Evolution of the Magnetic Field Distribution of Solar Active Regions

Sally Dacie

Department of Space and Climate Physics

University College London

Supervisors: Professor Lidia van Driel-Gesztelyi and Dr David Long

A thesis submitted for a Master of Philosophy degree (MPhil)

December 22, 2016

I, Sally Dacie, confirm that the work presented in this thesis is my own. Where information has been derived from other sources, I confirm that this has been indicated in the thesis.

_____

Abstract

Magnetic flux emergence through the solar photosphere in the form of an active region (AR) and the subsequent evolution of this flux are crucial processes in the solar cycle for their role in converting toroidal magnetic field to poloidal magnetic field. Although the temporal evolution of ARs is relatively well understood, the processes involved continue to be the subject of investigation. To better characterise how ARs emerge and disperse, the distribution of the vertical component of the photospheric magnetic field (flux density) of a series of ARs has been studied. A kernel density estimation plot of the magnetic field distribution was created on a log-log scale for each AR at each time step. The central portion of the distribution was found to be approximately linear and its slope was used to characterise the evolution of the magnetic field. The slopes were seen to evolve with time, becoming less steep as the fragmented emerging flux coalesces. The slopes reached a maximum value of -1.5 just before the time of maximum flux before becoming steeper during the decay phase towards the quiet Sun value of -3. This behaviour differs significantly from a classical diffusion model, which produces a slope of -1, suggesting that simple classical diffusion is not responsible for the observed changes in field distribution. Other processes must play a significant role in flux dispersion, one of which may be magnetic flux reprocessing by (super)granular convective cells. Moreover, comparisons of the observed distributions to those produced by data from simulations of AR emergence with and without convection suggest that convective processes also play an important role in magnetic flux emergence.

Acknowledgements

I would like to thank Lidia van Driel-Geszteyli and Dave Long for being two of the loveliest supervisors I could have asked for. It has been an absolute pleasure to work with them both and with Pascal Démoulin, who has contributed so much to my work and provided so much insight. Thanks to Pascal and to everyone else in the solar group at l'Observatoire de Paris for being so welcoming on my visits. I would like to thank Mark Linton and Tibor Török for disagreeing with me and then for their excellent explanations. I'd also like to thank Stuart Mumford for convincing me to use python and David Pérez Suárez without whom I wouldn't have succeeded. Thanks to David for always making time to help, knowing whenever I needed to talk about work and being there to bounce ideas off, as well as for fixing my punctures, his animal impressions, his Spanish omelette, trying to repair my window and coming swimming with me in December. Then, thanks to Steph Yardley for working together on the code and Alex James and Magnus Woods for their company in GO1 and all the solar (and non solar) conversations. Thanks also to Deb Baker, Lucie Green and Gherado Valori for their discussions and Francesco Zuccarello for helping me avoid IDL one last time. Thanks to Mark Cheung for giving me his beautiful simulation data, James Leake for sending me data when he was away travelling, Cooper Downs for his extrapolations, Marianno Poisson for his analysis of magnetic tongues and Xudong Sun for the vector magnetogram data. Thanks to my examiners, Vasilis Archontis and Sarah Matthews, for teaching me something new right at the end.

There are so many people who have helped make my time at MSSL. Thanks to Ravi Desai for helping me move house, Emma Gale for always being up for a chat, Zsuzsanna Vizi for her interesting life stories, Adam Lea for telling me when to cycle home to avoid the rain, Monica Tress Barojas for her continuous support, Annie Wellbrock, Vitor Botelho, Priyank Hirani and Sam Taylor for lunchtimes and walks, Si-Ting Xiong for showing me her drawings, Daisuke Katawa for his kindness, Jack Carlyle for his enthusiasm, Jamie Ryan for helping me move office, Mateo Inch for watching shooting stars, Laurence Hogan for inviting everyone to my house, Gethyn Lewis for letting me set off fireworks, Alex Camargio for giving me a ride on the back of his truck, Hubert Hu for badminton and letting me live with him, Roy Oldfield for always making me laugh and Georgie Graham, Denis Gonzalez, Jennifer Chan, Zak Tsiga, Lang Feng, Alice Foster, Ed Malina, Kirthika Mohan and Jason Hunt for being there and offering me somewhere to sleep.

I would like to thank my parents and my sister Ruth for helping me whenever they can and Anna Feledi for being a wonderful friend. Thanks also to Dan Challacombe and Markus Frembs for accompanying me to Paris and to Ellie Ringland, Dawn Brooks, Theresa Lederer and Rachael Collins for letting me stay with them.

Contents

1 Introduction **9**

1.1 The Sun's Magnetic Field and the Solar Cycle 9

1.2 Active Regions (ARs) 12

1.2.1 Asymmetries and evolution of ARs 12

1.2.2 Sunspots 13

1.2.3 Simulations of AR emergence 16

1.2.4 AR decay 19

1.3 Thesis Overview 21

2 Instrumentation and Data **22**

2.1 The Helioseismic Magnetic Imager (HMI) 22

2.1.1 Optical system 23

2.1.2 Line profile fitting 25

2.1.3 Derivation of the magnetic field 26

2.1.4 Errors 27

2.2 Extreme UltraViolet Imager (EUVI) 27

3 Evolution of the Magnetic Field Distribution of Active Regions **29**

3.1 Theory 29

3.1.1 Emergence and clustering 29

3.1.2 Diffusion of a magnetic polarity 31

3.1.3 Diffusion of a magnetic bipole 33

3.1.4 The induction equation 37

3.2 Data and AR Selection 38

3.2.1 Data and data treatment 38

3.2.2 Definition of AR area 38

3.3 Kernel Density Estimation Analysis 40

3.3.1 Method 40

3.3.2 Observed distributions 42

3.4 Observational Results 42

3.4.1 Temporal evolution 42

3.4.2 Evolution versus magnetic flux . . . 43
3.4.3 Decayed ARs . . . 45
3.4.4 Quiet Sun . . . 46
3.5 Possible Issues for the Derived Distributions . . . 47
3.6 Discussion and Conclusions . . . 47

4 Future Work **50**
4.1 A Larger Study . . . 50
4.2 The Effect of Spatial Resolution . . . 51
4.3 The Field Distribution of Simulation Data . . . 53
4.3.1 A simulation including convection . . . 53
4.3.2 Simulations without convection . . . 55
4.3.3 Comparison of the simulations . . . 56

5 Summary **58**

A Magnetohydrodynamics (MHD) **59**

B Magnetic Reconnection **62**

C Table of ARs **64**

References **66**

List of Figures

1.1 Poloidal to toroidal field. . . . 10
1.2 Magnetic butterfly diagram. . . . 11
1.3 Flux cancellation. . . . 11
1.4 Sunspot with fine structure. . . . 14
1.5 MMFs. . . . 15

2.1 HMI optical system. . . . 23
2.2 EUVI telescope. . . . 28

3.1 Superposition versus cancellation. . . . 33
3.2 Distribution for a diffusing bipole model. . . . 34
3.3 Flux decrease for superposing polarities. . . . 35
3.4 Projection effect. . . . 39
3.5 Area selection and dilation procedure. . . . 39
3.6 KDE plot. . . . 41
3.7 Slope against time. . . . 42
3.8 Peak flux against emergence time. . . . 43
3.9 Slope against normalised flux. . . . 44
3.10 Decaying region and quiet Sun examples. . . . 45
3.11 Slope against time for decaying regions. . . . 46
3.12 Slope against normalised flux for ARs near the central meridian. . . . 48
3.13 Evolution of the distribution. . . . 49

4.1 KDEs from magnetograms of different resolutions. . . . 51
4.2 Simulated magnetograms (Rempel & Cheung 2014) and their KDEs. . . . 53
4.3 Slope evolution for simulations (Rempel & Cheung 2014). . . . 54
4.4 Simulated magnetograms (James Leake and Mark Linton) and their KDEs. . . . 55
4.5 Slope evolution for simulations (James Leake and Mark Linton). . . . 56

B.1 Magnetic reconnection. . . . 62

Abbreviations

AR	Active Region
CCD	Charge-Coupled Device
CME	Coronal Mass Ejection
EUV	Extreme Ultra-Violet
EUVI	Extreme UltraViolet Imager
FWHM	Full Width Half Maximum
HMI	Helioseismic Magnetic Imager
KDE	Kernal Density Estimation
KDEs	Kernal Density Estimation plots
MDI	Michelson Doppler Imager
MHD	MagnetoHydroDynamics
MMF	Moving Magnetic Feature
NASA	National Aeronautics and Space Administration
NOAA	National Oceanic and Atmospheric Administration
PIL	Polarity Inversion Line
SDO	*Solar Dynamics Observatory*
STEREO	*Solar TErrestrial RElations Observatory*

Chapter 1

Introduction

1.1 The Sun's Magnetic Field and the Solar Cycle

The evolution of the Sun's magnetic field with time is crucial to the understanding of solar activity, particularly in the context of the 11 year sunspot cycle. The Sun's magnetic field undergoes an oscillation between a poloidal configuration at sunspot minimum and a toroidal configuration at sunspot maximum. Magnetic field is thought to be amplified at the tachocline; the base of the convection zone, where convective transport of energy starts to dominate over radiative transfer. Plasma in the convection zone undergoes differential rotation, with the rotational period near the surface being $\sim$ 25 days at the equator and $\sim$ 34 days at the poles (Choudhuri 2015). The period of rotation is fairly constant with depth in the convection zone, except at high latitudes (above $\sim 60^{\circ}$) where it decreases with depth (Wilson et al. 1997; Schou et al. 1998). In the convection zone, the ratio of plasma pressure to magnetic pressure $\beta > 0$, so the motion of the magnetic field lines is driven by plasma motion (see Alfvén's frozen-in-flux theorem and plasma beta in Appendix A). As a result, the differential rotation causes the field lines to wind up, converting the poloidal magnetic field to a toroidal one and leading to the build up of strong magnetic field in some areas (Figure 1.1).

Regions of magnetic field B have an associated magnetic pressure (an energy density) proportional to B^2. When the magnetic field is strong, the pressure is large, and when the magnetic field strength changes in space there is a gradient in the magnetic pressure. This gradient results in the magnetic pressure force, which acts in the direction of decreasing magnetic field strength. In order to maintain pressure balance, the plasma pressure within the strong field region must decrease. The high thermal conductivity of the convection zone means that any temperature difference between the strong and weak magnetic field regions is soon equalised. This, in turn, means that the reduced plasma pressure must be associated with a reduced plasma density. Thus, regions of strong magnetic field (flux tubes) have a lower plasma density than their surroundings, so they begin to rise. The integrity of flux tubes as they rise through the convection zone may only be maintained when they have a certain amount of twist (in which case they are more specifically called flux ropes [1]), as this prevents loss of the magnetic field from the flux tube due to erosion by vortices forming in its wake (Emonet & Moreno-Insertis 1998). The emergence of magnetic flux at the solar surface and its subsequent cancellation or eruption play a crucial role in converting the toroidal magnetic field to a

[1] The amount of twist required for a flux tube to be called a flux rope is a highly debated topic (Stephanie Yardley, personal communication). In this thesis, I refer to tube-like regions of strong magnetic field with the general term 'flux tubes'. If the flux tube is twisted, I may also use the more specific term of 'flux rope'.

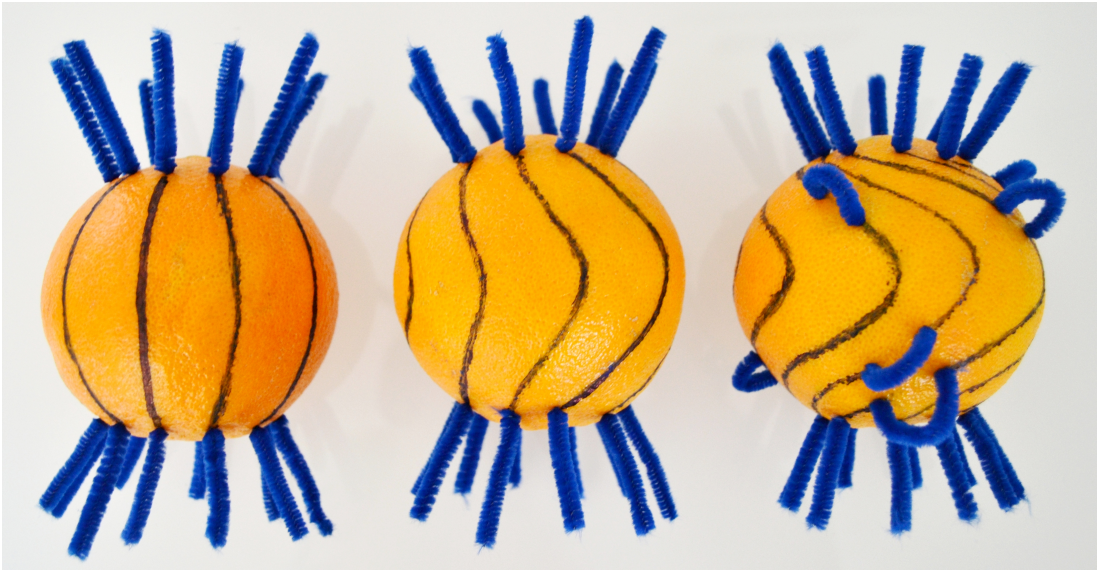

Figure 1.1: Poloidal to toroidal field. The dark lines drawn on the oranges show field lines at the tachocline, and the blue pipe cleaners show field lines that are rising up through the convection zone. Differential rotation causes the poloidal field at solar minimum (left) to wind up, producing a toroidal field component. As the field strengthens (indicated by the field lines becoming closer together), some regions rise as omega loops (right). The orientation of the magnetic field in these loops is opposite in the two hemispheres (Hale's law).

poloidal one.

When magnetic flux emerges into the solar atmosphere, a pair of sunspots (see Section 1.2.2), one of each magnetic polarity, form on the photosphere at the two footpoints of the flux tube. In the vast majority of cases (96%, Wang & Sheeley 1989), the following spot (as defined by the solar rotational direction) of the pair has opposite magnetic polarity to that of its hemisphere's polar (dipole) field polarity at the start of the cycle, in accordance with Hale's law. The orientation of the sunspot pair is described by Joy's law, which states that the following spot normally appears further from the equator than the leading spot. The poloidal magnetic field component of sunspot pairs following Joy's law is important for building the global poloidal field of the next cycle. The latitudes at which magnetic field rises to the solar surface depend on where the toroidal field is strongest. Near the start of the solar cycle, sunspots form in a band at $\sim 15^\circ$–30° latitude and as the cycle progresses the range of latitudes widens and moves equatorward, but without reaching the equator (Spörer's law, see Figure 1.2). Two to three years before sunspot forming ARs emerge, the first flux of the cycle emerges at latitudes poleward of $\sim 30^\circ$– 35° (Harvey 1993). The change in the latitude at which the magnetic flux emerges is related to the return flow of the meridional circulation at the base of the convection zone.

Granular motion (the convective motion and evolution of convection cells) near the top of the convection zone acts to disperse the photospheric magnetic field. The dispersion takes place in all directions, but its overall effect is to produce a diverging motion of the opposite polarities during the decay phase of the AR, because flux cancellation (Figure 1.3) occurs along the internal polarity inversion line (PIL) when opposite polarity fragments come together. The divergence increases the difference in latitude of the two polarities, with the leading polarity tending to move equatorward and the following polarity poleward. In addition, the meridional circulation near the photosphere acts to advect magnetic flux poleward, greatly increasing

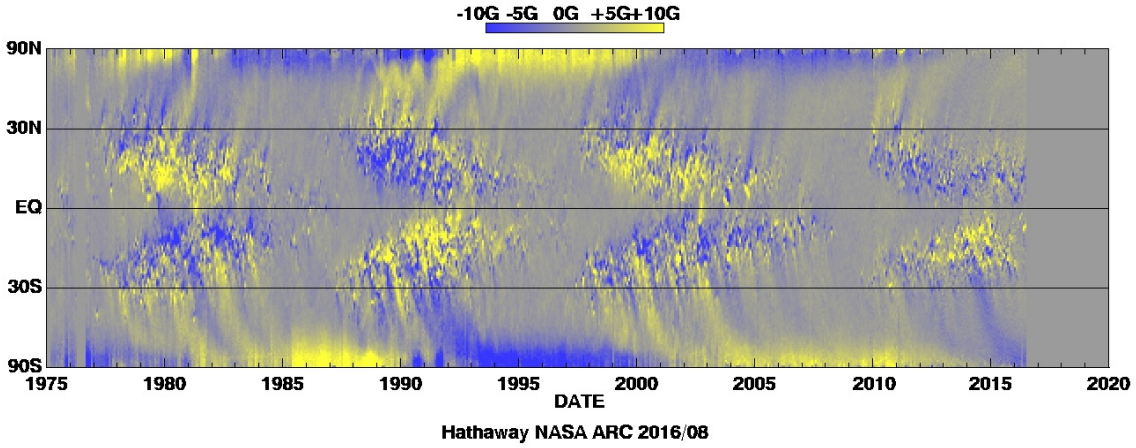

Figure 1.2: Magnetic butterfly diagram. Each vertical strip is the longitudinal average of the magnetic field strength in a synoptic map. This illustrates the latitudes of sunspot emergence and how these positions change over the course of the solar cycle. It also illustrates the dispersion and advection of magnetic field from the sunspots, which leads to the switch in orientation of the solar poloidal magnetic field at around or soon after the sunspot cycle maximum. Credit: David Hathaway NASA/ARC.

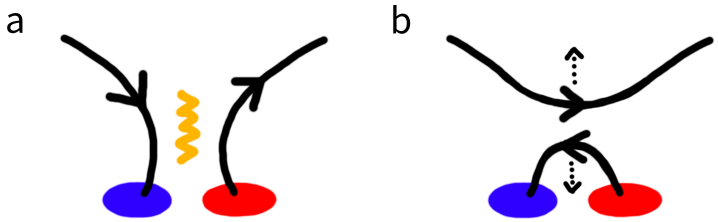

Figure 1.3: Diagram showing flux cancellation. Positive and negative flux fragments are shown in red and blue respectively, with field lines shown in solid black. Fragments of opposite polarity flux approach each other, producing a configuration similar to that of panel a. Reconnection (Appendix B) occurs low in the atmosphere (at the reconnection site indicated by the yellow zigzag representing a current sheet in panel a) between the two different magnetic flux systems forming new connectivities (panel b), namely long overlying field lines and a small loop between the two flux fragments. The small radius of curvature of the small loop provides a large magnetic tension, pulling the small loop down beneath the photosphere. This reduces the flux through the photosphere and is observed in magnetograms as the convergence and disappearance of opposite polarity fragments.

the rate at which the following polarity flux reaches the pole. The velocity of the meridional circulation increases with latitude, so it affects the following polarity more strongly than the leading one, contributing to the increasing latitudinal separation of the polarities. Flux cancellation occurs near the equator, where magnetic flux fragments from the leading spots in the northern hemisphere meet flux fragments of the other polarity from leading spots in the southern hemisphere. Cancellation also takes place at the poles between magnetic flux fragments of the following spots and the magnetic flux of the pole of the original dipole field. This reduces the Sun's magnetic dipole moment until each of the poles switch sign (which may not be at the same time) and the solar dipole reverses. The next 11 year cycle begins when the reversed field has strengthened forming a new poloidal field. This description of the solar activity cycle is the solar dynamo model. The emergence and dispersion of magnetic flux at the photosphere over the last few solar cycles is shown in Figure 1.2.

Eruptive events may also play an important role in the solar cycle. Coronal mass ejections (CMEs) release large amounts of energy and carry mass and magnetic field away from the Sun into the heliosphere. Magnetic helicity, a measure of the twist and writhe of magnetic field lines associated with a non-potential magnetic field configuration, is also ejected during eruptive events (*e.g.*, Rust 1994). Preceding the eruption, helicity accumulates due to the emergence of twisted magnetic field and photospheric motions associated with the cycle-invariant differential rotation. Eruptive events currently provide the best explanation for the removal of helicity (Chen 2011).

