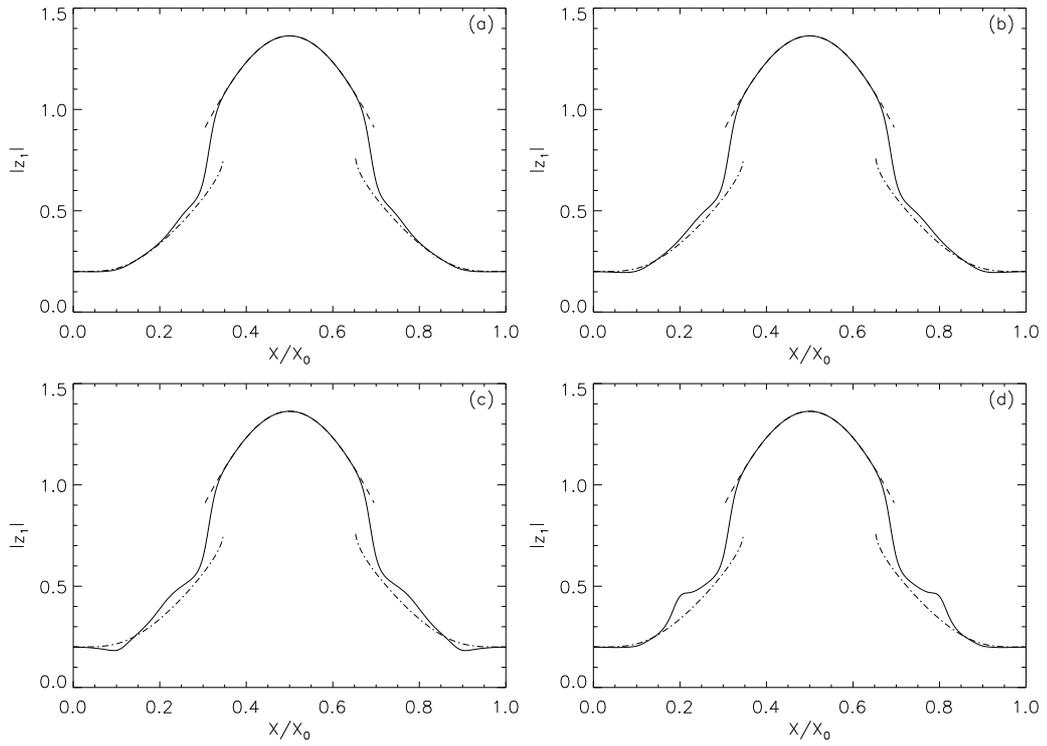


Figure 4.8: Plots of $|z_1|$ against X for the case $\tilde{c} = 0.2$. The four graphs (a)–(d) correspond to the bottom four frames from Figure 4.7.

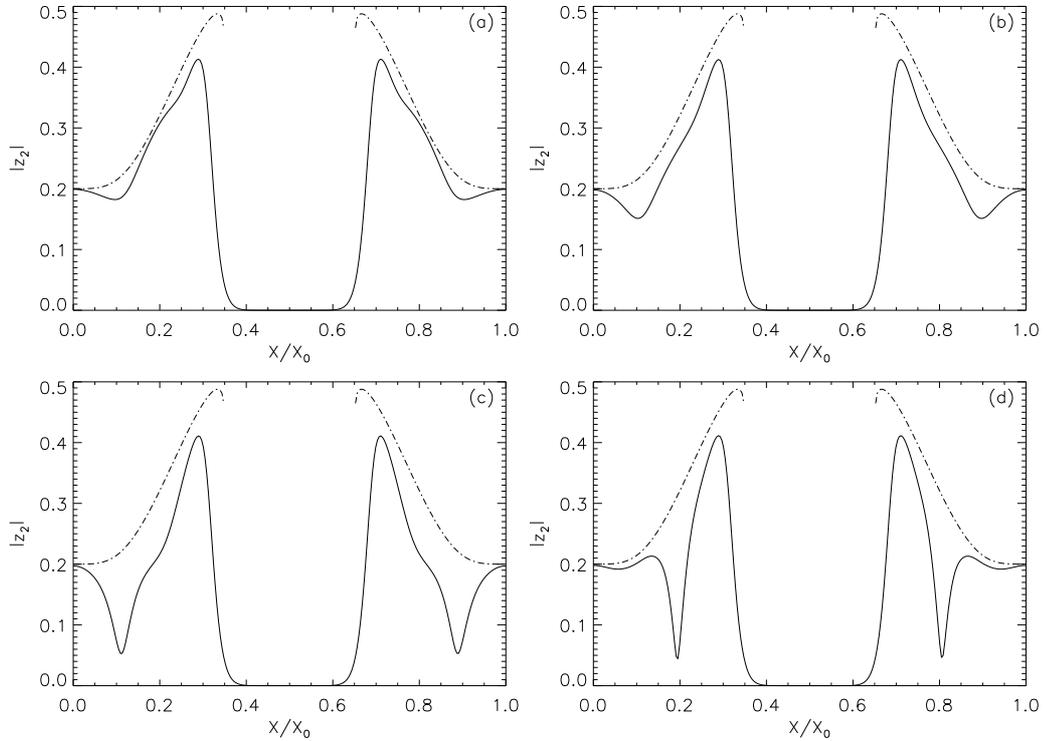


Figure 4.9: As Figure 4.8, but showing $|z_2| = |z_3|$ instead of $|z_1|$.

Equation (4.43) shows that θ increases linearly with time, at a rate proportional to $\tilde{c}\phi(x)$. This rate of increase depends on x , so the phase is ‘wound up’ over time, and gradients θ' develop.

Equation (4.42) shows that R evolves according to a diffusion equation, with a linear growth term proportional to $r_2 - D\theta^2$. Therefore, R usually grows exponentially, but if $|\theta'|$ is large enough it can start *decaying* exponentially.

Note that θ' represents a shift of the wavenumber of the basic pattern: the preferred value of the wavenumber is 1, but if θ' is nonzero then this is shifted to $1 + \theta'$. Therefore, if θ' is large, the pattern has been stretched or compressed by a large amount. (This corresponds to the ‘concertina effect’ mentioned above.) The wavenumber will be far away from its preferred value, and such modes will quickly decay.

Thus our picture is that the phases of z_2 and z_3 are ‘wound up’ with time, until eventually the wavenumber gets so far from critical that z_2 and z_3 decay sharply. This can be seen on Figures 4.9(c) and 4.9(d). By contrast, z_1 is not directly affected (since the $\tilde{c}\phi$ term does not appear in the \dot{z}_1 equation), so we see on Figure 4.8 that $|z_1|$ does not change a great deal during the evolution.

It can be seen, therefore, that this stretching and compression process has the effect of turning hexagons into rolls, with z_1 nonzero but z_2 and z_3 both very small. However, once we get close to this roll solution, we cannot stay there for long, because rolls are unstable to hexagons (at this value of X). Therefore the amplitudes $|z_2|$ and $|z_3|$ grow exponentially, and the system returns once again to hexagons (close to the critical wavenumber), restarting the cycle.

A similar pattern of motion is observed for all \tilde{c} values up to about 0.9. For smaller \tilde{c} , e.g. 0.1, the ‘concertina effect’ is observed at one value of X only, and the pattern repeats with a regular period; for larger \tilde{c} , the solution is more erratic and the ‘concertina effect’ is observed at multiple values of X .

For $\tilde{c} \gtrsim 0.9$ a different solution is observed. The system reaches a ‘quasi-steady’ state, where the amplitudes $|z_j|$ are all unchanging with time, and only the phases are time-varying. An example of such a solution (for $\tilde{c} = 1$) is shown in Figure 4.10. The hexagons in this picture travel, but only very slowly; they also expand slightly in wavelength as they move from right to left. As \tilde{c} is increased further, this state remains (values of \tilde{c} up to 20 were checked); the only change is that the region of hexagons shrinks slightly and the region of rolls expands slightly as \tilde{c} is increased.

Notice here that the rolls exist for a much larger range of X than they usually do, even

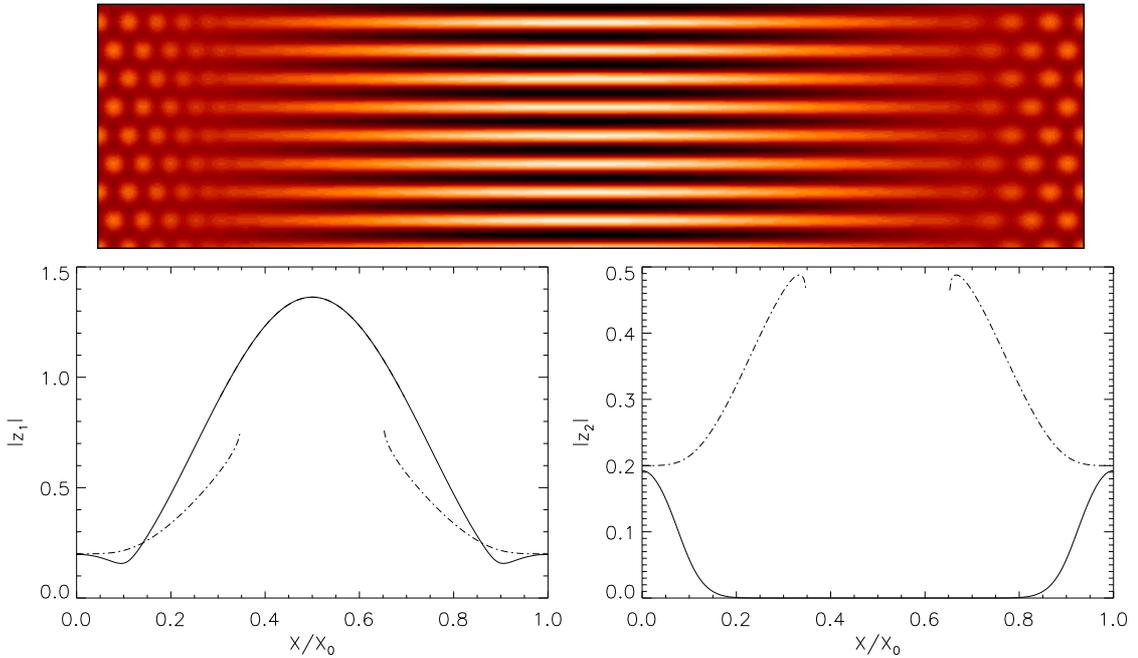


Figure 4.10: Quasi-steady solution for $\tilde{c} = 1$. The pattern is steady apart from a slow drift of the hexagons from right to left.

persisting in areas where only the hexagon branch is supposed to be stable. Although it is still true that perturbations to rolls would grow exponentially, these perturbations would also be affected by the $c_1\phi_1$ terms in equations (4.27) and (4.28). For large enough $c_1\phi_1$ it seems likely that the perturbations would be quickly ‘sheared’ away by these terms, before they had time to grow large enough to disrupt the roll solution. Therefore, it appears that the instability leading from rolls to hexagons cannot develop in regions where $c_1\phi_1$ is too large.

Additional results

Up until now we have used initial conditions consisting of purely real values for the amplitudes. We now consider the more general case of a complex initial perturbation. The results for three different values of \tilde{c} are shown in Figure 4.11. The first panel shows the result for $\tilde{c} = 0$, which is similar to the previous result (cf. Figure 4.4), except that the rolls are now slightly curved. This curvature is a relic of the initial condition and eventually fades away by diffusion, although it takes rather a long time to disappear completely, because the diffusion term in the z_1 equation is fourth order with a small diffusion coefficient.

The second panel in Figure 4.11 shows the case $\tilde{c} = 0.5$, which is again similar to

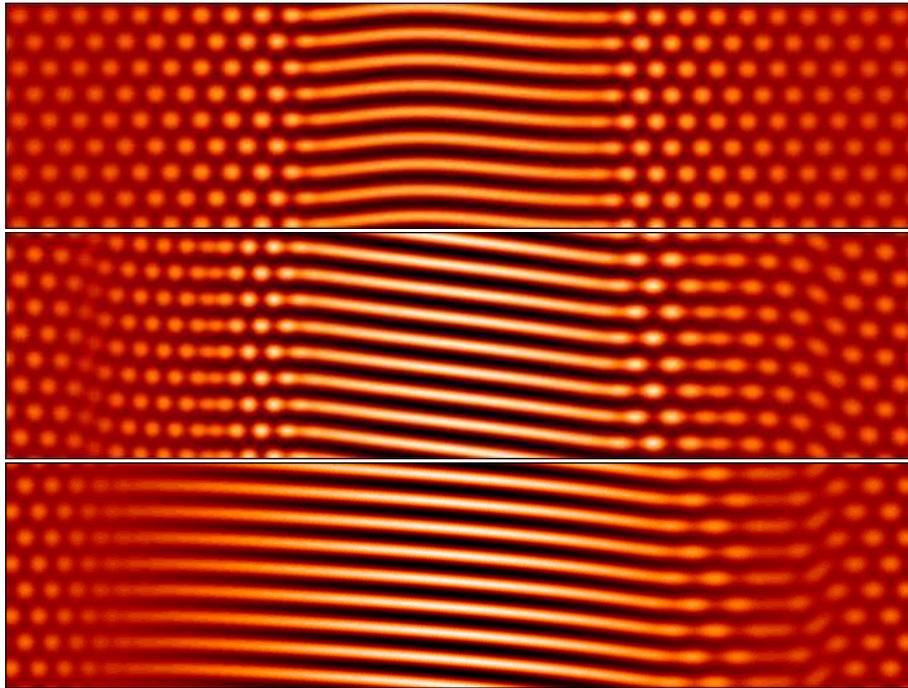


Figure 4.11: *The effect of using a complex initial condition instead of a real one. Top: $\tilde{c} = 0$, middle; $\tilde{c} = 0.5$; bottom: $\tilde{c} = 1$. As usual, the roll patterns are steady, but the hexagonal patterns travel from right to left.*

the previous results, although there is now a slightly more complicated pattern, with the possibility of curvature in both the rolls and the hexagons, as well as rolls which are not perfectly aligned with the x -axis. Finally, the bottom panel shows the case $\tilde{c} = 1$, which can be compared to Figure 4.10; the main difference here is that the rolls appear to ‘break up’ into a stretched-out cellular pattern at the right-hand side of the domain (with the left-hand side being largely unchanged).

We also briefly discuss what happens when the transition between rolls and hexagons is supercritical rather than subcritical. (This corresponds to case II, or equation 4.30, from section 4.3.1.) This is illustrated in Figure 4.12, which is the equivalent of Figure 4.4 for this case. Notice that the transition between rolls and hexagons is less sharply defined in this supercritical case (as expected); however, the difference is not all that noticeable, since the transition tends to be smoothed out by diffusion. (This does depend on our choice of ϕ though: if we chose ϕ to vary very slowly, then in the supercritical case we would see a very broad transition from rolls to hexagons, while in the subcritical case, the transition region would remain narrow. In other words, in the supercritical case the width of the transition depends directly on how quickly ϕ changes, whereas in the subcritical case the width of the transition depends only on the amount of diffusion

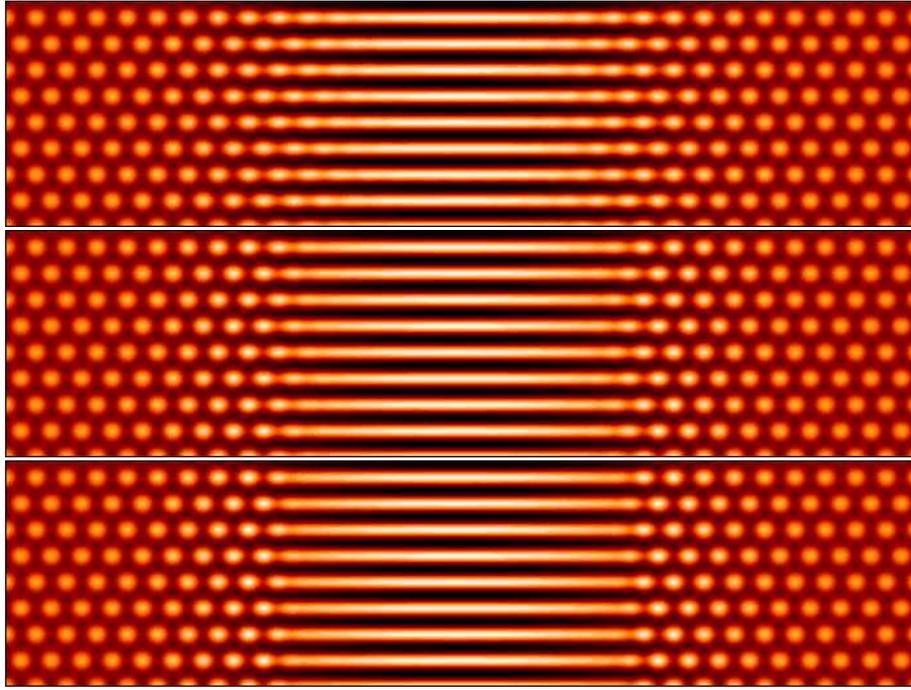


Figure 4.12: The equivalent of Figure 4.4 for the supercritical case. Notice how the transition is not quite so sharply defined as before.

in the system.)

The effects of varying ϕ

We have also looked at varying our function $\phi_1(X)$, which describes how the tilt angle varies with position. Instead of (4.31), we now use the following slightly more general definition:

$$\phi_1 = \frac{\phi_{\text{centre}}}{2} \left(1 - \cos \left(\frac{2\pi X}{X_0} \right) \right). \quad (4.44)$$

The previous definition corresponds to $\phi_{\text{centre}} = 1.5$.

So far all of our solutions have contained hexagons at the edges of the domain and rolls in the centre. This is the general situation for sufficiently large ϕ_{centre} . The other generic possibility occurs for small ϕ_{centre} and here we simply get hexagons throughout the entire domain.

There is also hysteresis between these two possibilities: both types of solution can be stable simultaneously for certain values of ϕ_{centre} . For example, this occurs for $\phi_{\text{centre}} = 1.18$, as illustrated in Figure 4.13. Here, both rolls and hexagons are stable solutions when $\phi = \phi_{\text{centre}}$, so it is possible to have either hexagons or rolls at the centre of the domain, while at the edges, the solution is always hexagons (since $\phi = 0$ at the edges of

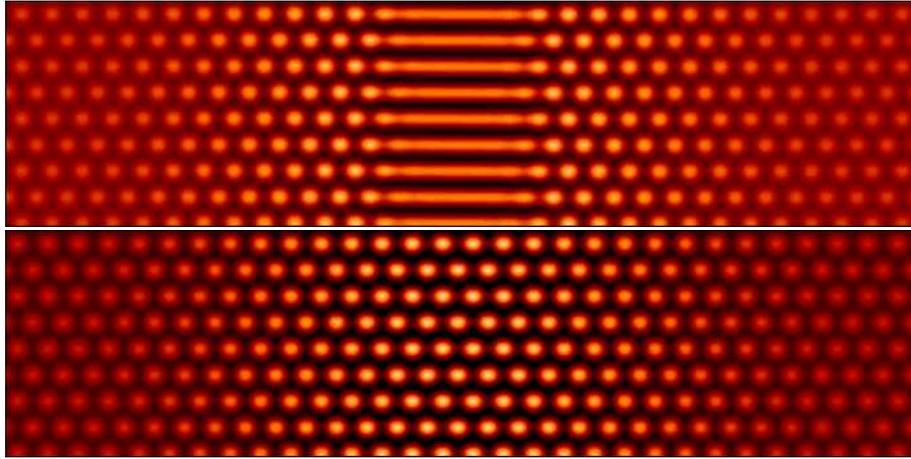


Figure 4.13: Two possible stable solutions with $\phi_{\text{centre}} = 1.18$, $r_2 = 0$ and $\tilde{c} = 0$. (These are both steady solutions.)

the domain). This leads to the two possible solutions shown in Figure 4.13; depending on the initial conditions, either of these two solutions can be obtained.

These results were obtained by setting $\tilde{c} = 0$. However, it appears that this hysteresis effect disappears once we take $\tilde{c} \neq 0$. In this case, only the upper solution of Figure 4.13 can be obtained (at least for long times). This is because of the distortion effect mentioned in section 4.3.3. As discussed there, the hexagons become stretched or compressed, because of the variation of the speed of travel with position, and eventually they are transformed into rolls. However, this time, rolls are actually stable in the centre of the domain, so the rolls do not turn back into hexagons again. Therefore, if we use the lower solution from Figure 4.13 as an initial condition, then it will be transformed into the upper solution from the same Figure, if \tilde{c} is non-zero. This is illustrated in Figure 4.14 for $\tilde{c} = 0.3$.

4.4 Numerical simulations of the full Swift-Hohenberg equation

We have also performed some simulations of the full Swift-Hohenberg equation (4.1). This model has more degrees of freedom than the Landau-Ginzburg equations (4.26)–(4.28), in a sense, because in the Swift-Hohenberg model the hexagonal lattice can be aligned at any angle, while in the Landau-Ginzburg model a particular orientation of the lattice has been assumed (see equation 4.19).

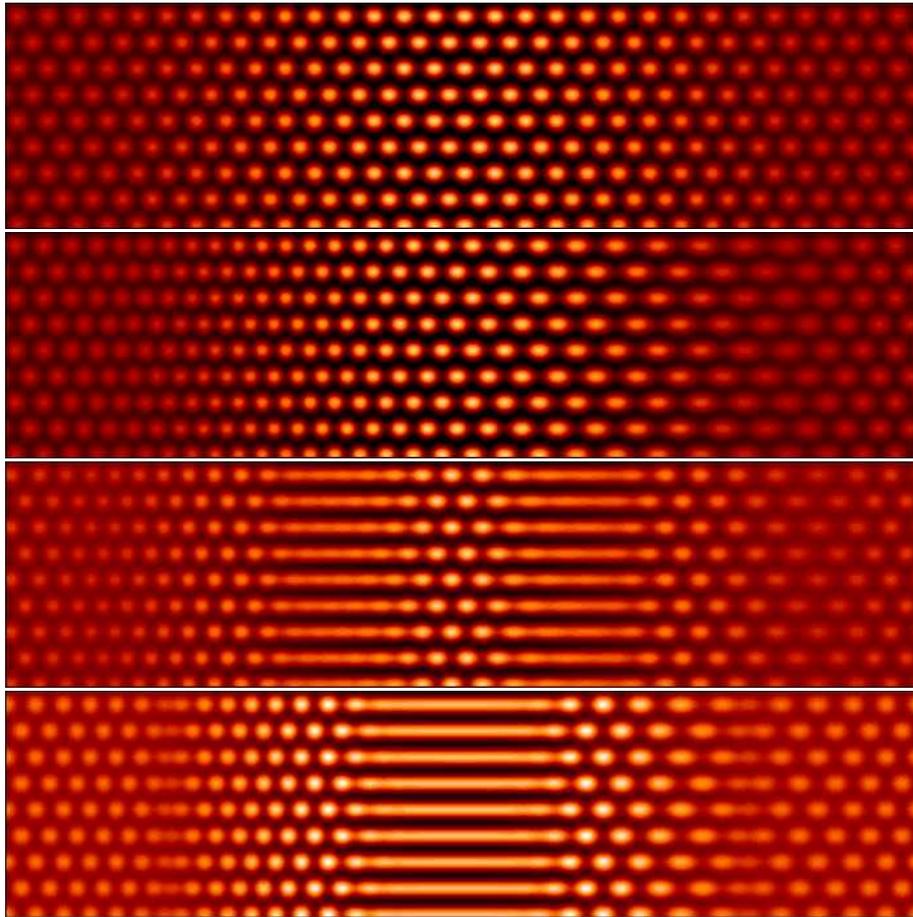


Figure 4.14: Sequence showing how the hexagons evolve into rolls when $\phi_{\text{centre}} = 1.18$, $r_2 = 0$ and $\tilde{c} = 0.3$. The top panel shows the initial condition and subsequent images show snapshots every 100 time units.

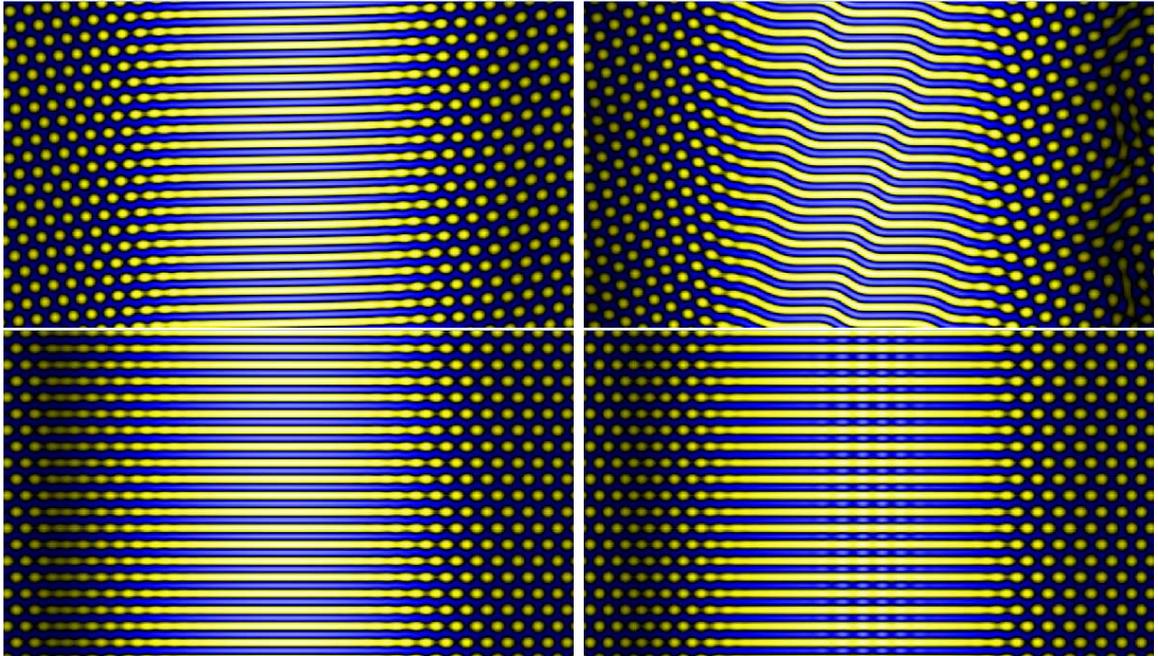


Figure 4.15: Four results from the full Swift-Hohenberg equation in rectangular geometry. The top two pictures correspond to a random initial condition, with $\tilde{c} = 0.2$ (left) and $\tilde{c} = 0.3$ (right). The bottom two pictures have a specially chosen initial condition in order to produce patterns aligned with the x axis (both with $\tilde{c} = 0.4$). The bottom right picture has a larger value of ϵ , illustrating the ‘cross rolls’ that appear at the centre in this case. Note that the hexagon patterns travel from right to left. The rolls are steady, except in the top-right hand case, where the ‘steps’ in the rolls travel leftwards.

We first look at the results in a rectangular (Cartesian) domain. In this case the results should agree with the Landau-Ginzburg model in the limit of small ϵ . In section 4.4.2 we extend the model to allow us to look at circular domains, in order to produce a better representation of real pores and sunspots (which are usually roughly circular in shape). The computations were all carried out using a modification of a code provided by S. Houghton.

4.4.1 Cartesian geometry

Figure 4.15 shows typical results in a Cartesian domain. (The parameters a and \tilde{b} were chosen as in ‘Case I’ described in section 4.3.1; we fixed $r = 0$ and looked at several values of \tilde{c} .) If ϵ is small then the results should agree with those found above (for the

Landau-Ginzburg model); the first three pictures in Figure 4.15 have $\epsilon = 0.3$ and we do indeed find similar results. However, for ϵ greater than about 0.4 to 0.5, ‘cross rolls’ start to appear at the centre of the domain (see bottom right hand picture in Figure 4.15, where $\epsilon = 0.5$). These were not seen earlier, and they might represent some sort of instability of the convection rolls. However, they could also be artifacts caused by the large value of ϵ (this model is only strictly speaking valid for small ϵ), so we must be careful in interpreting them.

Note that in the bottom two pictures, a specially chosen initial condition was used, designed to give patterns aligned with the x axis. This is the equivalent of taking a *real* initial condition in the Landau-Ginzburg model, and we indeed get good agreement with those cases. In the top two pictures of Figure 4.15, a general, random initial perturbation was used. This case does not directly correspond to the Landau-Ginzburg equations, since it tends to produce a randomly aligned lattice – and not one aligned with the x axis, as assumed in the previous model (see equation 4.19). Nevertheless, the results are similar, with a transition between hexagons at the edges and rolls in the centre of the domain, as before. One striking feature is that the rolls always manage to align themselves with the x axis (parallel to the tilt of the magnetic field), even sometimes producing a curious ‘stepped’ structure in order to achieve this (see top right picture of Figure 4.15).

4.4.2 Cylindrical geometry

We have also produced an extension of the model (4.1) in which the magnetic field’s tilt angle takes on a two-dimensional character. In other words, instead of simply imagining the field to be tilted in the x -direction, we allow the field to be tilted in any horizontal direction. This can be modelled in a simple way by allowing the ‘tilt angle’ (ϕ in the above) to become a two-dimensional vector rather than a scalar, indicating a direction as well as amount of tilt. Equation (4.1) then becomes

$$\frac{\partial\psi}{\partial t} = r\psi - (1 + \nabla^2)^2\psi + \alpha\psi^2 - \gamma\psi^3 + a|\phi|^2\psi + b(\phi \cdot \nabla)^2\psi + c\phi \cdot \nabla\psi. \quad (4.45)$$

This equation can now be simulated within a circular domain, with the magnetic tilt direction pointing radially outwards. This circular configuration is more like a sunspot or pore than our previous model, which captured the behaviour of the convection in inclined fields, but not the particular geometry of a sunspot. We expect this extension of the model to give qualitatively similar results, with a travelling pattern of hexagons

and a transition to rolls for larger tilt angles, although there might now be additional effects due to the new geometry.

It is unclear what boundary conditions should be taken around the circumference of the domain. However, since this is intended to be an idealized, qualitative model only, we will simply choose illustrative boundary conditions, namely either Dirichlet ($\psi = 0$) or Neumann ($\partial\psi/\partial r = 0$) conditions at the outer edge. The use of two different boundary conditions allows us to check which results are robust and do not depend on effects at the outer boundary.

Initial results

We chose ϕ to be a radial vector: $\phi = f(r)e_r$, with $f(r)$ varying linearly from 0 at the centre to 2.3 at the outer edge of the domain. This choice allows both rolls and hexagons to appear. (The other parameters were taken from ‘Case I’ described in section 4.3.1, with $r = 0$.) The results, for four different values of \tilde{c} , are shown in Figure 4.16. These results were obtained using the Dirichlet boundary condition, but we found no qualitative differences when the Neumann boundary condition was used instead.

The four results are actually fairly similar to each other. There is a transition from hexagons at the centre (where ϕ is small) to radial rolls around the edge (where ϕ is large), as expected. For the $\tilde{c} = 0$ case, the pattern is steady and time-independent (after an initial transient); for $\tilde{c} > 0$, the hexagons travel inward, while the rolls are more or less steady and unchanging. The rolls contain a number of dislocations, which arise because of the circular geometry. If we simply had a radial pattern of rolls, with no dislocations, then the wavelength of the rolls would be increasing with radius; however, the rolls have a preferred wavelength of 2π , and so this would not be an optimal solution. Instead, dislocations are formed, and extra rolls are created, to keep the wavelength close to 2π .

Notice that the boundary between rolls and hexagons is not a perfect circle at some particular radius. Instead, the boundary is slightly blurred and ill-defined, and is irregular in shape (particularly for smaller \tilde{c}). When $\tilde{c} > 0$ the boundary is also somewhat dynamic, with patterns near the boundary constantly changing from rolls to hexagons and back again.

This irregularly-shaped boundary is a little puzzling, given that our calculation in section 4.3.2 predicted that the front would be located at a particular, well-defined value of ϕ , which in the present model would correspond to a circle at a particular

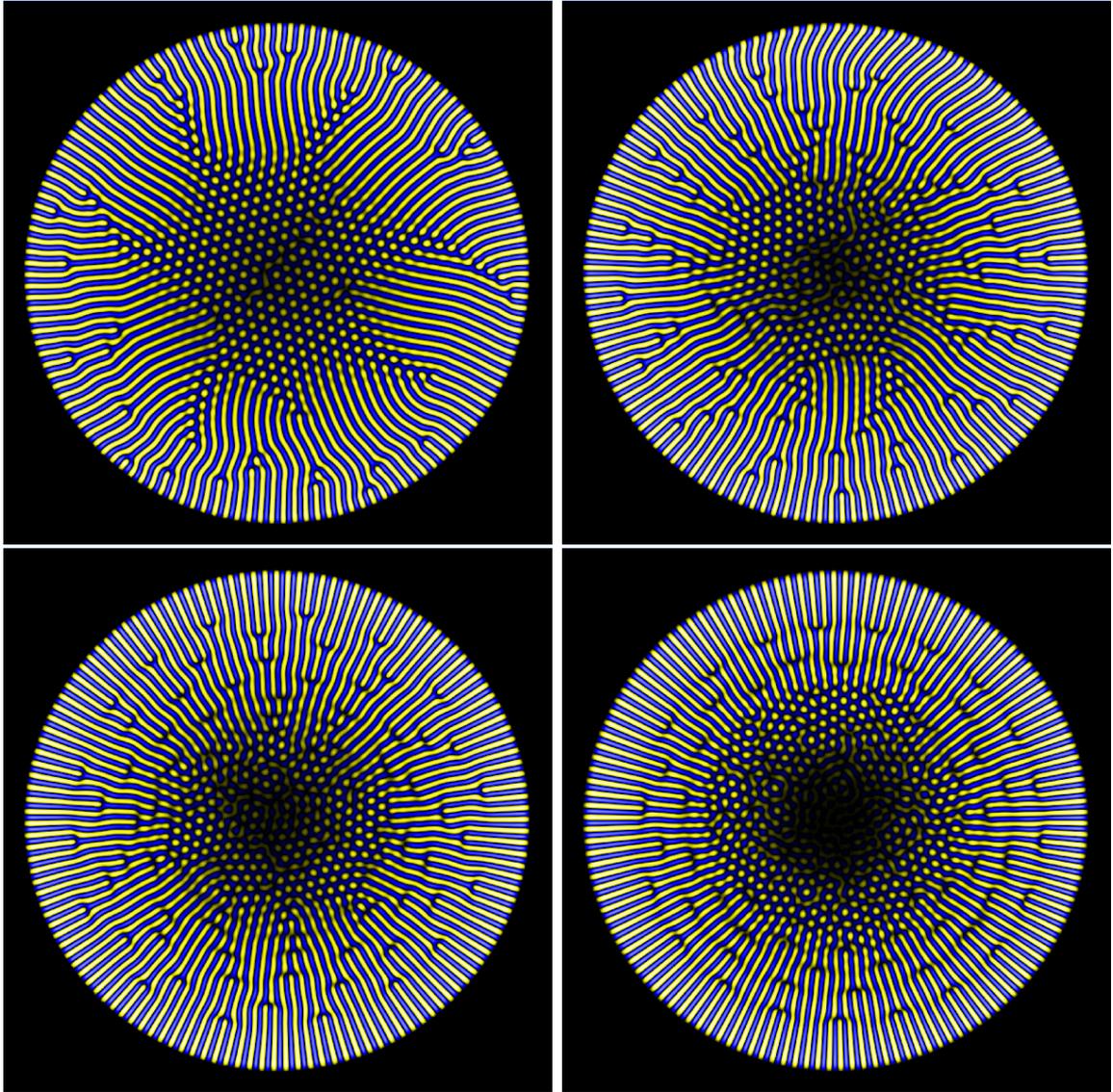


Figure 4.16: Results from the cylindrical model (equation 4.45). Top left: $\tilde{c} = 0$; top right: $\tilde{c} = 0.3$; bottom left: $\tilde{c} = 0.6$; bottom right: $\tilde{c} = 1.2$. In these pictures, the rolls are steady, while the hexagon pattern is time-dependent (with individual hexagons travelling towards the centre).

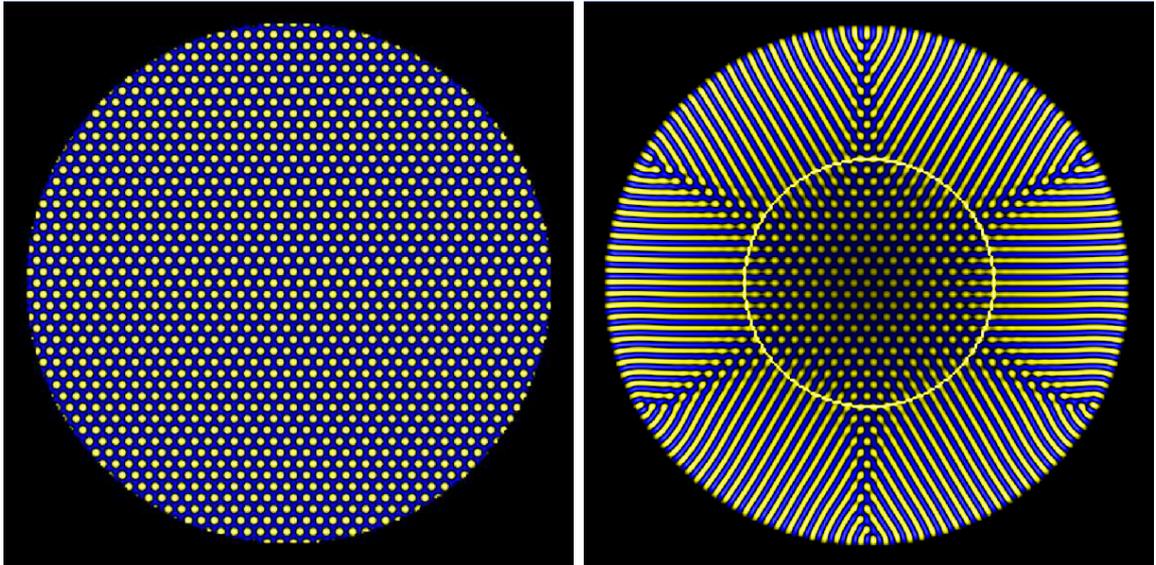


Figure 4.17: Results with $\tilde{c} = 0$ and with a specially chosen initial condition (left), designed to give a particular alignment of the hexagonal lattice. The final state is shown on the right. (These are both steady solutions.) The circle shows the expected position of the front between rolls and hexagons, based on the calculations in section 4.3.2.

radius. However, note that in that calculation, a certain particular alignment between the hexagonal lattice and the magnetic tilt vector was assumed. The rolls had to be aligned along the tilt vector and so did the hexagonal lattice (see equation 4.19). In the present results (e.g. Figure 4.16, top left picture), this condition is not always satisfied; the angle between the hexagonal lattice and the magnetic tilt direction changes with position. We can assume that the location of the front will also vary slightly, depending on this angle. Therefore, we end up with an irregularly shaped front, whose radius depends on position, as seen in the results.

As a way of illustrating this point, we have produced a result using a much more regular initial condition, shown in Figure 4.17, left-hand panel. (Note also that we set $\tilde{c} = 0$ for this result, in order to prevent any travelling wave effects from destroying the regularity of the pattern.) The initial condition is stable in the centre but unstable at the circumference of the domain, so the system rapidly evolves into the steady state shown in the right-hand panel of Figure 4.17.

Notice how the final pattern is now much more regular and symmetric than in the previous results. Correspondingly, the front between hexagons and rolls also has a much more regular structure than before. The position of the front now agrees well with the

predicted position (based on the Landau-Ginzburg calculation of section 4.3.2), which has been plotted as a circle on Figure 4.17.

Increasing the domain size with time

An interesting experiment is to allow the radius of the domain, as well as the corresponding tilt angle at the outer edge, to increase with time. This can be thought of as modelling a pore that is slowly growing with time; initially, it will contain cellular, hexagon-like convection, but once the radius goes above a certain value, the tilt angle at the outer edge will become large enough to allow rolls to appear. To model this situation we have taken our usual distribution for ϕ ($\phi = \phi_0 r e_r$, with ϕ_0 left constant) but have allowed r_{\max} to grow with time. We take a linear growth with $r_{\max} = r_0 + r_1 t$.

We ran this calculation using three different values of r_1 and with a fixed value for \tilde{c} of 0.3. A typical result is shown in Figure 4.18. Circles have also been plotted on the diagram, at radii corresponding to $\phi = \phi_{\text{pf}}$ and $\phi = \phi_{\text{sn}}$ (based on the calculations of the previous chapter), so that rolls are stable outside the outer circle, hexagons are stable inside the inner circle, and both patterns are stable in between the two circles. (Note however that the values of ϕ_{pf} and ϕ_{sn} will vary slightly, depending on the chosen orientation of the lattice – the variable θ from the previous chapter – and so the circles should be taken as a guide only.)

When the domain is small, we find a solution consisting only of hexagons (e.g. top left picture in each case), and when the domain grows larger, the solution has hexagons in the centre and rolls around the edge (e.g. bottom left and bottom right pictures). We might identify the former as a ‘pore-like’ solution and the latter as a ‘sunspot-like’ solution.

Note that one might expect the transition from hexagons to rolls to occur as ϕ increases through ϕ_{sn} , i.e. as the boundary moves through the outermost of the two circles plotted on the diagrams. In other words, we might expect the solution to follow the hexagon branch until the very end, only switching to rolls after the hexagons finally become unstable. However, this appears not to happen. Instead, the rolls seem to appear as soon as ϕ reaches ϕ_{pf} , i.e. as the boundary moves through the innermost circle, as seen in the top right picture of Figure 4.18. The reason for this is unclear, but it might be due to boundary effects which cause the roll pattern to be favoured near the boundary.

We have also performed runs in which the domain shrinks with time instead of ex-

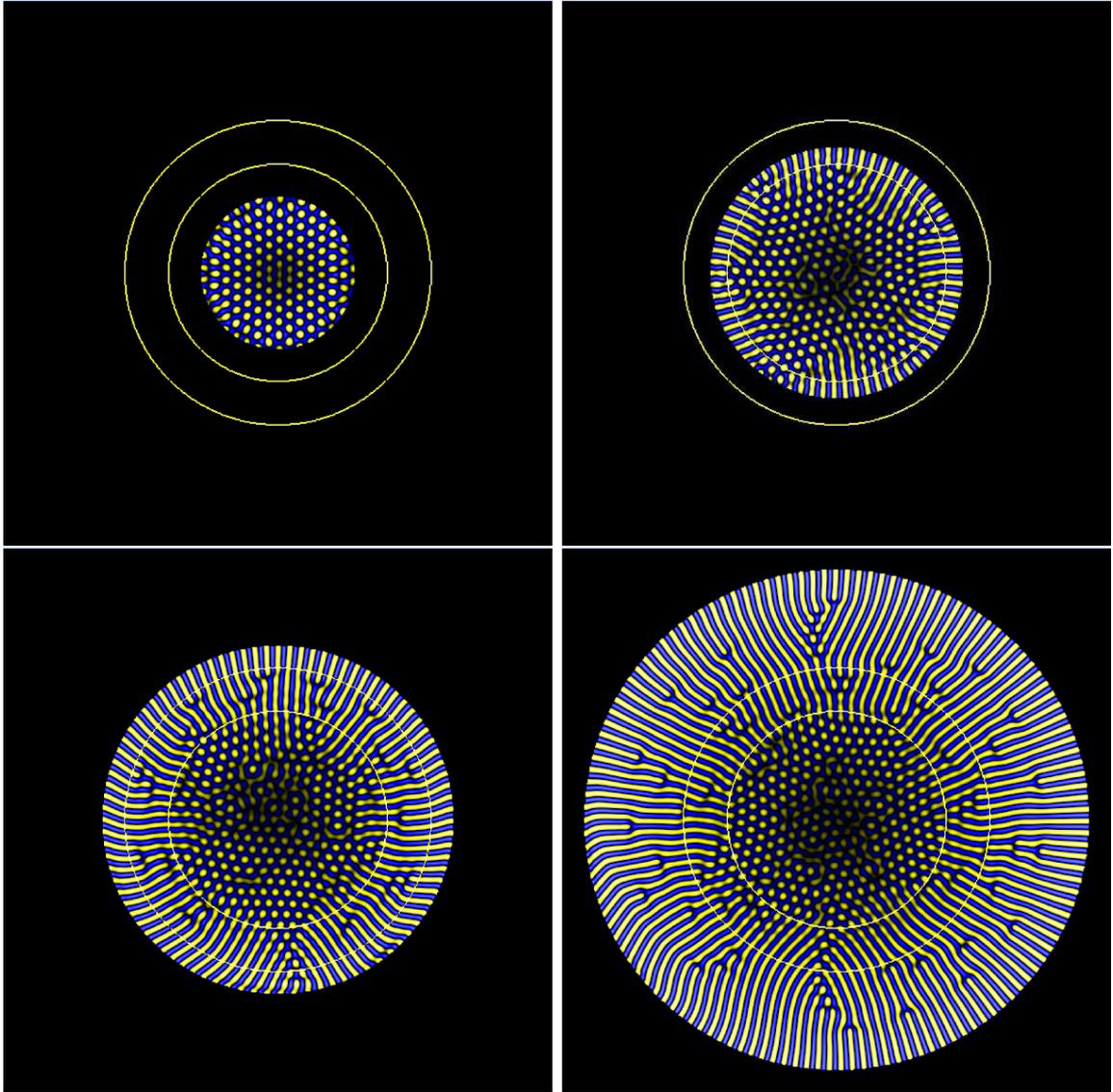


Figure 4.18: A run in which the radius of the domain has been allowed to increase slowly with time. (Snapshots are shown at four different times, with time increasing from left to right and top to bottom.) The Neumann boundary condition has been used. The two circles indicate the radii at which $\phi = \phi_{\text{pf}} = 1$ and $\phi = \phi_{\text{sn}} = 1.4$. Note that the rolls first appear (roughly) as the boundary expands through the innermost circle (see top right hand picture).

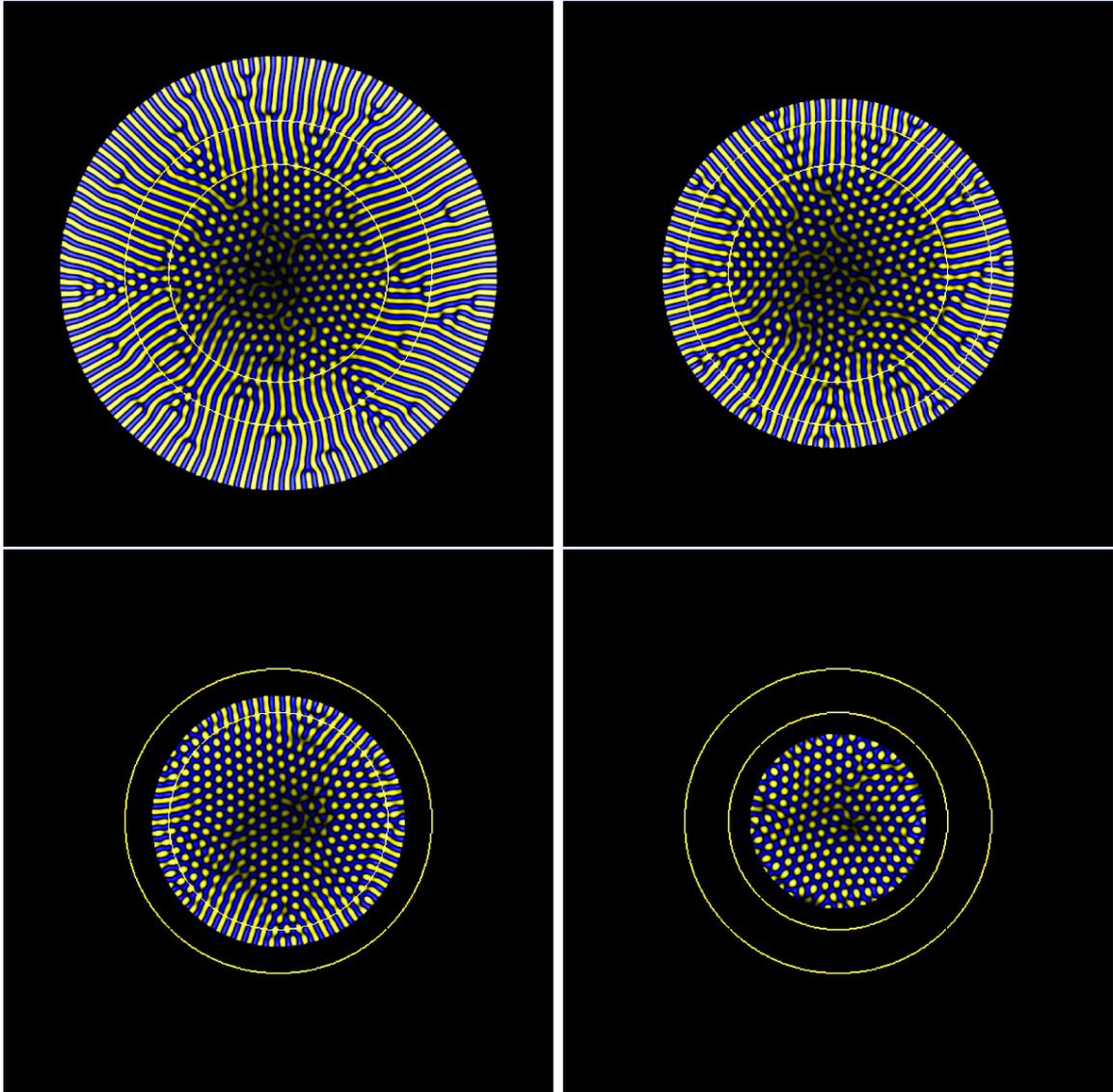


Figure 4.19: As Figure 4.18 except that the domain is now contracting, rather than expanding, with time. We see more or less the same sequence of events in reverse. In particular, the rolls disappear at (more or less) the same point at which they appeared, i.e. as the domain boundary passes through the innermost circle. In other words, there is little or no hysteresis between the point where the rolls appear in an ‘expanding’ run and where they disappear in a ‘contracting’ run.

pands (see Figure 4.19). Here we see more or less the same sequence of events in reverse, with the transition from ‘sunspot-like’ to ‘pore-like’ configurations again occurring as the boundary moves through the inner circle (corresponding to $\phi = \phi_{\text{pf}}$). In other words, the model shows no hysteresis between the two configurations as the domain is expanded and shrunk again.

Notice that the rolls in these images contain a number of dislocations along their length. We have already seen this phenomenon (e.g. Figure 4.16), and it can be explained as a way of maintaining the preferred wavelength of 2π despite the circular geometry. This pattern of rolls is seen in most of our runs, but we also find one other pattern in certain cases, and that is a spiral pattern; see Figure 4.20.

To obtain the spirals one must set the problem up quite carefully. In particular, they seem to arise only when the domain expands very slowly (and even then, only in some cases). Otherwise, the straight radial rolls will be seen. The spiral pattern is therefore an interesting quirk of the model, but may not be related to any particular phenomenon in sunspots.

Finally, note that there is a different time dependence between the two cases (Figures 4.18 and 4.20). In both cases, the hexagons travel inwards, but the behaviour of the rolls is different. In the former case (with straight radial rolls) the rolls are steady; in the latter case, the spiral pattern is observed to rotate with time. The spiral patterns can rotate because they break the rotational symmetry of the system. In addition, the pattern is drifting inwards at each point (for $c \neq 0$), which will create the impression that the entire pattern is rotating.

4.5 Conclusions

In this chapter we have constructed a modified Swift-Hohenberg equation that can be used to model convection in an inclined field where the inclination angle ϕ varies across the domain (in a predefined way). We have also derived a set of three coupled Landau-Ginzburg equations from this. This allowed the transition between rolls in one part of the domain and hexagons in another part to be investigated. The equations were studied in both Cartesian and circular domains – the former being easier to work with theoretically, and the latter being closer to a real sunspot.

The Cartesian results demonstrate two main points. The first is that there can indeed be a sharp transition, or front, between the hexagon and roll patterns (in the case where

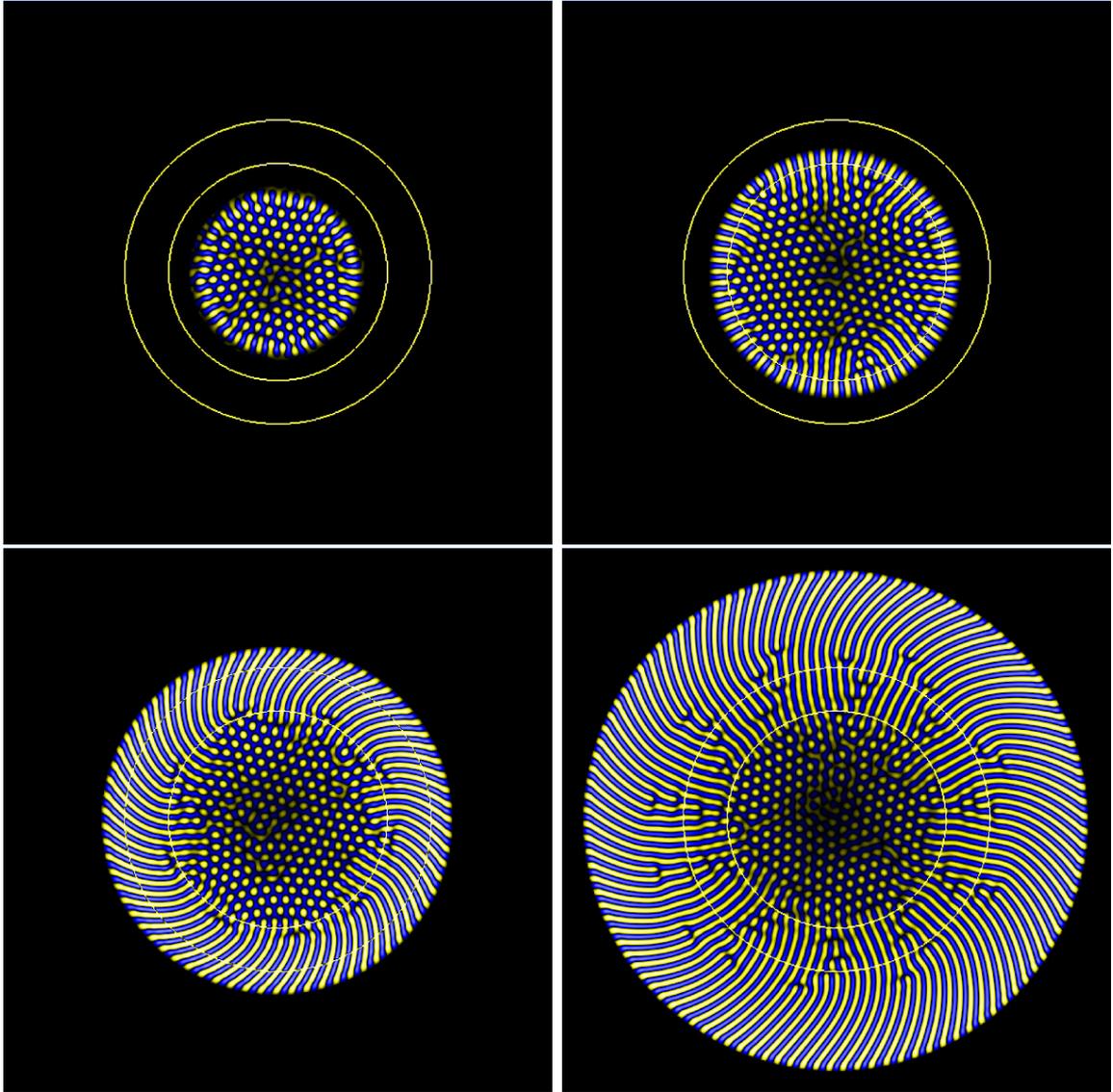


Figure 4.20: As Figure 4.18, except that here the domain has been expanded much more slowly than before (and also the Dirichlet boundary condition has been used). The rolls have formed a coherent spiral pattern, which appears to rotate with time (in a clockwise direction).

the bifurcation from hexagons to rolls is subcritical). The sharpness of the transition is limited only by the amount of diffusion in the system. The other effect demonstrated by the Cartesian results is that the hexagonal pattern travels, with a speed that depends on position; if ϕ decreases along the direction of travel, as in a sunspot umbra, then this causes the pattern to become compressed with time (since the ‘back’ of the pattern is moving faster than the ‘front’), eventually causing adjacent pairs of hexagons to be ‘squashed’ together.

More immediately applicable to sunspots are the results in the circular domains. Here, we find hexagonal convection in the centre of the domain, corresponding to the umbra of a sunspot, and rolls near the circumference, which would be the penumbra. The rolls are oriented radially, but because of the circular geometry, there cannot simply be a pattern of radial rolls without dislocations. Instead we find either a spiral pattern of curved rolls (which rotates with time, since the pattern is travelling inwards at each point), or, more commonly, a static pattern of straight radial rolls, with a number of dislocations.

These results are interesting because they show that certain features of sunspots can be understood using only a relatively simple model. In other words, whilst many sunspot phenomena undoubtedly require knowledge of the underlying solar physics for their explanation, there are nevertheless some aspects that can be understood in terms of simplified models using only minimal assumptions. Indeed, our model does not even use the MHD equations, relying instead on simple symmetry-based arguments, and yet it produces reasonably ‘sunspot-like’ results, if we interpret the hexagonal regions in our models as the ‘umbra’, and the regions containing rolls as the ‘penumbra’.

Among the features of real sunspots illustrated by the model is the fact that patterns tend to travel. Our model shows that the hexagonal patterns within the umbra tend to travel inwards; in real sunspots, travelling features are observed within both the umbra and the penumbra. This motion is naturally explained as a pattern drift, caused by symmetry breaking induced by the tilted magnetic field. Note that the direction of travel is not predictable from symmetry arguments alone – we know that patterns must travel, but we cannot say in which direction they will travel without knowing more about the underlying physics.

Another interesting feature of the model is that it shows quite clearly a transition in the pattern of convection from hexagons to rolls. This transition seems to be a fundamental property of magnetoconvection in inclined fields and its existence was

demonstrated in the weakly nonlinear models of the previous chapter. The results in this chapter illustrate how this can quite naturally lead to a two-component sunspot with both an umbra and a penumbra, if the magnetic field is sufficiently tilted at the edge of the spot. No special physics is needed; the transition occurs naturally as a result of the (minimal) assumptions built into the model.

As would be expected given the nature of the model, there are also many details of real sunspots that are *not* captured by our results. For example, the model only applies to weakly nonlinear convection; in the strongly nonlinear conditions found in the Sun, we would expect more complicated, and possibly more turbulent, versions of the hexagon and roll patterns to be present (although it seems likely that the transition from hexagonal to roll-like patterns would still exist in some form).

Another example is found in the penumbra: while our model ‘penumbrae’ consist only of simple radially-oriented convection rolls, real penumbrae have a more complicated structure, with distinct bright and dark filaments and an intricate interlocking-comb magnetic field configuration. Clearly, this magnetic fine structure requires a more complicated model for its explanation.

Indeed, our model does not allow for the possibility of azimuthal variations in the magnetic field inclination. Instead, the magnetic field configuration is specified in advance (by giving the vector $\phi(\mathbf{x})$) and is assumed to be unchanging during the calculation. In reality, we would expect the field to be rearranged by the convection. For example, Weiss et al. (2004) have suggested that the penumbral fine structure starts to form because of an instability leading to a fluted field configuration. If this is correct, then our model would represent a sort of ‘proto-sunspot’ configuration, at a stage before this instability has taken hold.

Another difference between our model and reality is that hysteresis between pores and sunspots is observed (the largest pores are larger than the smallest sunspots), but no similar effect is found in the model (see Figures 4.18 and 4.19). However, this is perhaps not surprising given that many important physical effects are not included in the model. For example, Weiss et al. (2004) suggest that flux pumping may be an essential part of the explanation. (See also Rucklidge et al. 1995 for a simple method of modelling this hysteresis.)

In summary, we see that the model presented in this chapter, a model based on simplified equations and including very little in the way of solar physics, can nevertheless produce results that are reasonably sunspot-like. It illustrates how some of the basic

properties of sunspots can be explained using only simple symmetry-related arguments – for example, a transition between umbra and penumbra seems almost inevitable, given the transition from hexagon-like convection to roll-like convection that must occur as the field becomes more tilted. Of course there remain many details that the model gets wrong, and that require more intricate explanations, based on detailed solar physics models: for example, the direction of travel of umbral and penumbral features (either inwards or outwards), or the complex fine-scale structure of the penumbra. These can ultimately only be explained by performing three-dimensional numerical simulations, and this is the subject that we will turn to in the next chapter.

Chapter 5

Numerical simulations

5.1 Introduction

Having looked at a number of simplified models, we are now ready to look at full three-dimensional numerical simulations. Although still idealized in many ways, these simulations will be a step closer to reality than the kind of models we have considered in previous chapters. Of course, this extra realism is balanced by the long run times required by three-dimensional computations, meaning that only a few isolated points in the parameter space can be examined, rather than being able to plot complete bifurcation diagrams, as we have done previously.

We will begin, in section 5.4, by looking at simulations with simple uniform magnetic fields. This will allow us to study the basic phenomena of magnetoconvection in inclined fields, and to look at what changes occur as the inclination angle is varied. These simulations will correspond to the type of situation considered in Chapters 2 and 3. Our aims will be to confirm the results from those chapters, and to examine the new behaviour that occurs in more strongly nonlinear regimes.

In a (relatively strong) vertical field, previous calculations have shown that a stationary hexagonal convection pattern will be found. (See section 1.5 from the Introduction for more information and references.) We have discussed in previous chapters the changes that would be expected to this pattern once the field becomes tilted. Of these, the main ones were that we would expect the cells to become elongated along the direction of tilt of the field for moderate tilt angles, and we would expect the pattern to travel, either towards or against the tilt. For larger tilt angles, a pattern of rolls would be expected.

The calculation of Hurlburt et al. (2000) appears to confirm these expectations. Their simulations show a travelling pattern of cells for $\phi = 22^\circ$, and a more roll-like pattern for $\phi = 67^\circ$. However, their simulations used relatively small boxes, and only three different runs were shown. In this chapter a larger number of runs is presented, and this reveals a clearer picture of which patterns occur in which areas of parameter space.

In sections 5.5 and 5.6 we move on to more complicated situations in which the angle of inclination of the field varies with position. This is motivated by the structure of sunspots, and in particular the sharp distinction observed between the umbra and the penumbra. We conjecture that this is ultimately caused by a change in the nature of convection that occurs once the field becomes sufficiently inclined to the vertical. In order to investigate this conjecture, therefore, we would like to answer the following question: Is it possible to produce a model in which there is a sharp distinction between two different forms of convection in two different regions of the computational domain, one where the field is near-vertical, and one where the field is near-horizontal?

5.2 Definition of the problem

In this chapter, we will use a standard model of compressible magnetoconvection, as described (for example) by Matthews et al. (1995); Bushby and Houghton (2005). This model assumes that the following quantities are all constant: the heat capacities c_p and c_v , the thermal conductivity K , the dynamic viscosity μ , the magnetic diffusivity η , the permeability μ_0 and the acceleration due to gravity g . The calculation is performed within a Cartesian box of depth d and with periodic boundary conditions in the horizontal directions. The lengths of the box in the x and y directions will be $\lambda_x d$ and $\lambda_y d$ respectively, where λ_x and λ_y are the (dimensionless) *aspect ratios*. The box is filled with conducting plasma which is initially static, with a known temperature T_0 and density ρ_0 at the top of the box.

The equations of motion will be expressed in non-dimensionalized form with lengths being scaled by d (the depth of the box), densities by ρ_0 , temperatures by T_0 and magnetic fields by the initial magnetic field strength, B_0 (the details of the initial magnetic field setup will be discussed further below). Time will be measured in units of the isothermal sound travel time at the top of the layer, $d/(R_* T_0)^{1/2}$, where R_* is the gas constant (equal to $c_p - c_v$).

The non-dimensionalized governing equations are as follows: the equation of mass conservation

$$\frac{\partial \rho}{\partial t} = -\nabla \cdot (\rho \mathbf{u}), \quad (5.1)$$

the momentum equation

$$\frac{\partial}{\partial t}(\rho \mathbf{u}) = -\nabla \left(p + \frac{F \mathbf{B}^2}{2} \right) + \nabla \cdot (F \mathbf{B} \mathbf{B} - \rho \mathbf{u} \mathbf{u} + \sigma \kappa \boldsymbol{\tau}) + \theta(m+1)\rho \mathbf{e}_z, \quad (5.2)$$

the induction equation

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \wedge (\mathbf{u} \wedge \mathbf{B}) + \zeta_0 \kappa \nabla^2 \mathbf{B} \quad (5.3)$$

and the heat equation

$$\frac{\partial T}{\partial t} = -\mathbf{u} \cdot \nabla T - (\gamma - 1)T \nabla \cdot \mathbf{u} + \frac{\gamma \kappa}{\rho} \nabla^2 T + \frac{\kappa(\gamma - 1)}{\rho} \left(\frac{\sigma \boldsymbol{\tau}^2}{2} + F \zeta_0 \mathbf{j}^2 \right), \quad (5.4)$$

together with the equation of state

$$p = \rho T \quad (5.5)$$

and Maxwell's equation

$$\nabla \cdot \mathbf{B} = 0. \quad (5.6)$$

(Note that we use a coordinate system in which z increases downwards.) The symbol \mathbf{j} represents the current density ($\mathbf{j} = \nabla \wedge \mathbf{B}$) while $\boldsymbol{\tau}$ is the viscous stress tensor, defined as follows:

$$\tau_{ij} = \frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i} - \frac{2}{3} \delta_{ij} \frac{\partial u_k}{\partial x_k}. \quad (5.7)$$

The non-dimensionalization introduces a number of dimensionless parameters, and these are summarized in Table 5.1.

In the absence of convection, these equations admit a static trivial solution where $\mathbf{u} = 0$ and ρ and T take the form of a polytrope of index m :

$$T = 1 + \theta z \quad (5.8)$$

$$\rho = (1 + \theta z)^m \quad (5.9)$$

$$p = (1 + \theta z)^{m+1}. \quad (5.10)$$

We will use this state as the initial condition for our simulations (together with a small random perturbation added to the temperature field). We have not yet specified the initial condition for \mathbf{B} , but note for now that we may take \mathbf{B} to be any field satisfying $\nabla \cdot \mathbf{B} = 0$ and $\nabla \wedge \mathbf{B} = 0$ (this ensures that there is no Lorentz force to disturb the static state, and that $\partial \mathbf{B} / \partial t = 0$ in the induction equation).

<i>Parameter</i>	<i>Symbol</i>	<i>Definition</i>	<i>Value</i>
Ratio of specific heats	γ	c_p/c_v	5/3
Prandtl number	σ	$\mu c_p/K$	1
Ratio of magnetic to thermal diffusivity (at $z = 0$)	ζ_0	$\eta c_p \rho_0 / K$	0.2
Polytropic index	m	$gd/R_*\Delta T - 1$	1
Dimensionless temperature difference	θ	$\Delta T/T_0$	10
Dimensionless thermal conductivity	κ	$K/d\rho_0 c_p (R_* T_0)^{1/2}$	variable
Dimensionless magnetic field strength	F	$B_0^2/R_* T_0 \rho_0 \mu_0$	variable

Table 5.1: Dimensionless parameters and their meanings. Note that ΔT is the imposed (dimensional) temperature difference between the top and the bottom of the layer.

We now discuss the values chosen for the parameters in Table 5.1. Note first of all that due to the large number of parameters, it would be impossible to conduct a full survey of the parameter space. Therefore, we have chosen fixed values for most of the parameters, and these values are shown in the final column of the Table. For example, γ is chosen to be the ratio of specific heats for an ideal monatomic gas (as appropriate for the Sun). The other parameters are largely arbitrary; we use values consistent with the work of Rucklidge et al. (2000). We note that the choice of $\theta = 10$ implies a fairly deep layer, in which the temperature varies by a factor of 11 from top to bottom. Note also that the effective value of ζ is a function of depth; the choice $\zeta_0 = 0.2$ corresponds to a value of 0.2 at the top of the layer, 1.2 at mid-layer and 2.2 at the bottom. Previous calculations (e.g. Weiss et al., 1990; Rucklidge et al., 2000) have shown that in this situation convection is steady, rather than oscillatory, at onset (although oscillatory behaviour can appear above onset via secondary bifurcations).

The remaining parameters – κ and F – are allowed to vary between different runs. It is more convenient to express these in terms of the Rayleigh number R and Chandrasekhar number Q , defined as follows:

$$R = \frac{(m+1)\theta^2}{\sigma\kappa^2\gamma}(m+1-m\gamma)(1+\theta/2)^{2m-1}, \quad (5.11)$$

$$Q = \frac{F}{\zeta_0\sigma\kappa^2}. \quad (5.12)$$

(Given values for R and Q , these equations can be inverted to find the appropriate

values for κ and F .)

At the top and bottom surfaces, we use standard impenetrable, stress-free boundary conditions, with $u_z = \partial u_x / \partial z = \partial u_y / \partial z = 0$. The temperature will be fixed ($T = 1$ at $z = 0$, and $T = 1 + \theta$ at $z = 1$). For the magnetic field, we will match to a potential field above the layer, while ‘tying’ the field lines at the lower boundary (i.e. the top boundary is insulating, and the lower boundary is conducting). As was discussed in Chapter 2, this seems an appropriate choice of boundary conditions for the Sun. (The equations for these magnetic boundary conditions are derived in Appendix A.)

5.3 Numerical methods

In this section we describe the methods used to solve (5.1)–(5.4) numerically.

Any numerical code for magnetohydrodynamics must have some method for ensuring that $\nabla \cdot \mathbf{B} = 0$. We can see from the induction equation that provided $\nabla \cdot \mathbf{B}$ is zero initially, it will, mathematically speaking, remain zero for all time. However, when implementing the equation on a computer, we will find, if we are not careful, that numerical errors in the simulation build up and lead to a non-zero $\nabla \cdot \mathbf{B}$ being produced after a certain time.