1.2 Active Regions (ARs)

Active regions (ARs) are regions of strong magnetic field, which arise when magnetic flux tubes from the convection zone break through the photosphere, as described in Section 1.1. They can be observed as strong magnetic concentrations in magnetograms at the photosphere and as bright loops in extreme ultra-violet (EUV) wavelengths. In white light, ARs appear as dark spots on the photosphere, known as sunspots (Section 1.2.2). When one or more of these spots appear, the AR is assigned a number by the National Oceanic and Atmospheric Administration (NOAA). One of the selection criteria for the ARs used in Section 3.2 was that they were given a NOAA AR number at some point during their passage across the solar disk.

ARs are often the source regions of flares and CMEs, as well as smaller scale activity such as brightenings (microflares) and jets. The amount of activity produced by an AR varies, as does the magnetic configuration, total flux and area of the region. The lifetime of ARs also varies greatly and is strongly dependent on the size of the AR with larger ARs living longer. Large ARs, containing magnetic flux on the order of 10^{22} Mx, can live for several months, while smaller regions, though still associated with sunspots, may decay in about a week (van Driel-Gesztelyi & Green 2015).

All of the characteristics mentioned above are affected to some extent by the background field into which the AR emerges. Interactions between the emerging flux and the background field produce lots of activity, particularly during the emergence phase of the AR. Reconnection with the background field can also lead to flux cancellation (Figure 1.3), thereby reducing the lifetime of the AR. ARs that emerge during solar minimum or early in the rise phase of the solar cycle, when the background field is relatively quiet, generally live longer than those emerging later in the cycle (Schrijver & Harvey 1994). Dispersed field from decayed ARs can be observed long after the sunspots have disappeared.

1.2.1 Asymmetries and evolution of ARs

ARs often show some asymmetries between the two polarities, originating as the omega-shaped flux tube rises through the convection zone. During the rise, plasma drains into the flux tube legs, with conservation of angular momentum causing more plasma to drain into the following leg. This distorts the shape of the flux tube and causes an increase (decrease) in magnetic pressure of the leading (following) leg, as a result of the total pressure balance. Additionally, the Coriolis force acts on the rising plasma, causing the flux tube to rotate slightly, which explains Joy's law.

When the flux tube reaches the base of the photosphere, its environment changes dramatically; in particular it is no longer buoyant. Magnetic field accumulates here until the undulatory instability (Spruit &

van Ballegooijen 1982) or convective upward motions allow fragments of the field to rise and break through the photosphere in a series of small magnetic loops (*e.g.*, Pariat et al. 2004, and references therein). The opposite polarities of these loops diverge and the like polarities of many of these small loops coalesce to form strong concentrated spots. The higher magnetic pressure of the leading flux tube leg means that the leading polarity forms a stronger, more compact spot than the following polarity. This process of fragmented emergence followed by coalescence has been well observed (*e.g.*, Zwaan 1978; Strous et al. 1996) and also modelled (*e.g.*, Cheung et al. 2010).

During the emergence phase, the two polarity centres diverge, with the separation of the strong magnetic concentrations (spots) reaching a plateau around the time the AR achieves its peak flux, indicating that the flux tube is no longer emerging (see Liu & Zhang 2006, who, in an observational study of NOAA 10488, took the constancy of separation between sunspots as a criteria for the active region's maturity). In addition, convective motions of (super)granular cells advect the region's field, breaking it apart and causing the magnetic field to disperse. Cancellation of flux fragments from the AR with those of the background contributes to the removal of magnetic flux from the photosphere, as well as cancellation between the two opposite polarities along the internal PIL of the AR. The decay phase of ARs is much longer than the emergence phase and can last for several weeks (*e.g.*, Hathaway & Choudhary 2008) or even months (*e.g.*, van Driel-Gesztelyi et al. 1999), with the weaker following spot decaying much faster than the leading spot.

1.2.2 Sunspots

Sunspots are the observational manifestation of ARs in white light images of the photosphere. They consist of a central dark core, known as the umbra, with an approximately radial magnetic field and a temperature of $\sim$ 4000 K (compared to the quiet Sun value of $\sim$ 5700 K), and a surrounding brighter penumbra, with a highly inclined magnetic field. The dark appearance and corresponding low temperatures of sunspots are due to the inhibition of convection by the strong magnetic field of the ARs, which may be as high as $\sim$ 3000 G in some parts of the umbra. Although convection is mostly suppressed in sunspots, some convective flows are necessary to maintain the temperature of the sunspots, which lose energy through radiation.

In the umbra, convection occurs in umbral dots, small regions with smaller field strengths and higher gas pressure than the rest of the umbra. The convection heats the umbral dots, giving them a bright appearance. Upflows have been observed at the centre of umbral dots (*e.g.*, Rimmele 2004; Socas-Navarro et al. 2004) and downflows at the edge of umbral dots in localised patches (Ortiz et al. 2010). Convection also occurs in light bridges, regions of relatively bright material that divide the dark umbra. Light bridges can be thought of as elongated umbral dots and they have many properties in common, with relatively low magnetic field strengths and associated upflows and downflows. Light bridges tend to have a dark central lane and similarly some umbral dots have dark cores. Upflows have been observed along the dark cores (Falco et al. 2016) and their dark appearance is thought to be caused by a build-up of the upflowing plasma, so that the observational height is increased to a region of cooler plasma (Schüssler & Vögler 2006). Numerical simulations have been able to reproduce the formation of umbral dots and light bridges, supporting their magneto-convective origin (*e.g.*, Schüssler & Vögler 2006; Cheung et al. 2010).

Some convection also occurs in the penumbra. Penumbra have a complex fine structure, consisting of narrow radial lanes with more vertical magnetic field known as spines interspersed with weaker and

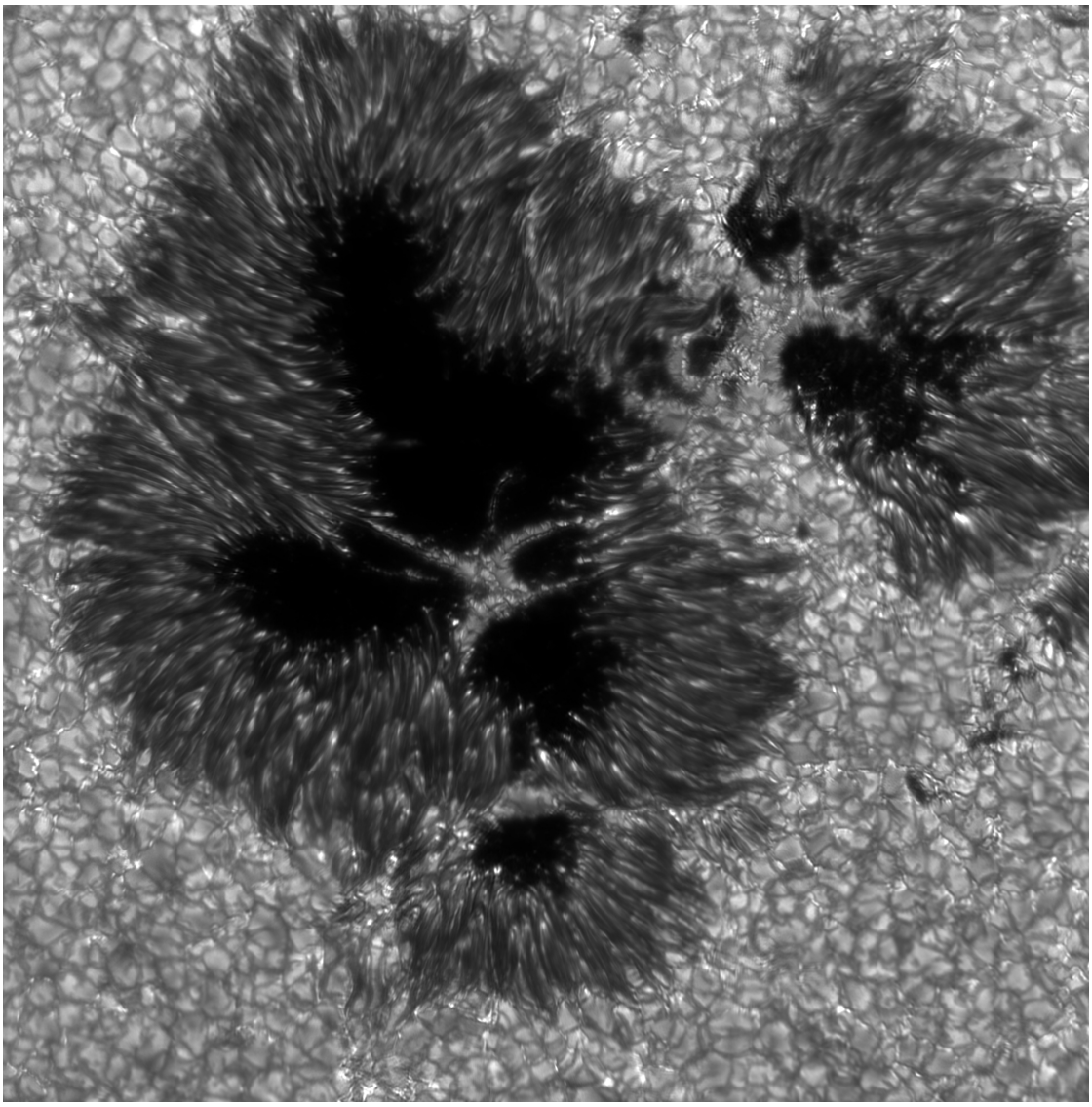

Figure 1.4: Image of the negative (leading) polarity of AR 10349. This image was taken and processed by Mats Löfdahl of the Institute for Solar Physics using the Swedish 1-m Solar Telescope (SST) at 430.5 nm on 2003-May-02. The fine structure of the penumbra, with its radial penumbral filaments and bright grains can be seen, along with light bridges dividing the dark umbra.

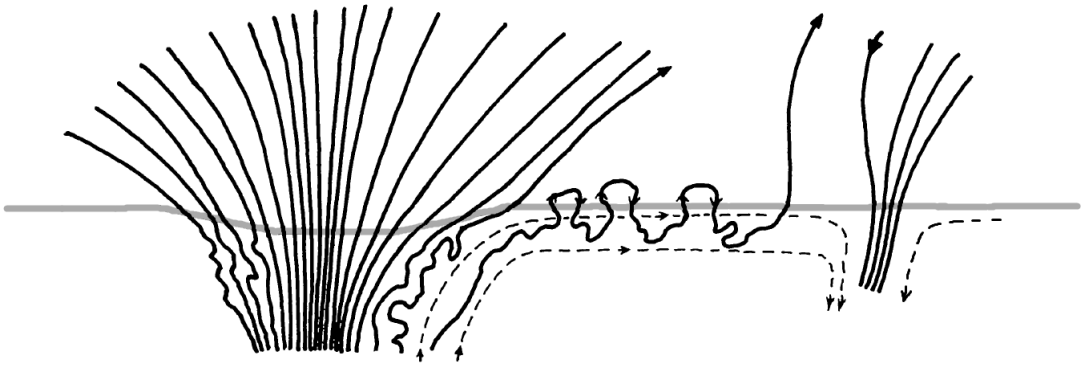

Figure 1.5: A diagram showing the possible magnetic field configuration of MMFs in the moat region (adapted from Figure 6 of Harvey & Harvey 1973). The photosphere is shown as a gray line, magnetic field lines as solid black lines and convective flows as dashed black lines. A sunspot is shown (left), with convective flows disturbing the field lines subsurface at the edge of the spot. These field lines are advected away from the spot towards the convective cell boundary (right), where cancellation with the network field can occur. Where the field lines break through the surface, MMFs are formed. Here, three bipolar MMFs are shown trailing one unipolar MMF, all of which have formed along the same field line.

more horizontal field known as intraspines. Collectively these radial lanes are also known as penumbral filaments. Radial convection is thought to occur along the intraspines, with upflows located at the inner penumbra, horizontal flows radially outward along the intraspines and downflows at the outer edge of the penumbra (*e.g.*, Borrero et al. 2011). The upflows are associated with heating and have been linked to the bright grains that can often be seen at the inner ends of the penumbral filaments (Figure 1.4). The horizontal flows are associated with the Evershed flow, which is observed as Doppler shifts in the light emitted from penumbra of sunspots near the solar limb and which Borrero & Solanki (2008) found to be concentrated in the intraspines. The downflows have been linked to fragments of opposite polarity flux at the edge of the penumbra (see Figure 34 of Borrero et al. 2011), as the penumbral field lines along which the plasma flows sink below the surface. Small-scale radially elongated bipolar structures have also been observed, which move radially outward, leading to their interpretation as field lines associated with the Evershed flow that have a sea-serpent structure (Sainz Dalda & Bellot Rubio 2008). Radial convection contributes to convective heating, however it is unclear whether this is enough to maintain the temperature of the penumbra. Azimuthal convective flows may also exist, which would be associated with field free regions, but these would likely be much smaller than the radial convective flows (Borrero et al. 2011).

For some mature or decaying ARs, there is an additional outflow region surrounding the penumbra known as the moat. The moat flow carries fragments of flux known as moving magnetic features (MMFs; Harvey & Harvey 1973) from the edge of the penumbra away from the spot, and this is thought to contribute to AR decay. The relationship between the Evershed flow and the moat flow is not yet understood, although there is some evidence to suggest that they are linked. Sainz Dalda & Bellot Rubio (2008) found that some of the bipolar structures formed in the penumbra left the spot forming MMFs in the sunspot moat. In this case, the MMFs were bipolar and thus did not reduce the magnetic flux of the sunspot. However, single polarity MMFs have also been observed. Harvey & Harvey (1973) found that the total flux transported by the moat flow was between two and ten times greater than the net flux. If the configuration of the MMFs is

as shown in Figure 1.5, this suggests that some unipolar MMFs travel alone while others are trailed by up to (or perhaps even more than) five bipolar MMFs. The moat flow has been linked to supergranular flows, because of their similar spatial scales and horizontal flow speeds, with Harvey & Harvey (1973) suggesting that the persistence of the moat flow is due to stabilisation of the supergranules by the sunspot magnetic field. One of the main differences between the moat flow and supergranular flows is that vertical flows in the moat region are always downward whereas supergranules contain upflows near their centres (Švanda et al. 2014). Another difference is that moat flows around the leading sunspots are asymmetric in the east-west direction due to the westward proper motion of the sunspots (Sobotka & Roudier 2007; Švanda et al. 2014), with larger outflow regions found on the eastern side. The radial extent of the moat region tends to end in network field regions, further supporting the idea that moat flows are a consequence of supergranular convective cells. Convective flows and their importance for sunspot decay will be discussed further in Section 1.2.4.

1.2.3 Simulations of AR emergence

Flux emergence simulations model the rise of flux tubes through the upper convection zone and as they break through the photosphere entering the solar atmosphere. There are many challenges associated with these simulations, due to the large change in plasma density with height and the different forces dominating the plasma's motion, which change from gas pressure and gravity in the interior to magnetic forces in the corona. To model AR emergence, the magnetohydrodynamic (MHD) equations (see Appendix A) and the associated assumptions are normally applied.

Magnetic flux tubes rise when their density is lower than that of the surrounding plasma. This is often achieved in simulations by enforcing pressure balance and assuming the tube and the background are in thermal equilibrium. Then the strength of the magnetic field is the determining factor. Different assumptions can be made, *e.g.*, that there is no heat transfer between the flux tube and the surrounding plasma. In this case, the stratification of the background, rather than the field strength of the tube, is important in determining whether the structure will rise. As the plasma inside the flux tube is considered isolated from the surrounding plasma by its magnetic field (see Alfvén's theorem in Appendix A), any changes it undergoes can be assumed to be adiabatic. If the density of the background plasma decreases less quickly with height than the adiabatic decrease in density associated with a small upward perturbation of the flux tube, then the flux tube is unstable and will rise. The stratification of the convection zone in simulations is often chosen to be superadiabatic (*i.e.* convectively unstable to adiabatic perturbations), but this is not the case at the photosphere, where the order of magnitude change in density is much greater. Simulations of the rising flux tubes tend to show at least a momentary pause of the flux tube beneath the photosphere, indicating that the simulations are robust and giving rise to the term 'two-step emergence' (Toriumi & Yokoyama 2011). When the flux tube reaches the base of the photosphere and can no longer rise, it expands, predominantly in the horizontal plane. Build-up of magnetic flux in this layer can increase the field strength, until it becomes buoyant again and magnetic loops can emerge into the higher atmospheric layers (*e.g.*, Archontis et al. 2004). Other instabilities initiated by turbulent perturbations are likely to contribute to this, resulting in complex sea-serpent like field structures above the photosphere.

In many models, the background plasma has a plane-parallel stratification, *i.e.* its properties change

only as a function of height. In the convection zone, the background plasma is often assumed to be in hydrostatic equilibrium[2] (MacTaggart & Hood 2009; Murray et al. 2006; Fan 2001) and is determined by imposing a temperature profile. In the interior there is no background magnetic field. Often these models (see Hood et al. 2012, for a review) include a twisted flux tube, a localised region of strong magnetic field, as the initial state (at the time when the simulation starts). The plasma pressure of the flux tube is commonly split into two components: one that balances the gravitational force (the background component) and the other that balances the Lorentz force produced by the magnetic field. Normally, the temperature is kept the same as that of the background and the density of the magnetised region is reduced, allowing it to rise. However, other approaches may also be used. Magara & Longcope (2001) used a flux tube with the same density as the background and a lower temperature to balance the pressures. Then they gave the flux tube some vertical momentum to initiate its rise. Different models are used for the initial state of the flux tube, the most common being a horizontal cylindrical flux rope and a toroidal shaped rope (Hood et al. 2012).

The formulation of the magnetic field of a twisted cylindrical flux tube (proposed by Fan 2001), consists of an axial component, $B_{\mathrm{y}} = B_0 * exp(-r^2/{r_0}^2)$, and an azimuthal component, $B_{\mathrm{theta}} = \alpha * r * B_{\mathrm{y}}$, where r is the distance from the flux tube axis, and r_0 and α indicate the flux tube radius and twist respectively. The density deficit formed by maintaining pressure balance and thermal equilibrium is normally modified to encourage the formation of an omega-shaped loop. One modification is multiplication by a factor $exp(-y^2/\lambda^2)$, where y is the horizontal distance from the centre of the flux tube and λ is a chosen scale length (Fan 2001). Instead of modifying the density deficit, a different geometry of magnetic structure can be used to form a toroidal shaped flux rope. For example, MacTaggart & Hood (2009) achieved this by using the same formulation for the magnetic field as for the cylindrical tube described above, but in a different coordinate system with a semicircular central axis (in a vertical plane) rather than a horizontal one.

Other studies (*e.g.*, Tortosa-Andreu & Moreno-Insertis 2009; Cheung et al. 2010) solve the MHD and radiative transfer equations self-consistently, by assuming local thermodynamic equilibrium. Before the introduction of the flux tube, a hydrodynamic simulation is performed until the radiatively driven convection reaches a statistical equilibrium; a state in which small fluctuations occur, but the average values of all the physical quantities are constant in time. This is a very different approach to the simulations discussed above which use a background with plane parallel stratification. A toroidal flux tube is then advected across the base of the computational domain (*e.g.*, Cheung et al. 2010; Rempel & Cheung 2014). The field strength and rise speed of the flux tube are chosen to be similar to that produced by simulations of thin flux tubes rising through the lower convection zone. When the base of the semitoroid reaches the computational domain, its speed is set to zero (but the rest of the tube can stretch up and continue to rise).

As the flux tube reaches the base of the photosphere, the conditions change dramatically, with a huge drop in the background plasma density. In the corona, the background component of plasma pressure is often taken as uniform, because the Lorentz force plays a much greater role than the weight of the overlying plasma in this region. The initial magnetic field in the corona is sometimes simply assumed to be uniform and either horizontal, slanted or vertical. In other studies, it is non-uniform, in which case either an analytical magnetic field is inserted or a simulation is used to produce it. Depending on its configuration, the

[2] In this context, hydrostatic equilibrium occurs when gravity is balanced by the pressure gradient force.

atmospheric magnetic field can act to suppress the horizontal expansion of the flux tube, forcing it to expand more vertically, thus assisting its rise (Krall et al. 1998). Horizontal expansion can also be reduced by the magnetic tension associated with the azimuthal field component of a twisted flux tube, so strongly twisted flux tubes may be more likely to break through the photosphere than those with weaker twist (Murray et al. 2006). Many simulations (*e.g.*, Hood et al. 2012) show that a large portion of the magnetic flux remains trapped beneath the photosphere. Depending on the initial geometry and magnetic field strength of the flux rope, the original axis of the flux rope may not make it to the corona. The height reached by the axis of cylindrical flux ropes increases with increasing initial magnetic field strength (Murray et al. 2006). This is also the case for toroidal flux ropes, and the axes of toroidal flux ropes with large magnetic field strengths are able to rise into the corona (MacTaggart & Hood 2009). Even if the original axis does rise into the corona, the geometry of the emerging field is generally very different to that of the original flux rope.

The horizontal and vertical expansion of the magnetic field as it reaches the solar atmosphere creates a low pressure region and possibly causes shearing[3] (Manchester IV 2001). Plasma flows into this low pressure region and reconnection occurs, which can create new flux ropes. These new flux ropes can form below the original axis, in which case the original flux rope rises to the corona, or above the original axis, trapping the original flux rope in the photosphere (MacTaggart & Hood 2009). If these processes continue, several new flux ropes can form. In some simulations (including that of Archontis et al. 2009), sigmoids (S-shaped structures, formed from twisted field lines) arise naturally from the shearing motions. As also found in observational studies (*e.g.*, Canfield et al. 1999), these structures are likely to lead to eruptions.

With regard to which flux rope structure provides the best model, MacTaggart & Hood (2009) found that toroidal flux ropes exhibit more realistic photospheric effects than cylindrical ones. This can be seen in the distance between the two polarity regions, which, in the toroidal flux rope models, reaches a maximum when the distance between them is the same as that between the footpoints of the flux loop (MacTaggart & Hood 2009). This is a good representation of observations, where sunspots are seen to drift apart only for a certain length of time (Liu & Zhang 2006). In contrast, in most simulations of cylindrical flux ropes (even with density deficits introduced to form omega-shaped loops), the separation of the two polarity regions continues through the whole computational domain (Hood et al. 2012). Toroidal flux ropes also have the advantage of allowing plasma to drain down the flux rope legs, helping to maintain the buoyancy of the tube as it rises through the atmosphere. Thus, for tubes of the same magnetic field strength, a toroidal one is more likely to reach the corona than a cylindrical one (MacTaggart & Hood 2009).

Turbulent convective flows (*e.g.*, in the model of Cheung et al. 2010; Rempel & Cheung 2014) interact with the rising flux tube, giving rise to both the undulatory and interchange instabilities. These are wave-like instabilities, with the undular mode having a wave vector along the field lines and the interchange mode being perpendicular to the field lines. In the undular mode, small scale plasma flows occur along field lines, producing alternating regions of reduced and increased plasma density. The regions with lower density rise, while those with higher density sink. The sinking regions create U-loops below the surface and magnetic reconnection within these loops can remove mass (as also demonstated by Tortosa-Andreu &

[3]Shearing involves the motion of opposite magnetic polarity elements in opposite directions (or different speeds) parallel to the PIL. It is an important process associated with an increase of twist and magnetic free energy.

Moreno-Insertis 2009). Then, the rising loops can break through the photosphere, producing a sea-serpent like field similar to that seen in observations of emerging flux. Convective flows have a similar effect to the plasma draining of MacTaggart & Hood (2009), in that they enhance the buoyancy of the tube. The removal of mass and the convective flows allow the small fragments of emerged magnetic field to coalesce. This forms larger concentrations of magnetic field on the surface, which become sunspots.

There are many different approaches to simulating magnetic flux emergence and they capture different aspects of the process, producing results similar to observations. They are able to capture the emergence of fragments of flux, which gradually coalesce forming strong bipoles that drift apart until they reach a certain separation. Models have shown the importance of twist in enabling the coherent rise of flux tubes through the convection zone (Emonet & Moreno-Insertis 1998). Comparisons of the results from cylindrical and toroidal flux rope models, with regard to bipole drift, suggest that rising flux ropes are toroidal, with another advantage of the toroidal geometry being that plasma can drain away more easily. A picture can be envisaged where rising flux tubes evolve from standing wave type instabilities at the tachocline. The success of toroidal flux rope models alters this image, making it unlikely that low wavenumber modes are responsible for producing the flux tubes. Studies allowing for convection (*e.g.*, Cheung et al. 2010; Rempel & Cheung 2014) provide insight into the role of turbulent flows, producing a more realistic image of fragmented magnetic flux at the photosphere, and providing an alternative method for plasma draining. In models with convective flows, the flux tube does not need to be twisted as it does not undergo bodily emergence, but instead emerges in fragments.

1.2.4 AR decay

Magnetic field evolution can be described by the induction equation (A.11 in Appendix A). The two processes that could act to disperse magnetic field are advection by plasma flows and diffusion of the magnetic field through the plasma. For the large length scales ($\sim$ 3000 km) associated with sunspots and their dispersion on the photosphere, magnetic diffusion would be too slow to explain the observed behaviour, so advection by plasma flows must play an important role. However, only weak and intermediate field strengths ($\leq$ 600 G, van Driel-Gesztelyi & Green 2015) can be advected in this way. One possible explanation is the turbulent erosion model of Petrovay & Moreno-Insertis (1997), in which a current sheet forms around the strong magnetic field region of the sunspot. Diffusion occurs in the current sheet, converting strong field to weak field at the edge of the spot. Then, convection can act to dissipate the weak magnetic field outside the current sheet. This model suggests that the sunspot decay rate is proportional to the perimeter of the strong flux region and this was supported by an observational study of the decay rate of umbral areas (Petrovay & van Driel-Gesztelyi 1997).

Another model (Borrero et al. 2011) envisages sunspots as a collection of closely packed narrow flux tubes, between which regions of weak field exist. Convection can occur in these regions, breaking the stronger magnetic flux into small fragments and allowing diffusion over small length scales to reduce the field strength. The small regions in which convection occurs correspond to umbral dots and light bridges (see Section 1.2.2). However, it is uncertain whether these observed phenomena contribute to sunspot decay by causing the sunspot to fragment or whether they are produced by the decay. In the modelling study of Rempel & Cheung (2014), which simulated AR evolution including the decay phase, convective flows

ultimately lead to the break up of the sunspots, and in some cases these flows also formed light bridges.

Outside of the central umbra, the field is weaker and advection by plasma flows dominates the decay. Unipolar MMFs, advected by the moat flow, carry flux radially outwards, dispersing the AR magnetic field. Kubo et al. (2008) compared the flux loss rate of a decaying sunspot to the horizontal flux transport rate, inferred using local correlation tracking. They concluded that MMFs are responsible for most of the flux lost from the sunspot and inner moat regions, transporting it to the outer boundary of the moat region where cancellation occurs. In their modelling study, Rempel & Cheung (2014) found that vertical flows were more important for AR dispersion than horizontal flows (which generally acted to increase the flux concentration in the spot), but noted that vertical flows acting on inclined field lines could appear as horizontal flows in local correlation tracking. Thus, their results do not contradict those of Kubo et al. (2008).

Advection by convective flows, granular or supergranular, plays a key part in AR decay and acts mainly along the perimeter of the sunspot, as suggested by the turbulent erosion model (Petrovay & Moreno-Insertis 1997). MMFs are the observational manifestation of the field line motions, but it is worth noting that many MMFs are bipolar (Figure 1.5) and these do not contribute to sunspot decay.

Another process that may be important for sunspot decay is the loss of connectivity of the sunspot magnetic field to the toroidal field layer at the base of the convection zone. During the emergence process, the two polarities are seen to diverge, but at some point their separation reaches a plateau. The horizontal magnetic tension that acts to force the polarities apart during the emergence phase is no longer playing a role. This may suggest that the field has become disconnected from the toroidal layer, and a possible mechanism for this is subsurface reconnection (Schrijver & Title 1999). However, this seems unlikely, as at the photosphere the two different polarities are observed to separate and reconnection requires the oppositely orientated field lines to come together. Alternatively, it is possible that the field is still connected but the magnetic forces in the high β convection zone are no longer strong enough to have a significant influence. If the legs of the flux tube where it crosses the photosphere were almost vertical (*e.g.*, if the flux tube had an omega shape with the base at the tachocline), then there would be no horizontal magnetic tension here. Instead, the horizontal magnetic tension would be lower in the convection zone, where it is less likely to play a role due to the small influence of the magnetic field in the dynamics of this high β region. Otherwise, if the orientation of the flux tube has a horizontal component near the surface the magnetic field strength here should be decreased. Schüssler & Rempel (2005) proposed a mechanism for this, invoking plasma upflows to increase the plasma pressure below the surface and weaken the magnetic field due to pressure balance. In the weak subsurface field resistive processes become important, *i.e.* Alfvén's theorem (see Appendix A) no longer holds, and convective flows can deform the field. At this point, the dynamical link between the subsurface magnetic field and the emerged flux is lost. It is difficult to confirm these models observationally, due to the limited observational data beneath the solar surface. However, from photospheric data of AR evolution, it is clear that subsurface processes soon break the dynamical ties between the emerged flux and its roots in the convection zone.

1.3 Thesis Overview

In the above I have described the evolution of ARs and the role they play in the solar cycle. This has included theory and models of their emergence and decay phases, as well as observations of sunspots at the photosphere. To gain a deeper understanding of AR evolution, I have studied the magnetic field distribution of ARs at the photosphere. Chapter 3 contains my research and includes distributions of simple models and a sample of ARs at various stages of their evolution. A description of the data I have used for the observational part of the study and of the instruments that made the observations is given in Chapter 2. In Chapter 4, I propose some future work, including extending the study to a larger sample of ARs and looking at the distributions provided by simulations. Finally, my findings are summarised in Chapter 5.

Chapter 2

Instrumentation and Data

The observational study of AR evolution in this thesis makes use of line-of-sight magnetograms from the Helioseismic Magnetic Imager (HMI, Section 2.1). For ARs that emerged on the far side of the Sun, data from the Extreme UltraViolet Imager (EUVI, Section 2.2) on board the *Solar TErrestrial RElations Observatory* (STEREO) was used to identify the time at which the AR began to emerge.

Section 2.1 provides an overview of the HMI instrument and an explanation of how line-of-sight magnetograms are produced. The information in Section 2.1 comes mainly from Schou et al. (2012), Couvidat et al. (2012) and Couvidat et al. (2016). A brief description of the STEREO spacecraft and the EUVI instrument is given in Section 2.2.

2.1 The Helioseismic Magnetic Imager (HMI)

The Helioseismic Magnetic Imager (HMI; Scherrer et al. 2012; Schou et al. 2012) is one of three instruments on board the *Solar Dynamics Observatory* (SDO; Pesnell et al. 2012), which was launched by NASA in February 2010 and began taking measurements in May 2010. SDO is in an inclined geosynchronous orbit, with a near-continuous view of the Sun.

HMI is a filtergraph, an instrument that takes a series of images at different wavelengths and polarisations[1], the combination of which allows physical parameters to be derived. The HMI instrument consists of two charge-coupled device (CCD) cameras, each of 4096×4096 pixels, with each pixel covering an area of 0.5×0.5 square arcseconds. One camera measures just the circularly polarised components of the light [2]. The other camera is devoted to calculating the full Stokes vector, for which it also needs to measure the components of light that are linearly polarised. The Stokes vector provides a complete description of the polarisation state of the light, which is necessary for inferring the vector magnetic field. Both cameras observe the full Sun at 6 positions in the Fe I 6173 Å spectral line, symmetric around the line centre at disk centre and at rest. Images are taken in a sequence, varying the wavelength and polarisation selection

[1] The polarisation of light describes the motion of the electric field vector in a plane perpendicular to the propagation of light. If the electric field vector has no preferred orientation of oscillation, the light is unpolarised. In linearly polarised light, the electric field vector oscillates along one axis. For circularly polarised light, the electric field vector rotates, anticlockwise when viewed by the receiver for left handed and clockwise for right handed. In a Cartesian coordinate system, if the circularly polarised light is propagating in the z direction, the components of the electric field vector E_x and E_y oscillate sinusoidally with a phase difference of $\pm\pi/2$.

[2] From May 2010 until April 2016, each of the two cameras took data to produce a set of products, with one camera providing data for the line-of-sight products and the other focusing on the vector magnetic field. Since then, the images from the two cameras have been used together, producing a full set of filtergrams (including those required for the vector field) at a higher cadence of 90 s.

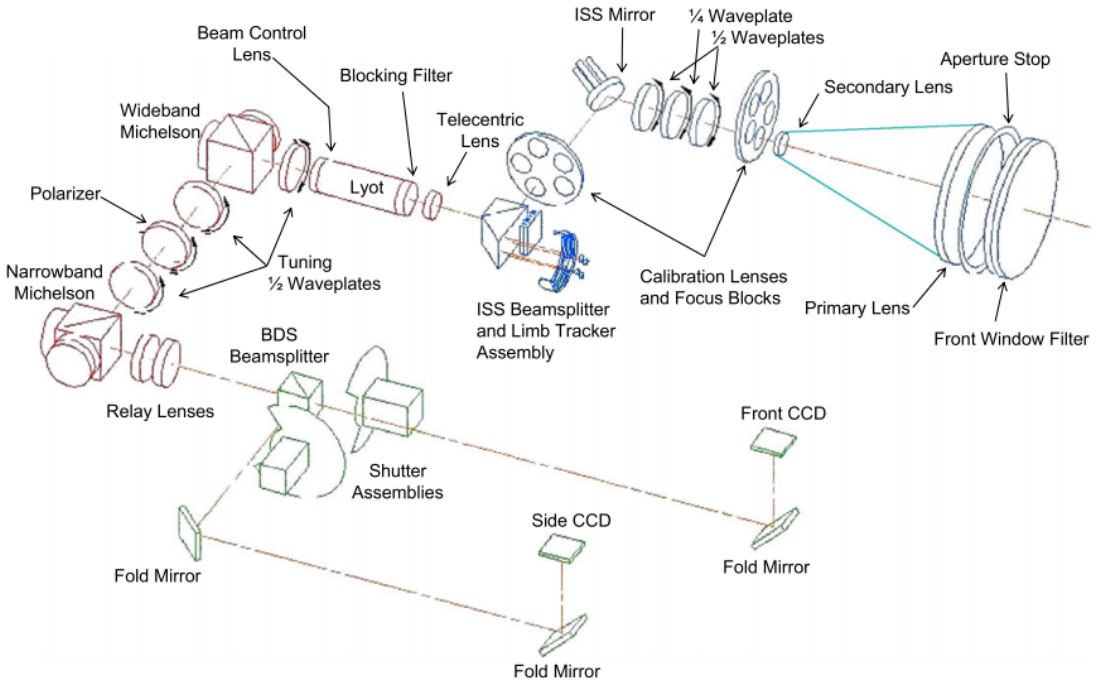

Figure 2.1: A diagram of the HMI optical system. From Schou et al. (2012).

between images and with each camera taking an image every ~ 3.75 seconds.

HMI produces full disk Dopplergrams (maps of the line-of-sight velocity of the solar surface), as well as full disk continuum filtergrams (images with a broad spectral resolution) and line-of-sight magnetograms. These products are produced with a time cadence of 45 seconds. Vector magnetograms are also produced, but at a lower time cadence due to the additional data required. A complete set of the required filtergrams is taken every 135 s, but these are normally averaged, providing vector magnetograms at a cadence of 720 s.

This research presented in this thesis makes use of the line-of-sight magnetograms, which show the sign and strength of the magnetic flux density in the line-of-sight direction as a function of position on the solar disk. The magnetic field value of each pixel is an average over the area of the photosphere covered by that pixel, and is given in Gauss (G).

Section 2.1.1 contains a description of how the HMI optical system works. The derivation of the line-of-sight magnetic field strength from the measured filtergrams is then given, with Section 2.1.2 describing how the Doppler velocities are calculated and Section 2.1.3 describing how the Doppler velocities of the different polarisations are used to produce the line-of-sight magnetograms.

2.1.1 Optical system

The HMI optical system (Figure 2.1) uses CCD cameras to measure the intensity of light at different wavelengths and polarisations. The desired wavelengths are selected by the HMI filter system, which consists of a front window, a blocking filter, a five-stage Lyot filter, one element of which is tunable, and two tunable Michelson interferometers. The front window is a 50 Å bandpass filter, blocking most of the spectrum to prevent the instrument from overheating. The light that passes through this is refracted by the two lenses which make up the main telescope.

Following this, and before the light reaches the other filters, the desired polarisation is selected. The light passes through three waveplates. Waveplates are made of birefringent material, meaning that they have two different refractive indices and the component of light with an electric field oscillating in one direction travels through the material at a different speed to that oscillating perpendicular to it. This retards the phase of one of the oscillating electric field components compared to the other component, thereby changing the polarisation of the light. The three waveplates can be individually rotated, allowing any single desired polarisation state to be converted to a fixed linear polarisation. This linearly polarised component is then selected by a polarising beamsplitter, which sends it to the rest of the filters, while sending the other components to the limb tracker assembly to determine the instrument's relative pointing. Polarising beamsplitters work by reflecting and transmitting different polarisation components of the light, which is incident on the surface of the beamsplitter at an angle to the normal. Light with an electric field vector oscillating in the plane of the surface is reflected, while light with an electric field perpendicular to the surface is transmitted.

The blocking filter is an interference filter: multiple reflections of the light within the filter interfere constructively at the desired wavelength and destructively at other wavelengths. Light that passes through this filter has an intensity profile with a full width half maximum (FWHM) of 8 Å.

The Lyot filter is made of a series of polarisers and birefringent plates. The birefringent plates, as in a waveplate, retard one component of the light with respect to the other, changing the polarisation state of the light. If the light is travelling in the z direction and is initially linearly polarised with components of equal magnitude in the x and the y directions, the electric field vector entering the birefringent plate (at $z = 0$) can be written as

$$\mathbf{E} = \mathrm{E}_0(\sin(\omega t)\,\hat{x} + \sin(\omega t)\,\hat{y})\,, \tag{2.1}$$

where t is time and $\omega = \frac{2\pi c}{\lambda}$ with λ being the wavelength of light and c the speed of light. This electric field has a time varying amplitude given by

$$a_{in} = \sqrt{2}\,\mathrm{E}_0 \sin(\omega t)\,. \tag{2.2}$$

With the refractive index for light polarised in the x direction differing to that in the y direction, the electric field of the light after passing through the birefringent plate can be written as

$$\mathbf{E} = \mathrm{E}_0(\sin(\omega t - kz_1)\,\hat{x} + \sin(\omega t - kz_2)\,\hat{y})\,, \tag{2.3}$$

where $k = \frac{2\pi}{\lambda}$ and z_1 and z_2 are the two different optical path lengths. z_1 and z_2 can be written in terms of the plate thickness d and the refractive indices μ_1 and μ_2 as $z_1 = \mathrm{d}\mu_1$ and $z_2 = \mathrm{d}\mu_2$. Each birefringent plate is followed by a polariser, allowing only the component of light oscillating along the axis defined by $y = x$ to be transmitted. Both the components in equation (2.3) have a projection onto this axis, calculated by multiplication by $\cos(\pi/4)$. After passing through the polariser, the two projected components interfere and the amplitude of the light can be written as a sum of these.

$$\begin{aligned} a_{out} &= \frac{\mathrm{E}_0}{\sqrt{2}}(sin(\omega t - kz_1) + \sin(\omega t - kz_2)) \\ &= \sqrt{2}\mathrm{E}_0\ \sin\left(\omega t - \frac{k(z_1+z_2)}{2}\right)\cos\left(\frac{k(z_1-z_2)}{2}\right). \end{aligned} \tag{2.4}$$

Comparing equations (2.2) and (2.4) and given that intensity is proportional to amplitude squared, we can see that the intensity of the light after passing through the birefringent plate and the polariser is reduced by a factor of $\cos^2(k(z_1 - z_2)/2)$. This factor can be written in terms of the wavelength λ and the difference in refractive indices $\delta\mu = \mu_1 - \mu_2$ as

$$\cos^2\left(\frac{\pi \, \mathrm{d} \, \delta\mu}{\lambda}\right). \tag{2.5}$$

The five elements of the Lyot filter have birefringent plates of widths differing by factors of two. The FWHM of the untuned part of the Lyot filter is 612 Å. The Lyot element with the smallest FWHM is tunable and is tuned depending on the radial velocity of SDO, which varies due to its geosynchronous orbit.