One way to combat this problem is to write the induction equation in conservative form, as $\partial \mathbf{B} / \partial t = -\nabla \wedge \mathbf{E}$. In theory, solving the equation this way means that $\nabla \cdot \mathbf{B}$ will be conserved to machine precision (because $\nabla \cdot (\nabla \wedge \mathbf{E}) = 0$). However, while the method works well within the interior of the layer, there can be problems at the top and bottom boundaries. These problems arise when numerically evaluating the required vertical derivatives of \mathbf{E} ($\partial E_x / \partial z$ and $\partial E_y / \partial z$) – if this is not done carefully, and in a manner consistent with the boundary conditions, then errors will be introduced and $\nabla \cdot \mathbf{B}$ will grow. For simple vertical field boundary conditions ($B_x = B_y = 0$), this will not be a problem (since we know $\partial E_x / \partial z = \partial E_y / \partial z = 0$ in such cases), but in general we will not be able to write down explicit boundary conditions for $\partial E_x / \partial z$ and $\partial E_y / \partial z$.

A more robust approach is to re-write the equations in poloidal-toroidal form. This basically means that \mathbf{B} is written as

$$\mathbf{B} = \nabla \wedge (B_T \mathbf{e}_z) + \nabla \wedge (\nabla \wedge (B_P \mathbf{e}_z)) \quad (5.13)$$

where B_T and B_P are the toroidal and poloidal potentials respectively. We can now derive induction equations describing the evolution of B_T and B_P , instead of B_x , B_y

and B_z (we also have to add a ‘mean field’ term, $\bar{\mathbf{B}}(z)$; see Appendix B for the details). The advantage of doing this is that it is now *impossible* for $\nabla \cdot \mathbf{B}$ to become non-zero – whatever values B_T and B_P take, the form of (5.13) is such that we will have $\nabla \cdot \mathbf{B} = 0$. The only disadvantage of this method is that the simulation will run slightly more slowly, since a few extra spatial derivatives need to be computed at each timestep.

Note that the other equations (the continuity, momentum and heat equations) do not change when using the poloidal-toroidal formulation. Formally, we should substitute (5.13) into (5.1)–(5.4), but in fact it is more efficient to compute \mathbf{B} once, using equation (5.13), and then simply to use that value wherever \mathbf{B} is needed in the other equations.

Once the equations have been rewritten in poloidal-toroidal form, they need to be discretized. For the spatial discretization we use a mixed finite difference/pseudospectral code. The term ‘pseudospectral’ refers, in our case, to the horizontal directions, where, since periodic boundary conditions are being used, all variables can be written as Fourier series. (The Fourier coefficients can be computed efficiently by using the Fast Fourier Transform algorithm, as described for example by Press et al. 1992.) Differentiation then becomes a simple matter of multiplying by ik (where k is the wavenumber) in Fourier space. The advantage of this method is that the spatial derivatives can be computed much more accurately than with a finite difference method, if the functions being differentiated are smooth (see e.g. Boyd 2001). However, this method cannot be used for the vertical derivatives, because the vertical boundary conditions are not periodic. Instead, a fourth order finite difference formula is used to compute the derivatives in this direction.

For the time discretization, the explicit third-order Adams-Bashforth method is used. The timestep is limited by stability constraints, derived both from the diffusive timescale and from the wave travel time across a mesh interval (i.e. the Courant condition; see Press et al. 1992). The computed maximum timestep is multiplied by a safety factor of 0.2. The scheme is stable provided that the diffusion coefficients are sufficiently large.

At the end of each timestep, boundary conditions are imposed. This is done simply by setting each variable to an appropriate value on the boundary. For example, the boundary values of u_z are set to zero, while the boundary values of u_x and u_y are set so as to ensure that $\partial u_x / \partial z$ and $\partial u_y / \partial z$ evaluate to zero at the boundary (when calculated using a one-sided finite difference formula). The boundary conditions are also utilized when calculating vertical derivatives near the boundary (where this is possible).

Also at the end of each timestep, we ‘dealias’ the variables, by applying the usual ‘2/3

rule’. This means that we filter out any modes with wavenumber (with respect to the horizontal directions) greater than two thirds of the Nyquist wavenumber (sometimes called ‘Nyquist frequency’, although in this case it is really a wavenumber rather than a frequency because it applies to space rather than time). This is easily implemented in Fourier space: the relevant Fourier amplitudes are simply set to zero.

The dealiasing is necessary because, although our computational grid can only represent modes of wavenumbers up to the Nyquist wavenumber (equal to π/λ , where λ is the grid spacing), modes of higher wavenumbers will in fact be generated by the nonlinear terms in the equations of motion. These modes do not simply disappear; instead their energy reappears at lower wavenumbers, in a process known as ‘aliasing’, and this can sometimes cause ‘aliasing instability’ (see e.g. Boyd 2001 for further explanation). The problem is avoided by filtering out the highest-frequency modes.

Note that the ‘2/3 rule’ applies to systems in which the nonlinearities are quadratic. This is true for most of our nonlinear terms, although there are some terms that involve divisions by ρ , for which the ‘2/3 rule’ is not strictly speaking appropriate. Nevertheless, we have found that in practice, this rule has been sufficient to prevent the aliasing instability from occurring.

One further problem with the code, as described so far, is that it is sometimes prone to numerical instabilities leading to fluctuations in the density. This can be a particular problem near the top boundary, where the lowest values of density can be found, and sometimes the density can even go negative, with disastrous results. We have found that this can often be solved by increasing the numerical resolution in the vertical direction. There are also two other ad hoc methods that we have used in order to improve the stability.

The first is to impose a boundary condition on the density. Strictly speaking, the density does not satisfy a boundary condition, but we can derive one by observing that $u_z = 0$ on the boundaries. This means that the z -component of the right-hand side of the momentum equation (5.2) will be zero on the boundary. The resulting equation can be interpreted as a boundary condition, as follows:

$$0 = -\frac{\partial}{\partial z} \left(\rho T + \frac{F \mathbf{B}^2}{2} \right) + \frac{\partial}{\partial x_j} (F B_j B_z + \sigma \kappa \tau_{jz}) + \theta(m+1)\rho. \quad (5.14)$$

In theory this equation should be satisfied automatically (provided that we are solving the momentum equation correctly), but in practice, numerical errors can cause the right-hand side of (5.14) to take on a small non-zero value. We correct this by explicitly

applying (5.14) as a boundary condition, i.e. by adjusting ρ on the top and bottom boundaries at the end of each timestep such that (5.14) is satisfied exactly. Doing this helps to ensure that the momentum equation is fully consistent with the boundary conditions, and this appears to improve the stability of the code in some cases.

Another method that can improve the stability of the density near the top boundary is to rewrite the continuity equation. Writing it as

$$\frac{\partial \rho}{\partial t} = -\rho \nabla \cdot \mathbf{u} - \mathbf{u} \cdot \nabla \rho, \quad (5.15)$$

instead of the conservative form (5.1), appears to improve stability (P. Bushby, private communication). One advantage of this formulation is that we no longer need to evaluate the vertical derivative of the density on the boundary (because $u_z = 0$ there). One disadvantage of not using the conservative form is that the method may no longer conserve mass exactly; however, in our simulations we typically find that mass is conserved to better than 1% accuracy, so this is not a serious problem.

All of these methods are implemented in a code which was originally written by Cattaneo et al. (1991) for the purely hydrodynamic case. Magnetic fields were added by Matthews et al. (1995), and the code has been worked on by various people since then (including P. Bushby, S. Houghton and N. Roxburgh). The changes made by the present author were the addition of the poloidal-toroidal field representation, and the addition of new magnetic boundary conditions. The modified code was tested by comparing it against linear theory.

Runs with resolutions of up to $256 \times 128 \times 144$ points were carried out. The code is parallelized (using MPI), so that several processors can be used for a single run. Smaller runs (of resolutions of around 64^3) were carried out on clusters of standard PCs, but for the larger runs we made use of the Cambridge-Cranfield High Performance Computing Facility, using up to 72 processors for each run.

5.4 Runs with uniform fields

In this section we take as our initial condition for \mathbf{B} a straight, uniform field, inclined at an angle ϕ to the vertical, as shown in Figure 5.1. We can then investigate the different types of behaviour that occur as the tilt angle is varied.

We choose in this section to keep B_z constant (as opposed to keeping $|\mathbf{B}|$ constant)

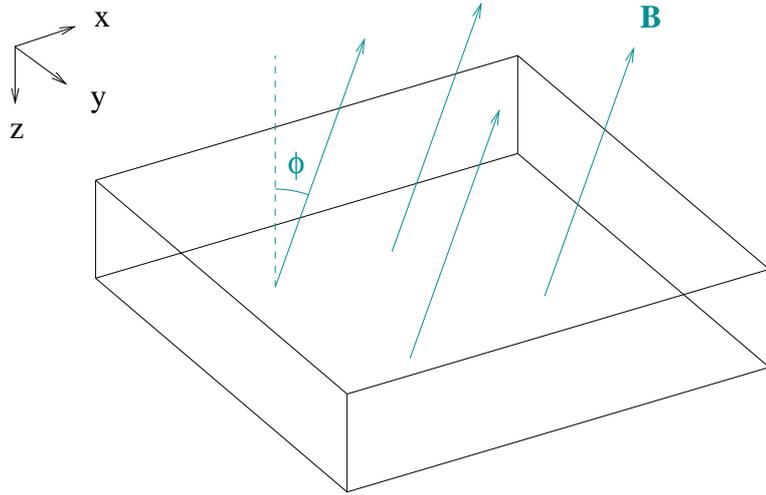


Figure 5.1: Diagram showing the initial field configuration for the ‘uniform field’ runs. Note the directions of the three coordinate axes. (The field is tilted towards the positive x -direction.)

as ϕ is varied. In other words, we take

$$\mathbf{B}_{\text{initial}} = \begin{pmatrix} -\tan \phi \\ 0 \\ 1 \end{pmatrix}. \quad (5.16)$$

This is different from the convention used in previous chapters (where we fixed $|\mathbf{B}|$). The motivation for this change is that, at least near onset, the amplitude of convection is controlled by the size of B_z , not $|\mathbf{B}|$. In other words, if we fix the Rayleigh number (and all other parameters) and vary ϕ , then for fixed B_z , the amplitude of convection stays roughly constant, while for fixed $|\mathbf{B}|$ the amplitude increases with increasing ϕ . The former seems more convenient, since we can control the amplitude of convection (and degree of nonlinearity) by varying R , independently of the changes in tilt angle.

For the numerical runs themselves, we chose to vary R and ϕ , while keeping Q fixed, at 2500 (for a vertical field, this would correspond to the ‘strong field’ regime of Weiss et al. 2002). We used aspect ratios of $\lambda_x = \lambda_y = 4$, with resolutions of 64 grid points in each horizontal direction, and between 40 and 80 grid points in the vertical.

The results are summarized in Figure 5.2. Each point on this graph represents the result of one simulation run, and the different symbols represent the different types of convection pattern that were found. Two basic patterns were found, namely rolls and cells. The cells also divided into three main types, which we have denoted ‘normal’, ‘wavy’ and ‘turbulent’. Finally, at the points marked ‘subcritical’, the trivial static

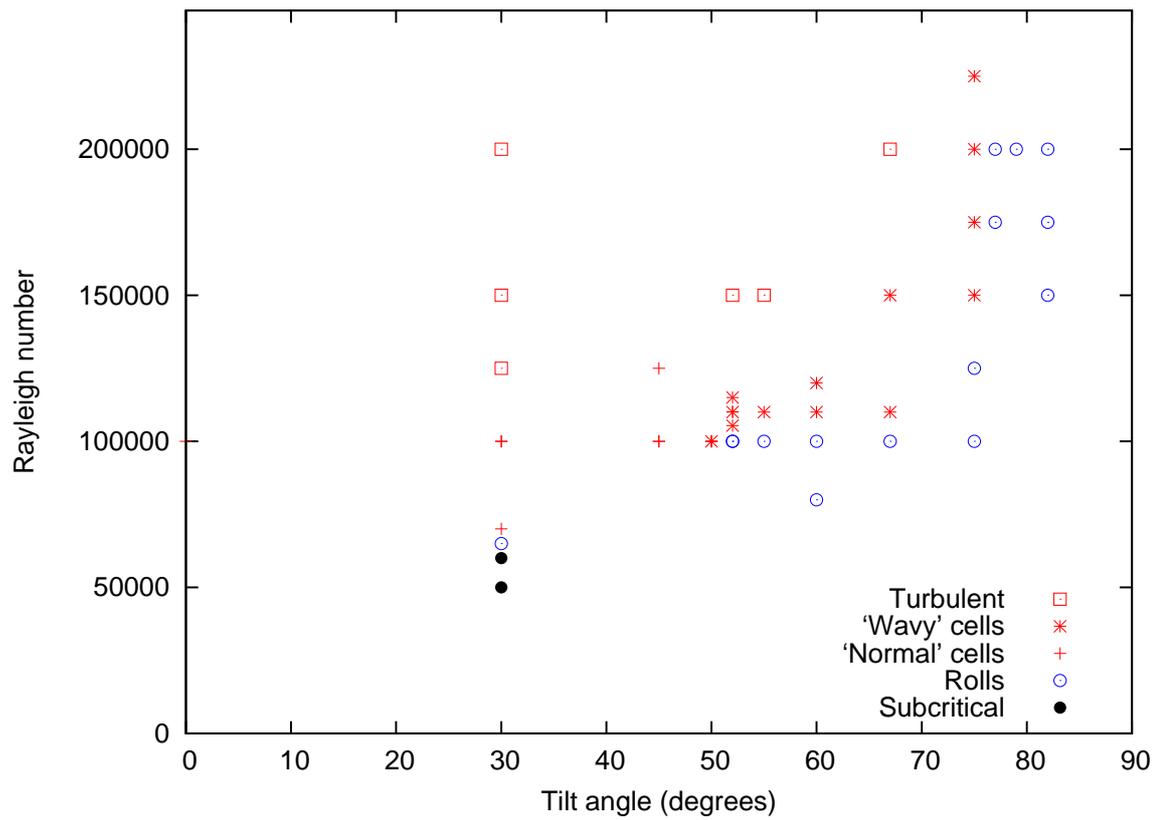


Figure 5.2: Diagram showing the solution types found at different values of R and ϕ , for $Q = 2500$.

solution was found to be stable, and no convection took place (since the simulations were started from scratch).¹

Some of the cellular solutions are illustrated in Figure 5.3. These runs all converged to uniformly drifting, but otherwise steady patterns (i.e., patterns which would be steady if viewed in an appropriately moving frame). The top two pictures in the Figure show what we have called ‘normal’ cellular solutions. Here the cells line up in a regular, more or less rectangular pattern. This is in contrast to the bottom two pictures, which show the ‘wavy’ cells. Here the convection cells form the same basic rectangular pattern, but this time it is subject to a zig-zag-like modulation.

There are a number of points to note about these solutions. Firstly, the basic pattern formed is one of rectangles, but the expected pattern in compressible convection is hexagonal. However, as pointed out by Matthews (1998), it is not possible to have a perfect hexagonal pattern in a square computational domain. In large domains we can obtain near-hexagonal patterns, but here we have used a relatively small aspect ratio (of 4) and this may be the reason that rectangular patterns are found instead.

Another observation is that the form of the solutions changes as the field becomes more inclined. For example, the cells become more and more elongated along the direction of tilt as ϕ increases. Moreover, the cells start to ‘run together’ forming more roll-like structures, particularly at the bottom of the box. At $\phi = 30^\circ$ (top left picture in Figure 5.3), there is a clear pattern of cells at both the top and the bottom of the box, but at $\phi = 45^\circ$ (top right picture), the cells are less clearly separated at the bottom of the box. By $\phi = 52^\circ$ (bottom left picture), the pattern at the bottom boundary is roll-like, although there are still clearly separated cells at the upper boundary. By $\phi = 75^\circ$ the cells at the top of the box have started to blend together as well.

The two other main types of solution are illustrated in Figure 5.4. The left-hand panel shows a roll solution; these are found mostly for very large ϕ , but they can also be found for smaller ϕ if R is only just above critical. For very large R we have found more turbulent solutions; a snapshot from one of these can be seen in the right-hand panel of Figure 5.4.

All of our solutions (except the straight parallel rolls) were found to travel. The

¹This means that these points are *linearly* subcritical. We did not attempt to look for *nonlinear* subcritical solutions here. For example, one might expect hexagons to be stable subcritically (see also Chapter 3); to see these, we would have had to run simulations starting from a (perturbed) hexagon solution, instead of starting from scratch.

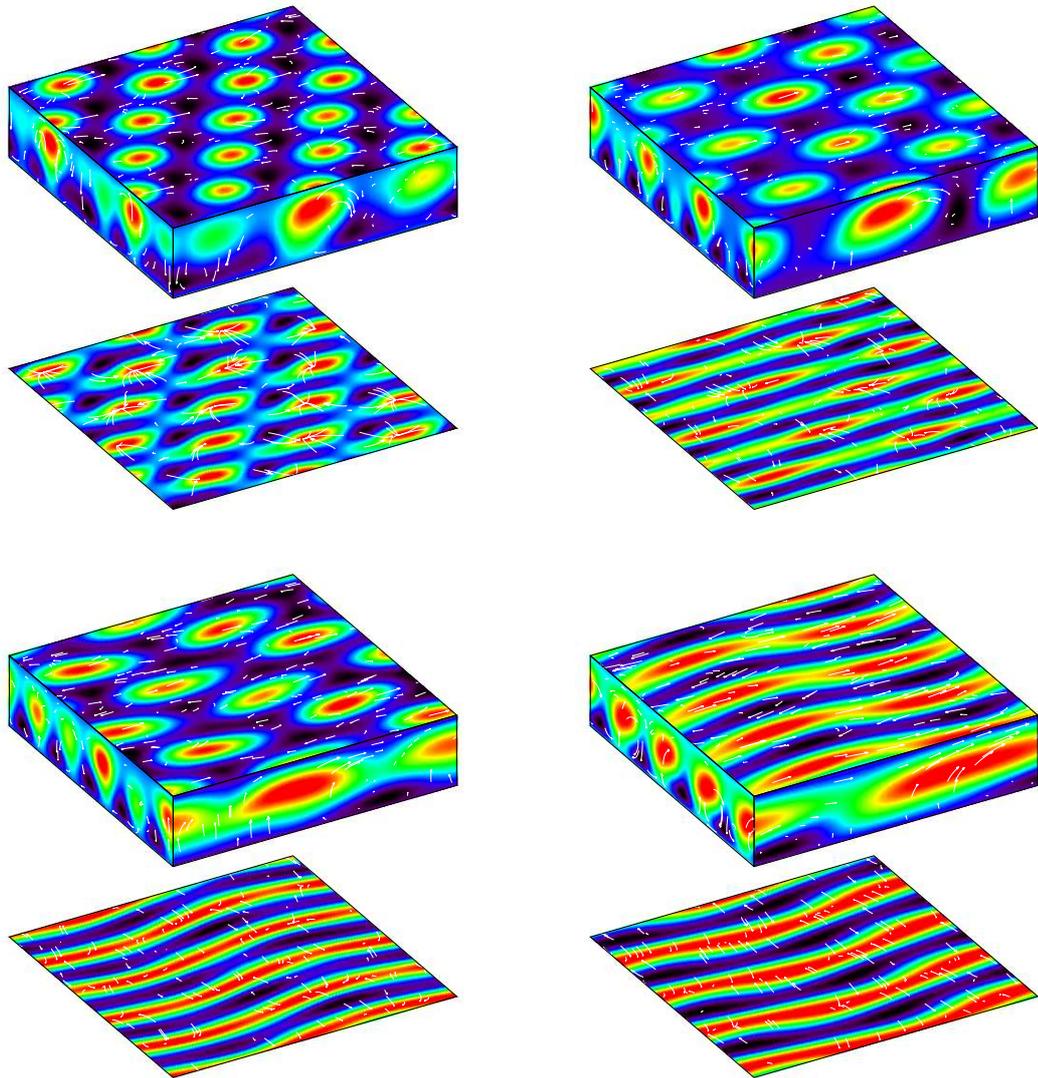


Figure 5.3: Solutions with uniform fields (all with $Q = 2500$). Top left: $R = 10^5$, $\phi = 30^\circ$; top right: $R = 10^5$, $\phi = 45^\circ$; bottom left: $R = 1.05 \times 10^5$, $\phi = 52^\circ$; bottom right: $R = 1.75 \times 10^5$, $\phi = 75^\circ$. The colours indicate temperature perturbation (with black, blue representing cool fluid, and green, yellow, red representing warm fluid), and the small arrows show the direction and magnitude of the fluid velocity (when projected onto each of the surfaces of the box). Note that each of these patterns drifts in the negative x -direction (against the direction of tilt).

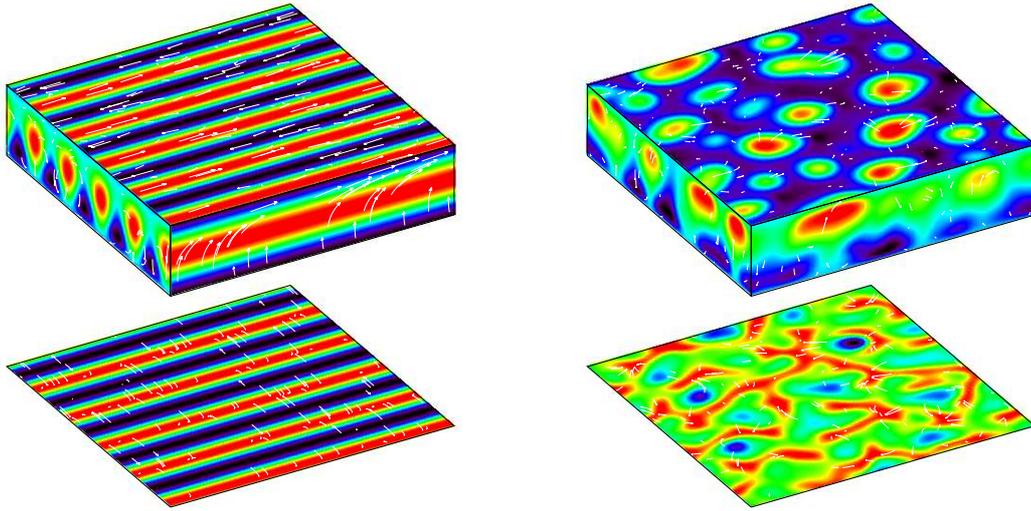


Figure 5.4: Additional solutions in a uniform field (with $Q = 2500$). Left-hand picture: $R = 10^5$, $\phi = 60^\circ$. This is a steady solution. Right-hand picture: $R = 2 \times 10^5$, $\phi = 30^\circ$. This solution is turbulent – the pattern is constantly evolving – but there is also a noticeable leftward drifting motion.

direction of travel was always against the tilt (i.e. in the negative x -direction) – there was no reversal of the direction of travel with increasing ϕ , although we do see a transition from drifting solutions to steady rolls as ϕ becomes large. We also find, in agreement with Hurlburt et al. (1996, 2000), that the mean *flow* at the surface is always in the opposite direction to this phase velocity (i.e. the pattern drifts to the left, but the mean flow at the surface is to the right).

5.5 Runs with non-uniform fields: ‘Arch’ initial condition

We now turn to runs with in which the initial condition involves a non-uniform magnetic field. As mentioned in the introduction to this chapter, these calculations are motivated by the difference in appearance between the umbra and the penumbra of a sunspot, and the aim will be to try to produce a model in which there is a sharp change between two different forms of convection. We begin with a simple ‘arch’-shaped field, and then move on (in the next section) to slightly more sophisticated initial configurations.

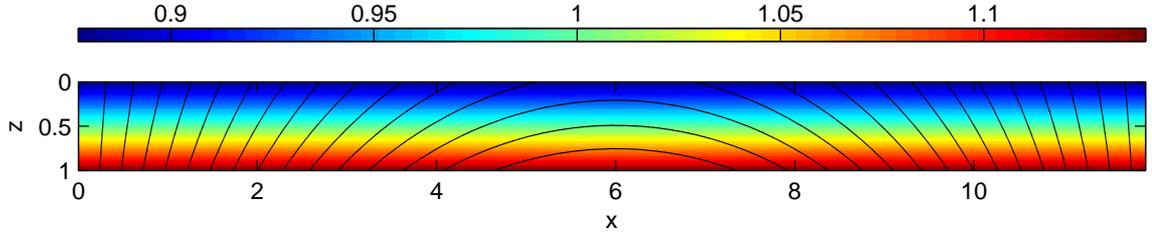


Figure 5.5: The ‘arch’ initial field (for $\lambda_x = 24$). The black curves are field lines, and the colours show the field strength $|\mathbf{B}|$.

We describe the ‘arch’-shaped field first. Here, the initial magnetic field configuration will be a two-dimensional field, given by

$$B_x(x, z) = -\frac{\partial\chi}{\partial z}, \quad B_y = 0, \quad B_z = \frac{\partial\chi}{\partial x} \quad (5.17)$$

where

$$\chi(x, z) = \frac{\lambda_x}{2\pi} \sin\left(\frac{2\pi x}{\lambda_x}\right) \exp\left(\frac{2\pi(z - 1/2)}{\lambda_x}\right). \quad (5.18)$$

Note that this field satisfies $\nabla \wedge \mathbf{B} = 0$, so that it is force-free and current-free, and therefore our trivial, motionless solution (5.8)–(5.10) is still an equilibrium state. Note also that

$$|\mathbf{B}| = \exp\left(\frac{2\pi(z - 1/2)}{\lambda_x}\right), \quad (5.19)$$

so that the field strength increases with depth; the normalization has been chosen such that $|\mathbf{B}| = 1$ at mid-layer ($z = 1/2$).

The field is illustrated in Figure 5.5. The Figure only shows the field for $x \in [0, \lambda_x/2]$; the second half (for $x \in [\lambda_x/2, \lambda_x]$) is identical to the first, but with a sign change. In fact, the entire field can be generated from the first quarter (for $x \in [0, \lambda_x/4]$), by using the following symmetries:

$$\mathbf{B}(x + \lambda_x/2, z) = -\mathbf{B}(x, z), \quad (5.20)$$

$$B_x(-x, z) = -B_x(x, z), \quad (5.21)$$

$$B_z(-x, z) = B_z(x, z). \quad (5.22)$$

Therefore, it is in fact possible to restrict attention to one quarter of a full period (i.e., $x \in [0, \lambda_x/4]$). By running simulations in only this subset of the box, we can in effect reduce the run time by a factor of 4. (This can be implemented by using so-called ‘quarter-wave’ Fourier transforms instead of ordinary ones.) We will use the symbol λ'_x to represent the width of the box actually used in the simulation (i.e. $\lambda'_x = \lambda_x/4$).

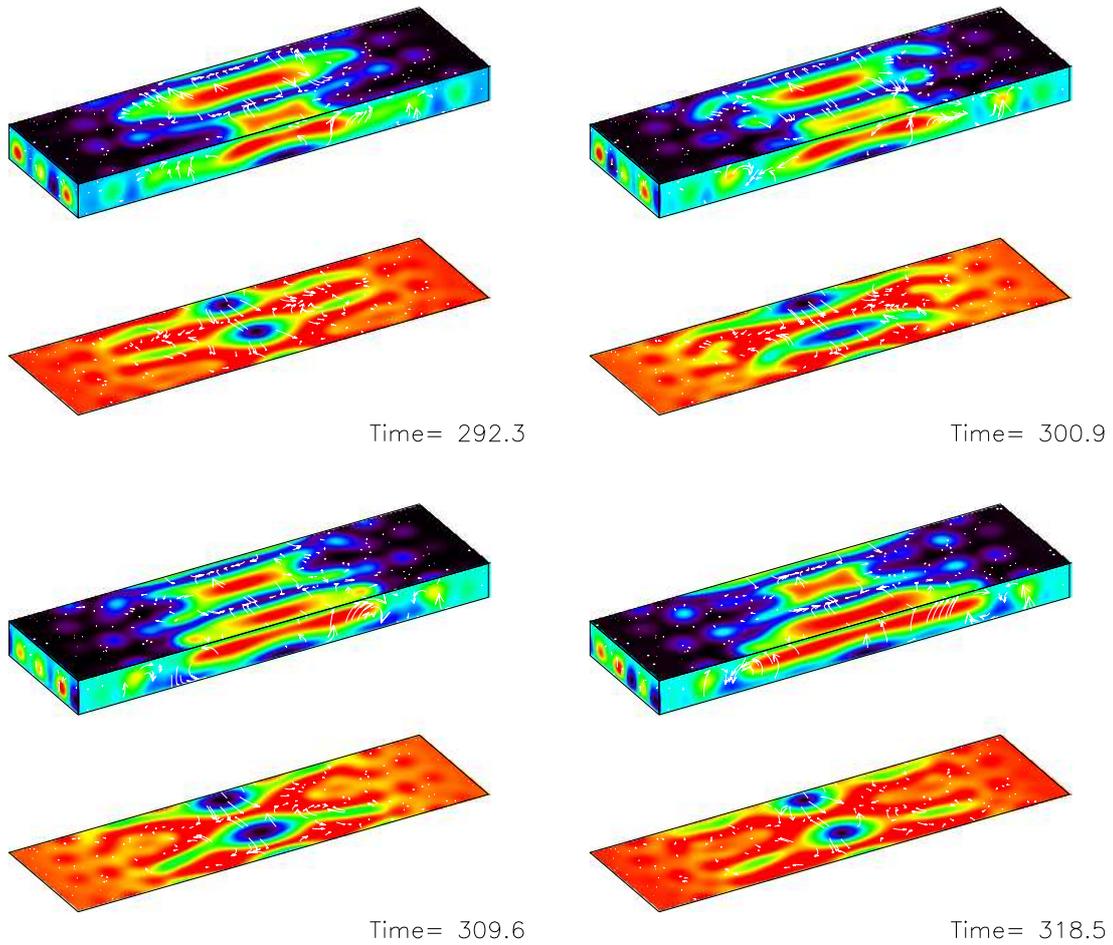


Figure 5.6: Snapshots at four different times during the ‘arch’ run with $Q = 1000$ and $R = 2.5 \times 10^4$, showing how the adjacent rolls wax and wane in amplitude. The colours show temperature perturbation while the arrows show the projection of the fluid velocity onto each surface of the box.

The reason for choosing the initial condition shown in Figure 5.5 was as a simple way of modelling a sunspot. The ‘sunspot’ is represented by the region with a near-vertical field – in fact, we have two ‘sunspots’ in our domain, one at each end of the box. This could also be seen as an infinite chain of ‘sunspots’, in view of the periodic boundary conditions. Note that the field is relatively strong everywhere – there is no region of weaker field to represent the ‘outside’ of a sunspot – so it would perhaps be more accurate to think of two ‘umbrae’ (the near-vertical fields near the edges of the box) connected by a shared ‘penumbra’ (the region of near-horizontal field in the centre).

We ran a simulation starting from this initial condition and using the parameters $R = 25000$, $Q = 1000$, $\lambda'_x = 6$, $\lambda_y = 3.5$, and a resolution of $65 \times 128 \times 48$. Note that a greater resolution is required in the y -direction than in the x -direction; this is

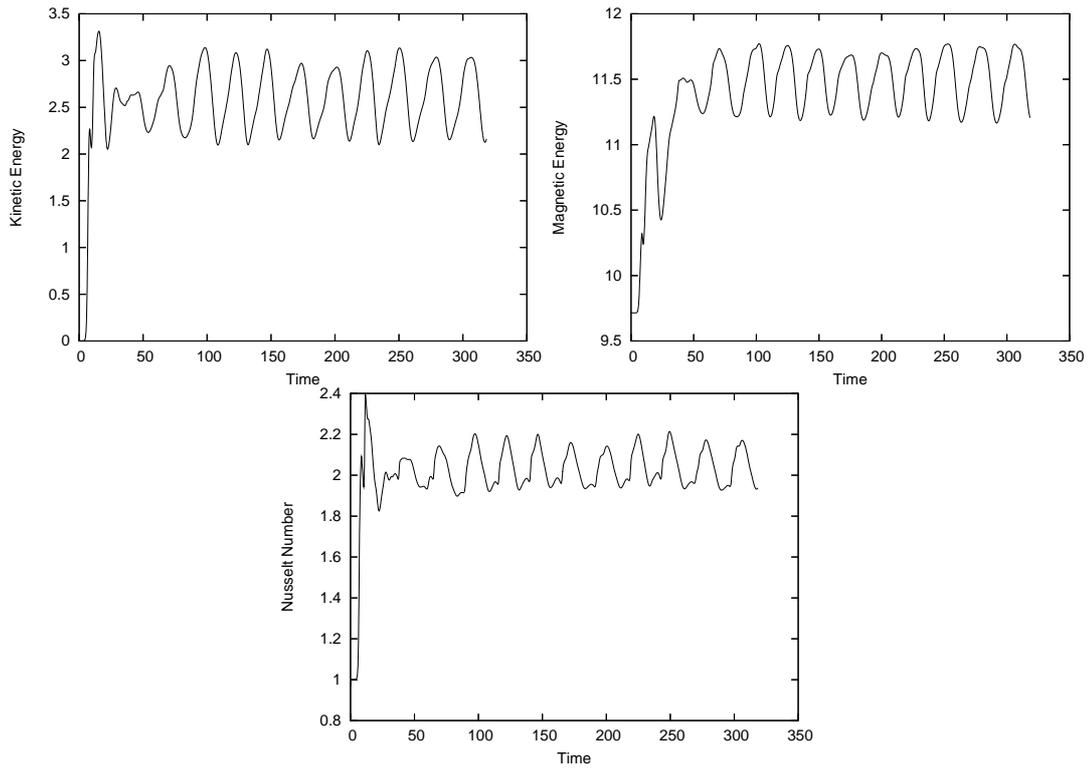


Figure 5.7: Plots of mean kinetic energy, mean magnetic energy and Nusselt number against time for the ‘arch’ run with $Q = 1000$ and $R = 2.5 \times 10^4$.

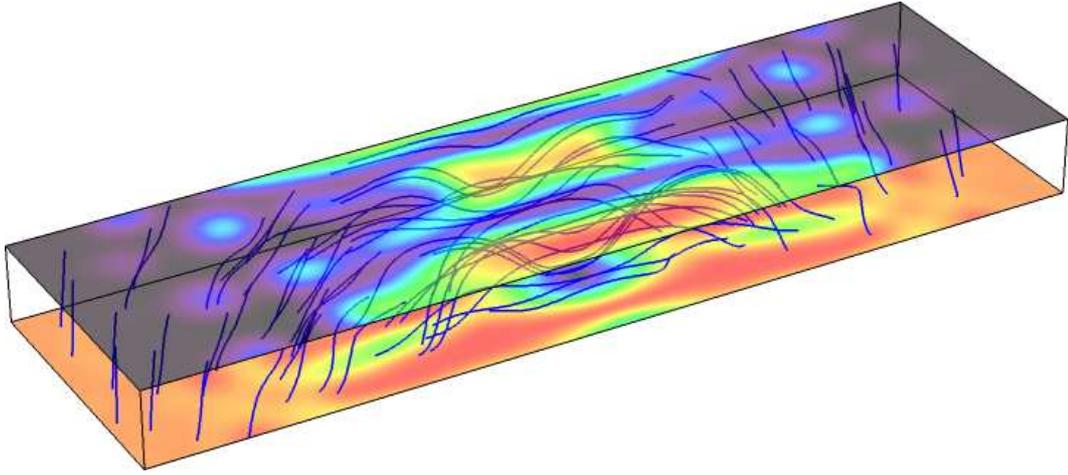


Figure 5.8: *Three-dimensional visualization of the magnetic field lines within the computational box (at the end of the calculation). For reference, the temperature perturbations on the top and bottom surfaces are also shown, where black, purple and blue indicate cooler fluid, and yellow, orange and red indicate warmer fluid.*

because the simulation produces rolls aligned with the magnetic field, which tend to have thin downflow sheets aligned in the x - z plane, and these downflows need to be properly resolved. (Note also that the vertical resolution is a little lower than what we have used for some of our other runs.)