Finally the light passes through two tunable Michelson interferometers. A Michelson interferometer contains a beamsplitter, which splits the light into two beams. These two beams travel different distances and are reflected back towards the beamsplitter where they interfere. Whether the interference is constructive or destructive depends on the ratio of the difference in distance travelled by the two beams and the wavelength of the light. Unwanted wavelengths can be made to interfere destructively, reducing the FWHM of the transmitted light.

The FWHM of the complete filter is 76 mÅ at the centre tuning position and can be tuned over a range of 690 mÅ. The light then passes to one of the CCD cameras, where the intensity is recorded. Filtergrams are taken at wavelengths separated by 70 mÅ.

Corrections are applied to the filtergrams to account for exposure time, dark current and gain in the electronics. A flat field correction is also applied, removing artifacts due to variations in sensitivity between pixels. Cosmic ray hits are identified, based on abrupt changes from one exposure to the next, and removed. These corrections make up the Level 1 processing, and it is these corrected filtergrams that are used to produce the HMI line-of-sight observables.

2.1.2 Line profile fitting

Multiple measurements of each polarisation state and at each wavelength position are combined, to interpolate the filtergrams to the same time and distance from the Sun. This is done even for the products which have a 45 s time cadence, so individual Dopplergrams and magnetograms are not completely independent of ones at neighbouring times.

By comparing the relative intensities of the combined measurements at the 6 different wavelength positions to the predicted Fe I 6173 Å line profile, red and blue Doppler shifts can be calculated.

The Fe I line was chosen for several reasons. One of these is that there is a clean continuum around the line, which is necessary because of the motion of SDO associated with its orbit, a daily variation of 6.5 km s^{-1}. Another reason is the high magnetic sensitivity of the line, with a small change in the field strength producing a large splitting, corresponding to a large apparent Doppler shift.

Fourier analysis is used to calculate the Doppler shift. The absorption line is a function of wavelength and can be written as a sum of sines and cosines:

$$I(\lambda) = \frac{a_0}{2} + \sum_{n=1}^{N}\left[a_n \cos\left(\frac{2\pi n\lambda}{T}\right) + b_n \sin\left(\frac{2\pi n\lambda}{T}\right)\right], \tag{2.6}$$

where T is the observational wavelength range and a_0, a_n and b_n are Fourier coefficients. The first Fourier

coefficients are given by

$$a_1 = \frac{2}{T}\int_{-T/2}^{T/2} I(\lambda)\cos\left(\frac{2\pi\lambda}{T}\right)\mathrm{d}\lambda$$
$$b_1 = \frac{2}{T}\int_{-T/2}^{T/2} I(\lambda)\sin\left(\frac{2\pi\lambda}{T}\right)\mathrm{d}\lambda\,. \tag{2.7}$$

These coefficients can be evaluated by making some assumptions about the form of $I(\lambda)$. Generally, a Gaussian form of the line profile is used (Couvidat et al. 2012):

$$I(\lambda) = I_c - I_d\, e^{-\frac{(\lambda-\lambda_0)^2}{\sigma^2}}\,, \tag{2.8}$$

where I_c is the continuum intensity, I_d the line depth, σ is a measure of the line width and λ_0 is the Doppler shift. Then, evaluating the Fourier coefficients and assuming that λ_0 is small provides

$$\lambda_0 \approx \frac{T}{2\pi}\arctan\left(\frac{b_1}{a_1}\right). \tag{2.9}$$

This corresponds to a Doppler velocity of

$$v = \frac{\mathrm{d}v}{\mathrm{d}\lambda}\,\frac{T}{2\pi}\arctan\left(\frac{b_1}{a_1}\right). \tag{2.10}$$

Here, $\mathrm{d}v/\mathrm{d}\lambda = 299792458.0\ \mathrm{m\ s^{-1}}/6173.3433\ \text{Å} = 48562.4\ \mathrm{m\ s^{-1}\,Å^{-1}}$ and $T = 6\times 68.8 = 412.8$ mÅ (*i.e.* six times the wavelength separation between two intensity measurements) (Couvidat et al. 2012). a_1 and b_1 can be approximated using the six intensity measurements I_j:

$$a_1 \approx \frac{2}{6}\sum_{j=1}^{6} I_j \cos\left(2\pi\frac{3.5-j}{6}\right)$$
$$b_1 \approx \frac{2}{6}\sum_{j=1}^{6} I_j \sin\left(2\pi\frac{3.5-j}{6}\right), \tag{2.11}$$

and substituting these approximations into equation (2.10), allows the Doppler velocity to be found. Corrections are then made, using a more realistic form of the pure line profile and taking into account the convolution by the transmission filters.

2.1.3 Derivation of the magnetic field

HMI measures the intensity of the right and left circularly polarised light separately at 6 wavelength positions across the 6173.3 Å spectral line. The Doppler shifts are calculated, as described in Section 2.1.2, and the difference in wavelength between the right and left circularly polarised light is found. This difference in wavelength arises due to the Zeeman effect and is approximately proportional to the magnetic field strength where the photon is emitted.

The Zeeman effect is the name for the splitting of a spectral line into several components in the presence of a magnetic field. The energy levels of the atom are split, due to the interaction of the magnetic moment of the atom with the external field. The perturbation to the energy states is proportional to the strength of the magnetic field and is given by $H' = -\vec{\mu}\cdot\vec{B}$, where $\vec{\mu}$ is the magnetic moment and $\vec{B}$ is the magnetic field. The magnetic moment of the atom has contributions from both the orbital and spin angular momentums. Here, the much smaller contribution from the atomic nucleus is ignored, and only the electronic momentums are considered. Hence, electronic states with different spin have different energies when placed in a magnetic field.

When an electron makes the transition from a high energy to a low energy state, the atom is said to relax and energy is released in the form of a photon. Depending on the spin of the initial and final electron states, the photon released will have a certain spin (positive or negative), which corresponds to a certain polarisation (left or right handed circular polarisation). The initial electron state can take different values of spin and, as described above, the energy of this state is split in a magnetic field. Therefore, photons released when the atom relaxes have different energies depending on which spin state the electron started in. As the spin state of the electron also determines the polarisation of the photon, light with different polarisations is released with slightly different energies. The energy of a photon is inversely proportional to its wavelength, so the different polarisation states have slightly different wavelengths.

When the different polarisations are measured by HMI, their differing wavelengths means that they are found to have different Doppler shifts. The difference in the Doppler shifts can be used to calculate the line-of-sight magnetic field strength, as follows:

$$\mathrm{B}_{LOS} = (V_{LCP} - V_{RCP})K_M\,, \tag{2.12}$$

where V_{LCP} and V_{RCP} are the Doppler velocities of the left and right circularly polarised components, respectively, and $K_M = 0.231$ Gm^{-1}s is a constant. The average of the Doppler shifts, $(V_{LCP} + V_{RCP})/2$, is used as an estimate for the Doppler velocity of the solar surface.

2.1.4 Errors

Although many corrections are applied to the raw data (Section 2.1.1), errors still exist in the HMI data products. The main source of these errors is the variation in velocity of SDO due to its geosynchronous orbit. This causes a significant shift in the Fe I line profile. A shift in the line profile is also experienced away from the central meridian position and increasing towards the solar limb, due to the rotation of the Sun.

Shifts in the line profile affect the sensitivity of the optical system, so the sensitivity varies both across the solar disk and with the time of day. Errors are generally larger towards the edge of the solar disk and 12 and 24 hour oscillations can be seen in the line-of-sight magnetogram data.

2.2 Extreme UltraViolet Imager (EUVI)

The *Solar TErrestrial RElations Observatory* (STEREO; Kaiser et al. 2008) mission, consisting of two spacecraft each containing several instruments including an Extreme UltraViolet Imager (EUVI; Wuelser et al. 2004) was launched by NASA in October 2006. One of the spacecraft, STEREO A, was placed in a heliocentric orbit slightly closer to the Sun than the Earth, while the other, STEREO B, was positioned slightly further from the Sun than the Earth. As a result (Kepler's third law), STEREO A moves with an angular velocity slightly greater than that of the Earth and STEREO B slightly less than that of the Earth, causing the satellites to drift apart. The angular separation of the two STEREO satellites reached 180° in February 2011, allowing us to observe the entire Sun. Contact with STEREO B was lost in October 2014 and regained in August 2016.

The EUVI instruments, part of the Sun Earth Connection Coronal and Heliospheric Investigation (SECCHI; Howard et al. 2008) instrument suite, are telescopes that observe the solar chromosphere and

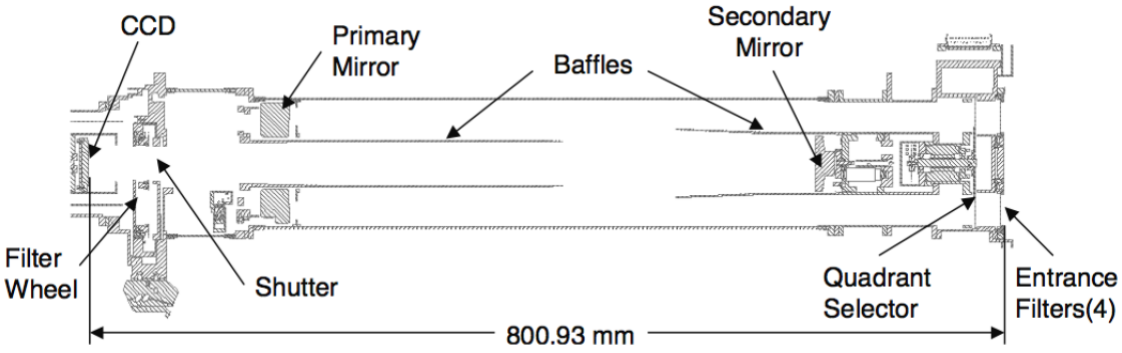

Figure 2.2: A diagram of the EUVI telescope. From Wuelser et al. (2004).

low corona in four narrow bandpass lines at wavelengths of 171, 195, 284 and 304 Å. A diagram of the telescope is shown in Figure 2.2. Light enters the telescope through a thin metal film filter, allowing only the EUV through and reducing the heat input. An aperture selector directs the light to one of the quadrants of the optical system. Each quadrant of the telescope's mirrors is coated with a multilayer reflective coating, optimised for one of the four lines. A CCD detector with 2048 $\times$ 2048 pixels records the image. The time cadence of the images differs for each of the four lines and has varied through the mission, decreasing as the spacecraft move further away from the Earth. Corrections are applied to the images to remove cosmic ray hits and bad pixels. Following this, the images are calibrated, taking into account several factors including the exposure time.

Chapter 3

Evolution of the Magnetic Field Distribution of Active Regions

I would especially like to thank Pascal Démoulin for collaborating with me on this research and in particular for his work on Section 3.1. The contents of this Chapter has been accepted for publication in *Astronomy & Astrophysics* by Dacie, S., Démoulin, P., van Driel-Gesztelyi, L., Long, D. M., Baker, D., Janvier, M., Yardley, S. L., and Pérez-Suárez, D.

To characterise AR emergence and dispersal and to gain insight into the mechanisms involved in these processes a series of ARs was studied, focusing on the distribution (probability density) of the vertical component of the photospheric magnetic field (flux density). For each AR at each time step, a log-log plot of the magnetic field distribution was created using kernel density estimation. The central portion of the distribution was found to be approximately linear and the slope of this portion was used to characterise the distribution. This is a new method of characterising the evolution of ARs, by looking at changes in the field distribution as the ARs evolve. Theoretical distributions were also considered and their slopes calculated, enabling comparisons to the observational data.

Section 3.1 describes the theoretical background of emergence, clustering (merging) and diffusion with a focus on the magnetic field distribution expected with these physical processes. The data used for the observational study are described in Section 3.2, along with the AR area selection code, which defines the pixels used to calculate the field distribution. The field distribution plots and their characterisation are explained in Section 3.3. Sections 3.4.1 and 3.4.2 show the statistical results of the characterisation of 37 ARs. The characterisation reflects the different evolutionary stages; fragmented emergence, coalescence to form strong sunspots and gradual dispersion of the AR. Then, in Sections 3.4.3 and 3.4.4 the dispersing phase of ARs and the quiet Sun distributions are analysed. Possible issues for the derived field distributions are discussed in Section 3.5. Finally, the observational and theoretical results are discussed and compared, allowing conclusions to be drawn in Section 3.6.

3.1 Theory

3.1.1 Emergence and clustering

The physical process of the emergence phase is too complex to derive an expected magnetic field distribution $\mathscr{P}(B_z)$ where B_z is the vertical magnetic field component. Numerical simulations of flux emergence

are typically set up with an axial field distributed with a Gaussian profile (Fan 2001; Toriumi & Yokoyama 2011), and if this structure did not change significantly upon emergence, a global Gaussian profile of the two polarities would be expected at the photosphere. This kind of distribution is discussed in the context of diffusion in Section 3.1.2, but it does not provide a good representation of the emergence phase. As the rising flux tube reaches the photospheric region it splits into tiny flux tubes, which emerge individually and reconnect with one another (Strous et al. 1996; Pariat et al. 2004; Cheung & Isobe 2014). Then, even if the axial field did have a Gaussian-like distribution in the convective zone, it is not known how this distribution would be transformed at the photospheric level.

During the later stages of emergence, small magnetic concentrations of like sign may merge to form strong polarities (sunspots). A simple description of such a flux concentration with axial symmetry is the magnetic field profile:

$$B_z(x,y) = B_{\max}((r/a)^2+1)^{-n/2}\,, \tag{3.1}$$

with $r^2 = x^2+y^2$, where x, y are the two orthogonal horizontal spatial coordinates and a defines a characteristic radius of the polarity. B_z is maximum at $r = 0$ and decreases as r^{-n} for $r \gg a$.

The case $n = 3$ corresponds to a magnetic source below the photosphere (see appendix of Démoulin et al. 1994) and can be derived as follows. The photosphere is taken as $z = 0$ and a magnetic point source of strength B_s is placed below the photosphere at $z = -a$. The magnetic field strength decreases away from the source following an inverse square law:

$$B(z,r) = \frac{B_s}{r^2+(z+a)^2}\,, \tag{3.2}$$

where $z+a$ is the distance in the vertical direction from the source and $r = \sqrt{x^2+y^2}$ is the horizontal distance from the source. The vertical component of the magnetic field strength B_z can be found by trigonometry, providing

$$\begin{aligned} B_z(z,r) &= B(z,r)\frac{z+a}{\sqrt{r^2+(z+a)^2}} \\ &= \frac{B_s\,(z+a)}{(r^2+(z+a)^2)^{3/2}}\,. \end{aligned} \tag{3.3}$$

At the photosphere,

$$B_z(r) = \frac{B_s\,a}{(r^2+a^2)^{3/2}}\,. \tag{3.4}$$

The maximum value of B_z at the photosphere,

$$B_{max} = \frac{B_s}{a^2}\,, \tag{3.5}$$

provides

$$B_z(r) = B_{\max}((r/a)^2+1)^{-3/2}\,. \tag{3.6}$$

Such field distributions have been used to model the magnetic field of ARs by using a series of flux concentrations with the parameters (position and intensity) fitted to the observed magnetograms. This allows the computation of the coronal magnetic topology.

Next, the probability $\mathscr{P}(B_z)\,\mathrm{d}B_z$ of having a magnetic field value of $B_z \pm \mathrm{d}B_z/2$ within the magnetic polarity is computed. This probability corresponds to the area $2\pi\, r\,\mathrm{d}r$ with r related to B_z by equation (3.1). Differentiation of equation (3.1) provides,

$$\begin{aligned} \mathrm{d}B_z &= -B_{\max}\frac{n}{a^2}\,\frac{1}{((r/a)^2+1)^{1+n/2}}\,r\,\mathrm{d}r \\ &= -B_{\max}\frac{n}{a^2}\,\left(\frac{B_z}{B_{\max}}\right)^{1+2/n}\,r\,\mathrm{d}r. \end{aligned} \tag{3.7}$$

By substituting equation (3.7) into $\mathscr{P}(B_z)\,\mathrm{d}B_z \propto 2\pi\, r\,\mathrm{d}r$ one gets

$$\mathscr{P}(B_z) = c\,B_z^{-1-2/n}, \tag{3.8}$$

where the proportionality constant, c, is found by setting the total probability to unity. This integral is calculated in the r-range $[0, r_{max}]$, corresponding to the B_z-range $[B_{z,\max}, B_{z,\min}]$, as follows.

$$\begin{aligned} \int_{B_{z,\min}}^{B_{z,\max}} \mathscr{P}(B_z)\,\mathrm{d}B_z &= c\int_{B_{z,\min}}^{B_{z,\max}} B_z^{-1-2/n}\,\mathrm{d}B_z \\ &= c\left[\frac{-n}{2}B_z^{-2/n}\right]_{B_{z,\min}}^{B_{z,\max}} \\ &= -c\,\frac{n}{2}\left(B_{z,\max}^{-2/n} - B_{z,\min}^{-2/n}\right). \end{aligned} \tag{3.9}$$

Setting this integral equal to one provides

$$c = \frac{2}{n}\,\frac{1}{(B_{z,\min}^{-2/n} - B_{z,\max}^{-2/n})}, \tag{3.10}$$

which gives

$$\mathscr{P}(B_z) = \frac{2}{n}\frac{B_z^{-1-2/n}}{(B_{z,\min}^{-2/n} - B_{z,\max}^{-2/n})}. \tag{3.11}$$

For $n = 3$, representing a magnetic source, $\mathscr{P}(B_z) \propto |B_z|^{-1.67}$, producing a slope of -1.67 when plotting $\mathscr{P}(B_z)$ against B_z on a log-log plot. At the limit of very large n values $\mathscr{P}(B_z) \propto |B_z|^{-1}$, as for the diffusion case of one polarity (*cf.* Section 3.1.2 and equation (3.24)). Although there is no clear physical interpretation of this, it is unsurprising that different B_z distributions, which depend on two spatial dimensions x and y, can produce the same $\mathscr{P}(B_z)$, which varies as a function of only one variable, namely B_z.

3.1.2 Diffusion of a magnetic polarity

Already during the emergence, and even more so during the decay phase, the AR magnetic field is affected by convective cells at various spatial scales (*i.e.* by granules and supergranules). The AR magnetic field is progressively dispersed in an ever increasing area nearly proportional to the time duration since the emergence started (*e.g.*, van Driel-Gesztelyi et al. 2003). This is the behaviour expected with classical linear diffusion. However, the dispersal of an AR also involves local mechanisms not included in classical diffusion, such as clustering of the field at the border of convective cells, cancellation of opposite-sign polarities and submergence of small scale loops. Classical diffusion can provide an explanation for the area evolution, but can it also explain the observed magnetic field distribution?

The classical diffusion of the vertical field component $B_z(x,y,t)$ is governed by

$$\frac{\partial B_z}{\partial t} = k\left(\frac{\partial^2 B_z}{\partial^2 x} + \frac{\partial^2 B_z}{\partial^2 y}\right), \tag{3.12}$$

where k is a constant coefficient and t is time.

For a concentrated initial field with one polarity of magnetic flux F, the solution of equation (3.12) is:

$$B_z(x,y,t) = \frac{F}{4\pi k\, t} \exp\left(-\frac{r^2}{4k\, t}\right), \tag{3.13}$$

with $r^2 = x^2 + y^2$. It can be shown that equation (3.13) is a solution of equation (3.12) as follows. Taking the partial derivative of equation (3.13) with respect to t gives

$$\begin{aligned} \frac{\partial B_z}{\partial t} &= -\frac{F}{4\pi k\, t^2} \exp\left(-\frac{r^2}{4k\, t}\right) + \frac{F}{4\pi k\, t}(-1) \exp\left(-\frac{r^2}{4kt}\right) \frac{\partial(\frac{r^2}{4k\, t})}{\partial t} \\ &= \frac{F}{4\pi k\, t}\left(-\frac{1}{t} + \frac{r^2}{4k\, t^2}\right) \exp\left(-\frac{r^2}{4k\, t}\right). \end{aligned} \tag{3.14}$$

Taking the first and then second order partial derivative of equation (3.13) with respect to x gives

$$\frac{\partial B_z}{\partial x} = \frac{F}{4\pi k\, t}\left(-\frac{2\, x}{4k\, t} \exp\left(-\frac{x^2}{4k\, t}\right)\right) \exp\left(-\frac{y^2}{4k\, t}\right) \tag{3.15}$$

$$\begin{aligned} \frac{\partial^2 B_z}{\partial^2 x} &= \frac{F}{4\pi k\, t} \exp\left(-\frac{y^2}{4k\, t}\right)\left[-\frac{2}{4k\, t} \exp\left(-\frac{x^2}{4k\, t}\right) - \frac{2\, x}{4k\, t}\left(-\frac{2\, x}{4k\, t}\right) \exp\left(-\frac{x^2}{4k\, t}\right)\right] \\ &= \frac{F}{4\pi k\, t}\left(-\frac{2}{4k\, t} + \frac{4x^2}{(4k\, t)^2}\right) \exp\left(-\frac{r^2}{4k\, t}\right). \end{aligned} \tag{3.16}$$

The equivalent can be found with respect to y:

$$\frac{\partial^2 B_z}{\partial^2 y} = \frac{F}{4\pi k\, t}\left(-\frac{2}{4k\, t} + \frac{4y^2}{(4k\, t)^2}\right) \exp\left(-\frac{r^2}{4k\, t}\right). \tag{3.17}$$

Multiplying by k and combining equations (3.16) and (3.17) gives

$$k\left(\frac{\partial^2 B_z}{\partial^2 x} + \frac{\partial^2 B_z}{\partial^2 y}\right) = \frac{F}{4\pi k\, t}\left(-\frac{1}{t} + \frac{r^2}{4k\, t^2}\right) \exp\left(-\frac{r^2}{4k\, t}\right), \tag{3.18}$$

and comparing this to equation (3.14), shows that equation (3.12) is satisfied.

As in Section 3.1.1, the probability $\mathscr{P}(B_z)\, \mathrm{d}B_z$ of having a magnetic field value of $B_z \pm \mathrm{d}B_z/2$ within the magnetic polarity is computed, which again corresponds to an area $2\pi\, r\, \mathrm{d}r$. In this case, r is related to B_z with equation (3.13). Differentiation provides

$$\begin{aligned} \mathrm{d}B_z &= \frac{F}{4\pi k\, t}\left(-\frac{2\, r}{4k\, t}\right) \exp\left(-\frac{r^2}{4k\, t}\right) \mathrm{d}r \\ &= -\frac{2\, r}{4k\, t} B_z \mathrm{d}r, \end{aligned} \tag{3.19}$$

and rearranging gives

$$2\pi\, r\, \mathrm{d}r = -\frac{4\pi k\, t}{B_z}\, \mathrm{d}B_z, \tag{3.20}$$

so

$$\mathscr{P}(B_z) = c\, \frac{1}{B_z}, \tag{3.21}$$

and the constant of proportionality, c, can again be found by normalisation of the total probability in the B_z-range $[B_{z,\max}, B_{z,\min}]$ as follows.

$$\begin{aligned} \int_{B_{z,\min}}^{B_{z,\max}} \mathscr{P}(B_z) \mathrm{d}B_z &= c \int_{B_{z,\min}}^{B_{z,\max}} \frac{1}{B_z} \mathrm{d}B_z \\ &= c\, [\ln B_z]_{B_{z,\min}}^{B_{z,\max}} \\ &= c\, (\ln B_{z,\max} - \ln B_{z,\min}) \\ &= c\, \ln\left(\frac{B_{z,\max}}{B_{z,\min}}\right). \end{aligned} \tag{3.22}$$

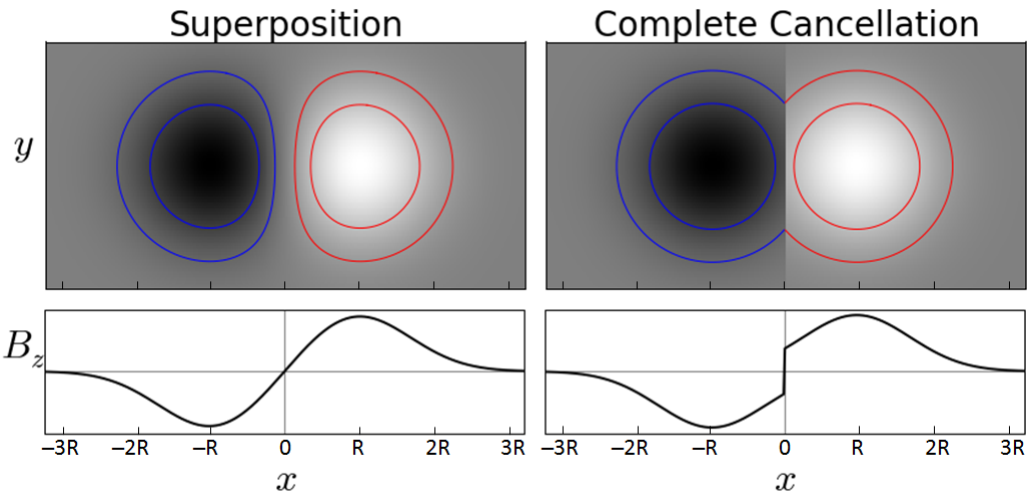

Figure 3.1: The two different scenarios for flux dispersion of a bipole as described in Section 3.1.3 are shown here. The left hand side of the plot shows superposition with no cancellation and the right hand side shows the case where complete cancellation occurs at $x = 0$. Two isocontours are also shown for both the negative and positive spots. The two plots in the bottom row show a cross section of the magnetic field values taken at $y = 0$.

Setting the integral equal to one provides

$$c = \frac{1}{\ln(B_{z,\max}/B_{z,\min})}, \tag{3.23}$$

which gives

$$\mathscr{P}(B_z) = \frac{1}{B_z \ln(B_{z,\max}/B_{z,\min})}. \tag{3.24}$$

Taking the log of equation (3.24),

$$\begin{aligned} \log(\mathscr{P}(B_z)) &= \log\left(\frac{c}{B_z}\right) \\ &= -\log(B_z) + \log(c), \end{aligned} \tag{3.25}$$

and comparing this to the equation of a straight line, it is clear that plotting $\mathscr{P}(B_z)$ against B_z in a log-log plot produces a slope of -1.

3.1.3 Diffusion of a magnetic bipole

To model the global evolution of an AR, a simple bipole is considered, with two polarities located at the fixed positions $x = \pm R, y = 0$. Since equation (3.12) is linear in B_z, the evolution of the bipole is the superposition of two solutions of equation (3.13) when shifted spatially by $\pm R$:

$$B_z(x,y,t) = \frac{F_0}{4\pi k\, t}\left(e^{-\frac{(x-R)^2}{4k\,t}} - e^{-\frac{(x+R)^2}{4k\,t}}\right) e^{-\frac{y^2}{4k\,t}}, \tag{3.26}$$

where F_0 is the total flux of the isolated positive polarity. This is illustrated in the top left panel of Figure 3.1, with the bottom left panel showing a cross section of the B_z values taken at $y = 0$.

The spatial distribution of B_z in equation (3.26) depends on the parameters F_0, $a = \sqrt{4k\,t}$ and R. F_0 defines the field strength while a defines the polarity size and both are scaling factors. The main parameter of equation (3.26) is R/a. For $R/a \gg 1$ the polarities are well separated and do not overlap, while for $R/a \ll 1$ the opposite polarities mostly cancel each other leaving a dipolar magnetic field.

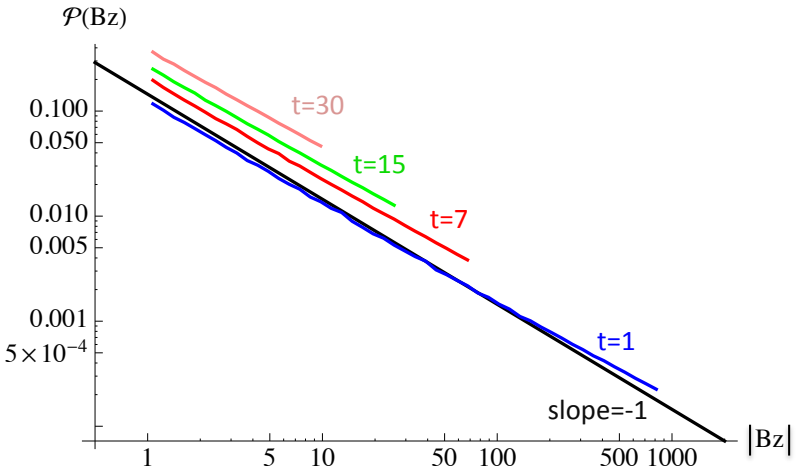

Figure 3.2: Log-log plot of the probability distribution of B_z versus B_z for a bipolar field diffusing with time. The B_z spatial distribution is described by equation (3.26). The time t is normalised by the initial time when $R/a = 1$ is selected (blue line). Only $B_z > 1$ is supposed to be detectable. The black straight line has a slope -1 for reference.

The probability distribution of B_z is computed numerically since the B_z isocontours are not simply circles ($\mathscr{P}(B_z)$ is computed by summation along these isocontours). In a log-log plot, the slope is in the range $[-1,-0.94]$ at all times during the diffusive evolution (Figure 3.2). More precisely, the slope ≈ -1 for $R/a > 3$ and converges rapidly to -1 when the separation between the polarities is larger, as expected from equation (3.24). The slope is maximum, ≈ -0.94, for $1 \leq R/a \leq 1.7$, while it shows a slight decrease to ≈ -0.96 for lower R/a values. When the polarities are interacting the slope is slightly less steep than with one polarity (slope $= -1$) because the other source acts to decrease B_z and even removes weak values surrounding the inversion line. This implies fewer cases with a given B_z value, particularly for low B_z, making the global slope of the distribution slightly less steep. Finally, the increase of probability values with time, as seen in Figure 3.2, is simply due to the normalisation of the total probability to 1 while the maximum of B_z is decreasing with time and the minimum of B_z is selected to be fixed (here to $B_z = 1$), modelling an instrumental detectability threshold.

The net flux evolution is computed by performing the integration on the positive polarity, where $x > 0$,

$$\begin{aligned} F(t) &= \int_0^\infty \mathrm{d}x \int_{-\infty}^\infty \mathrm{d}y \, \frac{F_0}{4\pi k\, t} \left(e^{-\frac{(x-R)^2}{4k\,t}} - e^{-\frac{(x+R)^2}{4k\,t}} \right) e^{-\frac{y^2}{4k\,t}} \\ &= \frac{F_0}{4\pi k\, t} \int_{-\infty}^\infty \mathrm{d}y \, e^{-\frac{y^2}{4k\,t}} \int_0^\infty \mathrm{d}x \left(e^{-\frac{(x-R)^2}{4k\,t}} - e^{-\frac{(x+R)^2}{4k\,t}} \right) \\ &= \frac{F_0}{4\pi k\, t}\, 2 \int_0^\infty \mathrm{d}y \, e^{-\frac{y^2}{4k\,t}} \left(\int_{-R}^\infty \mathrm{d}x \, e^{-\frac{x^2}{4k\,t}} - \int_R^\infty \mathrm{d}x \, e^{-\frac{x^2}{4k\,t}} \right) . \end{aligned} \tag{3.27}$$

Using the error function,

$$\mathrm{erf}(x) = \frac{2}{\sqrt{\pi}} \int_0^x e^{-u^2} \mathrm{d}u \,, \tag{3.28}$$

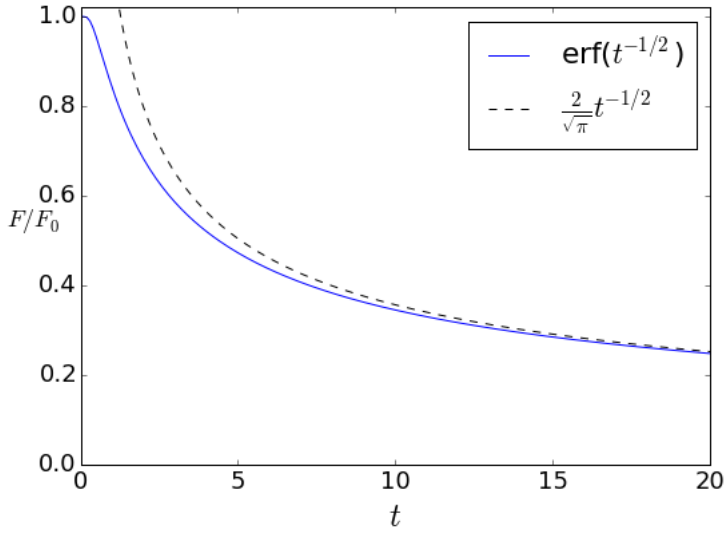

Figure 3.3: F/F_0 is shown with time for the case of two superposing polarities (blue), with the approximation for large times shown as a dashed black line.

with $\mathrm{erf}(\infty) = 1$, and making substitutions for x and y, $F(t)$ can be rewritten as

$$\begin{aligned} F(t) &= \frac{2F_0}{\pi}\int_0^\infty \mathrm{d}y\, e^{-y^2}\left(\int_{\frac{-R}{\sqrt{4k\,t}}}^\infty \mathrm{d}x\, e^{-x^2} - \int_{\frac{R}{\sqrt{4k\,t}}}^\infty \mathrm{d}x\, e^{-x^2}\right) \\ &= \frac{F_0}{\sqrt{\pi}}\mathrm{erf}(\infty)\left(\int_{\frac{-R}{\sqrt{4k\,t}}}^0 \mathrm{d}x\, e^{-x^2} + \int_0^\infty \mathrm{d}x\, e^{-x^2} - \int_{\frac{R}{\sqrt{4k\,t}}}^0 \mathrm{d}x\, e^{-x^2} - \int_0^\infty \mathrm{d}x\, e^{-x^2}\right) \\ &= \frac{F_0}{\sqrt{\pi}}\, 2\int_0^{\frac{R}{\sqrt{4k\,t}}} \mathrm{d}x\, e^{-x^2} \\ &= F_0\,\mathrm{erf}\left(\frac{R}{\sqrt{4k\,t}}\right). \end{aligned} \tag{3.29}$$

This is shown in Figure 3.3. For small diffusion times, $\sqrt{4k\,t} \ll R$, $F(t)$ is nearly constant. For large diffusion times, $R/\sqrt{4k\,t}$ is small, and

$$\mathrm{erf}\left(\frac{R}{\sqrt{4k\,t}}\right) \approx \frac{2}{\sqrt{\pi}}\frac{R}{\sqrt{4k\,t}}. \tag{3.30}$$

Then equation (3.29) can be approximated as

$$F(t) \approx F_0\,\frac{R}{\sqrt{\pi k\,t}}, \tag{3.31}$$

so the flux of the polarities decreases as $t^{-1/2}$. The cancellation of flux is due to the superposition of the opposite polarities in the x direction.

Equation (3.26) describes the diffusion of each polarity isotropically, *i.e.*, they do not cancel but just superpose spatially. This implies an apparent loss of magnetic flux by superposition of opposite B_z values when computing the total flux, however there is not a physical cancellation since the full amount of both polarities continue to diffuse with time. The other extreme behaviour, a complete cancellation at the inversion line, can be described by the field:

$$B_z(x,y,t) = \frac{F_0}{4\pi k\,t}\,\mathrm{ext}\left(e^{-\frac{(x-R)^2}{4\,k\,t}}, -e^{-\frac{(x+R)^2}{4\,k\,t}}\right)\, e^{-\frac{y^2}{4\,k\,t}}, \tag{3.32}$$

where the function ext(a,b) is defined as a if $|a| > |b|$, b if $|a| < |b|$, and 0 otherwise (then it is a signed extremum function except at $|a| = |b|$ where it vanishes to model immediate cancellation). This scenario is shown in the top right panel of Figure 3.1, with the bottom right panel showing the B_z values at $y = 0$.

As for the case of superposition, the flux evolution as a function of time can be calculated for the case of complete cancellation. The integration can again be performed just over the positive polarity, where $x > 0$ and

$$B_z(x > 0, y, t) = \frac{F_0}{4\pi k\, t}\, e^{-\frac{(x-R)^2}{4\,k\,t}}\, e^{-\frac{y^2}{4\,k\,t}}\,. \tag{3.33}$$

Then,

$$\begin{aligned} F(t) &= \int_0^\infty \mathrm{d}x \int_{-\infty}^\infty \mathrm{d}y \frac{F_0}{4\pi k\, t}\, e^{-\frac{(x-R)^2}{4k\,t}}\, e^{-\frac{y^2}{4k\,t}} \\ &= \frac{F_0}{4\pi k\, t} \int_{-\infty}^\infty \mathrm{d}y\, e^{-\frac{y^2}{4k\,t}} \int_0^\infty \mathrm{d}x\, e^{-\frac{(x-R)^2}{4k\,t}}\,. \end{aligned} \tag{3.34}$$

Making substitutions for x and y, and using the error function,

$$\begin{aligned} F(t) &= \frac{2F_0}{\pi} \int_0^\infty \mathrm{d}y\, e^{-y^2} \int_{-\frac{R}{\sqrt{4k\,t}}}^\infty \mathrm{d}x\, e^{-x^2} \\ &= \frac{2F_0}{\pi}\, \frac{\sqrt{\pi}}{2} \mathrm{erf}(\infty) \left(\int_{-\frac{R}{\sqrt{4k\,t}}}^0 \mathrm{d}x\, e^{-x^2} + \int_0^\infty \mathrm{d}x\, e^{-x^2} \right) \\ &= \frac{F_0}{\sqrt{\pi}} \left(\int_0^{\frac{R}{\sqrt{4k\,t}}} \mathrm{d}x\, e^{-x^2} + \frac{\sqrt{\pi}}{2} \mathrm{erf}(\infty) \right) \\ &= \frac{F_0}{2} \left(\mathrm{erf}\left(\frac{R}{\sqrt{4k\,t}} \right) + 1 \right)\,. \end{aligned} \tag{3.35}$$

For large diffusion times, equation (3.35) can be approximated as

$$F(t) \approx \frac{F_0}{2} \left(\frac{R}{\sqrt{\pi k\, t}} + 1 \right)\,, \tag{3.36}$$

which for very large times provides

$$\lim_{t\to\infty} F(t) = \frac{1}{2}\,. \tag{3.37}$$

In this case, half of the positive polarity flux remains in the $x > 0$ domain, unaffected by the cancellation process, while the other half has been completely cancelled at $x = 0$.

Equation (3.32) describes two magnetic polarities which are diffusing independently of each other except at the PIL ($x = 0$), where they cancel exactly, so there is no diffusion into the region of the other polarity. In the photosphere opposite polarities diffuse into each other as a result of advection by convective cells of small flux tubes, which at some point meet and cancel. This process is mostly driven by the supergranular flows and the typical distance of travel across the inversion line in the other polarity (the mean free path) is expected to be of the order of a supergranular cell (about 30 Mm). Equations (3.26) and (3.32) are the limits when the mean free path is infinite and zero, respectively.

The probability of B_z, for values where the corresponding isocontour is not touching the inversion line ($x = 0$), given by equation (3.32), has the same B_z dependence as equation (3.24). For the lower B_z values, some of the pixels are not present (those which would be located on the other side of the y axis if they had not undergone cancellation) so the probability is slightly lower. In summary, the distribution $\mathscr{P}(B_z)$ shown

in a log-log plot (as in Figure 3.2) has a slope of exactly -1 for large $|B_z|$ values and slightly above -1 (less steep) for low $|B_z|$ values.

All in all, this implies that $\mathscr{P}(B_z)$ is similar for an inefficient cancellation rate (equation (3.26)) and for a very efficient one (equation (3.32)). Thus, any cancellation rate, with a classical diffusion of the polarities is expected to imply a slope ~ -1 for $\mathscr{P}(B_z)$ drawn in a log-log plot.