Figure 5.6 shows snapshots of this run at four different times near the end of the calculation (after a quasi-steady state has been reached). The results clearly show a region of hexagon-like convection towards the outer edge of the domain, where the field is more vertical, with a transition to roll-like features in the centre, where the field is close to horizontal. Note the time dependence; the ‘rolls’ are undergoing a sort of oscillation, with adjacent rolls waxing and waning in amplitude. Figure 5.7 shows time series for kinetic and magnetic energy and the Nusselt number, which illustrate these oscillations more clearly. (The Nusselt number is a dimensionless measure of the superadiabatic temperature gradient, defined as follows:

$$N = \frac{d\bar{T}/dz - \theta(m+1)(1-1/\gamma)}{\theta - \theta(m+1)(1-1/\gamma)}, \quad (5.23)$$

where $d\bar{T}/dz$ is the mean temperature gradient at the bottom of the layer. Recall that $d\bar{T}/dz = \theta$ in the absence of convection; thus $N = 1$ if convection is not present, and $N - 1$ measures the convective contribution.)

Figures 5.8 and 5.9 show three-dimensional visualizations of the magnetic field struc-

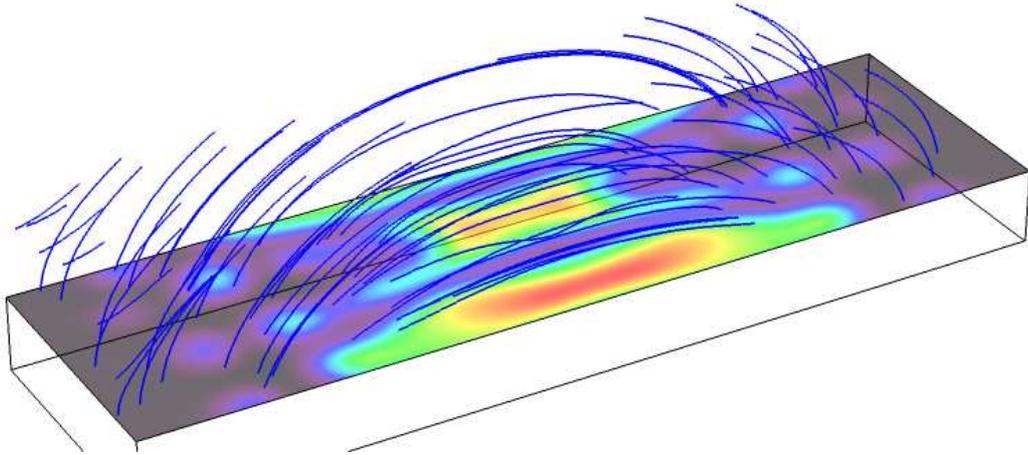


Figure 5.9: Three-dimensional visualization of a potential-field extrapolation of the field lines above the computational box. The temperature perturbation on the top of the box is also shown.

ture. The former shows the actual magnetic field structure within the box, while the latter shows the potential field extrapolation of the magnetic field into the region above the box.

In Figure 5.8, notice that at the edge of the box (in the region where hexagonal convection is found, and the field lines are near-vertical), the magnetic field structure is largely unaltered from its initial configuration. However, towards the centre of the box, where the near-horizontal fields can be found, the field lines are visibly pushed downwards compared to the initial condition. This can perhaps be seen more clearly in Figure 5.10 which shows the y -averaged magnetic field. Here, the depression of the field lines towards the right-hand end, and the increased field strength towards the bottom of the layer, are visible.

This seems to be evidence that *flux pumping* (Tobias et al., 2001; Thomas et al., 2002a; Weiss et al., 2004) is taking place within our simulation. This is a phenomenon in which magnetic field is transported downwards by the convection, due to the asymmetry – found only in compressible convection – between the fast-flowing downward plumes and the slow-moving upflows. In order to investigate this further, we have plotted the *centre of magnetic mass* as a function of both x and t . The centre of magnetic mass, z_{mm} , is in some sense a measure of the ‘average depth’ of the (horizontal) field, and we

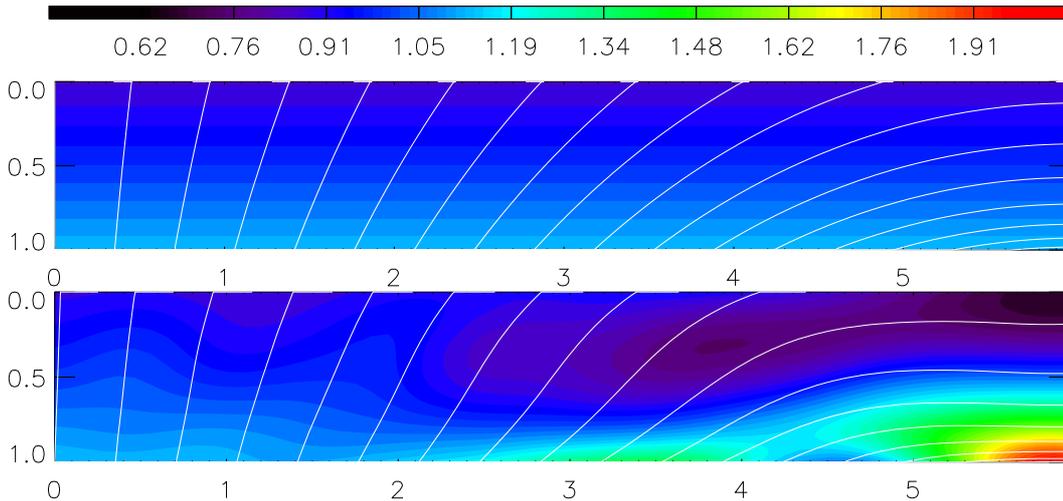


Figure 5.10: The magnetic field within the computational box, after averaging over the y -direction (thus obtaining a two-dimensional field). The colours indicate $|\mathbf{B}|$, and field lines are also plotted. The upper panel shows the initial condition, while the lower panel shows the final state (time-averaged over one oscillation period).

define it as follows:

$$z_{\text{mm}}(x, t) = \frac{\int_0^{\lambda_y} \int_0^1 z B_x dz dy}{\int_0^{\lambda_y} \int_0^1 B_x dz dy}. \quad (5.24)$$

(Note that flux pumping acts only on the horizontal flux, so we are only interested in the ‘centre of mass’ of the horizontal, not the vertical, field.)

We also define a version averaged over x as well as y :

$$\bar{z}_{\text{mm}}(t) = \frac{\int_0^{\lambda'_x} \int_0^{\lambda_y} \int_0^1 z B_x dz dy dx}{\int_0^{\lambda'_x} \int_0^{\lambda_y} \int_0^1 B_x dz dy dx}, \quad (5.25)$$

The first version is useful for describing how the effectiveness of pumping depends on position within the box, while the second version gives an average measure of the pumping over the entire box. (The second version corresponds to the definition adopted by Tobias et al. 2001, except that here we use B_x instead of B_y .)

Figure 5.11 shows z_{mm} as a function of x for our run with $R = 25000$ and $Q = 1000$. The dotted line shows the initial value, while the solid line shows the final value of z_{mm} ; the latter has been time-averaged over one oscillation period. The graph shows that at the right-hand end of the box, where the field is near-horizontal, the centre of magnetic mass has been pumped downwards significantly, but the pumping is less effective in regions where the field is more vertical (indeed, z_{mm} actually reduces in value near the left-hand end of the box).

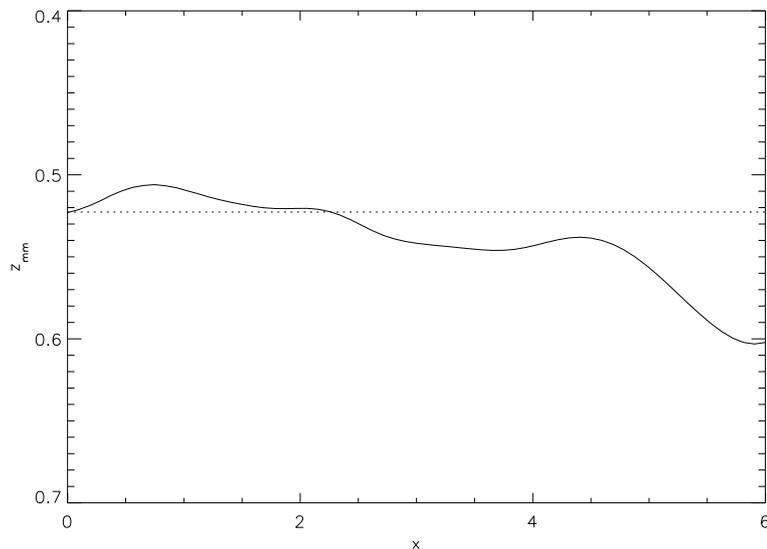


Figure 5.11: *The centre of magnetic mass $z_{\text{mm}}(x)$, time-averaged over one oscillation period. (The dotted line shows the value of z_{mm} at $t = 0$, for comparison.)*

Even at the right-hand end of the box, the change in z_{mm} is relatively small – only about 10% of the layer depth. In the calculation of Tobias et al. (2001), the horizontal field was pumped downward by a somewhat larger amount. The difference is that in their calculation, the computational domain was extended by adding a stable layer underneath the normal convecting region; the field could therefore be pumped down into this stable layer. In our calculation, we do not have this extra layer underneath the box, and so the field is blocked by the impenetrable lower boundary. This presumably limits the amount by which the field can be pumped downwards.

Figure 5.12 shows the depth of the centre of magnetic mass as a function of time. In agreement with Tobias et al. (2001), we find that there is a very rapid pumping phase at the beginning of the calculation, after which the centre of magnetic mass remains at a more or less constant level.

5.6 Runs with non-uniform fields: ‘Flux sheet’ initial condition

The above results, using the ‘arch’ initial condition, showed an interesting transition between hexagonal and roll-like convection. However, the results were a little unsatisfactory, because the length of the region of roll-like convection was rather small (compared

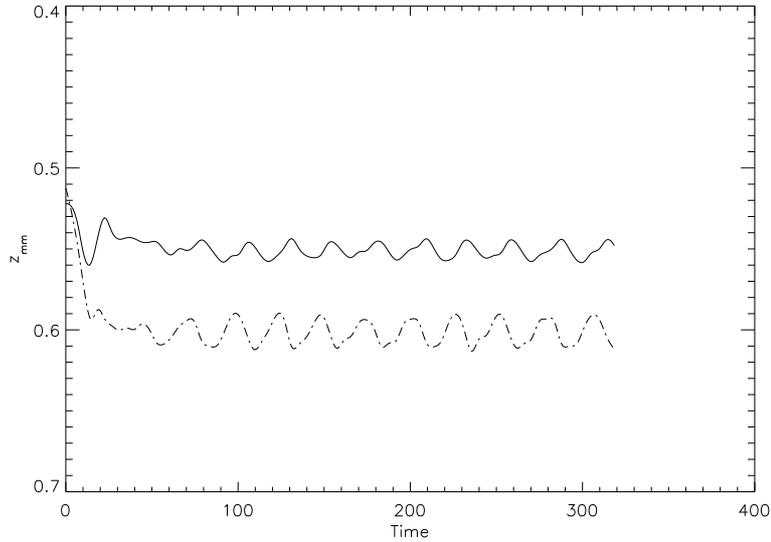


Figure 5.12: Centre of magnetic mass against time (for the ‘arch’ run). The solid line shows \bar{z}_{mm} (i.e. the centre of magnetic mass averaged over all x), while the dash-dotted line shows z_{mm} at the right-hand end of the box (where pumping is found to be most effective).

to its width), making it difficult to see precisely what was happening in this region.

We have therefore refined the choice of initial condition in order to allow greater control over the relative proportions of ‘umbra’ and ‘penumbra’ in the model (and hopefully to allow a wider region of roll-like convection in the results). The central observation is that we can specify B_z on the lower boundary; this, together with the requirement that the field within the layer be a potential field, is enough to uniquely specify the initial field.

We choose the following profile for B_z on the lower boundary ($z = 1$):

$$B_z = \begin{cases} \lambda'_x/x_0 & (0 \leq x \leq x_0) \\ 0 & (x_0 \leq x \leq \lambda'_x) \end{cases} \quad (5.26)$$

(This choice of scaling ensures that the average of B_z over the surface $z = 1$ is 1, or in other words, that the total flux of \mathbf{B} through the lower boundary is independent of the choice of x_0 .)

This field is sketched in Figure 5.13. The field below the box is assumed to continue downwards indefinitely, so that we have a column or tube of flux extending downwards in the region $x < x_0$, $z > 1$. (This choice was inspired by the flux tubes that are believed to be present underneath sunspots – although what we have here is more of a flux sheet than a flux tube, given that it extends indefinitely in the y -direction.) The potential

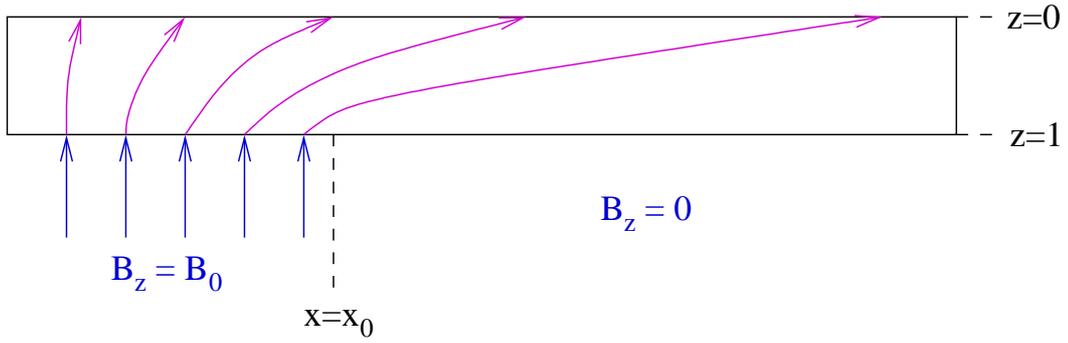


Figure 5.13: Sketch of the ‘flux sheet’ initial condition. Below the box, there is imagined to be a uniform sheet of flux in the region $x < x_0$ (blue field lines), with zero field for $x > x_0$. We match this onto a potential field (purple field lines), which forms the initial condition for the magnetic field within the computational box ($0 < z < 1$).

field above the surface $z = 1$ is what will form our actual initial condition.

Note that by varying x_0 , we can adjust the width of the flux sheet, and thus control the relative width of the ‘umbral’ and ‘penumbral’ regions within the simulation.

We begin constructing this initial condition by defining the *flux function* χ , as follows:

$$B_x = -\frac{\partial\chi}{\partial z}, \quad (5.27)$$

$$B_z = \frac{\partial\chi}{\partial x}. \quad (5.28)$$

The flux function defines a two-dimensional magnetic field satisfying $\nabla \cdot \mathbf{B} = 0$. If we also want χ to represent a potential field, then we must have $\nabla^2\chi = 0$. We may now write χ as a Fourier series as follows:

$$\chi(x, z) = \sum_{k \text{ odd}, k \geq 1} A_k \sin\left(\frac{\pi k x}{2\lambda'_x}\right) \exp\left(\frac{\pi k(z-1)}{2\lambda'_x}\right). \quad (5.29)$$

(As before, $\lambda'_x = \lambda_x/4$, i.e. λ'_x denotes the actual width of the box used in the simulation, as opposed to one full period.) Note that we have restricted the series to sin terms, with the cos terms being omitted, and we have also included only the odd-numbered modes. These restrictions arise because of the ‘quarter-wave’ symmetry that we want our field to obey – see equations (5.20)–(5.22) above.

In view of our choice (5.26) for B_z , we now have the following expression for χ at $z = 1$:

$$\chi(x, 1) = \begin{cases} x\lambda'_x/x_0 & (0 < x < x_0) \\ \lambda'_x & (x_0 < x < \lambda'_x) \end{cases} \quad (5.30)$$

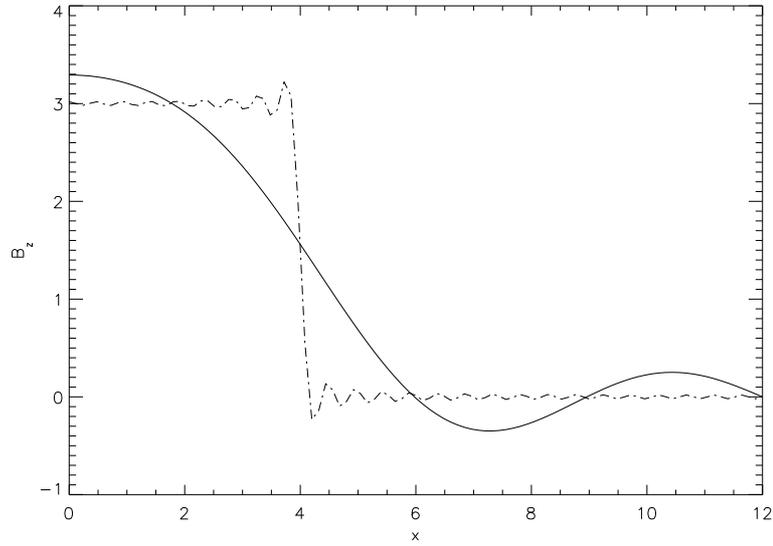


Figure 5.14: B_z as a function of x at the lower boundary, as given by the series (5.29), with A_k as defined in (5.34). Dot-dashed curve: series truncated at the Nyquist frequency. Solid curve: series truncated after 4 terms.

We must now compute values for A_k that make (5.29) and (5.30) consistent with one another. To do so, we proceed as follows:

$$\int_0^{\lambda'_x} \chi(x, 1) \sin\left(\frac{\pi k x}{2\lambda'_x}\right) dx \quad (5.31)$$

$$= \int_0^{\lambda'_x} \sum_{j \text{ odd}, j \geq 1} A_j \sin\left(\frac{\pi j x}{2\lambda'_x}\right) \sin\left(\frac{\pi k x}{2\lambda'_x}\right) dx \quad (5.32)$$

$$= \frac{A_k \lambda'_x}{2}. \quad (5.33)$$

To compute A_k , therefore, we need to evaluate (5.31), substituting for χ from (5.30), and multiplying the result by $2/\lambda'_x$. After a little algebra we find that

$$A_k = \frac{8\lambda_x'^2}{\pi^2 k^2 x_0} \sin\left(\frac{\pi k x_0}{2\lambda'_x}\right). \quad (5.34)$$

One unforeseen problem with this particular choice of initial condition is that it produces a number of undesirable oscillations in the value of B_z at $z = 1$. To see this, consider the representation (5.29). In this equation, we should in theory sum over all positive odd integers k ; however, in practice, only a finite number of Fourier modes can be represented in the calculation, and we must truncate the sum at some point. It is this truncation that generates the oscillatory behaviour; see Figure 5.14, dash-dotted curve, in which the series was truncated at the Nyquist frequency. The oscillations arise because Fourier series cannot represent discontinuities well (they are an example of the Gibbs phenomenon).

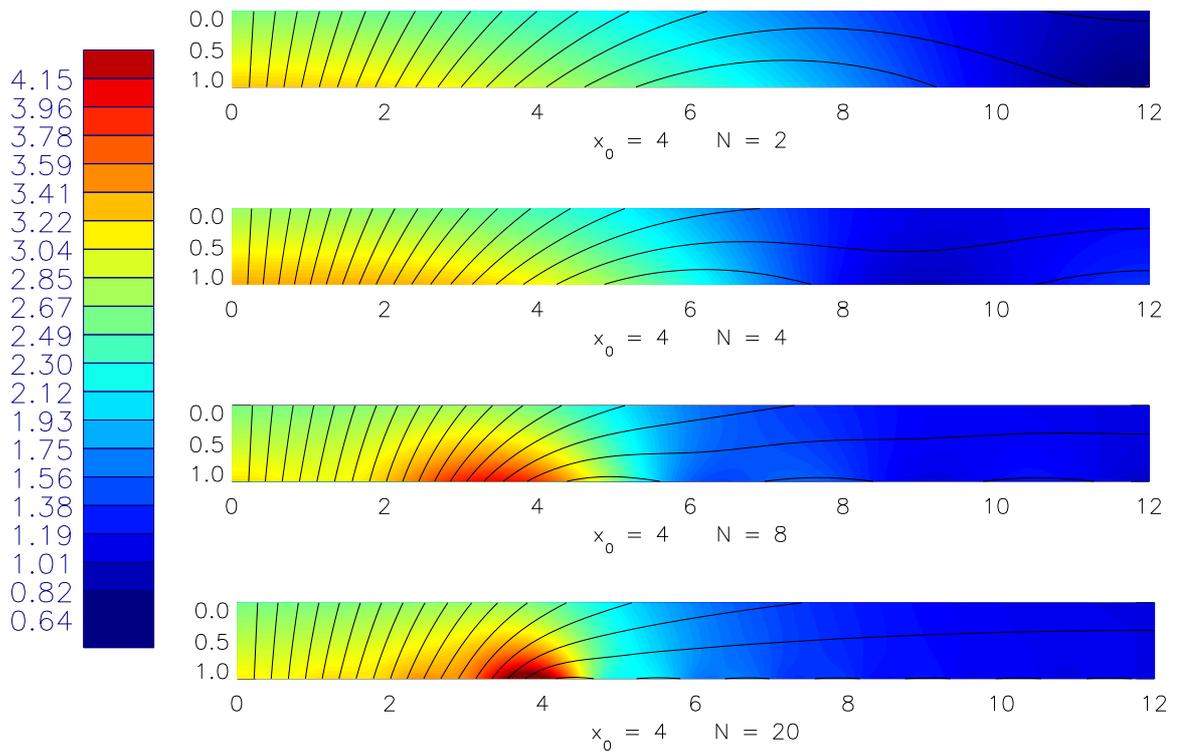


Figure 5.15: Illustration of the initial condition given by (5.29) and (5.34), plotted as a function of x and z . The colours indicate $|\mathbf{B}|$ and the curves are field lines.

The oscillations are undesirable because they lead to numerical instabilities. When using a ‘tied field’ boundary condition, the problem is worse, because the value of B_z is fixed on such a boundary; therefore, the oscillations can never smoothed away by diffusion, and they remain present for the entire simulation.

After discovering this problem, we resolved it in a fairly simple way: by truncating the sum at a lower value of k .² We will use N to represent the total number of modes used (i.e., the odd modes from $k = 1$ to $k = 2N - 1$ are included). The result when choosing $N = 4$ is shown on Figure 5.14 (solid curve); the sharp oscillations have disappeared, and have been replaced by a smooth curve.

Plots of the magnetic field for various N are shown in Figure 5.15. Here we have chosen $x_0 = 4$, since this gives a relatively large ‘penumbral’ region to the right-hand side, whilst still retaining a reasonable amount of ‘umbra’ at the left-hand end. Notice

²An alternative solution would be to choose a different profile for B_z in the first place, i.e. to choose a continuous function in place of the discontinuous (5.26). We will look at this possibility as well, later on in this chapter.

that choosing N too low results in a field which is not much different from the ‘arch’ field (which would correspond to $N = 1$), while choosing N too high produces a region of quite strong field at the base of the layer near $x = 4$ (which seems unrealistic). In our simulations, we have chosen to use the value $N = 4$, as a compromise between these two extremes.

5.6.1 Initial results

We began by running a simulation in a 12×2 box (i.e. $\lambda'_x = 12$ and $\lambda_y = 2$). The resolution was $257 \times 64 \times 144$, and we set $Q = 45$ and $R = 25000$. The simulation converged to the final state shown in Figure 5.16, with a region of hexagon-like cellular convection to the left, where the field is mainly vertical, and a region with long roll-like structures to the right, where the field is near-horizontal. The pattern is time-dependent: the hexagons travel to the left, while the rolls undulate and continually evolve. The pattern in the central part of the roll region also appears to travel leftwards, while the pattern near the right-hand end of the box does not seem to travel in any particular direction.

We must be careful in interpreting this result, since the box is rather narrow in the y -direction – with room for only a single convection roll – and the motion is therefore highly constrained. To reduce this effect we performed a second calculation using the same parameters but in an enlarged box with $\lambda_y = 4$ instead of 2. (The resolution in the y -direction was also correspondingly increased.) As an initial condition we used the result from the run in the narrower box, duplicated twice in the y -direction, and with a small perturbation added to the temperature field.

Initially, this new calculation proceeded similarly to the original run, but once enough time had passed, the original pattern of rolls decayed and broke up. This is shown in Figure 5.17; the top picture shows the situation when the rolls first begin to break up, while the bottom picture shows the situation some time later when almost all trace of the original roll pattern has disappeared. (The break-up begins from the right-hand end of the box, and proceeds leftwards.)

To understand the cause of this phenomenon we must look at the magnetic field structure. This is illustrated in Figure 5.18, which plots magnetic field lines both before and after the break-up of the roll pattern. In the earlier picture, although the field lines snake from side to side a little, they are still predominantly horizontal and aligned with

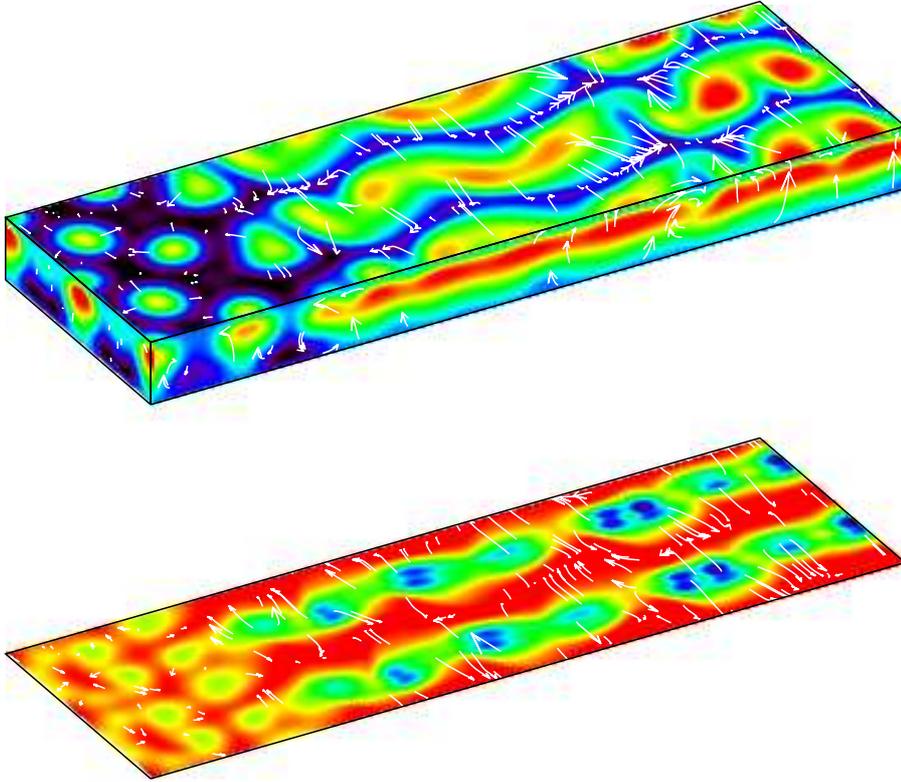


Figure 5.16: Temperature plot for the solution with the ‘flux sheet’ initial condition, in the narrow box ($\lambda_y = 2$). Note that in order to better show the periodicity, the diagram actually shows two copies of the box, placed side by side along the y -direction (so that the actual computational domain is half of what is shown above).

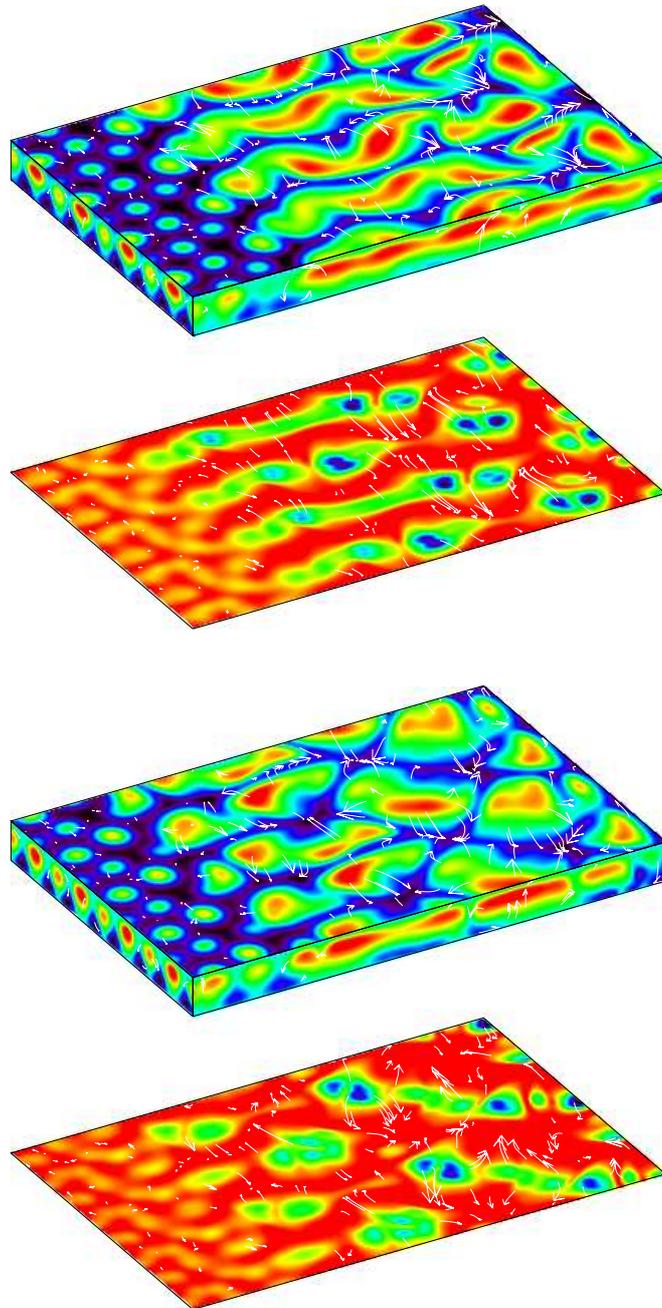


Figure 5.17: As Figure 5.16, but extended to a wider box (with $\lambda_y = 4$ instead of 2). Once again, two copies of the box are shown in the y -direction. The upper picture shows the situation mid-way through the calculation, and the lower diagram shows the final state.

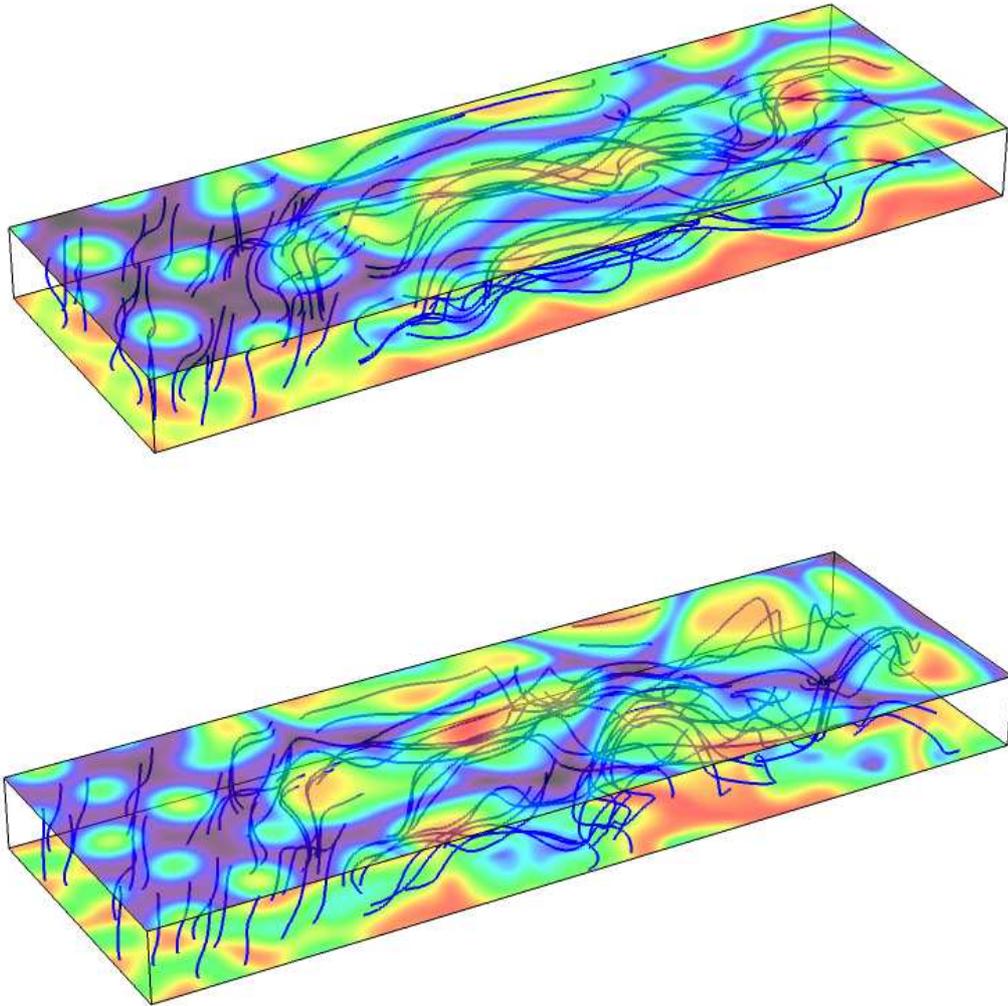


Figure 5.18: Plots of magnetic field lines within the (wider) box, with temperature shown in colour on the top and bottom surfaces. (Only one copy of the box is shown in this diagram.) The top picture shows the situation just after the width of the box was doubled and the calculation restarted. The bottom picture shows the final state.

the roll axes. In the later picture, the field lines are much more tangled. What appears to have happened is that part of the field has risen upwards (presumably by some sort of magnetic buoyancy process) and this has disrupted the roll pattern.

This can be seen clearly if one plots $\langle B_x \rangle$, where the angle brackets denote a horizontal average, against z and t . (This measures the amount of magnetic flux in the x -direction at each depth within the box.) See Figure 5.19, where $\langle B_x \rangle$ is plotted on the x -axis, with depth on the y -axis, and with different times represented by differently-coloured curves (lighter colours representing earlier times). The first graph represents the first calculation (the one with the narrower box). Here, the initially almost-uniform flux profile is quickly redistributed and pumped downwards. The total amount of flux in the box also increases during the calculation, indicating that flux is coming down into the domain through the top boundary. (Recall that although the lower boundary is a perfect conductor, the upper boundary is not, and therefore flux can enter or leave the computational domain via the top boundary.)

The second panel of Figure 5.19 shows what happens in the second calculation (after the width of the box has been increased). (The final, i.e. darkest, curve in the top picture therefore corresponds to the initial, i.e. lightest, curve in the bottom picture.) Initially, the profile remains fairly constant, but then the flux starts to move upwards (with $\langle B_x \rangle$ increasing for $z \gtrsim 0.4$ and decreasing for $z \lesssim 0.4$). In the final stages the flux seems to decrease at all depths, indicating that flux is being lost through the upper boundary. The corresponding reduction in field strength, together with the rearrangement of the field from a more-or-less coherent horizontal structure to a more tangled pattern, seems to be enough to prevent the long field-aligned convection rolls from forming.