3.1.4 The induction equation

In the above Sections, the dispersion of the magnetic field by random walk motions was considered, leading to a solution for B_z of the form of equation (3.13). However, the relationship between the magnetic field **B** and the velocity field **v** has not yet been considered. While resistivity likely plays an important role, in order to simplify the problem the ideal induction equation (A.12) is assumed, which in the stationary regime reduces to

$$0 = \nabla \times (\mathbf{v} \times \mathbf{B}) . \tag{3.38}$$

The curl of a vector **u** in cylindrical coordinates is given by

$$\nabla \times \mathbf{u} = \hat{\mathbf{r}} \left(\frac{1}{r} \frac{\partial u_z}{\partial \phi} - \frac{\partial u_\phi}{\partial z} \right) + \hat{\phi} \left(\frac{\partial u_r}{\partial z} - \frac{\partial u_z}{\partial r} \right) + \hat{\mathbf{z}} \frac{1}{r} \left(\frac{\partial (r\, u_\phi)}{\partial r} - \frac{\partial u_r}{\partial \phi} \right) . \tag{3.39}$$

With a purely radial flow $\mathbf{v} = v_r \hat{\mathbf{r}}$,

$$\mathbf{v} \times \mathbf{B} = -v_r B_z \hat{\phi} + v_r B_\phi \hat{\mathbf{z}} , \tag{3.40}$$

$$\nabla \times (\mathbf{v} \times \mathbf{B}) = \hat{\mathbf{r}} \left(\frac{1}{r} \frac{\partial}{\partial \phi} (v_r B_\phi) - \frac{\partial}{\partial z} (-v_r B_z) \right) - \hat{\phi} \frac{\partial}{\partial r} (v_r B_\phi) + \hat{\mathbf{z}} \frac{1}{r} \frac{\partial}{\partial r} (r\, (-v_r B_z)) . \tag{3.41}$$

Equation (3.38) means that all three components are equal to zero. Taking the $\hat{\mathbf{z}}$ component of equation (3.41) equal to zero provides

$$\begin{aligned} 0 &= \frac{\partial}{\partial r} (r\, (v_r B_z)) \\ &= v_r B_z + r \frac{\partial}{\partial r} (v_r B_z) . \end{aligned} \tag{3.42}$$

A solution to equation (3.42) is

$$B_z(r) \propto \frac{1}{r\, v_r} . \tag{3.43}$$

Supposing that v_r is independent of r away from the cell centre, at $r = 0$, and the cell boundary, implies

$$B_z \propto \frac{1}{r} . \tag{3.44}$$

As in Sections 3.1.1 and 3.1.2, B_z depends only on r, so $\mathscr{P}(B_z)\, \mathrm{d}B_z \propto 2\pi\, r\, \mathrm{d}r$. Differentiation of equation (3.44) provides

$$\mathrm{d}B_z \propto \frac{1}{r^2}\, \mathrm{d}r , \tag{3.45}$$

and rearranging gives

$$\mathrm{d}r \propto \frac{1}{B_z^2}\, \mathrm{d}B_z . \tag{3.46}$$

This leads to

$$\mathscr{P}(B_z) \propto \frac{1}{B_z^3} , \tag{3.47}$$

which produces a slope of -3 on a log-log plot. This rough approach assumes a very simple form for the convective flow and does not include other important physical processes such as the field concentration/cancellation at the cell boundary. Resistivity is known to be an important parameter in simulations of flux emergence (*e.g.*, Arber et al. 2007) and a more realistic model would include resistive processes as well as a more complex flow field, both of which may affect the slope.

3.2 Data and AR Selection

3.2.1 Data and data treatment

The magnetic field evolution of 37 ARs that emerged on the solar disk between June 2010 and the end of 2014 was studied. To ensure consistency and that the emergence time could be correctly identified, only those ARs that emerged in relatively magnetic field free regions (without detectable remnant of other ARs) were chosen. In addition, only ARs emerging prior to central meridian passage (CMP) were selected to analyse all, or at least most, of the emergence phase. The resulting ARs are given in Appendix C.

The ARs were studied using photospheric line-of-sight magnetograms from HMI (Section 2.1). The time at which each AR started to emerge was found by visual inspection of the line-of-sight magnetograms. Although this method is subjective and is influenced by the position of the AR on the solar disk (issues which are discussed in Fu & Welsch 2016), these effects were negligible as a result of the relatively low time cadence data (6 hours) used here to study the ARs.

For the decaying ARs discussed in Section 3.4.3, the ARs began to emerge on the far side of the Sun. In order to estimate the emergence start times in these cases, data from EUVI on the STEREO B spacecraft (Section 2.2) was used.

The magnetic field data were first treated using a cosine correction assuming that the magnetic field is locally vertical (a reasonable assumption for sunspot umbra, but not for penumbra, see Section 1.2.2 and Figure 3.4). This correction adjusts for the fact that the line-of-sight component of the radial field is smaller near the limb; this also means that the signal-to-noise ratio is smaller near the limb, in other words the measurements are less precise. Then a correction for the area was applied, which takes into account the fact that ARs near the limb are foreshortened, as each pixel of the detector measures the flux density of a larger photospheric area. In order to compare flux values, the flux density must be scaled by the photospheric area. It is common to 'derotate' a magnetogram, to produce an image, similar to that which would be seen if the detector were looking at the Sun from a position where the feature being studied was at disk centre. A derotation was applied to each of the magnetograms to the time at which the AR was seen to cross the central meridian.

3.2.2 Definition of AR area

A rectangular submap surrounding the maximum extent of the AR was manually defined, with the area of the AR itself defined using a semi-automated technique (adapted from that outlined in Yardley et al. 2016). This process was important to ensure an unbiased estimation of the field distribution of the AR, especially as the rectangular submap contained magnetic field from the quiet Sun as well as possibly flux from neighbouring regions which would have affected the results.

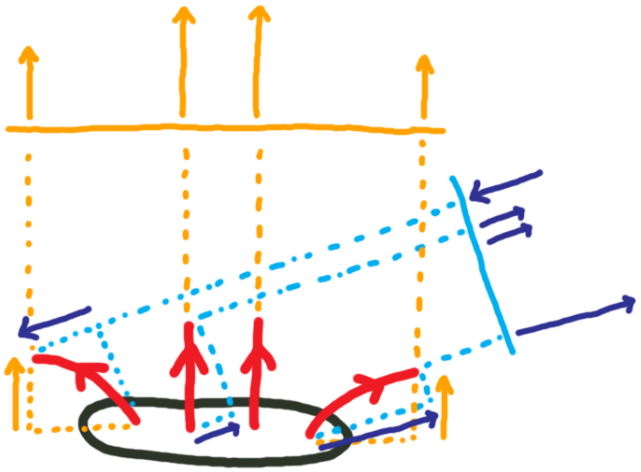

Figure 3.4: Diagram showing the projection effect on penumbral field lines. The sunspot is represented by the black oval shape, its field lines by the red arrows. The field at the centre is radial, while that at the edge of the spot is inclined. The orange arrows represent the amount of magnetic flux measured when the sunspot is at disk centre and the blue arrows when the sunspot is near the limb. Note that it is not the number of arrows that is important, but the length of the arrows and also the area of the detector over which the spot is observed, which is represented by the solid horizontal orange line and the tilted blue one. The field is normally assumed to be radial and the measured components are simply multiplied by a certain (positive) scaling factor, which depends on how close the spot is to disk centre. In the case where the sunspot is near the limb (blue), the magnetic field of the penumbra at the far side of the sunspot (left hand side of the diagram) is measured to be of opposite direction, and so appears to come from a region of opposite polarity. The value calculated for the magnetic field at the near side of the spot (the right hand side) is much larger than reality. This illustrates the problem with assuming a radial field.

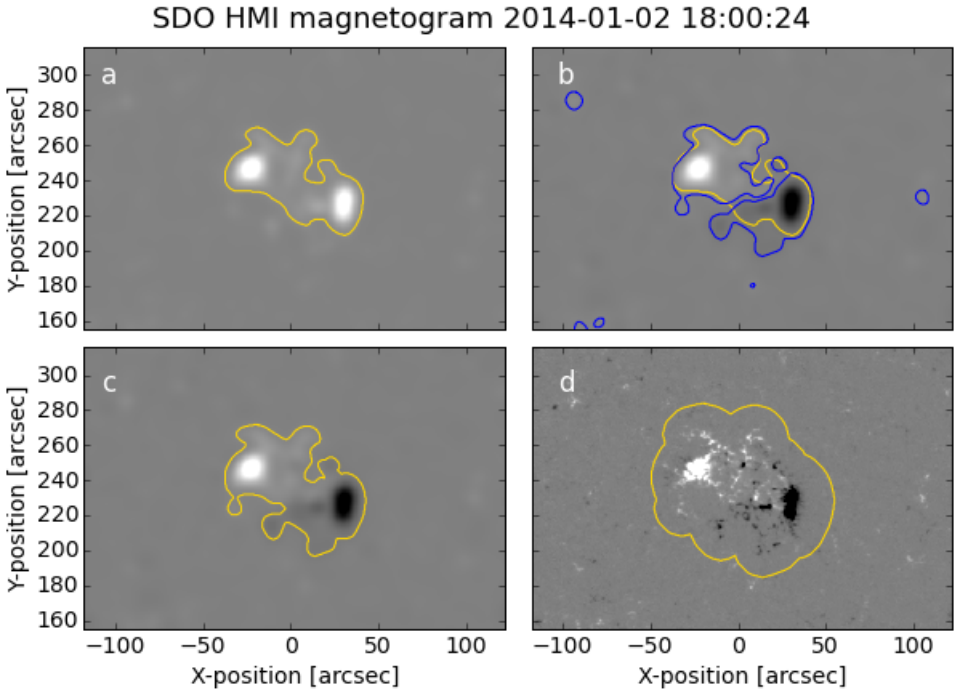

Figure 3.5: Example of the area selection and dilation procedure (Section 3.2.2) for NOAA AR 11945. Panel a shows the selection $>$ 40 G in the blurred map of absolute field values. Panel b shows the selection $>$ 20 G and $<$ -20 G in the smoothed signed map (blue), along with the selection from panel a (yellow). Panel c shows the selection after the field dependent dilation and panel d the final region selection after the field independent dilation.

To focus on the larger scale features rather than individual flux fragments, the absolute values of the magnetic field within the submap were smoothed using a Gaussian filter with a standard deviation (width) of 7 pixels. The AR extension was then automatically defined from the areas in which the pixels had smoothed values > 40 G (shown in Figure 3.5a). For the initial stages of flux emergence, the code searched for a bipolar region, defined as having between a fifth and five times the amount of positive flux as negative flux. If no bipole was found a search was made for positive and negative bubbles that were close to each other (within a distance of 20 pixels, approximately 10 arcseconds). To include regions of fragmented flux emergence, more of the original bubbles were accepted if they lay within 20 pixels of this selection or 10 pixels of these additional bubbles.

As ARs evolve, they change significantly, with the sea serpent structures coalescing into two main polarities which grow to cover a larger area of the solar surface. To reflect these changes, the AR area selection criteria changed after the unsigned flux of the region reached 8×10^{20} Mx, when the AR area was found to dominate the rectangular submap. The AR area was still chosen from the regions of pixels with smoothed values > 40 G, but the selection criteria were different, with larger regions and regions close to the large regions being chosen. To capture fragments of flux that broke away from the main polarities, this area was then extended to neighbouring regions (within 10 pixels).

If at any time step the automated selection was not satisfactory, for example if there were inconsistencies between neighbouring time steps (*i.e.* flux at the edge of the region was not included in the selection at one time step, but was included in both the preceding and the following time steps), the selection could be made manually from the smoothed submap.

After the region had been selected, dilations were applied to include the region immediately surrounding the selection, as this was found to have a different magnetic field distribution to that of the quiet Sun (see Section 3.5) and was therefore considered to be strongly influenced by the emerging flux. For the first dilation, the signed data were smoothed using a Gaussian filter of width 7 pixels. The pixels of smoothed values > 20 G or < -20 G were selected (shown in Figure 3.5b with blue contours). Each of these regions was compared to the previously selected region (yellow contour), and if there was some overlap, the pixels in that region were unmasked (panel c of Figure 3.5). Because the blurring was applied to the signed values, the dispersing parts of the region were likely to be accepted, while any diffuse neighbouring regions of opposite polarity should have been avoided. Finally two field-independent dilations were applied, both using a disk-shaped kernel with a radius of 12 pixels (providing a smoother edge than a single dilation with a larger kernel). All the pixels in the dilated region (shown Figure 3.5d) were analysed.

3.3 Kernel Density Estimation Analysis

3.3.1 Method

Kernel density estimation plots (KDEs; Wegman 1972; Silverman 1986, and references therein) were created for each time step, showing the distribution of pixels with respect to field value (flux density) on a log-log scale. KDEs are very similar to histograms, in that both types of plots show the distribution of data points. However, histograms rely on discretising the data points, using user-defined bins, which normally leads to a loss of information. Moreover, the arbitrary choice of bin positions can affect the subsequent

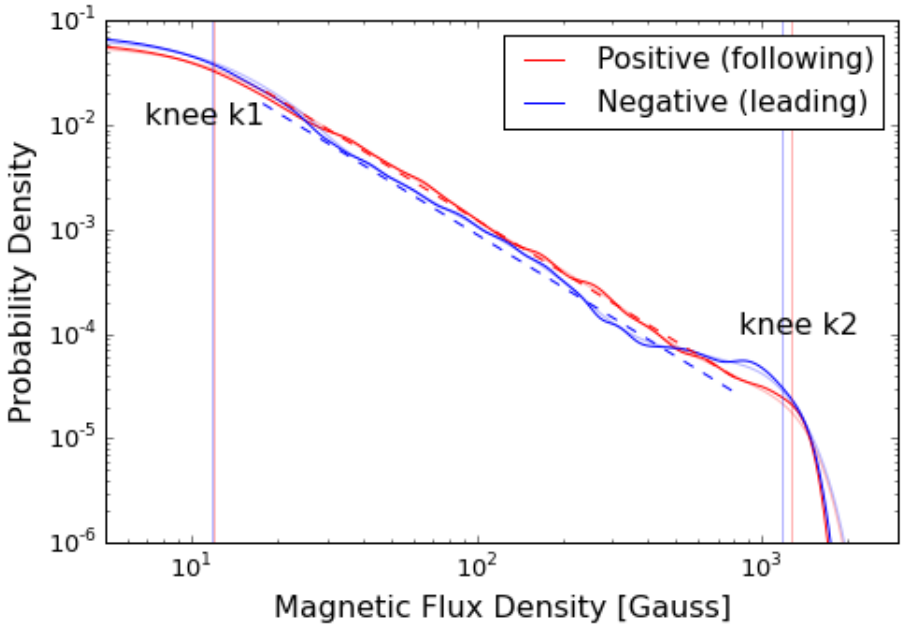

Figure 3.6: The field distribution of NOAA AR 11945 18 hours after the start of its emergence. The distribution is calculated from the pixels within the contour shown in panel d of Figure 3.5. The positions of the knees, k1 and k2 (vertical lines), and the lines of best fit to the middle section (dashed) are shown for both polarities.

interpretation. On the other hand, KDE analysis assigns a kernel of a certain width to each data point, effectively smoothing the finite number of data points. This KDE technique has been shown to be superior to histograms both theoretically and in applications (*e.g.*, de Jager et al. 1986; Vio et al. 1994), and KDEs have been used in a broad range of astrophysics domains (*e.g.*, Schulze-Hartung et al. 2012; Sui et al. 2014).

The applied kernels of KDE analysis can take on any shape and were chosen to peak at the data point value and to decay away from this. The kernel widths can also be varied depending on the data point value. A larger width can be used where there are fewer data points, enabling the noise in the distribution to be made roughly constant. While the smoothness of KDEs clearly depends on the kernel widths, similar to the dependence of histograms on their bin widths, unlike histograms there is no dependence on an arbitrary choice of bin position (the kernel is applied and centred on each data point).

The KDEs in this study[1] were created using a Gaussian kernel with standard deviation equal to 2.5 G for field values between -25 and $+25$ G and equal to $|B_z|/10$ for $|B_z| > 25$ G. This was chosen to give a plot with the noise level being roughly independent of B_z (as it would appear on a log-log plot). From the theoretical considerations of Section 3.1, a line with a slope of -1 would be expected for the KDE when plotted on a log-log scale. This means that the pixel frequency would be proportional to $1/|B_z|$. The optimal KDE plot should be built with a variable kernel width such that the kernel width is inversely proportional to the frequency at that point. Thus a kernel width proportional to $|B_z|$ was used. There are a few reasons why a lower limit on the kernel width was chosen for low $|B_z|$ values. One reason is to represent the instrumental errors, which reduce the precision of the measurement. Another reason is that the frequency does not increase indefinitely as the value $|B_z|$ decreases. Instead, it is expected that the distribution and its

[1]The code used to create and analyse KDEs in this study is available here: https://github.com/SallyDa/Sally_KDE

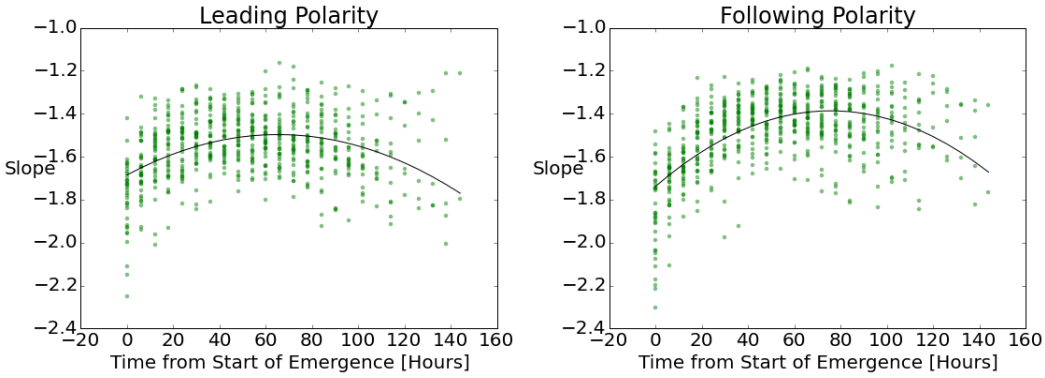

Figure 3.7: The slope of the best fit straight line to the middle section of the log-log field distribution plot was calculated as illustrated in Figure 3.6. This plot shows each slope value plotted against the age of the AR from the start of its emergence ($t = 0$) at each time step and for each of the 37 ARs. Data for leading and following sunspots were separated. Second order polynomials were least-square fitted, showing the general trend.

derivative are continuous across 0 G on a linear B_z scale. Indeed the KDEs do show that the distributions are flatter for $|B_z|$ values less than $\sim$ 25 G (*e.g.*, Figure 3.6).

3.3.2 Observed distributions

All the KDEs were found to have some common features, namely a flatter section below $\sim$ 10 G followed by a turning point (the first knee, at field value k1, indicated by a vertical line on Figure 3.6), a steeper section up to $\sim$ 1000 G and then another turning point (the second knee, at field value k2, also indicated by a vertical line) after which the frequency drops off rapidly. Although k2 varies greatly with the age and strength of the AR, it was not used to characterise the distribution as it is highly dependent on only a very small proportion of pixels.

To capture only the larger scale features, an extra smoothed KDE plot was created using Gaussian kernels of twice the width as before, and the positions of the knees (k1 and k2) were found from this. k1 was taken as the point of most negative second derivative before the slope goes to values less than -1. k2 was taken as the first time, moving along the KDE graph from right to left, that the slope becomes more positive (less steep) than -3. A best fit line was fitted to the original (not extra smooth) KDE plot between $1.5\times$ k1 and $\frac{2}{3}\times$ k2. The best fit line was calculated using the method of least squares, taking data points evenly spaced along the $\log(B_z)$ axis. Examples are shown (dashed lines) in Figure 3.6 for both the positive and negative polarity field distributions.

3.4 Observational Results

3.4.1 Temporal evolution

Figure 3.7 shows how the slope of the distributions changes with time from the start of AR emergence for all studied ARs (listed in Appendix C). The slopes start off steep and negative, in about the range $[-2.2, -1.5]$, peak in the range $[-1.7, -1.2]$, and then become steeper (more negative) again as the ARs decay.

The initially steep slope is related to the fragmented emergence. It is followed by the coalescence of flux elements to form stronger flux concentrations and indeed the slopes around $t = 60$ hours are comparable

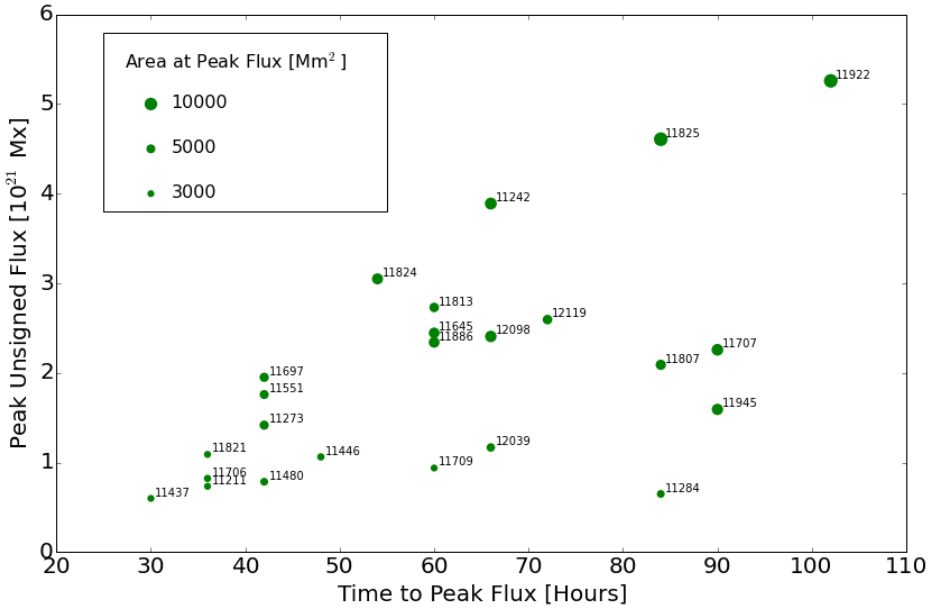

Figure 3.8: The peak unsigned flux reached by each AR is plotted against the time it took to reach that value from the start of the region's emergence. The size of the points gives an indication of the area covered by the region at that time. The points are labelled by their NOAA AR number.

to the slopes found for a magnetic source with $n = 3$ (≈ -1.67) as derived in Section 3.1.1, equation (3.11). The change in slope during the regions' decay is due to dispersion with the number of higher-field pixels decreasing and the proportion of low- to middle-field pixels increasing. The latter evolution goes away from the slope given by classical diffusion (≈ -1, Sections 3.1.2 and 3.1.3).

3.4.2 Evolution versus magnetic flux

The timescale of emergence varies broadly from one AR to another (*e.g.*, van Driel-Gesztelyi & Green 2015, and references therein). As a result, in Figure 3.7 ARs with various flux values and emergence durations are presented together, mixing different phases and field strength distributions of AR evolution (*e.g.*, ~ 80 hours after the start of flux emergence some ARs are still emerging while smaller ARs are already in their decay phase). In this Section, the emergence stage of an AR is characterised by the amount of emerged magnetic flux, normalised to its maximum value.

The maximum flux, $F_{\max}$, is defined as the maximum unsigned flux achieved during the AR evolution, so long as this value was not reached at the last time step. 24 regions from the original 37 reached peak flux and were included in this part of the study. The other regions were still in their emergence phase at the end of the observational time period.

Figure 3.8 shows that larger ARs generally do take a longer time to emerge, but that the relationship between emergence time and peak flux is not a simple one, with some regions taking a long time to emerge without reaching a particularly high peak flux. The rates of flux emergence were seen to vary, both during the emergence of single regions and from one AR to another in agreement with the results of Poisson et al. (2015, 2016) obtained on other sets of ARs. Some regions were seen to have an initial gradual emergence phase followed by a more rapid phase, while others emerged rapidly at the start and had a decreasing

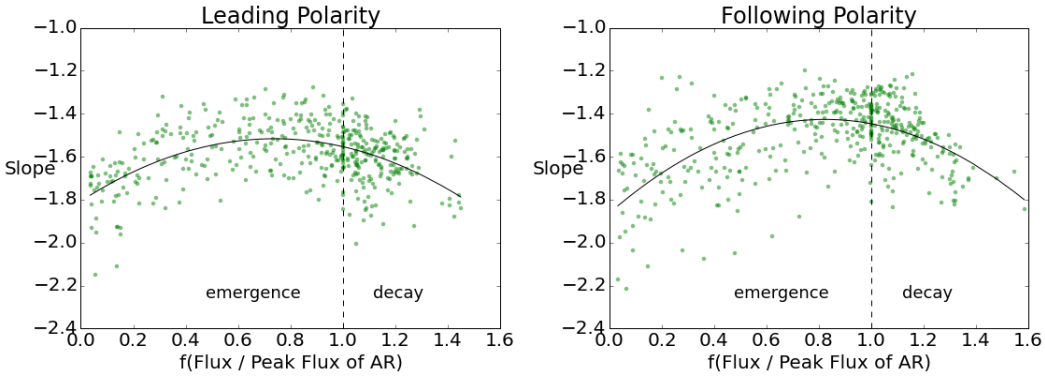

Figure 3.9: Similar to Figure 3.7, this plot shows the evolution of the slope using data from 24 ARs. Here, the slope is plotted against a function of the normalised magnetic flux. This normalised flux is defined as the ratio of the AR flux to the peak flux that region achieves. The peak flux and the ratios are found separately for the positive and negative magnetic polarities. To distinguish between emergence and decay phases, the function $f(F)$, defined by equation (3.48), is used in abscissa. The emergence phase is in the range $[0,1]$ and the decay phase is for abscissa > 1. Second order polynomials were least-square fitted, showing the general trend.

emergence rate later on.

The ratio of flux to peak magnetic flux is used in Figure 3.9 to study the emergence independent of the time it takes. Here the maximum flux is computed separately for the two magnetic polarities. Moreover to separate the emergence from the decay phase (which are in the same range of flux) the function $f(F)$, defined below, is used.

$$\begin{aligned} f(F/F_{\max}) &= F/F_{\max} \qquad \text{for } t \leq t_{\max} \\ &= 2 - F/F_{\max} \quad \text{for } t > t_{\max}\,, \end{aligned} \tag{3.48}$$

where t is the time since the emergence start, $F_{\max}$ is the maximum flux and $t_{\max}$ the duration of the emergence until $F_{\max}$ is reached. Despite the issues arising due to variable emergence rates, Figure 3.9 shows an improvement to Figure 3.7 in that the trends, particularly in the case of the leading spot, are more pronounced.

On average, there is an imbalance present in the slopes of the field distribution between the leading and following polarities. At the beginning of the emergence phase the field distribution slope is slightly more negative for the following than the leading polarity, becoming less negative as the peak value of flux is reached. However, these differences (best seen in the fitted polynomials) are comparable to the standard deviation of the residuals. To reduce the statistical noise and investigate whether these slopes are dependent upon parameters other than the normalised flux a larger sample of regions would need to be studied (Section 4.1).

In summary, Figure 3.9 shows one important aspect of the evolution of the magnetic field distribution, namely the transformation of a distribution mainly dominated by the weak fields (steeper slope), at the beginning of emergence, to one with more numerous stronger field pixels, around the maximum flux, and then back to a distribution dominated by weak field pixels. The typical slopes around the maximum flux correspond to the slope found with the simple model described in Section 3.1.1. In contrast, neither the initial emergence nor the decay phase conforms with the plausible theoretical model considered. In particular the

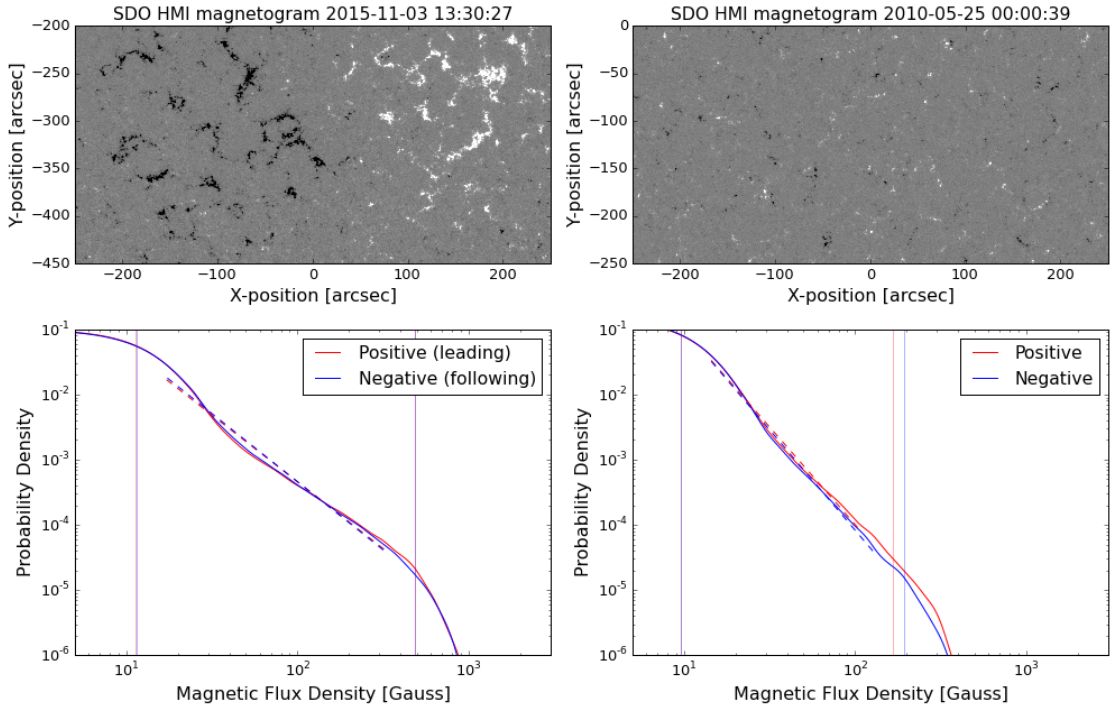

Figure 3.10: A decayed AR (left column) and a region of quiet Sun (right column) are shown in the top panels, with their magnetic field distributions in the bottom left and right panels respectively. The positions of the knees, k1 and k2 (vertical lines), and the lines of best fit to the middle section (dashed) are shown for both polarities. The slope for the decaying region is -2.07 and that for the quiet Sun region is -3.04, with positive and negative slopes having a difference of less than 0.1 for each case.

distribution during the beginning of the decay shows an evolution opposite to that expected from classical diffusion (with the slope becoming steeper rather than tending towards the expected value -1 as found in Sections 3.1.2 and 3.1.3).

3.4.3 Decayed ARs

The ARs selected for this study were ones that emerged on the solar disk and were tracked until they rotated too far ($\geq 60^o$) from the central meridian. As such, all the data came from relatively young ARs, up to just 6 days old. A preliminary study of decaying ARs was performed to see if the decreasing slope trend continues as ARs continue to age. Three relatively isolated decaying ARs (NOAA 12165, 12176 and 12414), in which no new flux emergence was observed, were selected. The ARs were chosen to be of approximately the same maximum flux ($\sim 5 \times 10^{21}$ Mx), as this affects how quickly the ARs decay and how quickly their distribution changes. I would also expect the distribution of a region with a larger maximum flux to have a more positive (less steep) slope, but further research would be needed to confirm this. In this study of the decay phase, two of the ARs started emerging on the far side of the Sun, so their emergence start times were identified by eye using 171 Å data from STEREO B. Two of the regions were observed for multiple solar rotations, providing data points at later times.

To calculate the field distributions of these regions, all of the pixels within a rectangular subplot centred on the AR were used. This simple procedure was used, as at these times the ARs are quite dispersed and difficult to distinguish from the quiet Sun field. An example of one of the decaying ARs and the rectangular subplot that defined the region is shown in the top left panel of Figure 3.10, with the AR's field distribution

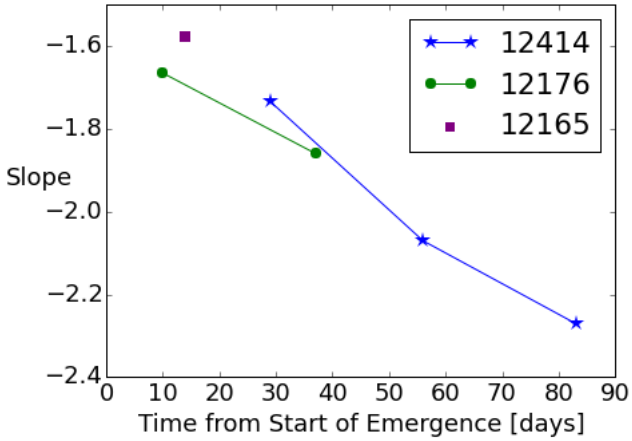

Figure 3.11: The slopes for older decaying ARs are shown against time from the start of the AR's emergence. A downward trend can be seen.

shown in the bottom left panel. The slopes of the three decaying ARs at the different times since they started emerging are shown in Figure 3.11. Here, an average of the positive and negative slope values was used, as the slopes for the two different polarities were consistently very similar, in each case differing by less than 0.1. This may not be the case for an AR where the leading polarity retains a coherent spot for longer than the following polarity.

The results shown in Figure 3.11 indicate that the downward (steepening) trend of the slope during AR decay continues as the regions continue to age. A deeper study is needed to confirm this result and to find the dependence on maximum flux (Section 4.1).

3.4.4 Quiet Sun

At the end stages of AR decay, the dispersing magnetic field becomes part of the quiet Sun. Four regions were studied to find the slope of the quiet Sun field distribution. As for the decaying ARs in Section 3.4.3, all the pixels within a selected rectangular area were used to build the field distribution. One of the four regions and its distribution is shown as an example in the right column of Figure 3.10. The regions studied were taken at the central meridian on 2010-May-25 (shown in Figure 3.10), 2015-Nov-25, 2015-Dec-22 and 2016-Jan-18. The latter three regions were centred on -75 arcseconds in Solar Y.

The average of the positive and negative distribution slopes were -3.0, -3.0, -2.9 and -3.3, with the difference between the positive and negative being less than 0.3 for each region. Combining this result with that of Figure 3.11, it implies that the slope of the AR field distribution continues to decrease (steepen) as the AR decays, until reaching the quiet Sun value ~ -3.

A possible origin of the ~ -3 slope arises from the relationship between the magnetic field and a diverging flow field, as described in Section 3.1.4. A detailed analysis of observations comparing the magnetic field distributions to Doppler velocity maps and/or numerical simulations could be used to test if the ~ -3 slope found in the quiet Sun is mainly due to the advection of the magnetic field by the diverging flow of supergranules.

3.5 Possible Issues for the Derived Distributions

This Section describes some of the tests done to ensure that the derived slopes in emerging ARs are not affected by the inclusion of surrounding quiet Sun or by projection effects. With regard to the area selection process (described in Section 3.2.2), the area bordering the ARs was analysed by selecting a ribbon-like region around the ARs and by studying the field distributions for various ribbon thickness. In the bordering region the distribution was found to differ from that of the quiet Sun. As such, the decision was made to include these neighbouring pixels in the selection by applying dilations. Care was taken to ensure that the dilations were not too large, avoiding the inclusion of decaying field from old ARs nearby, which could be identified by eye.

The quiet Sun region was also used to study possible effects arising from the position of the region on the solar disk with respect to distance from the central meridian. Contrary to ARs, the distribution of the quiet Sun is expected not to evolve on a time scale of weeks, so any observed change of the observed distribution indicates a viewing point bias as the region shifts away from the central meridian. The main change found in the distribution was a shift of k1 to higher field values. One reason for this is the larger errors in the line-of-sight magnetic field measurement closer to the solar limb (Hoeksema et al. 2014). These errors are then exaggerated by the cosine correction under the radial field assumption. Even if there were no errors in the measured field values, the validity of the assumption that the measurements are of a purely radial field decreases for both low field strength regions (less vertical magnetic field) and away from the disk centre, where the measured field contains a larger component of the transverse field. This also introduces errors in the field distribution.

A shift in k1 has implications for the slope, with a larger k1 in general giving rise to a steeper slope. The shift in k1 becomes more pronounced $\sim$ 3 - 4 days from the central meridian ($\sim$ 45 - 60 degrees). In the AR sample analysed in Sections 3.4.1 and 3.4.2, if k1 was seen to increase significantly as the region moved away from the central meridian, then this and any following time steps were removed from the analysis.

To further ensure that the trends seen in Figures 3.7 and 3.9 did not result from the position of the AR on the disk in terms of area foreshortening effects and interpolations associated with derotation, Figure 3.12 includes only data from ARs within 10 hours, $\sim$ 6 degrees, of their central meridian passage. The regions were in various stages of their evolution as they passed the central meridian. Figure 3.12 shows the same trends as were obtained with the larger number of data points in Figure 3.9, with a slightly more marked difference between the leading and following polarities. Thus, the trends in slope are related to the real field distributions of the ARs.

3.6 Discussion and Conclusions

This study of 37 emerging ARs has shown that there is a relationship between the slope of the vertical component of the photospheric field distribution and the age of an AR. This is summarised in Figure 3.13. At the beginning of a region's emergence, the slope is steep and negative. The slope becomes less steep, which indicates the coalescence of the fragmented flux that emerges. Later, the slope reaches a maximum just before the region achieves its peak flux value (at $\sim 0.75 - 0.8$ peak flux), before the decay processes become dominant. The slope becomes more negative as the region disperses and this decreasing trend

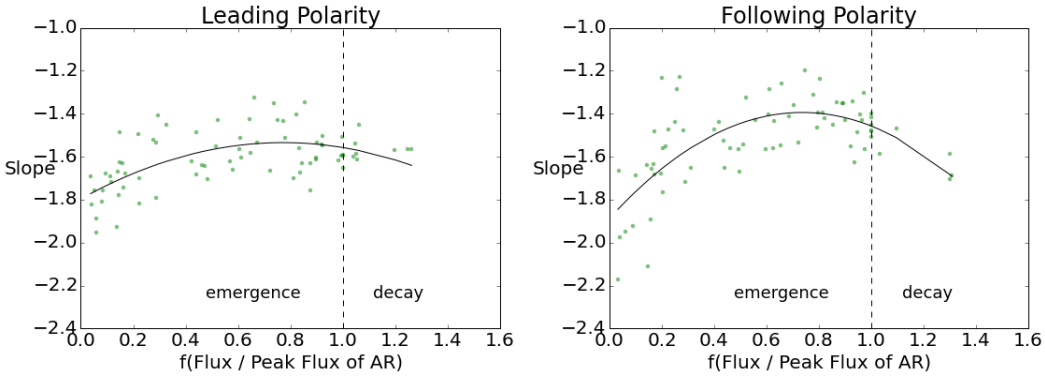

Figure 3.12: Similar to Figure 3.9, but in this case only using data points from the 24 ARs within 10 hours, ~ 6 degrees, of their central meridian passage. To distinguish between emergence and decay phases, the function $f(F)$, defined by equation (3.48), is used in abscissa. Second order polynomials were least-square fitted, showing the general trend.

continues towards the quiet Sun slope value of ~ -3 (Figure 3.10).

A comparison between the observational and theoretical results shows that a simple model of magnetic concentrations can describe the field (flux density) distribution in emerging ARs during the coalescence phase when smaller flux concentrations merge to form larger ones, leading to sunspot formation. The model predicts a slope of ≈ -1.67 for $n = 3$, in good agreement with the slope values found in observations of the coalescence phase (Figures 3.7 and 3.9).

However, later on there is a major deviation from the classical-diffusion model in the decay phase, indicating that AR magnetic fields do not disperse by simple diffusion. The latter predicts that after reaching peak flux, the field strength distribution should be characterised by a slope which is evolving from the range $[-1.6, -1.4]$ towards the diffusion exponent value of -1. However, as Figures 3.9 and 3.12 clearly demonstrate, once ARs pass their peak flux and start decaying, their field strength distribution slopes evolve quite differently from these expectations: they start to attain higher negative values. Furthermore, ARs measured in the later decay phase display slopes in the range of $[-2.3, -1.6]$, as shown in Figure 3.11, while the quiet Sun, which can be regarded as the end-product of AR decay, shows a slope ≈ -3. How can this behaviour be understood, when it is so clearly opposite to the classical diffusion scenario?

The importance of the flow field and its interaction with the magnetic field was suggested in Section 3.1.4, where a slope of -3 was found using the induction equation and assuming a radially directed flow with constant magnitude. This is an oversimplified model and it is likely that multiple processes contribute to the observed slope.