This can also be seen by looking at plots of the centre of magnetic mass \bar{z}_{mm} (equation 5.25) against time, as shown in Figure 5.20. The left-hand picture shows the plot for the initial run, in the narrower box; here, the centre of magnetic mass descends quite rapidly at the beginning of the simulation, showing that flux is pumped downwards. The right-hand picture shows the plot for the second run, in the wider box, and here we see that the initial flux pumping is reversed (to a certain extent) with flux rising back out of the top of the box again.

In summary, therefore, we see that in the narrow box, flux is brought down from above, and this strengthens the field within the penumbral region. In the wider box, the reverse happens and flux escapes from the domain, causing a somewhat weaker field in the penumbral region. In the former case a pattern of rolls is seen, while in

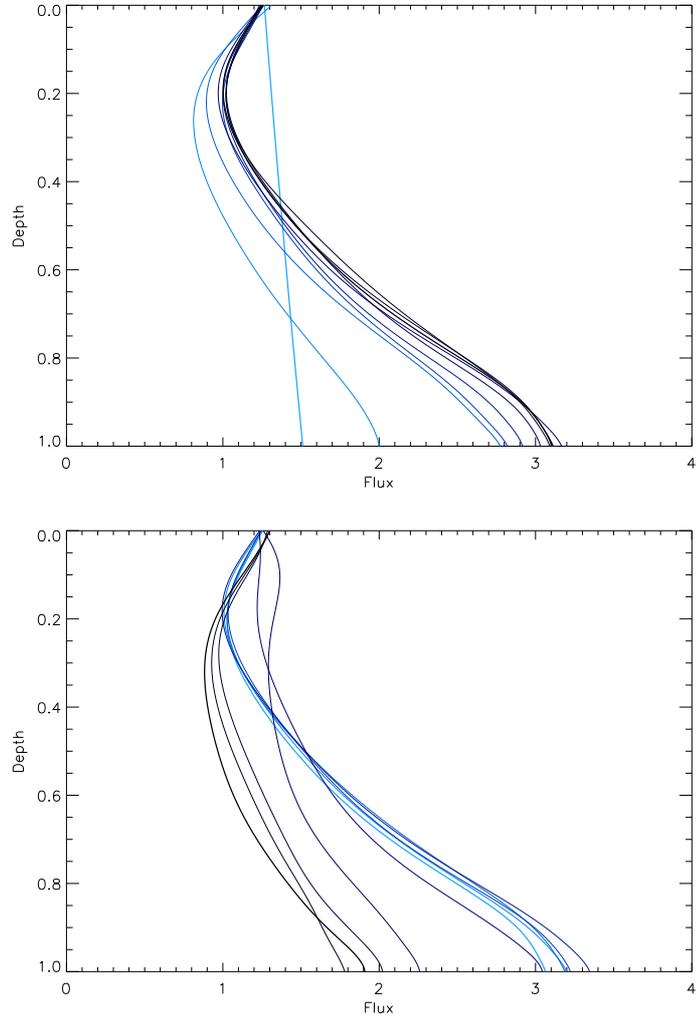


Figure 5.19: Plots of horizontal flux (i.e. $\langle B_x \rangle$) against z for various different times. Graphs are shown at equally spaced time intervals, with the lighter colours indicating earlier times, and darker colours later times. The top picture shows the initial calculation (in the narrow box) while the bottom picture shows the second calculation (with the wider box).

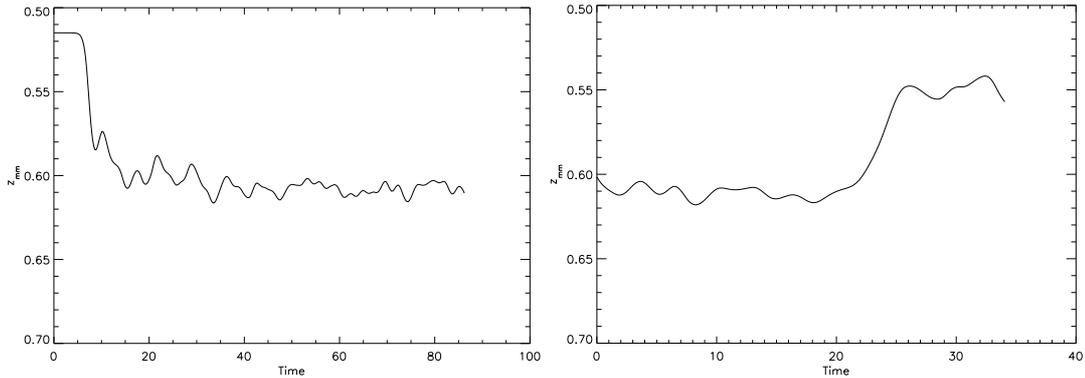


Figure 5.20: *The centre of magnetic mass \bar{z}_{mm} against t , for the original run (left-hand picture) and for the run in the widened box (right-hand picture).*

the latter case a cellular structure is found. This is somewhat unfortunate, since it is the wider box that corresponds to the more realistic situation (the narrower box being rather strongly constrained by the periodicity in the y -direction); however, the pattern found within the wider box is a rather uninteresting cellular pattern, looking a lot like field-free granulation. To model real penumbrae it seems clear that we will require a somewhat stronger field, in order to induce a transition to a new form of ‘penumbral’ convection. (This is what happened in the narrower box, but we would like to find a way of obtaining that pattern, or one like it, within a wider, less constrained computational box.)

5.6.2 Preventing the flux from escaping

Therefore, we would like to find a way to prevent the flux from escaping through the upper boundary. In this section we look at one way of achieving this aim.

Since the simulations take a considerable amount of time to run, we decided, before proceeding any further, to speed things up by using a reduced numerical resolution (of $129 \times 64 \times 144$, instead of the previous resolution of $257 \times 128 \times 144$) for all of our runs from this point onwards. This decision allows more runs to be performed, at the expense of possibly losing some accuracy. In this case we felt it was more important to be able to perform multiple runs than to have the highest accuracy possible, because we were mainly interested in the general patterns produced rather than the detailed nature of the solutions. In particular, we wanted to be able to search through parameter space (as far as possible) for any evidence of a ‘sunspot-like’ regime, in which a difference between ‘umbral’ and ‘penumbral’ convection could be seen, as was discussed above.

A second factor of note is that, we encountered problems (in some of the runs presented below) in which numerical instabilities would develop in the vertical direction. In particular, high-wavenumber vertical oscillations would appear (especially in the variables ρ and u_z). Where this occurred, it was corrected fairly straightforwardly, by occasionally applying a ‘smoothing’ procedure, in which high-wavenumber modes were filtered out (similarly to the dealiasing already being applied in the horizontal directions).

Imposing a perfectly conducting upper boundary

With these points in mind, we then looked for a way to run a simulation in which the flux was somehow constrained not to escape from the simulation domain. The simplest way to achieve this is to change the upper boundary condition to a perfectly conducting boundary (instead of matching to a potential field), which would prevent flux from either entering or leaving the box. We chose to run a simulation with this modified boundary condition (and with all other parameters being unchanged). This new run was started from scratch (i.e. from the motionless equilibrium state plus a small temperature perturbation) and the larger box size was used.

The final state for this new run is shown in Figures 5.21 (showing the temperature) and 5.22 (showing y -averaged magnetic field). In this result, we do *not* see a roll-like pattern in the penumbra; instead, a cellular pattern, similar to the previous result, Figure 5.17 (lower panel), is seen. So, although our change to the boundary conditions did prevent flux from escaping, it appears that this in itself was not sufficient; if we want to see more anisotropic patterns, then we will also need to arrange for there to be more flux within the domain at the beginning of the calculation.

The simplest way to do this would be to increase the parameter Q . However, this would increase the field strength within both the umbral and penumbral regions of the box, whereas we only want to increase the field strength in the penumbra. The problem with increasing the field strength in the umbra is that the field is already rather strong here, and we are in danger of shutting down the convection completely if Q is much increased (whereas ideally we want a simulation showing convection within both the umbra and the penumbra). To see this, note that the ‘effective Q ’ (i.e. $Q|\mathbf{B}|^2$, where $|\mathbf{B}|$ refers to the dimensionless field strength) is approximately 500 in the umbral region (although this does vary depending on exact position), compared to a value of about 50 to 100 in the penumbral region. Since the Rayleigh number is only around 25000, we

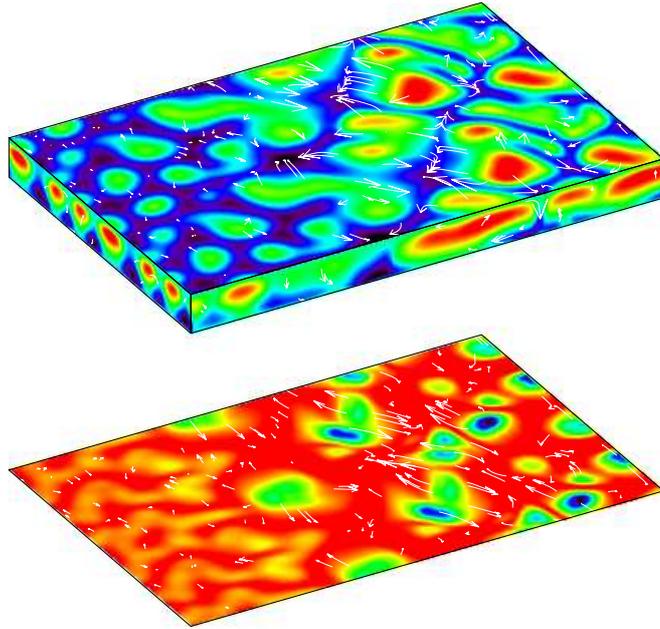


Figure 5.21: Temperature plot for the solution with the ‘flux sheet’ initial condition, in the wider box ($\lambda_y = 4$), and with the upper boundary condition changed to be perfectly conducting. (As usual, two copies of the domain are shown in the y -direction.) Note that the result shows a cellular pattern at the right-hand end (similar to Figure 5.17) rather than a roll pattern (as was seen in Figure 5.16).

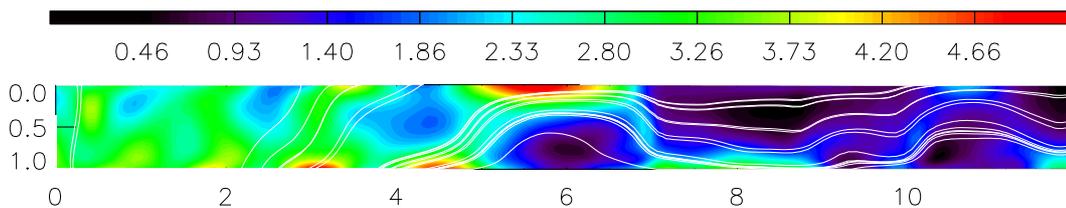


Figure 5.22: Plot of the magnetic field for the solution shown in Figure 5.21 (after averaging in the y -direction). The colours indicate $|\mathbf{B}|$ and the white lines are field lines. Notice how flux accumulates near the top boundary (in certain places); flux is still rising upwards as before, but this time it cannot actually move up and out of the box, because of the new boundary condition.

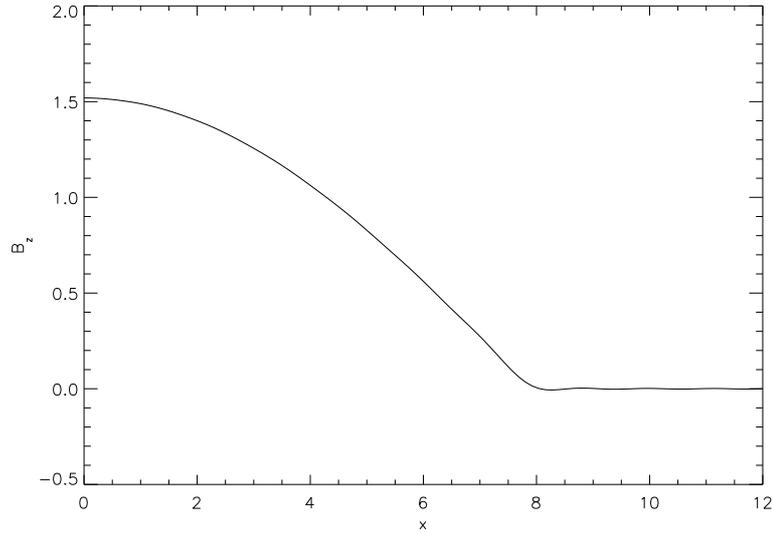


Figure 5.23: B_z as a function of x at the lower boundary, for the initial condition given by equations (5.35)–(5.36). (Compare Figure 5.14 above.)

would expect the convection to be shut down for an effective Q of about 1000 or above (very roughly); therefore, we cannot increase Q all that much without shutting down the convection in the umbra.

A new initial condition

To mitigate these issues, we looked at reworking the initial condition slightly, in order to reduce the contrast in field strength between umbra and penumbra in the model. This can be done by increasing x_0 in (5.26) above. In addition, we took this opportunity to slightly modify the initial condition (in order to reduce the contrast in field strength still further), by replacing (5.26) with the following new profile for B_z at the base of the layer:

$$B_z = \begin{cases} (\lambda'_x/x_0) \cos(\pi x/2x_0) & (0 \leq x \leq x_0) \\ 0 & (x_0 \leq x \leq \lambda'_x) \end{cases}. \quad (5.35)$$

This leads to the following replacement for equation (5.34) for A_k :

$$A_k = \frac{4\lambda'_x}{\pi^2} \left(\frac{2}{k} - \frac{1}{k - \lambda'_x/x_0} - \frac{1}{k + \lambda'_x/x_0} \right) \cos \left(\frac{\pi k x_0}{2\lambda'_x} \right). \quad (5.36)$$

The new choice (5.35) for B_z is similar to the previous one, but it gives a slightly less sharp transition between the umbral and penumbral fields (e.g. the previous ‘overshoot’ in B_z , where B_z became negative for a short distance, has been eliminated), and reduces slightly the difference in field strength between the two.

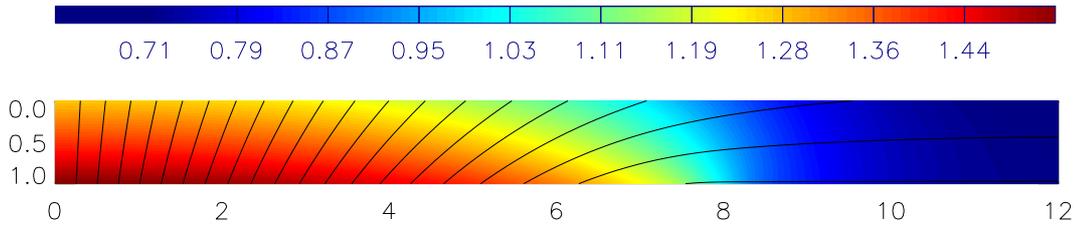


Figure 5.24: *Illustration of the initial condition given by (5.35)–(5.36). The colours indicate $|\mathbf{B}|$ and the curves are field lines. (Compare Figure 5.15 above.)*

The new B_z is plotted in Figure 5.23 (for the particular choice $x_0 = 7.9$), which may be compared to Figure 5.14 (solid curve) above. Note that the Gibbs phenomenon does not affect this new initial condition (because equation 5.35 does not contain a discontinuity) and so the number of modes N is irrelevant (provided it is sufficiently large).

The new field configuration itself is shown in Figure 5.24. Notice that the umbral region is somewhat wider than before, but in return we have a smaller contrast in the field strength between umbra and penumbra, meaning that we can have higher field strengths in the penumbra without necessarily needing excessive field strengths within the umbral region.

Using this new field configuration, runs for three different values of Q were performed, namely 500, 1000 and 1500. The results are shown in Figures 5.25–5.28. (Note that the Q values for these results are not directly comparable with the Q values for the previous results, because the new initial condition for \mathbf{B} is scaled differently.)

In the case $Q = 500$ (Figure 5.25), we see a hexagonal pattern of convection in the umbra (as usual). In the penumbra, a time-dependent, constantly changing pattern is seen. Sometimes the pattern looks fairly roll-like (lower panel of the Figure), and sometimes it looks more cellular (upper panel), albeit with some of the cells being stretched out along the x -direction. Therefore, our new initial condition has met with some success, because the penumbral pattern is now showing signs of anisotropy (and we have not lost the hexagonal umbral convection). However, in this regime the field is not quite strong enough to give a sustained pattern of rolls (or roll-like structures) in the penumbra.

When we increase Q to 1000, a different pattern is seen. The umbral convection is shut down for a field this strong. The penumbral convection initially takes the form of

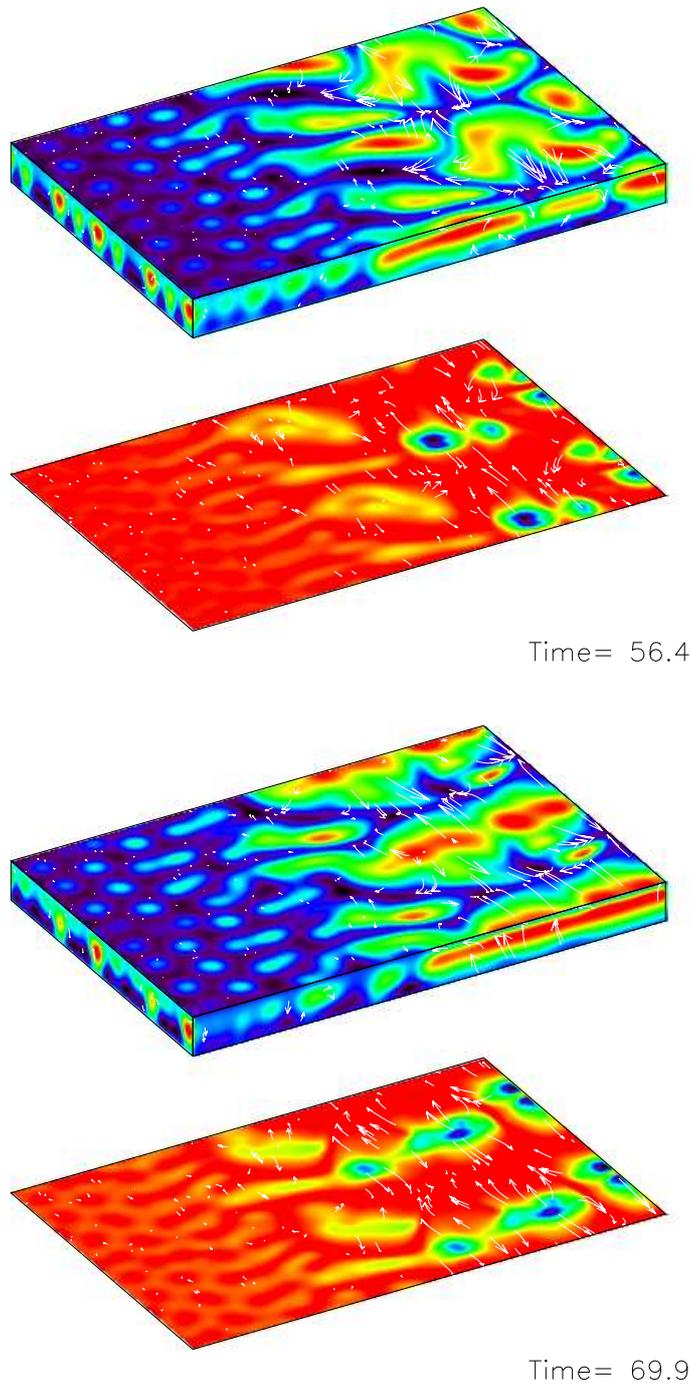


Figure 5.25: Results for the initial condition depicted in Figure 5.24, and with the perfectly conducting upper boundary condition. The result for $Q = 500$ is shown. The solution is time-dependent; snapshots at two different times are shown.

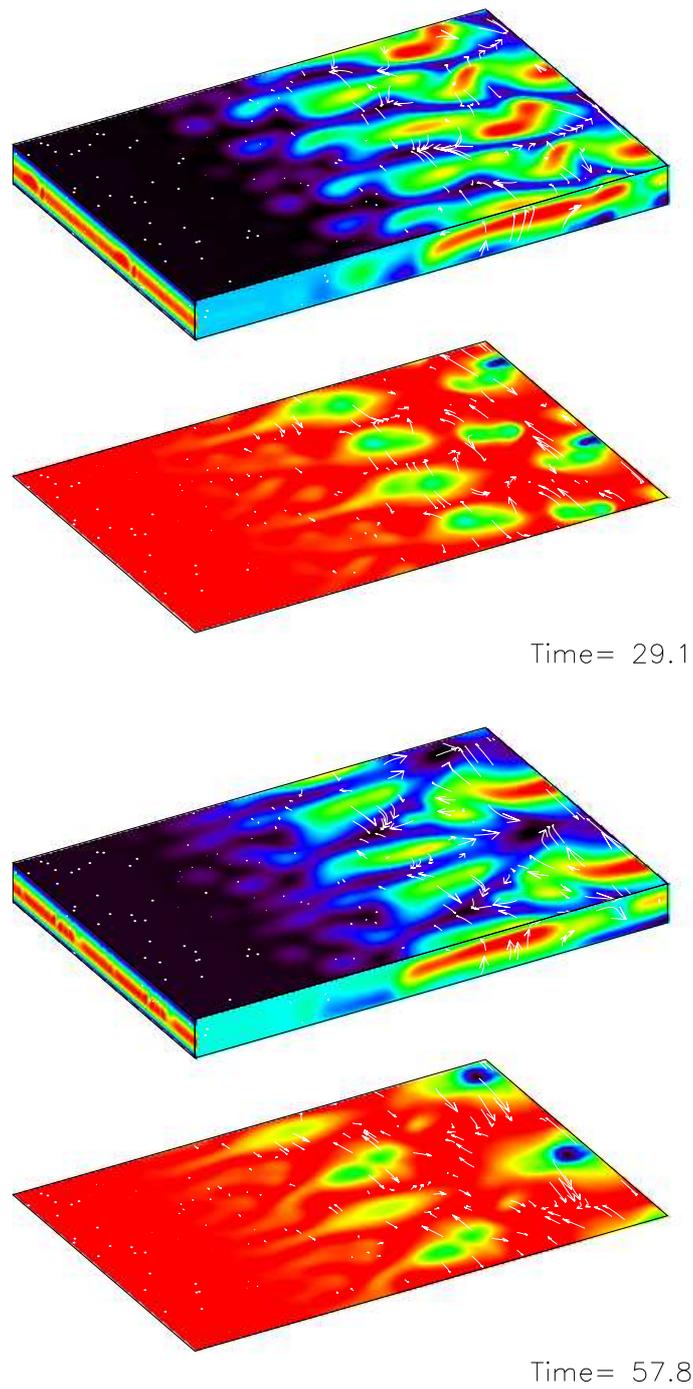


Figure 5.26: As Figure 5.25, but for $Q = 1000$. The upper panel shows the initial (transient) state and the lower panel shows the new state that appears after a certain time.

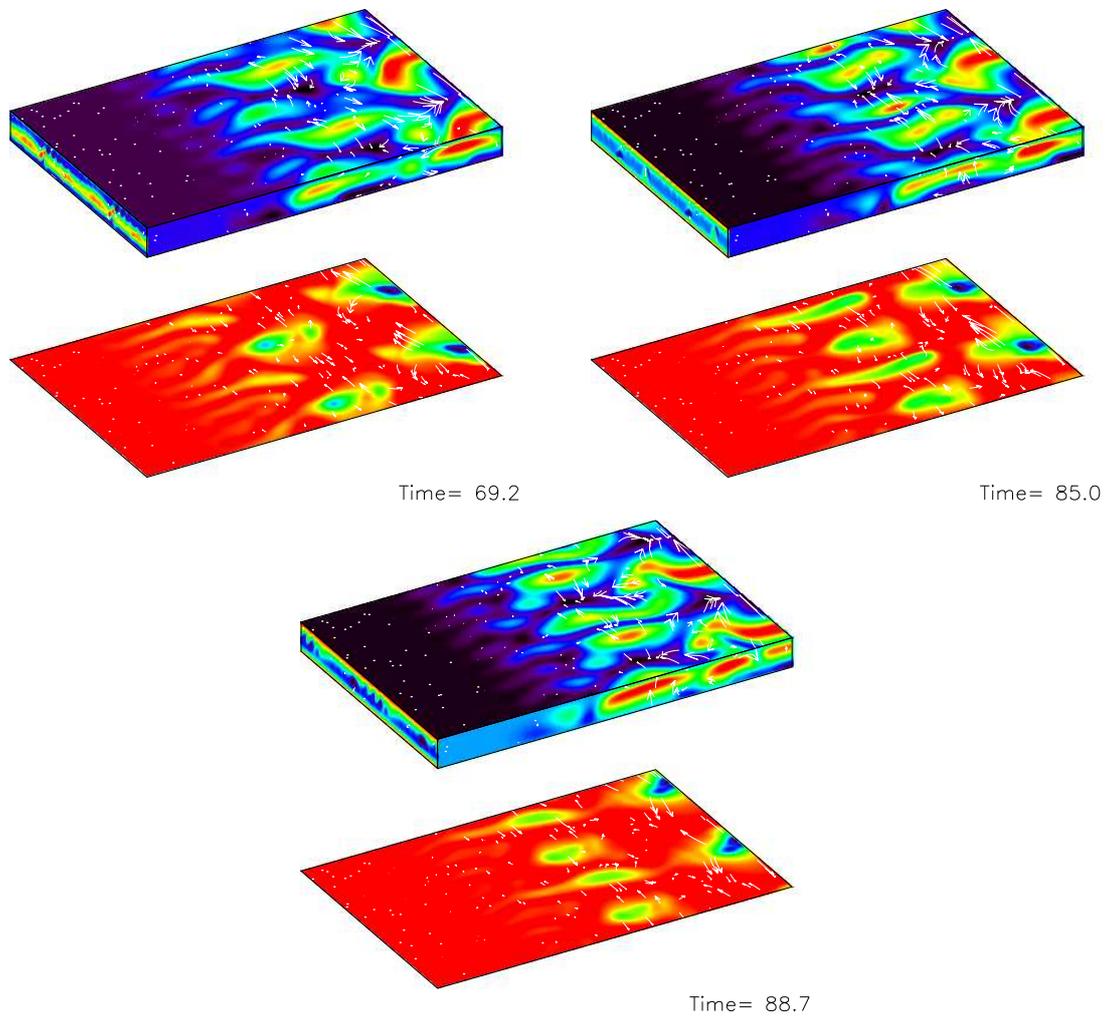


Figure 5.27: The continuation of the simulation for $Q = 1000$ beyond Figure 5.26. Snapshots are shown at three different times. The basic pattern (from Figure 5.26, lower panel) persists, although it is time-dependent in the details. In particular the rolls near the centre appear to wax and wane in amplitude in a chaotic manner.

rolls (Figure 5.26, top panel), but this is only a transient state and after a short time, the rolls break up into two separate groups (bottom panel, same Figure). We now see four smaller rolls near the centre, and two larger roll-like structures at the right-hand end of the domain (actually, there are really only two of the smaller rolls and one of the larger rolls, because the computational domain is only half of what is shown in the Figure).

This basic pattern persists throughout the rest of the calculation. The fine details are however time-dependent: the smaller rolls (near the middle of the box) wax and wane in amplitude, while the larger rolls (at the right-hand end) undulate from side to side somewhat. These changes appear to be chaotic – there is no well-defined period of oscillation, for example; rather the rolls wax and wane in a seemingly random fashion. This time-dependence is illustrated in Figure 5.27, which shows snapshots of this final phase of the solution at three different times.

When $Q = 1500$ (Figure 5.28) we see much the same behaviour as for $Q = 1000$. The only major difference is that the evolution is somewhat slower this time, and it takes longer for the initial roll pattern (upper panel of Figure 5.28) to change into the final pattern (lower panel).

Unfortunately, we did not have time to run the calculation for $Q = 1500$ for quite as long as we would have liked; the calculation was stopped soon after the roll pattern seen in the lower panel of Figure 5.28 first appeared. We cannot say whether this pattern will continue to evolve in a similar way to the previous case (see Figure 5.27) or in some new way; although one might perhaps expect a less vigorous time-dependence in this new case, because of the higher field strength. (Indeed one might expect a transition from the chaotic time-dependence found for $Q = 1000$ to something more regular as Q is increased, or alternatively as R is decreased; however, we did not have time to investigate this idea.)

Note that for $Q = 1000$ and above, the field within the umbra has become too strong for convection to take place there. This is not what we intended (since a model of sunspot convection should ideally contain convection within both the umbra and penumbra), but it is still worth investigating these higher values of Q , since they show what happens to the ‘penumbral’ convection when the field strength is increased.

It seems therefore that our model has two different regimes. The first of these, found at lower values of Q , has convection within the umbra, and a time-dependent pattern within the penumbra. The penumbral pattern does show some anisotropy, but

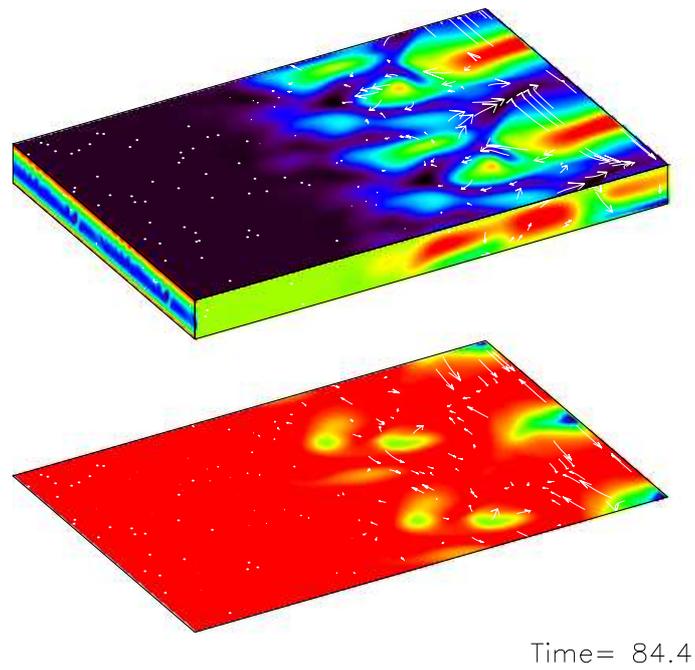
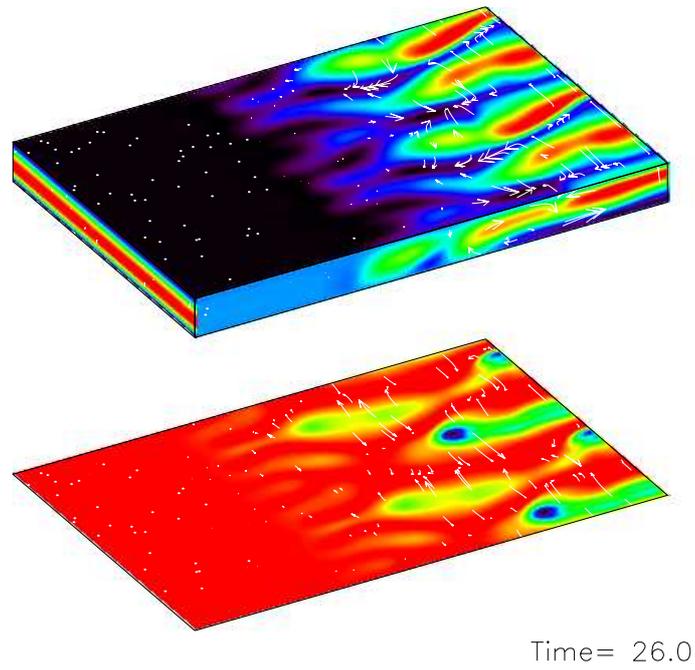


Figure 5.28: As Figure 5.25, but for $Q = 1500$. The upper panel shows the initial (transient) state and the lower panel shows the final state.

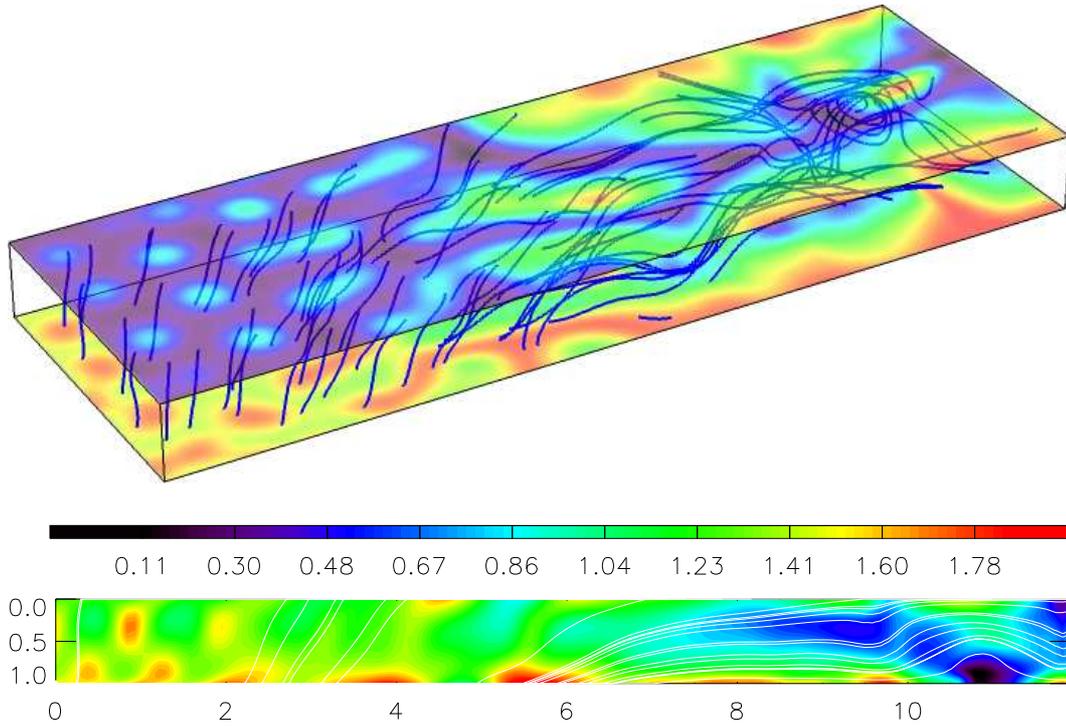


Figure 5.29: Magnetic field configuration for the $Q = 500$ run. (This is a snapshot corresponding to the lower panel of Figure 5.25.) The top panel shows the field lines in 3D, with temperature plots shown at the top and bottom surfaces for reference (with red-yellow-green indicating warm fluid and blue-black indicating cool fluid). The bottom panel shows the y -averaged magnetic field as a 2D plot. (The colour bar indicates the scale for $|\mathbf{B}|$ on the 2D plot.)

no coherent roll-like structures are found. In the second regime, found for higher Q , the penumbral field is strong enough to produce more recognizable field-aligned rolls in the penumbra, but convection is no longer present within the umbra.

The magnetic field structures for the cases $Q = 500$ and $Q = 1000$ are shown in Figures 5.29 and 5.30 (as both three-dimensional plots and y -averaged two-dimensional plots). The field is slightly more tangled and disordered in the former case than in the latter (as might be expected). Notice that in both cases, a significant amount of flux has risen to the upper boundary (visible as a bright green region near the top of the y -averaged plots); this indicates that flux is still rising upwards in the simulations, although this time it is prevented from leaving the box, and instead comes to rest against the upper boundary.

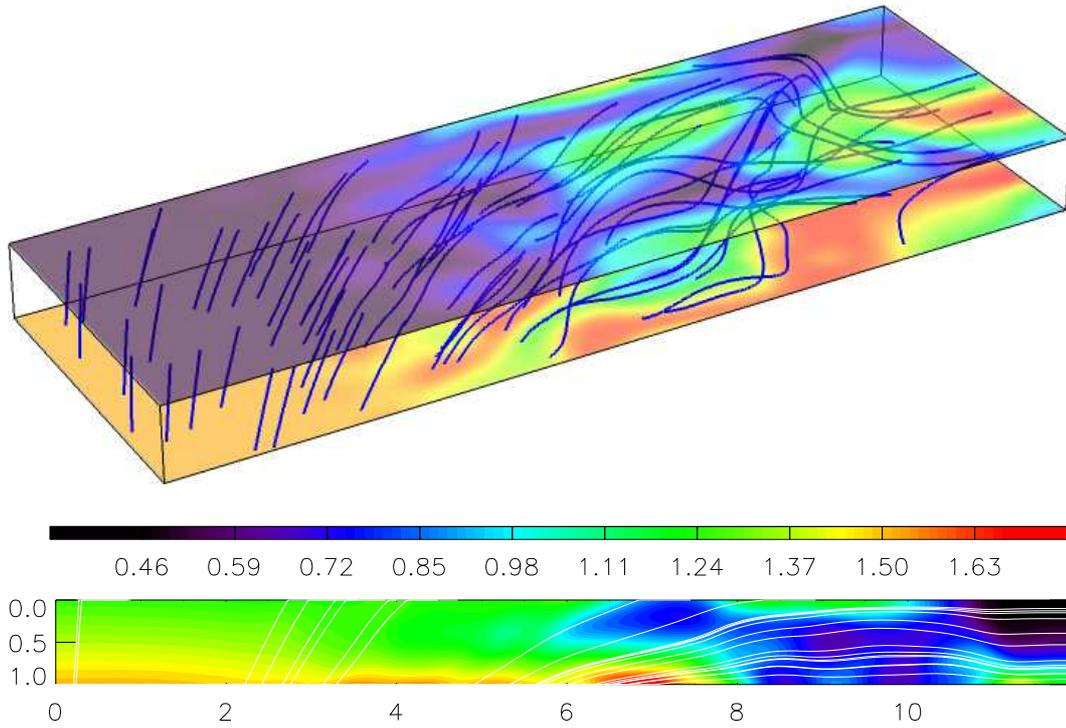


Figure 5.30: As Figure 5.29, but for the case $Q = 1000$. This Figure is a snapshot, taken at a time corresponding to the lower panel of Figure 5.26.

5.7 Conclusions

In this chapter we have looked at three-dimensional magnetoconvection simulations, with both uniform and non-uniform fields.

5.7.1 Runs with uniform fields

Initially we looked at runs where the magnetic field was uniformly inclined at an angle ϕ to the vertical. The aim was to look for changes in the convective pattern with varying ϕ . Although this has been briefly considered before by Hurlburt et al. (2000), they presented only a limited number of results, and used fairly small computational boxes. Here we have made a more thorough investigation, and built up a picture of the transitions that occur as both R and ϕ are varied (although, for reasons of time, we have considered only a single value of Q).

The results clearly show a transition from hexagonal convection for small ϕ , to convection in elongated cells (stretched out along the direction of tilt) for moderate ϕ , to rolls (aligned with the tilt) for larger ϕ . This was more or less in line with what we expected from the weakly nonlinear results of Chapter 3. One difference is that while

hysteresis (bistability) between rolls and hexagons was predicted in the weakly nonlinear analysis, we do not find anything similar in the fully nonlinear model of this chapter.

This can be explained in more detail by comparing Figure 5.2 (on page 161) with the lower part of Figure 3.9(b) (on page 83). Although these diagrams are not identical, we do see a region of rolls in Figure 5.2 for large ϕ or small R (corresponding to region R_1 in Figure 3.9), and a region of cells above this (corresponding to region H in Figure 3.9). According to Figure 3.9, the region of bistability would be found just above this; however, in Figure 5.2, we find only turbulent cellular solutions in this area. This may indicate that we have already left the weakly nonlinear regime before we ever get to the region where bistability was predicted, and hence we do not see the bistability in the fully nonlinear model. (Note however that we only looked at a limited range of parameters – and in particular only one value of Q – so it is always possible that bistability might be present at some other set of parameter values, that we did not investigate.)

Another feature that we find in the nonlinear simulations of this chapter is that the cellular patterns are sometimes subject to a ‘wavy’ modulation (e.g. Figure 5.2, bottom row). This kind of pattern could not be detected by the weakly nonlinear models of Chapter 3, because it involves modulations to the basic wavenumber of the pattern. It appears that we are seeing an instability of the basic, ‘straight’ cellular patterns; but we do not at present have an explanation for why this occurs.

One additional feature worth mentioning is that the results often show a difference between the convective pattern at the top of the box and the pattern at the bottom. See for example Figure 5.3, bottom left, where individual cells are seen near the top, but rolls are seen at the bottom. This is an interesting result since it illustrates that the pattern of convection at the surface of a sunspot (which is what we observe) may not be the same as the pattern in the deeper, sub-surface layers – something that must be borne in mind when interpreting observations of penumbrae.

In summary, the main result of the runs with uniform fields is to illustrate how the convective pattern changes for different values of ϕ , and in particular how it varies from a hexagonal or cellular pattern for small ϕ , to a pattern of rolls when ϕ is large.

The value of Q used to obtain these results corresponded to a strong field regime, and hence strongly anisotropic patterns were found, with the patterns aligning themselves with the field. An interesting future project might be to investigate how the results vary when smaller values of Q are used. In such a case we would expect to see a transition to more isotropic patterns, eventually leading to the granulation-like patterns that are

found in field-free convection, and it would be interesting to investigate the details of the transition between these two extremes.

5.7.2 Runs with non-uniform fields

The runs using non-uniform fields have shown some interesting results. Here we have run a number of simulations in a sunspot-like field configuration, in which the field varies from vertical at one end of the box, representing the umbra, to horizontal at the other end, representing the penumbra. The aim was to see whether a transition similar to the umbra-penumbra transition in a real sunspot could be observed, and to look at the kind of convective patterns found within each of these two regions of the simulation.

Summary of the results

We started off by using an ‘arch’-shaped magnetic field (section 5.5). The results here did indeed show a sharp transition between hexagonal patterns in the ‘umbra’ of the model and roll-like patterns in the ‘penumbra’. The hexagonal patterns correspond to what is expected in a near-vertical field (see e.g. Rucklidge et al. 2000 or Weiss et al. 2002), while the roll-like patterns can be compared to those found by Hurlburt et al. (2000).

Unfortunately, the region of roll-like patterns turned out to be fairly small in our ‘arch’ model, whereas we would actually like a model with a relatively large penumbral region, since this is where the more complex and interesting phenomena are found in real sunspots.

Therefore, we switched to using an initial condition based on a ‘flux sheet’ configuration (section 5.6). The results again showed a sharp distinction between ‘umbral’ and ‘penumbral’ forms of convection. (In fact, the presence of this transition was a robust feature that was seen in all of our simulation runs.) However, this time the actual pattern seen within the penumbra was different. Instead of roll-like structures, we saw a more disordered, almost granulation-like pattern (Figure 5.17). This was because the model had a somewhat reduced field strength in the penumbra, compared to the ‘arch’ model; this was not helped by the fact that much of the flux actually escaped through the upper boundary of the simulation (see e.g. Figure 5.19).

So, although the ‘flux sheet’ model does have a larger penumbral region (as desired), it also has a much stronger difference in the field strength between the umbra and the

penumbra, and consequently only a relatively weak magnetic field within the penumbra. It seems likely that a stronger field would be needed before more ‘penumbra-like’ (more anisotropic) forms of convection would be seen.

Our next step, therefore, was to look into ways of strengthening the magnetic field within the penumbral region. Two changes were made. Firstly, the potential field boundary condition at the top of the layer was replaced by a perfectly conducting condition; this prevented magnetic flux from escaping through the top boundary. Secondly, the initial condition was slightly rearranged, in order to strengthen the field within the penumbral region of the domain a little.

These changes seemed to have the desired effect, at least to a certain extent. For example, with a relatively low Q (of 500), we found that the penumbral convection pattern was more anisotropic than what was previously seen; however, the pattern was also somewhat disordered and time-dependent, with roll-like features continually appearing and disappearing. When Q was increased (to 1000 and above) a roll-like pattern was found in the penumbra, but unfortunately this time there was no convection within the umbra (because of the high field strength there). The penumbral pattern in this case showed an interesting ‘split’ structure with a pair of smaller rolls near the centre of the domain, and one larger roll-like structure at the right-hand edge (where the field was slightly weaker). The penumbral pattern was also somewhat time-dependent with rolls waxing and waning irregularly as a function of time. It would have been interesting to look at these roll structures in greater detail, but unfortunately we did not have time to do so.

Strengths and limitations of the model, and relation to other work

Although the simulations presented above present a reasonable first step towards a ‘unified’ model of sunspot convection (i.e. one containing both umbra-like and penumbra-like regions), there still remains much more work to be done. Perhaps the best way to illustrate the shortcomings of our results is to enumerate the features that we would like an ‘ideal’ model to have:

- convection within both the umbral and penumbral regions (this is important if we wish to investigate the interaction between the two forms of convection);
- a relatively strong field within the penumbral region, so that patterns do become noticeably anisotropic there;

- a relatively large penumbral region (compared to the size of the umbral region), so that we can investigate the more interesting penumbral behaviour in greater detail.

The ‘arch’ model does not satisfy the third point on this list. The ‘flux sheet’ model is better in this respect, since it allows control over the relative width of the umbral and penumbral regions. However, this model has a strong contrast in the field strength between the umbra and penumbra, which means that either the first or the second point on the above list would not be satisfied. The closest that we came to satisfying both points was the runs shown in Figures 5.25 and 5.26, but even here, one must choose between either a fairly weak field strength within the penumbra (Figure 5.25), or a field strength too strong for convection in the umbra (Figure 5.26).

Unfortunately, we have not found a model satisfying all three of the above points, although we hope that future, more complex models might be able to address this. One approach might simply be to look at higher Rayleigh numbers, which would allow us to increase Q to higher values while still having convection within the umbra. (Our runs have all been at relatively low Rayleigh numbers.) We did not do this because we encountered numerical problems; it was found that higher resolutions were required when higher Rayleigh numbers were used.

An alternative method would be to find another initial condition which produces a smaller contrast in field strength between umbra and penumbra. This is quite difficult to do if one restricts oneself to potential field initial conditions (as we have done). However, if one is willing to start from a non-equilibrium initial condition, then more interesting results might be obtained. (The downside of this would be that since the initial condition is not an equilibrium, it would rearrange into some new configuration right from the beginning of the calculation; although one could argue that this was already happening to some extent in our own results, as the magnetic field would eventually be rearranged somewhat once the convection got started.)

There is only one other published model so far that attempts to include both umbral and penumbral regions of convection within the same simulation; this is the cylindrical model described by Hurlburt et al. (2000).³ Here, the authors start with a ‘flux

³Here we are excluding two-dimensional or axisymmetric models, such as those of Hurlburt et al. (2002) or Botha et al. (2006), since, although these do a good job of describing the overall structure of umbrae and penumbrae, they do not show the detailed three-dimensional convection patterns, which is what we are mainly interested in here.

tube’ structure (the cylindrical equivalent of our ‘flux sheet’ models) and run a non-axisymmetric simulation within a wedge (consisting of 1/12 of a full cylinder). The results are similar to our later ‘flux sheet’ runs, in that a pattern of rolls is found within the penumbra, in this case appearing as radial spokes. These are not dissimilar to the rolls found in our own simulations, although our rolls are more disordered and time-dependent than those found by Hurlburt et al. (2000). In the umbral region of the model of Hurlburt et al. (2000), it appears that no significant convective motion is seen; in this sense their model suffers from the same weakness as our model, in that the first point on the above list is not satisfied. (It should be pointed out that the results of Hurlburt et al. 2000 are preliminary. According to Botha et al. 2006, the authors intend to carry out more detailed simulations of this kind, as the next stage of their research project.)

Magnetic buoyancy, flux pumping and boundary conditions

Another very important point concerns the maintenance of the magnetic field configuration against magnetic buoyancy. In our simulations, we set up a carefully constructed magnetic field at the beginning of the simulation; but there is no guarantee that this field structure will be maintained (in any recognizable form) once convection starts up.

Indeed, we found in some of our later simulations that a significant amount of flux escaped through the upper boundary of the computational box. This was the motivation for introducing a perfectly conducting upper boundary condition. However, real sunspots do not have perfectly conducting ‘lids’ over them, so the use of this boundary condition is somewhat unrealistic. On the other hand, we do need *some* method of ensuring that our chosen field structure is maintained against magnetic buoyancy. A proper solution (in the long term) would be to construct a model in which the flux could be held down self-consistently; but in the short term, the use of the conducting upper boundary condition allowed us to investigate the convective patterns that can occur within a sunspot-like field structure, without becoming overly concerned about how to maintain that structure against magnetic buoyancy.

This does of course raise the question of how fields within real sunspots are maintained against buoyancy. It has been suggested (e.g. Weiss et al. 2004) that turbulent flux pumping plays a key role here. In our simulations, we did find that flux was pumped downwards in some cases (for example, our ‘arch’ runs), but not in others.⁴ Clearly the

⁴It should also be pointed out that placing a stable layer underneath the main convecting region, as

balance between magnetic buoyancy and flux pumping is something that requires further investigation.

In fact, this is an idea that is currently being looked at in a calculation by Brummell et al. (2006). This calculation uses a similar field configuration to our own; the differences are that the Rayleigh number is much higher, and the box is narrower in the x direction, but deeper in the z direction (with a convectively stable layer being added underneath the main convecting layer). The aim of the calculation is to determine how effective flux pumping is at holding down the magnetic field against buoyancy. The convection in this calculation takes the form of field-aligned rolls, somewhat similar in appearance to the rolls seen in Figure 5.16 above; and preliminary results indicate that the magnetic flux is indeed pumped downwards. These results should help to clarify under what circumstances flux pumping is and is not effective within this kind of field configuration.

Filamentary structure

As a final point, we have not yet mentioned the intricate structure that is found within real penumbrae, such as the bright and dark filaments or the interlocking-comb magnetic structure. Nothing like this was seen in our model; instead of different structures representing bright and dark filaments, we see only the one ‘roll’ pattern within our penumbral region.

One theory that is emerging is that the filamentary structure is caused by an instability that leads to fluting around the edge of a sunspot and to the beginnings of a filamentary structure (a recent description of this idea is given by Weiss et al. 2004). For example, the calculation of Tildesley (2003a,b) (see also Tildesley and Weiss 2004) found an instability within an idealized Boussinesq configuration which led to a filamentary structure. The calculation of Hurlburt et al. (2000) also found a kind of fluting instability (this time within cylindrical geometry).

Therefore, it might be fruitful to look for similar instabilities (along the y -direction) within our model (or one like it). One point to note here is that Tildesley’s instability had a fairly long wavelength in the y -direction, which may indicate that a fairly wide computational box would be needed before any such instability would be seen.

was done by Tobias et al. (2001) for example, might change these results. For example, in cases where the flux was pumped downwards, we found a tendency for flux to build up against the bottom boundary (see e.g. Figure 5.10). If the flux could instead have been pumped down ‘through’ the bottom boundary into a convectively stable region underneath, then the simulation might have had very different results.

Chapter 6

Conclusions and future work

The problem of sunspots remains far from solved. It is true that the basic physical properties and large-scale structure of sunspots can be described reasonably well by considering axisymmetric models (see also section 1.3 in the Introduction). However, the more interesting problem is to explain the fine-scale structure of sunspots, and this is far more challenging. In order to go about this, one needs a detailed understanding of magnetoconvection, and in particular the ways in which the form taken by the convection might vary within different parts of the sunspot.

One of the most fundamental variations between different parts of a sunspot is the difference between the umbra (the central region) and the penumbra (the outer region). This is presumably the result of some new form of convection appearing once the magnetic field becomes sufficiently inclined to the vertical. However, in order to make this idea more concrete, we need to produce models or simulations of magnetoconvection that demonstrate the transitions that occur as the angle of inclination of the field is varied. In this thesis, we have produced several models that do exactly this, ranging from simplified linearized models within uniform fields to three-dimensional numerical simulations involving non-uniform field configurations.

We began in Chapter 2 with an idealized, linear Boussinesq model, which showed the asymmetry between left- and right-travelling waves. In the oscillatory regime ($\zeta < 1$), different types of rolls (parallel, perpendicular and oblique) were possible as the tilt was varied, whereas in the steady regime ($\zeta > 1$), the parallel rolls were always preferred near onset.

In Chapter 3 we moved on to look at a number of weakly nonlinear models. The simplest of these looked at the steady regime ($\zeta > 1$) and considered solutions on a

hexagonal lattice. This model is also perhaps the most relevant to sunspots (since compressible simulations, intended to model sunspots, have indicated that convection is steady near onset – see for example Weiss et al. 1990, 1996).

This hexagonal model allows us to investigate the transition between hexagons, for small tilt angle ϕ , and field-aligned rolls, for larger ϕ ; this transition has been known about for some time (Danielson, 1961), and is likely to be responsible (at least in part) for the difference in appearance between umbra and penumbra. Our model has shown that this transition can be associated with hysteresis.

In Chapter 5 we presented a number of numerical simulations of compressible magnetoconvection within inclined fields. The results (Figure 5.2, page 161) may be compared with the predictions of the weakly nonlinear model (Figure 3.9b, page 83). The transition between hexagons and rolls was observed, as was the lengthening of the hexagonal cells in the direction of tilt for larger ϕ , as predicted. A new pattern of ‘wavy’ cells was also found, as well as turbulent solutions for larger R . However, we did not find anything analogous to the hysteresis observed in the weakly nonlinear model. (Of course, this does not prove that the hysteresis is not present, since it may be that it is found only at parameter values that were omitted from our investigation.)

One aspect that we did not investigate is the dependence of these results on the value of Q (which measures the field strength). Our results have mostly concentrated on the strong-field regime, where the anisotropy has a strong effect, so that the patterns tend to align themselves with the field. However, when Q is small or zero, we would of course expect isotropic patterns. It would be interesting to investigate what happens for intermediate values of Q , since this would reveal the details of the transition between these two extremes.

The remaining work in this thesis has been concerned with overall sunspot structure, and the transition between umbral and penumbral forms of convection, which we have studied by using models in which the magnetic field inclination is a function of position. In Chapter 4 this was done within the context of a weakly nonlinear model, while in Chapter 5 full compressible simulations were attempted.

Chapter 4 demonstrated that a fairly simple model, using only the basic symmetries of the problem and no specific knowledge of solar physics, can nevertheless reproduce many of the basic properties of sunspots. Of particular note is the relatively sharp transition from hexagonal convection (in the umbra) to roll-like convection (in the penumbra) found in the model. This demonstrates that the umbra-penumbra transition (viewed as

a transition from one type of convection to another) is a fundamental property of this kind of convection, and does not rely on solar physics as such, but simply the (minimal) assumptions regarding symmetries of the situation. The model does not however produce anything resembling the complex filamentary structure found in real sunspots. This can be interpreted as saying that the filamentary structure is less fundamental and *does* require more complex solar physics for its understanding.

Finally, in Chapter 5 we aimed to reproduce these kinds of results within a fully non-linear, compressible simulation. Here, we set up a magnetic field varying from vertical at one end of the domain to horizontal at the other, and then ran the simulation to see what form of convection would be produced. The main result of these simulations was that a relatively sharp transition between two different forms of convection, one in the umbra and one in the penumbra, was found.

One problem with this model was that there was, within our chosen initial field setup, quite a large contrast in the field strength between the umbra and penumbra. We typically found that either the field within the umbra was too strong (preventing any umbral convection from occurring) or that the field within the penumbra was too weak (meaning that the field did not exert a very strong influence on the resulting convection pattern). Another issue was that we were forced to use a perfectly conducting upper boundary condition (while a potential field condition would have been more realistic) in order to prevent flux from escaping via magnetic buoyancy. We hope that future work will be able to produce an improved model that solves some of these problems, and also to investigate the various convective regimes in more detail than we were able to here.

In conclusion, this thesis has looked at different ways of modelling the changes in the form of magnetoconvection that would be expected as the tilt angle of the magnetic field (ϕ) is varied. In particular, we conclude that a sharp transition in the form of convection, from hexagons to more roll-like structures, would be expected once ϕ exceeds some critical value (as indicated by the results of Chapters 4 and 5); this transition is a robust effect, and is therefore likely to be at least partly responsible for the distinction between the umbra and the penumbra of a sunspot.

We say only ‘partly responsible’ because a change in the form of convection would not in itself be enough to account for all of the rich and complex structure that is observed in real penumbrae. Rather, the transition from hexagonal to roll-like convection is likely to act as a starting point which triggers the formation of the more complex magnetic structures and flow patterns that make up a real penumbra. We hope that the models

presented here (and especially the compressible models of Chapter 5) will provide a suitable starting point for more detailed investigations into these enigmatic phenomena.

Appendix A

Magnetic boundary conditions

Here we derive the boundary conditions on \mathbf{B} for perfectly electrically conducting and perfectly electrically insulating boundary surfaces.

A.1 Perfectly conducting boundaries (tied field lines)

Our preferred boundary condition at the bottom of the layer is to imagine that the material below the layer is a perfect electrical conductor. We can then use the condition that $\mathbf{E} = 0$ within a perfect conductor. The tangential components of \mathbf{E} are always continuous across a boundary between two materials, so E_x and E_y must both be zero at the bottom of the fluid layer. Now, Ohm's law tells us that

$$\mathbf{j} = \sigma(\mathbf{E} + \mathbf{u} \wedge \mathbf{B}) \quad (\text{A.1})$$

while one of Maxwell's equations gives us (neglecting the displacement current)

$$\mathbf{j} = \frac{1}{\mu_0} \nabla \wedge \mathbf{B}. \quad (\text{A.2})$$

Equating these two expressions for \mathbf{j} , and setting E_x and E_y to zero, gives

$$u_y B_z = \eta \left(\frac{\partial B_z}{\partial y} - \frac{\partial B_y}{\partial z} \right) \quad (\text{A.3})$$

$$-u_x B_z = \eta \left(\frac{\partial B_x}{\partial z} - \frac{\partial B_z}{\partial x} \right) \quad (\text{A.4})$$

(where we have also used that $u_z = 0$ on the boundary). Here $\eta = 1/(\mu_0 \sigma)$ as usual. Note that after non-dimensionalizing, η becomes ζ (in the scaling used in Chapter 2) or $\zeta_0 \kappa$ (in the scaling used in Chapter 5).

A.2 Perfectly insulating boundaries (matching to a potential field)

Above the layer, we wish to match our magnetic field onto a potential field, which is equivalent to assuming that the material above the layer is perfectly insulating, with $\mathbf{j} = 0$. This implies that $\nabla \wedge \mathbf{B} = 0$, or equivalently $\mathbf{B} = \nabla \Omega$ for some scalar Ω (the ‘potential’).

Assume that the top of our layer is at $z = 0$, with z increasing downwards. Suppose that $\mathbf{B} = \mathbf{B}^+$ (a known function) within $z > 0$, and $\mathbf{B} = \mathbf{B}^-$ (to be determined) in $z < 0$. We assume periodic boundary conditions in the horizontal, so that \mathbf{B}^\pm can be expanded as Fourier series, as follows:

$$\mathbf{B}^\pm = \sum_{m,n} \tilde{\mathbf{B}}_{mn}^\pm(z) e^{i\mathbf{k}_{mn} \cdot \mathbf{x}} \quad (\text{A.5})$$

(real part understood). Here m and n are non-negative integers which number the different Fourier modes, with corresponding wavevectors $\mathbf{k}_{mn} = (k_{mn}, l_{mn}, 0)$.

We now consider the form of \mathbf{B}^- . This must satisfy the following:

$$\nabla \cdot \mathbf{B}^- = 0 \quad (\text{A.6})$$

$$\nabla \wedge \mathbf{B}^- = 0 \quad (\text{A.7})$$

$$\mathbf{B}^- \rightarrow \mathbf{B}_\infty \quad \text{as } z \rightarrow -\infty \quad (\text{A.8})$$

where \mathbf{B}_∞ is a constant. The first two of these are just Maxwell’s equations (with $\mathbf{j} = 0$), and the third is an additional boundary condition at infinity, which we will require for uniqueness.

We now show that (A.6)–(A.8) are satisfied if and only if \mathbf{B}^- takes the following form:

$$\mathbf{B}^- = \sum_{(m,n) \neq (0,0)} \beta_{mn} e^{i\mathbf{k}_{mn} \cdot \mathbf{x}} e^{|\mathbf{k}_{mn}|z} + \mathbf{B}_\infty, \quad (\text{A.9})$$

where the β_{mn} are defined by

$$\beta_{mn} = C_{mn} \begin{pmatrix} ik_{mn}/|\mathbf{k}_{mn}| \\ il_{mn}/|\mathbf{k}_{mn}| \\ 1 \end{pmatrix}, \quad (\text{A.10})$$

for (arbitrary) complex constants C_{mn} .

It can be shown trivially that (A.9)–(A.10) imply (A.6)–(A.8), simply by substituting the one into the other.

To show (A.6)–(A.8) imply (A.9)–(A.10), we must solve the former for \mathbf{B}^- . Equation (A.5) gives:

$$\mathbf{B}^- = \sum_{m,n} \tilde{\mathbf{B}}_{mn}^-(z) e^{i\mathbf{k}_{mn} \cdot \mathbf{x}}. \quad (\text{A.11})$$

Now

$$\nabla \cdot \mathbf{B}^- = 0, \quad \nabla \wedge \mathbf{B}^- = 0 \quad \Rightarrow \quad \nabla^2 \mathbf{B}^- = 0 \quad (\text{A.12})$$

$$\Rightarrow -|\mathbf{k}_{mn}|^2 + \frac{d^2 \tilde{\mathbf{B}}_{mn}^-}{dz^2} = 0 \quad (\text{A.13})$$

$$\Rightarrow \tilde{\mathbf{B}}_{mn}^- = \alpha_{mn} e^{-|\mathbf{k}_{mn}|z} + \beta_{mn} e^{|\mathbf{k}_{mn}|z} \quad (\text{A.14})$$

where α_{mn} and β_{mn} are constants of integration. These equations apply for all $(m, n) \neq (0, 0)$. Applying the boundary condition that \mathbf{B}^- is bounded as $z \rightarrow -\infty$, we conclude that $\alpha_{mn} = 0$. Meanwhile for $(m, n) = (0, 0)$ (i.e. the constant term in the Fourier series) we find $\tilde{\mathbf{B}}_{00}^- = \gamma z + \delta$, and from the boundary conditions, $\gamma = 0$ and $\delta = \mathbf{B}_\infty$. Hence \mathbf{B}^- is of the form (A.9), but we have yet to show that β_{mn} has the required form. This can be done by applying $\nabla \wedge \mathbf{B}^- = 0$ to obtain the following equations:

$$il_{mn}\beta_{z,mn} = |\mathbf{k}_{mn}|\beta_{y,mn} \quad (\text{A.15})$$

$$ik_{mn}\beta_{z,mn} = |\mathbf{k}_{mn}|\beta_{x,mn} \quad (\text{A.16})$$

$$ik_{mn}\beta_{y,mn} = il_{mn}\beta_{x,mn}. \quad (\text{A.17})$$

Upon solving these we find that the β_{mn} are as defined in (A.10), as required.

Now that we know the form of \mathbf{B}^- , we can apply boundary conditions to match \mathbf{B}^- and \mathbf{B}^+ together. We know that B_z will be continuous across the boundary; B_x and B_y will also be continuous if there is no surface current (which is expected to be the case at an insulating boundary). Furthermore, since $\nabla \cdot \mathbf{B} = 0$, we infer that dB_z/dz will also be continuous across the boundary.

For $(m, n) \neq (0, 0)$, continuity of B_z tells us that

$$C_{mn} = \tilde{B}_{z,mn}^- = \tilde{B}_{z,mn}^+. \quad (\text{A.18})$$

Then from the continuity conditions $B_x^+ = B_x^-$, $B_y^+ = B_y^-$ and $dB_z^+/dz = dB_z^-/dz$ (all

evaluated at $z = 0$), we derive the boundary conditions:

$$\tilde{B}_{x,mn} = \frac{ik_{mn}}{|\mathbf{k}_{mn}|} \tilde{B}_{z,mn} \quad (\text{A.19})$$

$$\tilde{B}_{y,mn} = \frac{il_{mn}}{|\mathbf{k}_{mn}|} \tilde{B}_{z,mn} \quad (\text{A.20})$$

$$\frac{d\tilde{B}_{z,mn}}{dz} = |\mathbf{k}_{mn}| \tilde{B}_{z,mn}. \quad (\text{A.21})$$

(these apply to both \mathbf{B}^+ and \mathbf{B}^- at $z = 0$). For the mean field components, with $(m, n) = (0, 0)$, we simply have that $\mathbf{B}_{00}^+ = \mathbf{B}_{00}^- = \mathbf{B}_\infty$ on the boundary.

We can also derive the equivalent condition for a potential field *below* the layer (i.e. a potential field in $z > 1$). The same derivation as above can be used; the results are the same except that a minus sign must be inserted on the right-hand side of each of (A.19), (A.20) and (A.21).

Appendix B

Representations for \mathbf{B}

In this Appendix we consider two different alternative representations for the magnetic field: the vector potential (\mathbf{A}) and poloidal and toroidal potentials (B_P and B_T). These are useful when programming numerical simulations for MHD, since they provide ways of ensuring that Maxwell's equation $\nabla \cdot \mathbf{B} = 0$ is satisfied.

B.1 The vector potential

Since $\nabla \cdot \mathbf{B} = 0$, we may write \mathbf{B} as the curl of a vector potential \mathbf{A} , as follows:

$$\mathbf{B} = \nabla \wedge \mathbf{A}. \quad (\text{B.1})$$

We also know from another of Maxwell's equations that

$$\nabla \wedge \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \quad (\text{B.2})$$

$$= -\frac{\partial \nabla \wedge \mathbf{A}}{\partial t}, \quad (\text{B.3})$$

from which

$$\mathbf{E} = -\nabla\Phi - \frac{\partial \mathbf{A}}{\partial t}. \quad (\text{B.4})$$

Here Φ is the *scalar potential* for the electric field.

We can derive an evolution equation for \mathbf{A} by making use of Ohm's law, which gives another expression for \mathbf{E} :

$$\mathbf{E} = -\mathbf{u} \wedge \mathbf{B} + \eta \nabla \wedge \mathbf{B}. \quad (\text{B.5})$$

Equating the two expressions for \mathbf{E} and rearranging, we derive the evolution equation

$$\frac{\partial \mathbf{A}}{\partial t} = \mathbf{u} \wedge \mathbf{B} - \eta \nabla \wedge \mathbf{B} - \nabla\Phi \quad (\text{B.6})$$

for \mathbf{A} .

Note that this representation for \mathbf{E} and \mathbf{B} is not unique. We can make a *gauge transformation*

$$\mathbf{A} \rightarrow \mathbf{A} + \nabla\chi \quad (\text{B.7})$$

$$\Phi \rightarrow \Phi - \frac{\partial\chi}{\partial t} \quad (\text{B.8})$$

(for any function χ) without changing the physical fields \mathbf{E} and \mathbf{B} .

We would now like to make a gauge transformation to eliminate the $\nabla\Phi$ term from (B.6), since that will allow us to eliminate Φ completely from the governing equations for our system. For example, this can be done by choosing χ as follows:

$$\chi = \int \Phi dt, \quad (\text{B.9})$$

which will set Φ to zero everywhere. Note that this gauge condition (that $\nabla\Phi$ should be identically zero) does not quite specify \mathbf{A} and Φ uniquely; for example, χ could be chosen to be a function of position alone (and not time), which would change \mathbf{A} but leave Φ (and hence $\nabla\Phi$) unaltered.

B.1.1 Boundary conditions

Using (B.6) (and an appropriate gauge transformation so that $\nabla\Phi \equiv 0$) we can now evolve our system numerically by using \mathbf{A} as a variable instead of \mathbf{B} . We now need to show how boundary conditions on \mathbf{B} can be rewritten as boundary conditions on \mathbf{A} . It turns out that this is readily done for simple boundary conditions (such as a vertical field condition) but is more troublesome for more complex conditions (e.g. potential or tied field conditions).

Vertical field condition

Here we require $B_x = B_y = 0$ on some horizontal surface ($z = \text{const}$). In terms of \mathbf{A} we have

$$\frac{\partial A_z}{\partial x} = \frac{\partial A_x}{\partial z}, \quad (\text{B.10})$$

$$\frac{\partial A_z}{\partial y} = \frac{\partial A_y}{\partial z}. \quad (\text{B.11})$$

However for a simulation we would require three boundary conditions, one for each component of \mathbf{A} . The third condition will come from the choice of gauge. We claim

that it is possible to choose a gauge in which (a) $A_z = 0$ on the boundary (so that the boundary conditions simply become $A_z = \partial A_x / \partial z = \partial A_y / \partial z = 0$) and (b) $\nabla\Phi \equiv 0$ (so that the evolution equation (B.6) is simplified).

This can be demonstrated as follows. First of all we make the transformation given above (equation B.9) to satisfy condition (b). Now we have

$$\frac{\partial \mathbf{A}}{\partial t} = -\mathbf{E} \quad (\text{B.12})$$

and it may be verified from (B.5) that $E_z = 0$ on the boundary (if \mathbf{B} is vertical there). Hence in this gauge A_z will remain constant in time on the boundary (although it may vary in space).

Now we will make a further transformation that sets A_z to zero on the boundary whilst preserving the condition $\nabla\Phi = 0$ everywhere. First let $a(x, y)$ be the value of A_z on the boundary in the old gauge. Now make the gauge transformation with

$$\chi = -za(x, y); \quad (\text{B.13})$$

we have

$$A_z^{\text{new}} = A_z^{\text{old}} + \frac{\partial \chi}{\partial z} \quad (\text{B.14})$$

$$= A_z^{\text{old}} - a(x, y), \quad (\text{B.15})$$

which is zero on the boundary as required. In addition χ does not depend on time so Φ and hence $\nabla\Phi$ are not changed.

This argument can be generalized to two horizontal boundary surfaces instead of one (e.g. using a gauge transformation of the form $\chi = -za(x, y) + z^2b(x, y)$).

Other boundary conditions

Unfortunately it is difficult to generalize this argument to other types of boundary condition. Take, for example, the tied field boundary condition, which states that the horizontal components of \mathbf{E} must be zero on the boundary surface. If we take $\nabla\Phi = 0$, then it is readily seen from (B.4) that the A_x and A_y will be constant (in time) on the boundary, but A_z will vary with time. Therefore, we have boundary conditions on A_x and A_y , but not A_z .

As above, the boundary condition on A_z will come from the gauge condition. However, we are constrained because we have already chosen to use a gauge in which $\nabla\Phi = 0$.

We must derive a boundary condition on A_z which is consistent with this gauge condition. For the vertical field case, this proved straightforward since we could make a gauge transformation to set $A_z = 0$. However, that gauge transformation will not work in this case because A_z is time-dependent. It is difficult to see how to derive a boundary condition for A_z that is consistent with the condition $\nabla\Phi = 0$.

Another option would be to abandon the choice $\nabla\Phi = 0$. For example, if we used the well-known *Coulomb gauge*, in which $\nabla \cdot \mathbf{A} = 0$, we would straight away have a boundary condition for $\partial A_z / \partial z$ in terms of derivatives of the (known) functions A_x and A_y . However, in this gauge $\nabla\Phi \neq 0$ in general, and so we have to deal with the extra complexity in the induction equation (B.6).

Note that in two dimensions, the problems disappear. In this case, we usually have $\mathbf{A} = A(x, z)\mathbf{e}_y$, and $\Phi \equiv 0$, and it can be seen that the only necessary boundary condition is that A is constant (in time) along the boundary.

B.2 Poloidal/toroidal decomposition

In view of these difficulties with boundary conditions, it seems preferable to use a different approach, in which the gauge-related ambiguities are removed. This can be done by decomposing \mathbf{B} into separate poloidal and toroidal potentials, as follows:

$$\mathbf{B} = \nabla \wedge (B_T \mathbf{e}_z) + \nabla \wedge (\nabla \wedge (B_P \mathbf{e}_z)). \quad (\text{B.16})$$

The components of \mathbf{B} are now represented as

$$\mathbf{B} = \begin{pmatrix} \partial B_T / \partial y + \partial^2 B_P / \partial x \partial z \\ -\partial B_T / \partial x + \partial^2 B_P / \partial y \partial z \\ -\nabla_H^2 B_P \end{pmatrix} \quad (\text{B.17})$$

where $\nabla_H^2 \equiv \partial^2 / \partial x^2 + \partial^2 / \partial y^2$.

The induction equation for \mathbf{B} is

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \wedge (\mathbf{u} \wedge \mathbf{B}) + \eta \nabla^2 \mathbf{B}. \quad (\text{B.18})$$

We take the scalar product of this equation with \mathbf{e}_z to obtain an equation for B_P :

$$\frac{\partial B_P}{\partial t} = \eta \nabla^2 B_P - \nabla_H^{-2} [\mathbf{e}_z \cdot \nabla \wedge (\mathbf{u} \wedge \mathbf{B})] \quad (\text{B.19})$$

(note that the ∇_H^{-2} operator can easily be computed in Fourier space). Furthermore, taking the curl of the induction equation and then forming the scalar product with \mathbf{e}_z

produces the following equation for B_T :

$$\frac{\partial B_T}{\partial t} = \eta \nabla^2 B_T - \nabla_H^{-2} [\mathbf{e}_z \cdot \nabla \wedge \nabla \wedge (\mathbf{u} \wedge \mathbf{B})]. \quad (\text{B.20})$$

Let

$$\boldsymbol{\mathcal{E}} = \mathbf{u} \wedge \mathbf{B} = \begin{pmatrix} u_y B_z - u_z B_y \\ u_z B_x - u_x B_z \\ u_x B_y - u_y B_x \end{pmatrix}. \quad (\text{B.21})$$

Then

$$\mathbf{e}_z \cdot \nabla \wedge \boldsymbol{\mathcal{E}} = \frac{\partial \mathcal{E}_y}{\partial x} - \frac{\partial \mathcal{E}_x}{\partial y} \quad (\text{B.22})$$

and

$$\mathbf{e}_z \cdot \nabla \wedge \nabla \wedge \boldsymbol{\mathcal{E}} = \frac{\partial^2 \mathcal{E}_x}{\partial x \partial z} + \frac{\partial^2 \mathcal{E}_y}{\partial y \partial z} - \nabla_H^2 \mathcal{E}_z \quad (\text{B.23})$$

so

$$\frac{\partial B_T}{\partial t} = \eta \nabla^2 B_T + \mathcal{E}_z - \nabla_H^{-2} \left(\frac{\partial^2 \mathcal{E}_x}{\partial x \partial z} + \frac{\partial^2 \mathcal{E}_y}{\partial y \partial z} \right) \quad (\text{B.24})$$

$$\frac{\partial B_P}{\partial t} = \eta \nabla^2 B_P + \nabla_H^{-2} \left(\frac{\partial \mathcal{E}_x}{\partial y} - \frac{\partial \mathcal{E}_y}{\partial x} \right). \quad (\text{B.25})$$

B.2.1 A note about periodicity

Simulations are usually run with periodic boundary conditions in x and y ; however, if we require B_T and B_P to be periodic, then we find that we cannot represent a uniform magnetic field. The solution is to treat the mean part of the field separately:

$$\mathbf{B} = \nabla \wedge (B_T \mathbf{e}_z) + \nabla \wedge (\nabla \wedge (B_P \mathbf{e}_z)) + \bar{\mathbf{B}} \quad (\text{B.26})$$

where B_T and B_P are periodic and $\bar{\mathbf{B}}$ is the mean field. We can assume without loss of generality that the horizontal means of both B_T and B_P are zero. \bar{B}_z is constant because vertical flux is conserved, while \bar{B}_x and \bar{B}_y will be functions of z and t .

This change does not explicitly affect the evolution equations for B_T and B_P . However, a new evolution equation for $\bar{\mathbf{B}}$ must be added. This can be found by averaging the induction equation:

$$\frac{\partial \bar{\mathbf{B}}}{\partial t} = \nabla \wedge (\overline{\mathbf{u} \wedge \mathbf{B}}) + \eta \nabla^2 \bar{\mathbf{B}} \quad (\text{B.27})$$

(where the bar represents an average over x and y).

B.2.2 Boundary conditions

Vertical field

Imposing a vertical field boundary condition is straightforward, since we require only $B_x = B_y = 0$, which gives

$$B_T = \frac{\partial B_P}{\partial z} = 0; \quad \bar{B}_x = \bar{B}_y = 0 \quad (\text{B.28})$$

at the boundaries.

Potential field

For a potential field above the layer we require (in Fourier space)

$$B_x = \frac{ik}{|\mathbf{k}|} B_z \quad (\text{B.29})$$

$$B_y = \frac{il}{|\mathbf{k}|} B_z \quad (\text{B.30})$$

$$\frac{\partial B_z}{\partial z} = |\mathbf{k}| B_z. \quad (\text{B.31})$$

The last of these implies

$$\frac{\partial B_P}{\partial z} = |\mathbf{k}| B_P \quad (\text{B.32})$$

and then (B.29) and (B.30) together imply

$$B_T = 0. \quad (\text{B.33})$$

The mean field on the boundaries is simply set to the value of the potential field at infinity.

Tied field

For tied field lines we use the fact that

$$E_x = E_y = 0 \quad (\text{B.34})$$

on the boundary. Substituting for \mathbf{E} from (B.5), we obtain

$$-\mathcal{E}_x + \eta \frac{\partial^2 B_T}{\partial x \partial z} - \eta \frac{\partial \nabla^2 B_P}{\partial y} = 0 \quad (\text{B.35})$$

$$-\mathcal{E}_y + \eta \frac{\partial^2 B_T}{\partial y \partial z} + \eta \frac{\partial \nabla^2 B_P}{\partial x} = 0 \quad (\text{B.36})$$

where \mathcal{E} is as defined in (B.21). Differentiating (B.35) w.r.t. y and (B.36) w.r.t. x , and subtracting the results, gives

$$-\frac{\partial \mathcal{E}_x}{\partial y} + \frac{\partial \mathcal{E}_y}{\partial x} - \eta \nabla^2 \nabla_H^2 B_P = 0 \quad (\text{B.37})$$

which, from the induction equation, is equivalent to

$$\frac{\partial B_P}{\partial t} = 0. \quad (\text{B.38})$$

(In other words this is a Dirichlet condition with B_P specified on the boundary.) We can derive a condition on B_T by differentiating (B.35) w.r.t. x and (B.36) w.r.t. y and adding the results:

$$-\frac{\partial \mathcal{E}_x}{\partial x} - \frac{\partial \mathcal{E}_y}{\partial y} + \eta \frac{\partial \nabla_H^2 B_T}{\partial z} = 0 \quad (\text{B.39})$$

which can be rearranged to give a condition on $\partial B_T / \partial z$:

$$\frac{\partial B_T}{\partial z} = \frac{1}{\eta} \nabla_H^{-2} \left(\frac{\partial \mathcal{E}_x}{\partial x} + \frac{\partial \mathcal{E}_y}{\partial y} \right) \quad (\text{B.40})$$

The boundary conditions for the mean field must be found separately. From substituting $\bar{E}_x = \bar{E}_y = 0$ into (B.5) we find

$$\frac{\partial \bar{B}_x}{\partial z} = -\frac{1}{\eta} \overline{u_x B_z} \quad (\text{B.41})$$

$$\frac{\partial \bar{B}_y}{\partial z} = -\frac{1}{\eta} \overline{u_y B_z} \quad (\text{B.42})$$

(using $u_z = 0$ on the boundary).

Appendix C

Weakly nonlinear derivations

In this Appendix we will show how the various solution branches to our weakly nonlinear models, including the oscillatory hexagonal and the rhombic models, were found, indicating how the existence and stability results were calculated.

C.1 The formula for A_c in the steady hexagonal model

Here we show how to derive equation (3.15) from chapter 3. This formula applied to the steady hexagonal model with $\theta = 30^\circ$, so that $A_1 > A_2 = A_3 = 1$.

We derive this equation by using the results of Malomed et al. (1994), and in particular their Figure 3. Their γ corresponds to our $\mu_{2r} = \mu_{3r}$, and their γ_3 corresponds to our μ_{1r} . When $r = 0$ we are essentially following the path

$$\begin{aligned}\gamma_3 &= A_1\phi^2 \\ \gamma &= \phi^2\end{aligned}$$

which is the straight line

$$\gamma = (1/A_1)\gamma_3. \tag{C.1}$$

Their curve FHK represents the pitchfork bifurcation at which R_1 rolls become stable. (This curve has the equation $\gamma = \beta\gamma_3 - \sqrt{\gamma_3}$.) If the line (C.1) crosses this curve to the left of the point H on their diagram, then the pitchfork is supercritical; if it crosses to the right of H, the pitchfork is subcritical. The point H corresponds to

$$\gamma_3 = \gamma_3^+ \equiv \frac{(2\beta + \sqrt{2\beta + 2})^2}{4(1 + \beta - 2\beta^2)^2} \tag{C.2}$$

while the ‘crossing point’ corresponds to

$$\frac{1}{A_1}\gamma_3 = \beta\gamma_3 - \sqrt{\gamma_3} \quad (\text{C.3})$$

which has the non-trivial solution

$$\gamma_3 = \frac{1}{\left(\frac{1}{A_1} - \beta\right)^2}. \quad (\text{C.4})$$

A supercritical bifurcation will therefore occur if and only if

$$\frac{1}{\left(\frac{1}{A_1} - \beta\right)^2} < \frac{(2\beta + \sqrt{2\beta + 2})^2}{4(1 + \beta - 2\beta^2)^2}. \quad (\text{C.5})$$

Since $\beta > 1$ and $1/A_1 < 1$ by assumption, we have that $1/A_1 - \beta < 0$, while $2\beta + \sqrt{2\beta + 2}$ is certainly positive. It can also be shown that $1 + \beta - 2\beta^2$ is always negative for $\beta > 1$. Therefore the inequality (C.5) is equivalent to

$$\frac{1}{\beta - 1/A_1} < \frac{2\beta + \sqrt{2\beta + 2}}{-2(1 + \beta - 2\beta^2)} \quad (\text{C.6})$$

which can be rearranged to give

$$\frac{1}{A_1} < \beta + \frac{2(1 + \beta - 2\beta^2)}{2\beta + \sqrt{2\beta + 2}}. \quad (\text{C.7})$$

If the right-hand side is positive, then this corresponds to a supercritical bifurcation for $A_1 > A_c$, and a subcritical bifurcation for $A_1 < A_c$, where A_c is the reciprocal of the right-hand side of (C.7). If A_c turns out to be negative, however, then the bifurcation is always *subcritical*.

C.2 Solving the weakly nonlinear amplitude equations: Some predefined equation systems

We now turn to the derivations of the various solution branches to the amplitude equations from Chapter 3, and in particular the oscillatory models (on both the rhombic and the hexagonal lattices).

There are a few systems of equations that will appear many times during the following analysis. We will refer to these as ‘system 1’ through to ‘system 6’. It will save time if we present the solutions (and stability criteria) for these common equation systems now, rather than repeating the results later. The reader may prefer to skip ahead to the next section, and refer back as needed to the equation systems defined here.

C.2.1 System 1

This is a linear system given by

$$\dot{x} = Ax + Be^{i\omega t}y \quad (\text{C.8})$$

$$\dot{y} = Be^{-i\omega t}x + Cy \quad (\text{C.9})$$

where A and B are complex and ω is real. We can solve it by substituting $z = e^{i\omega t}y$:

$$\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} A & B \\ B & C + i\omega \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \quad (\text{C.10})$$

If the eigenvalues of the above matrix both have negative real parts, then the solutions are exponentially decaying; otherwise they are exponentially growing.

C.2.2 System 2

System 2 will be the normal form equations for the Hopf bifurcation problem with $O(2)$ symmetry, as follows:

$$\dot{x} = Ax + C|x|^2x + D|y|^2x \quad (\text{C.11})$$

$$\dot{y} = By + C|y|^2y + D|x|^2y \quad (\text{C.12})$$

This is a standard problem. We are interested in the solution with x and y both nonzero, for which we may decompose the system into amplitude and phase equations by substituting $x = Re^{i\theta}$, $y = Se^{i\phi}$; this leads to the following solution for R , S , θ and ϕ :

$$|x|^2 = \frac{B_r D_r - A_r C_r}{C_r^2 - D_r^2} \quad (\text{C.13})$$

$$|y|^2 = \frac{A_r D_r - B_r C_r}{C_r^2 - D_r^2} \quad (\text{C.14})$$

$$d/dt(\arg x) = A_i + C_i|x|^2 + D_i|y|^2 \quad (\text{C.15})$$

$$d/dt(\arg y) = B_i + C_i|y|^2 + D_i|x|^2. \quad (\text{C.16})$$

The condition for existence of these solutions is that $|x|^2$ and $|y|^2$ both be positive; the condition for stability is that both $C_r < 0$ and $|C_r| > |D_r|$.

C.2.3 System 3

This is a linear system defined as follows:

$$\dot{x} = Ax + Be^{i\omega t}\bar{y} \quad (\text{C.17})$$

$$\dot{y} = Cy + Be^{i\omega t}\bar{x} \quad (\text{C.18})$$

To solve, substitute $z = e^{i\omega t}\bar{y}$:

$$\begin{pmatrix} \dot{x} \\ \dot{z} \end{pmatrix} = \begin{pmatrix} A & B \\ \bar{B} & \bar{C} + i\omega \end{pmatrix} \begin{pmatrix} x \\ z \end{pmatrix} \quad (\text{C.19})$$

The equations have exponentially decaying solutions if and only if the eigenvalues of this matrix both have negative real parts.

C.2.4 System 4

System 4 is the following:

$$\dot{x} = Ax + \alpha_3 e^{i(-\omega_1 + \omega_2)t}y + \alpha_2 e^{i(-\omega_1 + \omega_3)t}z \quad (\text{C.20})$$

$$\dot{y} = By + \alpha_1 e^{i(-\omega_2 + \omega_3)t}z + \alpha_3 e^{i(\omega_1 - \omega_2)t}x \quad (\text{C.21})$$

$$\dot{z} = Cz + \alpha_2 e^{i(\omega_1 - \omega_3)t}x + \alpha_1 e^{i(\omega_2 - \omega_3)t}y \quad (\text{C.22})$$

where A , B and C are complex and ω_1 , ω_2 and ω_3 are real.

To solve, substitute $X = e^{i\omega_1 t}x$, $Y = e^{i\omega_2 t}y$, $Z = e^{i\omega_3 t}z$, to obtain

$$\begin{pmatrix} \dot{X} \\ \dot{Y} \\ \dot{Z} \end{pmatrix} = \begin{pmatrix} i\omega_1 + A & \alpha_3 & \alpha_2 \\ \alpha_3 & i\omega_2 + B & \alpha_1 \\ \alpha_2 & \alpha_1 & i\omega_3 + C \end{pmatrix} \begin{pmatrix} X \\ Y \\ Z \end{pmatrix}. \quad (\text{C.23})$$

The three eigenvalues of this matrix must have negative real parts for stability.

C.2.5 System 5

System 5 is the following nonlinear system:

$$\dot{x} = Ax + D|x|^2x + E(|y|^2 + |z|^2)x \quad (\text{C.24})$$

$$\dot{y} = By + D|y|^2y + E(|x|^2 + |z|^2)y \quad (\text{C.25})$$

$$\dot{z} = Cz + D|z|^2z + E(|x|^2 + |y|^2)z \quad (\text{C.26})$$

As usual, we are interested in solutions with constant non-zero amplitudes $|x|$, $|y|$ and $|z|$, but time-varying phases. These will be as follows:

$$|x|^2 = \frac{E_r(B_r + C_r) - (D_r + E_r)A_r}{(D_r - E_r)(D_r + 2E_r)} \quad (\text{C.27})$$

$$|y|^2 = \frac{E_r(A_r + C_r) - (D_r + E_r)B_r}{(D_r - E_r)(D_r + 2E_r)} \quad (\text{C.28})$$

$$|z|^2 = \frac{E_r(A_r + B_r) - (D_r + E_r)C_r}{(D_r - E_r)(D_r + 2E_r)} \quad (\text{C.29})$$

with

$$\frac{d}{dt}(\arg x) = A_i + D_i|x|^2 + E_i(|y|^2 + |z|^2) \quad (\text{C.30})$$

$$\frac{d}{dt}(\arg y) = B_i + D_i|y|^2 + E_i(|x|^2 + |z|^2) \quad (\text{C.31})$$

$$\frac{d}{dt}(\arg z) = C_i + D_i|z|^2 + E_i(|x|^2 + |y|^2) \quad (\text{C.32})$$

For stability, note that the amplitude and phase equations decouple, so we only need to consider stability for the amplitude equations. The Jacobian of the three amplitude equations is as follows:

$$\begin{pmatrix} A_r + E_r(|y|^2 + |z|^2) + 3D_r|x|^2 & 2E_r|x||y| & 2E_r|x||z| \\ 2E_r|x||y| & B_r + E_r(|x|^2 + |z|^2) + 3D_r|y|^2 & 2E_r|y||z| \\ 2E_r|x||z| & 2E_r|y||z| & C_r + E_r(|x|^2 + |y|^2) + 3D_r|z|^2 \end{pmatrix}. \quad (\text{C.33})$$

The solutions are stable if the three eigenvalues of this matrix have negative real parts.

C.2.6 System 6

This is a generalization of system 5, defined as follows:

$$\dot{x} = Ax + D|x|^2x + E|y|^2x + F|z|^2x \quad (\text{C.34})$$

$$\dot{y} = By + D|y|^2y + E|x|^2y + F|z|^2y \quad (\text{C.35})$$

$$\dot{z} = Cz + D|z|^2z + F(|x|^2 + |y|^2)z \quad (\text{C.36})$$

To solve this, we can set $\text{Re}(\dot{x}/x) = \text{Re}(\dot{y}/y) = \text{Re}(\dot{z}/z) = 0$, which results in a matrix equation which can be solved numerically to obtain the three amplitudes. The solutions will be stable if the eigenvalues of the following matrix both have negative real parts:

$$\begin{pmatrix} A_r + 3D_r|x|^2 + E_r|y|^2 + F_r|z|^2 & 2E_r|x||y| & 2F_r|x||z| \\ 2E_r|x||y| & B_r + E_r|x|^2 + 3D_r|y|^2 + F_r|z|^2 & 2F_r|y||z| \\ 2F_r|x||z| & 2F_r|y||z| & C_r + 3D_r|z|^2 + F_r(|x|^2 + |y|^2) \end{pmatrix}. \quad (\text{C.37})$$

C.3 Rhombic model

For the oscillatory model of section 3.5, the equations to be solved are:

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_2|^2 + b|w_1|^2 + 2a|z_2|^2)z_1 + bw_2z_2\bar{w}_1 \quad (\text{C.38})$$

$$\dot{z}_2 = \mu z_2 + (a|z_2|^2 + b|w_1|^2 + b|w_2|^2 + 2a|z_1|^2)z_2 + bw_1z_1\bar{w}_2 \quad (\text{C.39})$$

$$\dot{w}_1 = \mu' w_1 + (a|w_1|^2 + b|z_2|^2 + b|z_1|^2 + 2a|w_2|^2)w_1 + bw_2z_2\bar{z}_1 \quad (\text{C.40})$$

$$\dot{w}_2 = \mu' w_2 + (a|w_2|^2 + b|z_1|^2 + b|z_2|^2 + 2a|w_1|^2)w_2 + bw_1z_1\bar{z}_2 \quad (\text{C.41})$$

We look for solutions in which the amplitudes are constant and the phases are periodic functions of time. We do this by splitting into different cases depending on how many of the amplitudes are non-zero.

C.3.1 One non-zero amplitude

If three of the four complex amplitudes are zero we get travelling rolls. We will take z_1 to be non-zero, to obtain TRo^R (TRo^R with z_2 non-zero are equivalent, and TRo^L, with either w_1 or w_2 non-zero, can be found analogously). The equation for z_1 is

$$\dot{z}_1 = \mu z_1 + a|z_1|^2 z_1 \quad (\text{C.42})$$

This has the solution

$$|z_1|^2 = -\mu_r/a_r \quad (\text{C.43})$$

$$\frac{d}{dt}(\arg z_1) = \mu_i + a_i|z_1|^2 \quad (\text{C.44})$$

The solutions are stable to perturbations in z_1 if $a_r < 0$.

The perturbations to the other variables decouple:

$$\delta \dot{z}_2 = (\mu + 2a|z_1|^2)\delta z_2 \quad (\text{C.45})$$

$$\delta \dot{w}_1 = (\mu' + b|z_1|^2)\delta w_1 \quad (\text{C.46})$$

$$\delta \dot{w}_2 = (\mu' + b|z_1|^2)\delta w_2 \quad (\text{C.47})$$

For stability, these must all have exponentially decaying solutions, which is true iff $\mu'_r/\mu_r < b_r/a_r$.

C.3.2 Two non-zero amplitudes

Standing rolls (SRo)

For example, z_1 and z_2 non-zero. The equations for these variables are

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_1|^2)z_1 \quad (\text{C.48})$$

$$\dot{w}_1 = \mu' w_1 + (a|w_1|^2 + b|z_1|^2)w_1 \quad (\text{C.49})$$

This corresponds to our ‘system 2’ (see page 217), with $A = \mu$, $B = \mu'$, $C = a$, $D = b$, $x = z_1$ and $y = w_1$. The stability conditions are: $a_r < 0$ and $|a_r| > |b_r|$.

Perturbations to the other variables give

$$\delta \dot{z}_2 = (\mu + b|w_1|^2 + 2a|z_1|^2)\delta z_2 + b w_1 z_1 \delta \bar{w}_2 \quad (\text{C.50})$$

$$\delta \dot{w}_2 = (\mu' + b|z_1|^2 + 2a|w_1|^2)\delta w_2 + b z_1 w_1 \delta \bar{z}_2 \quad (\text{C.51})$$

This corresponds to system 3 (defined on page 218), with $A = \mu + b|w_1|^2 + 2a|z_1|^2$, $B = b|w_1||z_1|$, $C = \mu' + b|z_1|^2 + 2a|w_1|^2$, $\omega = d/dt(\arg z_1) + d/dt(\arg w_1) = \mu_i + \mu'_i + (a_i + b_i)(|z_1|^2 + |w_1|^2)$. We have stability when the eigenvalues of the following matrix both have negative real parts:

$$M = \begin{pmatrix} \mu + b|w_1|^2 + 2a|z_1|^2 & b|w_1||z_1| \\ \bar{b}|w_1||z_1| & \bar{\mu}' + \bar{b}|z_1|^2 + 2\bar{a}|w_1|^2 + i\omega \end{pmatrix}. \quad (\text{C.52})$$

We can simplify this: by eliminating μ_r and μ'_r in favour of $|z_1|^2$ and $|w_1|^2$, we obtain

$$M = \begin{pmatrix} a|z_1|^2 + i(\mu_i + b_i|w_1|^2 + a_i|z_1|^2) & b|w_1||z_1| \\ \bar{b}|w_1||z_1| & \bar{a}|w_1|^2 + i(\mu_i + a_i|z_1|^2 + b_i|w_1|^2) \end{pmatrix}. \quad (\text{C.53})$$

We can now add $-i(\mu_i + b_i|w_1|^2 + a_i|z_1|^2)$ times the identity matrix to M ; this does not affect the real parts of its eigenvalues. The resulting matrix is

$$M' = \begin{pmatrix} a|z_1|^2 & b|w_1||z_1| \\ \bar{b}|w_1||z_1| & \bar{a}|w_1|^2 \end{pmatrix}. \quad (\text{C.54})$$

The trace of this matrix is $a(|z_1|^2 + |w_1|^2)$ which has negative real part (since $a_r < 0$ is already a stability condition, see above). The determinant is $(|a|^2 - |b|^2)|w_1|^2|z_1|^2$, which is negative and has the sign of $|a|^2 - |b|^2$. From this information we see that the two eigenvalues of M' (and therefore also M) either both have negative real parts, if $|a|^2 > |b|^2$, or both have positive real parts, if $|a|^2 < |b|^2$.

Overall, therefore, SRo are stable if $a_r < 0$, $|a_r| > |b_r|$ and $|a| > |b|$.

Leftward- and rightward-travelling rectangles (TRe^L and TRe^R)

We look at TRe^R with z_1 and z_2 non-zero. The equations for z_1 and z_2 are

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + 2a|z_2|^2)z_1 \quad (\text{C.55})$$

$$\dot{z}_2 = \mu z_2 + (a|z_2|^2 + 2a|z_1|^2)z_2 \quad (\text{C.56})$$

This corresponds to system 2 (page 217), with $C = a$ and $D = 2a$. Notice that the stability condition for system 2 is not satisfied with these values for C and D . (The same applies for TRe^R.) Therefore left- or right-travelling rectangles can never be stable.

Perpendicular travelling rectangles (TRe[⊥])

For example, z_1 and w_2 non-zero. The equations are

$$\dot{z}_1 = \mu z_1 + (a|z_1|^2 + b|w_2|^2)z_1 \quad (\text{C.57})$$

$$\dot{w}_2 = \mu' w_2 + (a|w_2|^2 + b|z_1|^2)w_2. \quad (\text{C.58})$$

This corresponds to system 2 (page 217) with $A = \mu$, $B = \mu'$, $C = a$, $D = b$, $x = z_1$ and $y = w_2$. The stability condition is $a_r < 0$ and $|a_r| > |b_r|$.

The perturbations to z_2 and w_1 satisfy

$$\delta \dot{z}_2 = (\mu + b|w_2|^2 + 2a|z_1|^2)\delta z_2 + bz_1 \bar{w}_2 \delta w_1 \quad (\text{C.59})$$

$$\delta \dot{w}_1 = (\mu' + b|z_1|^2 + 2a|w_2|^2)\delta w_1 + bw_2 \bar{z}_1 \delta z_2. \quad (\text{C.60})$$

This corresponds to system 1 (page 217), with $A = \mu + b|w_2|^2 + 2a|z_1|^2$, $B = b|z_1||w_2|$, $C = \mu' + b|z_1|^2 + 2a|w_2|^2$, $\omega = \mu_1 - \mu'_1 + (a_1 - b_1)(|z_1|^2 - |w_2|^2)$. We use similar methods as in section C.3.2: we take the stability matrix from system 1, eliminate μ and μ' , and add an imaginary multiple of the identity matrix, eventually obtaining the following matrix:

$$M = \begin{pmatrix} a|z_1|^2 & b|z_1||w_2| \\ b|z_1||w_2| & a|w_2|^2 \end{pmatrix}. \quad (\text{C.61})$$

We now show that if $a_r < 0$ and $|a_r| > |b_r|$, then both eigenvalues of M always have negative real parts. We first of all note that without loss of generality, we can rescale such that $a_r = -1$ and $|z_1| = 1$. The eigenvalues λ of M are now given by

$$(a - \lambda)(aY - \lambda) - b^2 Y = 0 \quad (\text{C.62})$$

where Y is a shorthand for $|w_2|^2$. We proceed by looking for bifurcation points where $\lambda_r = 0$, i.e. where $\lambda = i\omega$ (ω real). Equation (C.62) becomes

$$(-1 + i(a_i - \omega))(-Y + i(a_i Y - \omega)) - (b_r^2 - b_i^2)Y - 2ib_r b_i Y = 0, \quad (\text{C.63})$$

which may be split into real and imaginary parts, yielding the following two equations:

$$Y - (a_i - \omega)(a_i Y - \omega) - (b_r^2 - b_i^2)Y = 0 \quad (\text{C.64})$$

$$-2a_i Y + \omega Y + \omega - 2b_r b_i Y = 0. \quad (\text{C.65})$$

Equation (C.65) gives an expression for ω , which can then be substituted into (C.64) to obtain (after some simplification):

$$(1 - b_r^2 + a_i^2 + b_i^2) + 2b_r b_i a_i \left(1 - 4\frac{Y}{(1+Y)^2}\right) - 4(b_r^2 b_i^2 + a_i^2) \frac{Y}{(1+Y)^2} = 0. \quad (\text{C.66})$$

The left hand side is minimized with respect to Y (under the constraint $Y > 0$) when $Y = 1$, when it takes the value $(1 - b_r^2)(1 + b_i^2)$. Thus, if $|b_r| < 1$, then equation (C.66) can never be satisfied; therefore there are no bifurcation points in the region $|b_r| < 1$. Since the eigenvalues of M can readily be shown to both be negative at some point within this region (for example if a and b are both real), then it follows, by continuity, that they must both be negative throughout the whole region. (If this were not true then there would have to be a bifurcation point somewhere at which one of the eigenvalues could change sign.)

Therefore, we conclude that TRe^\perp are stable iff $a_r < 0$ and $|a_r| > |b_r|$.

C.3.3 Three nonzero amplitudes

This is not possible – for example, if only w_2 is zero, then an inspection of the \dot{w}_2 equation shows that $\text{Re } \dot{w}_2$ will not be zero (in general), so this will not be a solution (of the type we are looking for here).

C.3.4 Four nonzero amplitudes

If all four amplitudes are nonzero we can write the equations in terms of amplitudes and phases, by setting $z_j = R_j \exp(i\theta_j)$, $w_j = S_j \exp(i\phi_j)$. This yields the following set

of equations:

$$\dot{R}_1 = \mu_r R_1 + (a_r R_1^2 + b_r S_2^2 + b_r S_1^2 + 2a_r R_2^2) R_1 + |b| \cos(\psi_0 - \psi) S_2 S_1 R_2 \quad (\text{C.67})$$

$$\dot{R}_2 = \mu'_r S_2 + (a_r S_2^2 + b_r R_1^2 + b_r R_2^2 + 2a_r S_1^2) S_2 + |b| \cos(\psi_0 + \psi) R_1 S_1 R_2 \quad (\text{C.68})$$

$$\dot{R}_3 = \mu'_r S_1 + (a_r S_1^2 + b_r R_2^2 + b_r R_1^2 + 2a_r S_2^2) S_1 + |b| \cos(\psi_0 - \psi) R_1 S_2 R_2 \quad (\text{C.69})$$

$$\dot{R}_4 = \mu_r R_2 + (a_r R_2^2 + b_r S_1^2 + b_r S_2^2 + 2a_r R_1^2) R_2 + |b| \cos(\psi_0 + \psi) R_1 S_2 S_1 \quad (\text{C.70})$$

$$\begin{aligned} \dot{\psi} = & a_i (-R_1^2 + S_2^2 - S_1^2 + R_2^2) + \left(\frac{S_2 S_1 R_2}{R_1} + \frac{R_1 S_2 R_2}{S_1} \right) |b| \sin(\psi_0 - \psi) \\ & - \left(\frac{R_1 S_1 R_2}{S_2} + \frac{R_1 S_2 S_1}{R_2} \right) |b| \sin(\psi_0 + \psi) \end{aligned} \quad (\text{C.71})$$

where $\psi = \theta_1 - \theta_2 + \phi_1 - \phi_2$.

The equations are now in a form suitable for investigation with AUTO.

C.4 Oscillatory hexagonal model

For the oscillatory hexagonal model, we wish to solve the following equations:

$$\dot{z}_1 = [\mu_1 + a|z_1|^2 + b|w_1|^2 + c(|z_2|^2 + |z_3|^2) + d(|w_2|^2 + |w_3|^2)]z_1 + f(z_2 w_2 + z_3 w_3) \bar{w}_1 \quad (\text{C.72})$$

$$\dot{z}_2 = [\mu_2 + a|z_2|^2 + b|w_2|^2 + c(|z_3|^2 + |z_1|^2) + d(|w_3|^2 + |w_1|^2)]z_2 + f(z_3 w_3 + z_1 w_1) \bar{w}_2 \quad (\text{C.73})$$

$$\dot{z}_3 = [\mu_3 + a|z_3|^2 + b|w_3|^2 + c(|z_1|^2 + |z_2|^2) + d(|w_1|^2 + |w_2|^2)]z_3 + f(z_1 w_1 + z_2 w_2) \bar{w}_3 \quad (\text{C.74})$$

$$\dot{w}_1 = [\mu'_1 + a|w_1|^2 + b|z_1|^2 + c(|w_2|^2 + |w_3|^2) + d(|z_2|^2 + |z_3|^2)]w_1 + f(z_2 w_2 + z_3 w_3) \bar{z}_1 \quad (\text{C.75})$$

$$\dot{w}_2 = [\mu'_2 + a|w_2|^2 + b|z_2|^2 + c(|w_3|^2 + |w_1|^2) + d(|z_3|^2 + |z_1|^2)]w_2 + f(z_3 w_3 + z_1 w_1) \bar{z}_2 \quad (\text{C.76})$$

$$\dot{w}_3 = [\mu'_3 + a|w_3|^2 + b|z_3|^2 + c(|w_1|^2 + |w_2|^2) + d(|z_1|^2 + |z_2|^2)]w_3 + f(z_1 w_1 + z_2 w_2) \bar{z}_3 \quad (\text{C.77})$$

For simplicity we assume that the real parts of the coefficients a - d and f are all negative. (This ensures that all solutions bifurcate supercritically at onset.)

To solve these, we consider cases in which a number of the z_j and w_j are set equal to zero. We start with the case where only one of these amplitudes is non-zero (which gives travelling rolls), and move up to the situation where all six of them are non-zero. As

before, our approach will be to look for solutions in which the amplitudes are constant, although the phases may vary with time.

In the following sections we will calculate the solutions and their stability by making use of the various equation systems already examined in section C.2. The findings will be presented in a number of tables, each of which will show the equation systems to be solved together with the relevant parameters and information on how the results are to be interpreted. (We have also written a FORTRAN program to automate these existence and stability calculations.)

C.4.1 One non-zero amplitude

This corresponds to travelling rolls. We investigate ‘ z_1 rolls’ as an example; the others can be obtained by cyclic permutations.

The amplitude z_1 satisfies

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1. \quad (\text{C.78})$$

This has the following solution in which $|z_1|$ is constant:

$$|z_1|^2 = -\mu_{1r}/a_r \quad (\text{C.79})$$

$$\frac{d}{dt}(\arg z_1) = \mu_{1i} + a_i|z_1|^2. \quad (\text{C.80})$$

To investigate stability, we now introduce perturbations $z_1 \rightarrow z_1 + \delta z_1$, $z_2 \rightarrow \delta z_2$, etc., and discard terms which are quadratic or higher in the perturbations. After doing this, the equation for \dot{z}_1 decouples from the other five equations.

The solutions of the \dot{z}_1 equation will be stable if they bifurcate for positive μ_{1r} (i.e. $a_r < 0$), and unstable otherwise.

The other five equations, when linearized, yield the following:

$$\delta \dot{z}_2 = (\mu_2 + c|z_1|^2)\delta z_2 \quad (\text{C.81})$$

$$\delta \dot{z}_3 = (\mu_3 + c|z_1|^2)\delta z_3 \quad (\text{C.82})$$

$$\delta \dot{w}_1 = (\mu'_1 + b|z_1|^2)\delta w_1 \quad (\text{C.83})$$

$$\delta \dot{w}_2 = (\mu'_2 + d|z_1|^2)\delta w_2 \quad (\text{C.84})$$

$$\delta \dot{w}_3 = (\mu'_3 + d|z_1|^2)\delta w_3 \quad (\text{C.85})$$

For stability we require the solutions of these equations to decay, which is the case when all of the following conditions are met:

$$\frac{\mu_{2r}}{\mu_{1r}} < \frac{c_r}{a_r}, \quad \frac{\mu_{3r}}{\mu_{1r}} < \frac{c_r}{a_r}, \quad \frac{\mu'_{1r}}{\mu_{1r}} < \frac{b_r}{a_r}, \quad \frac{\mu'_{2r}}{\mu_{1r}} < \frac{d_r}{a_r}, \quad \frac{\mu'_{3r}}{\mu_{1r}} < \frac{d_r}{a_r} \quad (\text{C.86})$$

(In deriving these, we have assumed that $\mu_{1r} > 0$; but this is a requirement for stability anyway.)

C.4.2 Two non-zero amplitudes

Standing rolls (SRo)

When two amplitudes are non-zero there are three different cases, depending on which of the amplitudes are taken to be non-zero. If two ‘opposite’ amplitudes, e.g. z_1 and w_1 , are non-zero, then we have standing rolls (SRo).

After setting $z_2 = z_3 = w_2 = w_3 = 0$, we obtain:

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + b|w_1|^2 z_1 \quad (\text{C.87})$$

$$\dot{w}_1 = \mu'_1 w_1 + a|w_1|^2 w_1 + b|z_1|^2 w_1 \quad (\text{C.88})$$

(These are unchanged even when linearized perturbations to z_2 , z_3 , w_2 and w_3 are allowed.) This corresponds to system 2 (page 217).

The linearized equations for perturbations to the other four variables decouple into two groups:

$$\delta \dot{z}_2 = (\mu_2 + c|z_1|^2 + d|w_1|^2) \delta z_2 + f z_1 w_1 \delta \bar{w}_2 \quad (\text{C.89})$$

$$\delta \dot{w}_2 = (\mu'_2 + c|w_1|^2 + d|z_1|^2) \delta w_2 + f z_1 w_1 \delta \bar{z}_2 \quad (\text{C.90})$$

and

$$\delta \dot{z}_3 = (\mu_3 + c|z_1|^2 + d|w_1|^2) \delta z_3 + f z_1 w_1 \delta \bar{w}_3 \quad (\text{C.91})$$

$$\delta \dot{w}_3 = (\mu'_3 + c|w_1|^2 + d|z_1|^2) \delta w_3 + f z_1 w_1 \delta \bar{z}_3. \quad (\text{C.92})$$

These correspond to system 3 (page 218). See Table C.1.

Travelling rectangles, type 1 (TRe1)

These have z_1 and z_2 non-zero. As before, the equations decouple into three groups. We may first solve for z_1 and z_2 , determining stability with respect to perturbations δz_1 and δz_2 ; we can then consider perturbations to the other variables separately.

The equations for z_1 and z_2 are

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + c|z_2|^2 z_1 \quad (\text{C.93})$$

$$\dot{z}_2 = \mu_2 z_2 + a|z_2|^2 z_2 + c|z_1|^2 z_2 \quad (\text{C.94})$$

Equation system	Parameters	Results
2	$A = \mu_1, B = \mu'_1, C = a, D = b$	The solutions for z_1 (from x) and w_1 (from y), plus two stability eigenvalues
3	$A = \mu_2 + c z_1 ^2 + d w_1 ^2,$ $B = f z_1 w_1 ,$ $C = \mu'_2 + c w_1 ^2 + d z_1 ^2,$ $\omega = d/dt(\arg z_1 + \arg w_1)$	Two stability eigenvalues
3	$A = \mu_3 + c z_1 ^2 + d w_1 ^2,$ $B = f z_1 w_1 ,$ $C = \mu'_3 + c w_1 ^2 + d z_1 ^2,$ $\omega = d/dt(\arg z_1 + \arg w_1)$	Two stability eigenvalues

Table C.1: Existence and stability calculation for SRO. The first column refers to equation systems from section C.2, which are to be solved given the parameter values in the second column. The third column shows what to do with the results. For example, in this case, system 2 is to be solved to determine whether the solution exists, and if so, the values of x and y (from equations C.13–C.16) will give the appropriate solutions for z_1 and z_2 respectively. Also, two stability eigenvalues may be obtained from system 2, and four from system 3, as indicated; the solution is stable if the real parts of all of these eigenvalues are negative.

Equation system	Parameters	Results
2	$A = \mu_1, B = \mu_2, C = a, D = c$	The solutions for z_1 (from x) and z_2 (from y), plus two stability eigenvalues
1	$A = \mu'_1 + b z_1 ^2 + d z_2 ^2, B = f z_1 z_2 ,$ $C = \mu'_2 + b z_2 ^2 + d z_1 ^2,$ $\omega = d/dt(\arg z_2 - \arg z_1)$	Two stability eigenvalues

The other two stability eigenvalues are: $\mu_3 + c(|z_1|^2 + |z_2|^2)$ and $\mu'_3 + d(|z_1|^2 + |z_2|^2)$.

Table C.2: Existence and stability calculation for *TRe1*.

This is equivalent to system 2 (page 217).

The perturbations $\delta\dot{z}_3$ and $\delta\dot{w}_3$ evolve according to

$$\delta\dot{z}_3 = [\mu_3 + c(|z_1|^2 + |z_2|^2)]\delta z_3 \quad (\text{C.95})$$

$$\delta\dot{w}_3 = [\mu'_3 + d(|z_1|^2 + |z_2|^2)]\delta w_3. \quad (\text{C.96})$$

These can both be solved trivially, providing two more stability eigenvalues.

The final two equations needed are

$$\delta\dot{w}_1 = (\mu'_1 + b|z_1|^2 + d|z_2|^2)\delta w_1 + f\bar{z}_1 z_2 \delta w_2 \quad (\text{C.97})$$

$$\delta\dot{w}_2 = (\mu'_2 + b|z_2|^2 + d|z_1|^2)\delta w_2 + f z_1 \bar{z}_2 \delta w_1 \quad (\text{C.98})$$

This corresponds to system 1 (page 217). See Table C.2.

Travelling rectangles, type 2 (TRe2)

These arise when z_1 and w_3 are nonzero. The equations for z_1 and w_3 themselves decouple, and are

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + d|w_3|^2 z_1 \quad (\text{C.99})$$

$$\dot{w}_3 = \mu'_3 w_3 + a|w_3|^2 w_3 + d|z_1|^2 w_3. \quad (\text{C.100})$$

This corresponds to system 2 (page 217).

The equations for small perturbations δz_2 and δw_2 are

$$\delta\dot{z}_2 = (\mu_2 + c|z_1|^2 + d|w_3|^2)\delta z_2 \quad (\text{C.101})$$

$$\delta\dot{w}_2 = (\mu'_2 + c|w_3|^2 + d|z_1|^2)\delta w_2 \quad (\text{C.102})$$

Equation system	Parameters	Results
2	$A = \mu_1, B = \mu'_3, C = a, D = d$	The solutions for z_1 (from x) and w_3 (from y), plus two stability eigenvalues
1	$A = \mu_3 + b w_3 ^2 + c z_1 ^2,$ $B = f z_1 w_3 ,$ $C = \mu'_1 + b z_1 ^2 + c w_3 ^2,$ $\omega = d/dt(\arg z_1 - \arg w_3)$	Two stability eigenvalues

The other two stability eigenvalues are: $\mu_2 + c|z_1|^2 + d|w_3|^2$ and $\mu'_2 + c|w_3|^2 + d|z_1|^2$.

Table C.3: Existence and stability calculation for TRe2.

These can be solved trivially, giving two further stability eigenvalues.

The equations for δz_3 and δw_1 are

$$\delta \dot{z}_3 = (\mu_3 + b|w_3|^2 + c|z_1|^2)\delta z_3 + f z_1 \bar{w}_3 \delta w_1 \quad (\text{C.103})$$

$$\delta \dot{w}_1 = (\mu'_1 + b|z_1|^2 + c|w_3|^2)\delta w_1 + f \bar{z}_1 w_3 \delta z_3 \quad (\text{C.104})$$

This corresponds to system 1 (page 217). See Table C.3.

C.4.3 Three non-zero amplitudes

Oscillating triangles (OT)

These are found when z_1 – z_3 are nonzero. (Of course, there is a similar solution with w_1 – w_3 nonzero, which can be found by exchanging z_j and w_j , and μ_j and μ'_j .)

The equations for z_1 – z_3 decouple from the linearized equations for δw_1 – δw_3 . The former set of equations is as follows:

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + c(|z_2|^2 + |z_3|^2) z_1 \quad (\text{C.105})$$

$$\dot{z}_2 = \mu_2 z_2 + a|z_2|^2 z_2 + c(|z_3|^2 + |z_1|^2) z_2 \quad (\text{C.106})$$

$$\dot{z}_3 = \mu_3 z_3 + a|z_3|^2 z_3 + c(|z_1|^2 + |z_2|^2) z_3. \quad (\text{C.107})$$

This corresponds to system 5 (page 218).

Equation system	Parameters	Results
5	$A = \mu_1, B = \mu_2, C = \mu_3, D = a,$ $E = c$	The solutions for z_1 (from x), z_2 (from y) and z_3 (from z), plus three stability eigenvalues
4	$A = \mu'_1 + b z_1 ^2 + d(z_2 ^2 + z_3 ^2),$ $B = \mu'_2 + b z_2 ^2 + d(z_3 ^2 + z_1 ^2),$ $C = \mu'_3 + b z_3 ^2 + d(z_1 ^2 + z_2 ^2),$ $\alpha_1 = f z_2 z_3 , \alpha_2 = f z_1 z_3 ,$ $\alpha_3 = f z_1 z_2 , \omega_1 = d/dt(\arg z_1),$ $\omega_2 = d/dt(\arg z_2), \omega_3 = d/dt(\arg z_3)$	Three stability eigenvalues

Table C.4: Existence and stability calculation for OT.

The equations for the δw 's are:

$$\delta \dot{w}_1 = (\mu'_1 + b|z_1|^2 + d(|z_2|^2 + |z_3|^2))\delta w_1 + f\bar{z}_1 z_2 \delta w_2 + f\bar{z}_1 z_3 \delta w_3 \quad (\text{C.108})$$

$$\delta \dot{w}_2 = (\mu'_2 + b|z_2|^2 + d(|z_3|^2 + |z_1|^2))\delta w_2 + f z_1 \bar{z}_2 \delta w_1 + f \bar{z}_2 z_3 \delta w_3 \quad (\text{C.109})$$

$$\delta \dot{w}_3 = (\mu'_3 + b|z_3|^2 + d(|z_1|^2 + |z_2|^2))\delta w_3 + f z_1 \bar{z}_3 \delta w_1 + f z_2 \bar{z}_3 \delta w_2. \quad (\text{C.110})$$

This corresponds to system 4 (page 218). See Table C.4.

New solution branch

As mentioned in Chapter 3, there is an additional solution to these equations (not found by Roberts et al. 1986), which is found by choosing z_1 , z_2 and w_3 to be non-zero. The equations for these three quantities are then

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + c|z_2|^2 z_1 + d|w_3|^2 z_1 \quad (\text{C.111})$$

$$\dot{z}_2 = \mu_2 z_2 + a|z_2|^2 z_2 + c|z_1|^2 z_2 + d|w_3|^2 z_2 \quad (\text{C.112})$$

$$\dot{w}_3 = \mu'_3 w_3 + a|w_3|^2 w_3 + d(|z_1|^2 + |z_2|^2)w_3 \quad (\text{C.113})$$

This corresponds to our system 6 (page 219).

Equation system	Parameters	Results
6	$A = \mu_1, B = \mu_2, C = \mu'_3, D = a,$ $E = c, F = d$	The solutions for z_1 (from x), z_2 (from y) and w_3 (from z), plus three stability eigenvalues
4	$A = \mu'_1 + b z_1 ^2 + c w_3 ^2 + d z_2 ^2,$ $B = \mu'_2 + b z_2 ^2 + c w_3 ^2 + d z_1 ^2,$ $C = \mu_3 + b w_3 ^2 + c(z_1 ^2 + z_2 ^2),$ $\alpha_1 = f z_2 w_3 , \alpha_2 = f z_1 w_3 ,$ $\alpha_3 = f z_1 z_2 , \omega_1 = d/dt(\arg z_1),$ $\omega_2 = d/dt(\arg z_2), \omega_3 = d/dt(\arg w_3)$	Three stability eigenvalues

Table C.5: Existence and stability calculation for the new solution branch (section C.4.3).

The linearized equations for perturbations to the other three variables are

$$\delta\dot{z}_3 = (\mu_3 + b|w_3|^2 + c(|z_1|^2 + |z_2|^2))\delta z_3 + f z_1 \bar{w}_3 \delta w_1 + f z_2 \bar{w}_3 \delta w_2 \quad (\text{C.114})$$

$$\delta\dot{w}_1 = (\mu'_1 + b|z_1|^2 + c|w_3|^2 + d|z_2|^2)\delta w_1 + f \bar{z}_1 z_2 \delta w_2 + f \bar{z}_1 w_3 \delta z_3 \quad (\text{C.115})$$

$$\delta\dot{w}_2 = (\mu'_2 + b|z_2|^2 + c|w_3|^2 + d|z_1|^2)\delta w_2 + f \bar{z}_2 w_3 \delta z_3 + f \bar{z}_2 z_1 \delta w_1 \quad (\text{C.116})$$

This corresponds to our system 4 (page 218). See Table C.5.

C.4.4 Four non-zero amplitudes

Only certain combinations are possible here. For example, if $w_2 = w_3 = 0$, but the other four amplitudes are nonzero, then an inspection of the equations reveals that \dot{w}_2 and \dot{w}_3 would be nonzero in general, so these two amplitudes would immediately become nonzero, which is a contradiction. More generally, we cannot have three z 's and one w , or three w 's and one z , nonzero; we must have two of each being nonzero.

Without loss of generality, therefore, we can assume that $z_1 z_2 \neq 0$ and $z_3 = 0$. For \dot{z}_3 to be zero we would require $w_3 = 0$ (the other possibility is $w_1 = w_2 = 0$, but by assumption we must have four nonzero amplitudes, not three). We conclude that z_1, z_2, w_1 and w_2 are nonzero in this case. (All other cases with four nonzero amplitudes can now be generated by cyclic permutations.)

The equations for the four nonzero amplitudes in this case are

$$\dot{z}_1 = \mu_1 z_1 + a|z_1|^2 z_1 + b|w_1|^2 z_1 + c|z_2|^2 z_1 + d|w_2|^2 z_1 + f z_2 w_2 \bar{w}_1 \quad (\text{C.117})$$

$$\dot{z}_2 = \mu_2 z_2 + a|z_2|^2 z_2 + b|w_2|^2 z_2 + c|z_1|^2 z_2 + d|w_1|^2 z_2 + f z_1 w_1 \bar{w}_2 \quad (\text{C.118})$$

$$\dot{w}_1 = \mu'_1 w_1 + a|w_1|^2 w_1 + b|z_1|^2 w_1 + c|w_2|^2 w_1 + d|z_2|^2 w_1 + f z_2 w_2 \bar{z}_1 \quad (\text{C.119})$$

$$\dot{w}_2 = \mu'_2 w_2 + a|w_2|^2 w_2 + b|z_2|^2 w_2 + c|w_1|^2 w_2 + d|z_1|^2 w_2 + f z_1 w_1 \bar{z}_2 \quad (\text{C.120})$$

These can be re-written as amplitude and phase equations:

$$\dot{R}_1 = (\mu_{1r} + a_r R_1^2 + b_r S_1^2 + c_r R_2^2 + d_r S_2^2) R_1 + |f| \cos(\arg f - \psi) R_2 S_2 S_1 \quad (\text{C.121})$$

$$\dot{R}_2 = (\mu_{2r} + a_r R_2^2 + b_r S_2^2 + c_r R_1^2 + d_r S_1^2) R_2 + |f| \cos(\arg f + \psi) R_1 S_1 S_2 \quad (\text{C.122})$$

$$\dot{S}_1 = (\mu'_{1r} + a_r S_1^2 + b_r R_1^2 + c_r S_2^2 + d_r R_2^2) S_1 + |f| \cos(\arg f - \psi) S_2 R_2 R_1 \quad (\text{C.123})$$

$$\dot{S}_2 = (\mu'_{2r} + a_r S_2^2 + b_r R_2^2 + c_r S_1^2 + d_r R_1^2) S_2 + |f| \cos(\arg f + \psi) S_1 R_1 R_2 \quad (\text{C.124})$$

$$\begin{aligned} \dot{\psi} &= \mu_{1i} - \mu_{2i} + \mu'_{1i} - \mu'_{2i} \\ &+ (a_i + b_i - c_i - d_i)(R_1^2 + S_1^2 - R_2^2 - S_2^2) \\ &+ |f| \sin(\arg f - \psi) \left(\frac{R_2 S_1 S_2}{R_1} + \frac{R_1 R_2 S_2}{S_1} \right) \\ &- |f| \sin(\arg f + \psi) \left(\frac{R_1 S_1 S_2}{R_2} + \frac{R_1 R_2 S_1}{S_2} \right). \end{aligned} \quad (\text{C.125})$$

Here $R_j = |z_j|$ and $S_j = |w_j|$ ($j = 1, 2$), and $\psi = \arg z_1 - \arg z_2 + \arg w_1 - \arg w_2$. The value of ψ determines which of two possible solution types occurs: $\psi = 0$ corresponds to standing rectangles (SRe), and $\psi = \pi$ corresponds to wavy rolls of the first kind (WR1).

The other two equations needed are

$$\delta \dot{z}_3 = [\mu_3 + c(|z_1|^2 + |z_2|^2) + d(|w_1|^2 + |w_2|^2)] \delta z_3 + f(z_1 w_1 + z_2 w_2) \delta \bar{w}_3 \quad (\text{C.126})$$

$$\delta \dot{w}_3 = [\mu'_3 + c(|w_1|^2 + |w_2|^2) + d(|z_1|^2 + |z_2|^2)] \delta w_3 + f(z_1 w_1 + z_2 w_2) \delta \bar{z}_3 \quad (\text{C.127})$$

These equations can all be placed into AUTO to determine existence/stability for particular cases.

C.4.5 Five non-zero amplitudes

This is impossible. For example, if $z_1 = 0$ but all the other amplitudes are nonzero, it is clear from the equations that \dot{z}_1 will be nonzero, which contradicts the assumption that z_1 will be zero for all time.

C.4.6 Six non-zero amplitudes

In this case we have the full complexity of equations (C.72)–(C.77) to deal with. We can break the system down into amplitude and phase equations by writing $z_j = R_j \exp(i\theta_j)$, $w_j = S_j \exp(i\phi_j)$. The amplitude equations give the following:

$$\begin{aligned} \dot{R}_1 &= \mu_{1r} R_1 + (a_r R_1^2 + b_r S_1^2 + c_r(R_2^2 + R_3^2) + d_r(S_2^2 + S_3^2)) R_1 \\ &\quad + |f| \cos(\arg f + \psi_3) R_2 S_2 S_1 + |f| \cos(\arg f - \psi_2) R_3 S_3 S_1 \end{aligned} \quad (\text{C.128})$$

$$\begin{aligned} \dot{S}_1 &= \mu'_{1r} S_1 + (a_r S_1^2 + b_r R_1^2 + c_r(S_2^2 + S_3^2) + d_r(R_2^2 + R_3^2)) S_1 \\ &\quad + |f| \cos(\arg f + \psi_3) S_2 R_2 R_1 + |f| \cos(\arg f - \psi_2) S_3 R_3 R_1 \end{aligned} \quad (\text{C.129})$$

with the equations for \dot{R}_2 and \dot{R}_3 being obtained by cyclic permutation. Here $\psi_1 = \arg z_3 + \arg w_3 - \arg z_2 - \arg w_2$; ψ_2 and ψ_3 are defined by cyclic permutations of this. The three ψ_j add up to zero so only two of them need to be kept track of at any one time. The evolution equation for ψ_1 is

$$\begin{aligned} \dot{\psi}_1 &= \mu_{3i} + \mu'_{3i} - \mu_{2i} - \mu'_{2i} \\ &\quad + a_i(R_3^2 + S_3^2 - R_2^2 - S_2^2) + b_i(S_3^2 + R_3^2 - S_2^2 - R_2^2) \\ &\quad + c_i(R_2^2 + S_2^2 - R_3^2 - S_3^2) + d_i(S_2^2 + R_2^2 - S_3^2 - R_3^2) \\ &\quad + |f| \left(\frac{R_1 S_1 S_3}{R_3} + \frac{S_1 R_1 R_3}{S_3} \right) \sin(\arg f + \psi_2) + |f| \left(\frac{R_2 S_2 S_3}{R_3} + \frac{S_2 R_2 R_3}{S_3} \right) \sin(\arg f - \psi_1) \\ &\quad - |f| \left(\frac{R_3 S_3 S_2}{R_2} + \frac{S_3 R_3 R_2}{S_2} \right) \sin(\arg f + \psi_1) - |f| \left(\frac{R_1 S_1 S_2}{R_2} + \frac{S_1 R_1 R_2}{S_2} \right) \sin(\arg f - \psi_3) \end{aligned} \quad (\text{C.130})$$

and again, the equations for $\dot{\psi}_2$ and $\dot{\psi}_3$ can be obtained by cyclic permutation. Depending on the values of these phase variables, we can produce either standing hexagons, standing regular triangles, twisted rectangles, or wavy rolls of the second kind.

The equations in this form can be analysed using the program AUTO.

Bibliography

- G. J. J. Botha, A. M. Rucklidge, and N. E. Hurlburt. Converging and diverging convection around axisymmetric magnetic flux tubes. *Monthly Notices of the Royal Astronomical Society*, 369:1611–1624, 2006.
- J. P. Boyd. *Chebyshev and Fourier Spectral Methods*. Dover, New York, 2001. Full text available online at http://www-personal.engin.umich.edu/~jpboyd/BOOK_Spectral2000.html.
- N. H. Brummell, S. M. Tobias, J. H. Thomas, and N. O. Weiss. Flux pumping in the outer penumbra of a sunspot. *Astrophysical Journal*, 2006. In preparation.
- P. J. Bushby and S. M. Houghton. Spatially intermittent fields in photospheric magnetoconvection. *Monthly Notices of the Royal Astronomical Society*, 362:313–320, 2005.
- F. H. Busse and R. M. Clever. Finite amplitude convection in the presence of an inclined magnetic field. *European Journal of Mechanics, B/Fluids*, 9:225–238, 1990.
- J. R. Cash and D. R. Moore. A high order method for the numerical solution of two-point boundary value problems. *BIT Numerical Mathematics*, 20:44–52, 1980.
- F. Cattaneo, N. H. Brummell, J. Toomre, A. Malagoli, and N. E. Hurlburt. Turbulent compressible convection. *Astrophysical Journal*, 370:282–294, 1991.
- S. Chandrasekhar. *Hydrodynamic and Hydromagnetic Stability*. Clarendon, Oxford, 1961.
- T. Clune and E. Knobloch. Pattern selection in three-dimensional magnetoconvection. *Physica D*, 74:151–176, 1994.
- S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176:430–455, 2002.

- R. E. Danielson. The structure of sunspot penumbras. II. Theoretical. *Astrophysical Journal*, 134:289–311, 1961.
- E. J. Doedel. AUTO: a program for the automatic bifurcation analysis of autonomous systems. *Congressus Numerantium*, 30:264–284, 1981.
- M. Golubitsky, J. W. Swift, and E. Knobloch. Symmetries and pattern selection in Rayleigh-Bénard convection. *Physica D*, 10:249–276, 1984.
- N. E. Hurlburt, D. Alexander, and A. M. Rucklidge. Complete models of axisymmetric sunspots: Magnetoconvection with coronal heating. *Astrophysical Journal*, 577:993–1005, 2002.
- N. E. Hurlburt, P. C. Matthews, and M. R. E. Proctor. Nonlinear compressible convection in oblique magnetic fields. *Astrophysical Journal*, 457:933–938, 1996.
- N. E. Hurlburt, P. C. Matthews, and A. M. Rucklidge. Solar magnetoconvection. *Solar Physics*, 192:109–118, 2000.
- N. E. Hurlburt and A. M. Rucklidge. Development of structure in pores and sunspots: flows around axisymmetric magnetic flux tubes. *Monthly Notices of the Royal Astronomical Society*, 314:793–806, 2000.
- K. Jahn. Current sheet as a diagnostic for the subphotospheric structure of a spot. *Astronomy and Astrophysics*, 222:264–292, 1989.
- K. Jahn and H. U. Schmidt. Thick penumbra in a magnetostatic sunspot model. *Astronomy and Astrophysics*, 290:295–317, 1994.
- K. Julien, E. Knobloch, and S. M. Tobias. Nonlinear magnetoconvection in the presence of strong oblique fields. *Journal of Fluid Mechanics*, 410:285–322, 2000.
- K. Julien, E. Knobloch, and S. M. Tobias. Highly supercritical convection in strong magnetic fields. In A. Ferriz-Mas and M. Núñez, editors, *Advances in Nonlinear Dynamical Systems*, volume 9 of *The Fluid Mechanics of Astrophysics and Geophysics*, pages 195–223, London, 2003. Taylor and Francis.
- K. D. Leka and A. Skumanich. The evolution of pores and the development of penumbrae. *Astrophysical Journal*, 507:454–469, 1998.

- B. A. Malomed, A. A. Nepomnyashchy, and A. E. Nuz. Nonequilateral hexagonal patterns. *Physica D*, 70:357–369, 1994.
- P. C. Matthews. Hexagonal patterns in finite domains. *Physica D*, 116:81–94, 1998.
- P. C. Matthews, N. E. Hurlburt, M. R. E. Proctor, and D. P. Brownjohn. Compressible magnetoconvection in oblique fields: linearized theory and simple nonlinear models. *Journal of Fluid Mechanics*, 240:559–569, 1992.
- P. C. Matthews, M. R. E. Proctor, and N. O. Weiss. Compressible magnetoconvection in three dimensions: planforms and nonlinear behaviour. *Journal of Fluid Mechanics*, 305:281–305, 1995.
- P. C. Matthews and A. M. Rucklidge. Travelling and standing waves in magnetoconvection. *Proceedings of the Royal Society A*, 441:649–658, 1993.
- F. Meyer, H. U. Schmidt, and N. O. Weiss. The stability of sunspots. *Monthly Notices of the Royal Astronomical Society*, 179:741–761, 1977.
- F. Meyer, H. U. Schmidt, P. R. Wilson, and N. O. Weiss. The growth and decay of sunspots. *Monthly Notices of the Royal Astronomical Society*, 169:35–57, 1974.
- E. N. Parker. Sunspots and the physics of magnetic flux tubes. I – The general nature of the sunspot. II – Aerodynamic drag. *Astrophysical Journal*, 230:905–923, 1979.
- W. Pesch and L. Kramer. Nonlinear analysis of spatial structures in two-dimensional anisotropic pattern forming systems. *Zeitschrift für Physik B*, 63:121–130, 1986.
- S. L. Pollicott, P. C. Matthews, and S. M. Cox. Instability of convection in a fluid layer rotating about an oblique axis. *Physical Review E*, 67:016301 (14 pages), 2003.
- W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling. *Numerical Recipes in FORTRAN 77*. Cambridge University Press, 1992.
- M. R. E. Proctor and N. O. Weiss. Magnetoconvection. *Reports on Progress in Physics*, 45:1317–1379, 1982.
- M. Roberts, J. W. Swift, and D. H. Wagner. The Hopf bifurcation on a hexagonal lattice. In M. Golubitsky and J. Guckenheimer, editors, *Multiparameter Bifurcation Theory*, volume 56 of *Contemporary Mathematics*, pages 283–318, Providence, RI, USA, 1986. American Mathematical Society.

- P. H. Roberts. *An Introduction to Magnetohydrodynamics*. Longmans, London, 1967.
- L. H. M. Rouppe van der Voort, M. G. Löfdahl, D. Kiselman, and G. B. Scharmer. Penumbra structure at 0.1 arcsec resolution. I. General appearance and power spectra. *Astronomy and Astrophysics*, 414:717–726, 2004.
- A. Roxin and H. Riecke. Rotating convection in an anisotropic system. *Physical Review E*, 65:046219 (6 pages), 2002.
- A. Roxin and H. Riecke. Erratum: Rotating convection in an anisotropic system. *Physical Review E*, 69:019901(E) (1 page), 2004.
- A. M. Rucklidge, H. U. Schmidt, and N. O. Weiss. The abrupt development of penumbrae in sunspots. *Monthly Notices of the Royal Astronomical Society*, 273:491–498, 1995.
- A. M. Rucklidge, N. O. Weiss, D. P. Brownjohn, P. C. Matthews, and M. R. E. Proctor. Compressible magnetoconvection in three dimensions: pattern formation in a strongly stratified layer. *Journal of Fluid Mechanics*, 419:283–323, 2000.
- B. J. Sams, L. Golub, and N. O. Weiss. X-ray observations of sunspot penumbral structure. *Astrophysical Journal*, 399:313–317, 1992.
- A. Schlüter and St. Temesváry. The internal constitution of sunspots. In B. Lehnert, editor, *Electromagnetic Phenomena in Cosmical Physics*, IAU Symp. No. 6, pages 263–271, Cambridge, 1958. Cambridge University Press.
- R. Schmitz and W. Zimmermann. Deformed hexagonal patterns in a weakly anisotropic system. *Journal de Physique II*, 7:677–683, 1997.
- M. Schüssler and A. Vögler. Magnetoconvection in a sunspot umbra. *Astrophysical Journal*, 641:L73–L76, 2006.
- M. Silber, H. Riecke, and L. Kramer. Symmetry-breaking Hopf bifurcation in anisotropic systems. *Physica D*, 61:260–278, 1992.
- M. Sobotka and A. Hanslmeier. Photometry of umbral dots. *Astronomy and Astrophysics*, 442:323–329, 2005.
- S. K. Solanki. Sunspots: An overview. *Astronomy and Astrophysics Review*, 11:153–286, 2003.

- J. Swift and P. C. Hohenberg. Hydrodynamic fluctuations at the convective instability. *Physical Review A*, 15:319–328, 1977.
- L. Tao, M. R. E. Proctor, and N. O. Weiss. Flux expulsion by inhomogeneous turbulence. *Monthly Notices of the Royal Astronomical Society*, 300:907–914, 1998.
- G. Thiessen. The structure of the sunspot-umbra. *The Observatory*, 70:234–235, 1950.
- J. H. Thomas and N. O. Weiss. Fine structure in sunspots. *Annual Review of Astronomy and Astrophysics*, 42:517–548, 2004.
- J. H. Thomas, N. O. Weiss, S. M. Tobias, and N. H. Brummell. Downward pumping of magnetic flux as the cause of filamentary structures in sunspot penumbrae. *Nature*, 420:390–393, 2002a.
- J. H. Thomas, N. O. Weiss, S. M. Tobias, and N. H. Brummell. Magnetic flux pumping and the structure of a sunspot penumbra. *Astronomische Nachrichten*, 323:383–386, 2002b.
- S. D. Thompson. Magnetoconvection in an inclined magnetic field: linear and weakly non-linear models. *Monthly Notices of the Royal Astronomical Society*, 360:1290–1304, 2005.
- M. J. Tildesley. *Modelling Filamentary Structure in Sunspots*. PhD thesis, University of Cambridge, 2003a.
- M. J. Tildesley. On the origin of filamentary structure in sunspot penumbrae: linear instabilities. *Monthly Notices of the Royal Astronomical Society*, 338:497–507, 2003b.
- M. J. Tildesley and N. O. Weiss. On the origin of filamentary structure in sunspot penumbrae: non-linear results. *Monthly Notices of the Royal Astronomical Society*, 350:657–670, 2004.
- S. M. Tobias, N. H. Brummell, T. L. Clune, and J. Toomre. Transport and storage of magnetic field by overshooting turbulent compressible convection. *Astrophysical Journal*, 549:1183–1203, 2001.
- A. Vögler, S. Shelyag, M. Schüssler, F. Cattaneo, T. Emonet, and T. Linde. Simulations of magneto-convection in the solar photosphere: Equations, methods and results of the MURaM code. *Astronomy and Astrophysics*, 429:335–351, 2005.

- D. Walgraef and C. Schiller. Anisotropy effects on pattern selection near dynamical instabilities. *Physica D*, 27:423–432, 1987.
- N. O. Weiss, D. P. Brownjohn, N. E. Hurlburt, and M. R. E. Proctor. Oscillatory convection in sunspot umbrae. *Monthly Notices of the Royal Astronomical Society*, 245:434–452, 1990.
- N. O. Weiss, D. P. Brownjohn, P. C. Matthews, and M. R. E. Proctor. Photospheric convection in strong magnetic fields. *Monthly Notices of the Royal Astronomical Society*, 283:1153–1164, 1996.
- N. O. Weiss, M. R. E. Proctor, and D. P. Brownjohn. Magnetic flux separation in photospheric convection. *Monthly Notices of the Royal Astronomical Society*, 337:293–304, 2002.
- N. O. Weiss, J. H. Thomas, N. H. Brummell, and S. M. Tobias. The origin of penumbral structure in sunspots: downward pumping of magnetic flux. *Astrophysical Journal*, 600:1073–1090, 2004.
- A. R. Winebarger, E. E. DeLuca, and L. Golub. Apparent flows above an active region observed with the Transition Region and Coronal Explorer. *Astrophysical Journal*, 553:L81–L84, 2001.
- C. Zwaan. The evolution of sunspots. In J. H. Thomas and N. O. Weiss, editors, *Sunspots: Theory and Observations*, volume 375 of *NATO ASI Series C: Mathematical and Physical Sciences*, pages 75–100, Dordrecht, 1992. Kluwer Academic Publishers.