Magnetic flux reprocessing by convective cells may be responsible for the observed evolution of field distributions. Magnetic flux is being gnawed away by granular and supergranular convective cells, in agreement with the turbulent diffusion model (*e.g.*, Petrovay & Moreno-Insertis 1997, see Section 1.2.4), which carry away flux concentrations from the strong-field area of ARs. The turbulent diffusion model is also consistent with the removal of AR magnetic field by MMFs (*e.g.*, Kubo et al. 2008), the movement of which is a result of convective flows. The advected field fragments become concentrated along the boundaries of supergranular cells, where they occasionally meet and cancel with opposite polarity field. The cancelled flux submerges and is re-processed by convection. At the same time, flux emerges in the centre of supergranular

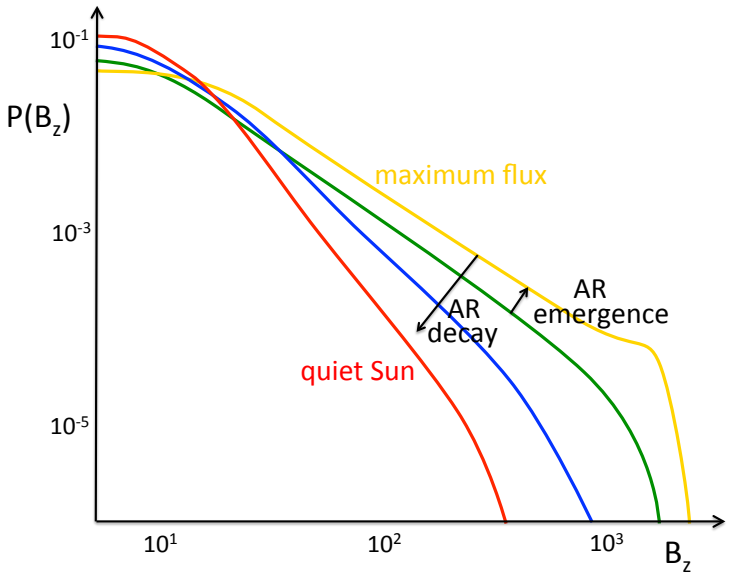

Figure 3.13: Schema summarising the evolution of the B_z distribution with a log-log plot. It shows the difference in the distribution between a newly emerging AR (green), one at around the time of peak flux (yellow) and one during the decay phase (blue). After this time, the field distribution of the AR evolves towards that of the quiet Sun (red). Since both magnetic polarities have very similar distributions at a given time (*e.g.*, Figures 3.6 and 3.10) they are not differentiated in this schema.

cells as intranetwork flux. The emerging flux tubes may be too small to be fully resolved by HMI, so the area in which they emerge appears to have weak flux values. The process leads to a removal of strong field pixels at the cell boundary and an increase in weaker field pixels in the cell centres, making the slope of the field distribution of a decaying AR steeper (more negative). Convection also carries the intranetwork field to the boundary of the supergranular cells where it builds up (becoming better resolved by HMI), and this is a counter-mechanism to the former scenario. Then the -3 slope found in the quiet-Sun field should be the result of the combination of these two mechanisms.

Although the evolution of the magnetic field in a decaying AR is very different from what is expected in classical diffusion (slope of -1), the AR area has previously been found to increase approximately linearly with time (*e.g.*, van Driel-Gesztelyi et al. 2003), in accordance with the classical diffusion model. Is there a contradiction here? The increase of AR area may be due to the non-stationarity of the supergranular convective cells, which have a lifetime of about one day. This non-stationarity creates a random walk of the flux tubes, which is the underlying physics of classical diffusion. Therefore the AR evolution can be seen as classical diffusion regarding area increase, while the magnetic field distribution is governed by magneto-convection.

Chapter 4

Future Work

4.1 A Larger Study

There are many advantages to having a larger data set, the main ones being that the statistical noise is reduced and we can have more confidence in any results. The most obvious future project would therefore be to extend the data set used in Section 3.4 to further confirm how the magnetic field distribution of an AR varies with the AR's age. Looking at a larger number of ARs could also allow investigations into whether the slope values are affected by other factors, such as the latitude of AR emergence and the period of the solar cycle. Figure 3.9 shows that the slope depends on the evolutionary stage of the AR *i.e.* the instantaneous ratio of the magnetic flux to the peak flux, but how does the evolution of the slope depend on the maximum flux? Other AR characteristics, such as magnetic twist may also influence the slopes. Is there then a link to the activity of the region?

In Sections 3.4.1 and 3.4.2, only ARs that were observed from the start of their emergence were used. One of the results was that the trends in slope evolution are better seen when the slopes were plotted against the normalised flux, as defined in equation (3.48), than against time. Thus it is important that the AR is observed at its maximum flux, but the start of emergence does not need to be observed. The AR selection criteria could be adjusted accordingly, allowing a larger number of ARs to be analysed.

Another way to increase the number of ARs would be to use data from the Michelson Doppler Imager (MDI; Scherrer et al. 1995), on board the *SOlar and Heliospheric Observatory* (SOHO; Domingo et al. 1995). MDI was the predecessor to HMI and started taking data in 1996, 14 years before the launch of SDO, providing a much longer data set covering more than one solar cycle. MDI magnetograms were produced from observations of the Doppler shift in left and right circularly polarised light calculated from filtergrams at 4 wavelength positions across the Ni I 6768 Å absorption line. They have a time cadence of 90 minutes, lower than HMI but more than sufficient for a study similar to that of Section 3.4. The spatial resolution of MDI magnetograms is also lower than that of HMI; one pixel covers an area roughly 16 times that of a HMI pixel. To determine whether it would be plausible to directly compare ARs observed by MDI with those observed by HMI, a feasibility study has been undertaken using an emerging AR observed by HMI and comparing its distribution to that produced when the resolution is lowered (Section 4.2).

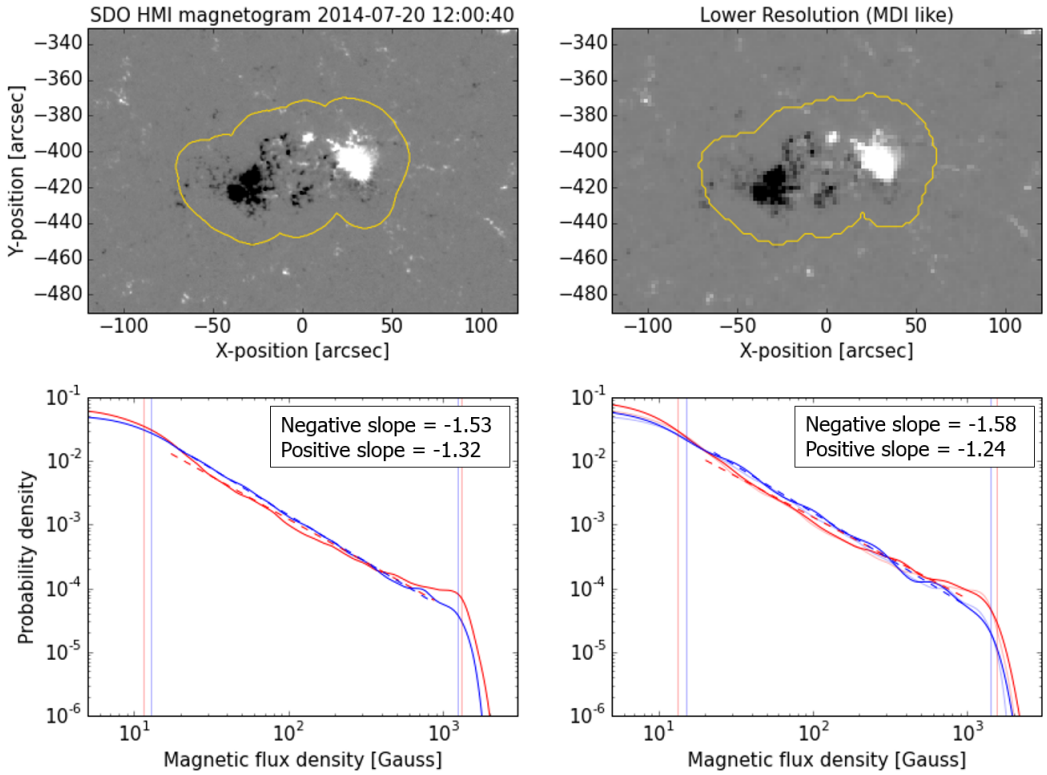

Figure 4.1: The effect of image resolution on the KDE plot. The left column contains the image at its original resolution and the corresponding KDE, created as described in Section 3.3. The right column contains a lower resolution image of the same magnetogram, with its resolution matching that of MDI magnetograms. The corresponding KDE (bottom right) has been created with a kernel width twice that of the original resolution KDE, to compensate for the reduced number of data points. The lower resolution KDE plot also shows the original KDE as faint lines. The distributions at the two different resolutions are remarkably similar.

4.2 The Effect of Spatial Resolution

It is important to know whether the field distribution depends on the resolution of the magnetogram, to know whether data sets at different resolutions could be compared or whether there would be systematic differences between them. Additionally, studying the effect of changing the magnetogram resolution may provide insight into important spatial scales on the photosphere.

In this Section, the effect of image resolution on the KDEs has been investigated for NOAA 12119, one of the emerging ARs studied in Section 3.4, at 20 time steps. Two different resolutions were compared: the original and a resolution comparable to that of MDI data. The resolution was reduced by taking the average value of 16 pixels (a 4×4 square) of the original data and assigning this to a single pixel in the lower resolution magnetogram. This is equivalent to a magnetogram taken by a lower resolution instrument, as each magnetic field value is the average over the area of the photosphere covered by that pixel (see Section 2.1). Figure 4.1 shows the NOAA 12119 and its field distribution at one time step, for both the original resolution data (left) and when the resolution of the image has been reduced (right). For this study, the AR extent (yellow contour) was defined from the original resolution image, and pixels in this region were taken also for the lower resolution KDE analysis. The KDEs were created as in Section 3.3, with the kernel width doubled for the lower resolution data, to compensate for the reduced number of magnetic field values. The characterisation of each KDE using the knee positions and slope of the best fit line to the central portion of

the distribution, was done as described in Section 3.3 for both the original and lower resolution cases.

For the positive polarity, the maximum absolute difference between the original and lower resolution KDE slopes was 0.09, and the mean absolute difference was 0.05, with the lower resolution slope being slightly greater (less steep) than the higher resolution slope at all but the first time steps. Similar results were obtained for the negative slope, when the first time step was discounted (due to the very low number of pixels producing a noisy KDE), with maximum and mean absolute differences of 0.15 and 0.04 respectively. The lower resolution slopes were slightly greater than the corresponding original slopes for 14 of the 19 time steps. The slopes for both the original and lower resolution varied in the range $[-1.9, -1.3]$, so even the largest difference between the different resolutions was comparatively small.

The positions of the knees play a role in determining the slopes, and indeed it was a large shift in the second knee which caused the large change in slope of the negative polarity at the first time step. The knee positions also differed at later time steps and maximum differences of 2.8 G and 260 G were found for the first and second knees respectively, with both knees generally shifted to higher values for the lower resolution KDEs. The fractional changes in the knee positions were sometimes in excess of 20%, but in general it seems these changes did not have a large effect on the slopes.

In a study of ARs observed by MDI, the method used to create and characterise the KDE plots could remain essentially the same as that for HMI magnetograms, with a simple increase in the width of the KDE kernel. Similarly, a few factors would need to be changed in the selection process of the AR extent, described in Section 3.2.2. Most notably, for MDI data, less smoothing would need to be applied; a Gaussian with a smaller standard deviation, in terms of pixel coordinates, could be used for this. Once the MDI regions had been selected and the AR extent defined, direct comparisons could be made with the observational study of Section 3.4. As suggested in Section 4.1, this would enable more ARs to be studied, hopefully confirming the previous results and enabling the investigation of other possible influencing factors, such as the peak flux, or the latitude of AR emergence.

The result that the KDE slopes of an emerging AR do not exhibit large changes when the resolution is reduced by a factor of 16, was not necessarily expected and is interesting - not only from the point of view of enabling further studies using MDI data. Further research into the effect of changing the resolution of the magnetograms could give some indication of important spatial scales. At some point, lowering the resolution would change the distribution, as the highest field values would be reduced by averaging with neighbouring lower field pixels. Low flux pixels would also be affected, with neighbouring positive and negative pixels averaging to a lower absolute field value. How the distribution changes and which parts change first would depend on the area scales associated to the regions of different field strengths. This could give some indication of the spatial scales involved in flux emergence, coalescence and dispersion. In addition, results from such a study could be compared to other studies (*e.g.*, Abramenko & Yurchyshyn 2010), which characterise the AR magnetic field by its degree of intermittency. The distributions of regions with highly intermittent small scale structure would be strongly affected by reducing the resolution of the data so that the small scale structures are no longer resolved. There is a limit to how much the resolution could be reduced whilst still producing a meaningful KDE, as this requires a reasonable number of data points. As such, small ARs may not be suited to this study. On the other hand, it may be possible to look at

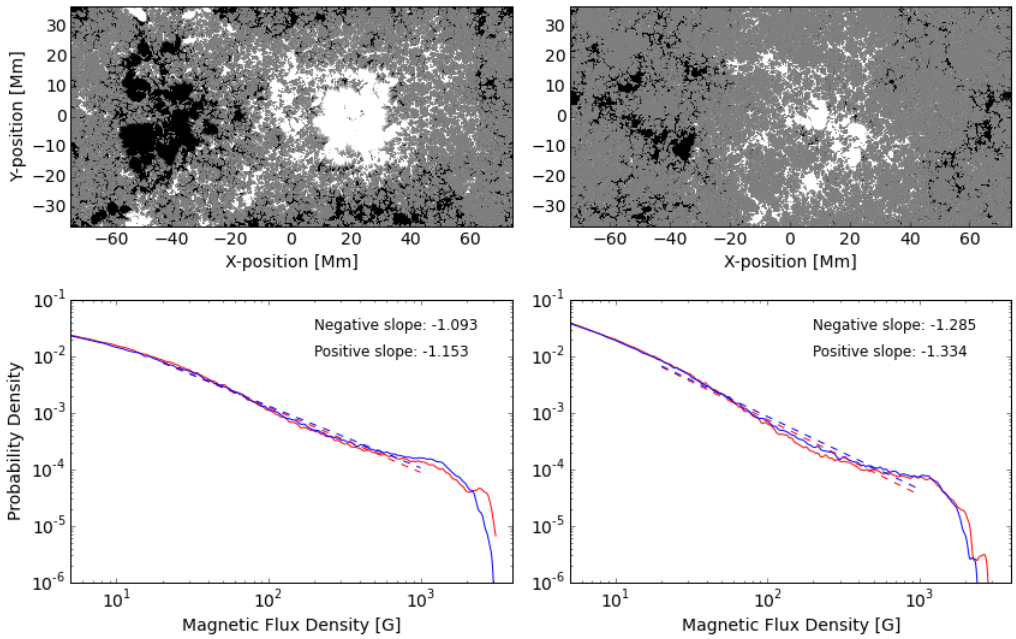

Figure 4.2: Images of the simulated magnetograms Rempel & Cheung (2014) and their corresponding KDEs at around the time of maximum flux (left) and at the end of the simulation run (right).

quiet Sun regions or even the solar disk as a whole.

4.3 The Field Distribution of Simulation Data

By performing KDE analysis on simulated magnetic field data at the photosphere and comparing this to the observational results found in Chapter 3, it may be possible to identify which processes are involved in producing the observed distributions. As discussed in Sections 1.2.3 and 1.2.4, the simulations of Rempel & Cheung (2014) include convection, whereas many other simulations of AR emergence do not. Comparing the magnetic field distributions from these simulations could indicate the importance of convection for AR emergence. Additionally, comparing simulations with different parameter values could give information about how the slopes are affected by the total flux, area (cross section of the flux tube), twist, and the time scale of emergence. An initial study comparing simulation data from Rempel & Cheung (2014) and from James Leake and Mark Linton of the Naval Research Laboratory (NRL; private communication) suggests that convection is important for the low flux values, both at the start of the emergence and when the AR is more established. However, it may not be necessary for the change in slopes at the higher B_z end of the distribution.

4.3.1 A simulation including convection

The KDEs produced from the Rempel & Cheung (2014) data (Figure 4.2) have a very similar shape to the observed distributions, but with less pronounced knees. For each time step, a best fit straight line was fitted to the distribution between 20 and 1000 G and the slope of this was used to characterise the distribution as before. The slope becomes less steep as emergence progresses, reaching a maximum value of ~ -1.1 at around the time of maximum flux and steepens again as the region decays (Figure 4.3). This trend is the

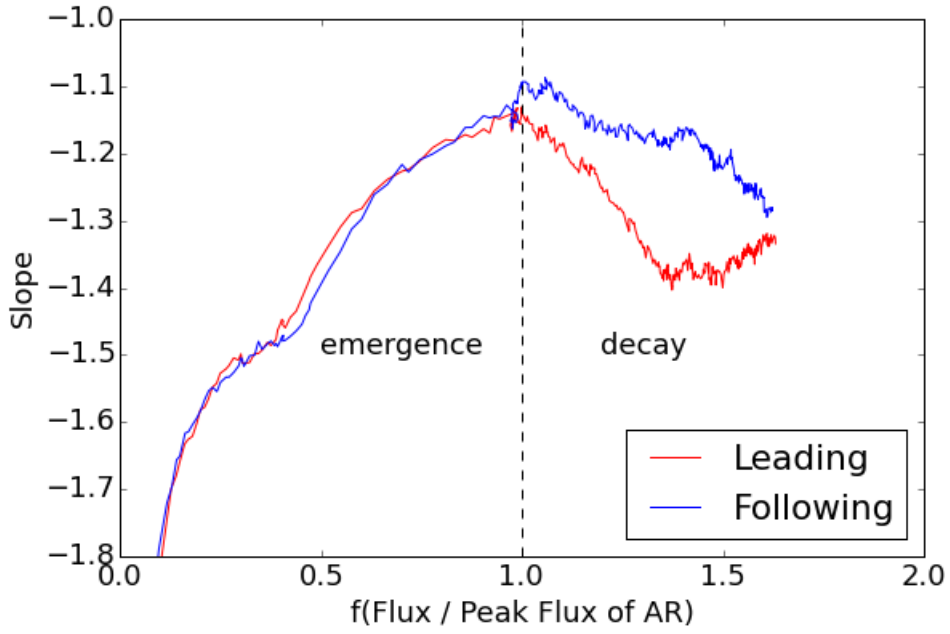

Figure 4.3: The slope of the field distribution of a simulated AR (Rempel & Cheung 2014) plotted against the normalised flux, defined in equation 3.48.

same as that found from observations, and the slope values are close, with observational slopes from some regions peaking at ~ -1.2 (Section 3.4).

There are a few differences in the procedure as compared to that used in Sections 3.2.2 and 3.3. For the simulated data, the slopes were defined between fixed points (20 and 1000 G), because the knees of the simulated distributions are not as clear as in the observational distributions. This should not have a big effect on the slopes, except at the beginning of the emergence where the slopes will appear steeper (more negative) than if a lower upper limit was used to define the best fit line.

Another difference is that there was no area selection process. For most of the simulation, the selection process used in Section 3.2.2 would have included the entire field of the simulation. The periodic boundary conditions of the simulation mean that negative flux appears on the right hand side and this should probably be excluded from the distribution. Using a rectangular box to exclude the region nearest the boundary (within 40 pixels, 7.68 Mm of each side) did not have a large effect on the distributions. For the positive polarity, the slope was increased by ~ 0.1 throughout the decay phase. For the negative polarity, the time of peak flux occurred later, but the negative slope values were very similar to those found using the entire field.

The resolutions of the simulated and observed magnetograms also differ. The horizontal grid spacing of the simulation data is 192 km, which is a lower resolution than that used in the simulation (96 km) to save space, but higher than that of the magnetograms produced by HMI. To determine whether this influenced the distribution, the simulated data was analysed at resolutions of one half (pixel width of 384 km, comparable to observations from HMI at disk centre), one quarter and one eighth (comparable to MDI) of the resolution of the saved data. As the resolution was decreased, the slope values showed an initial increase before

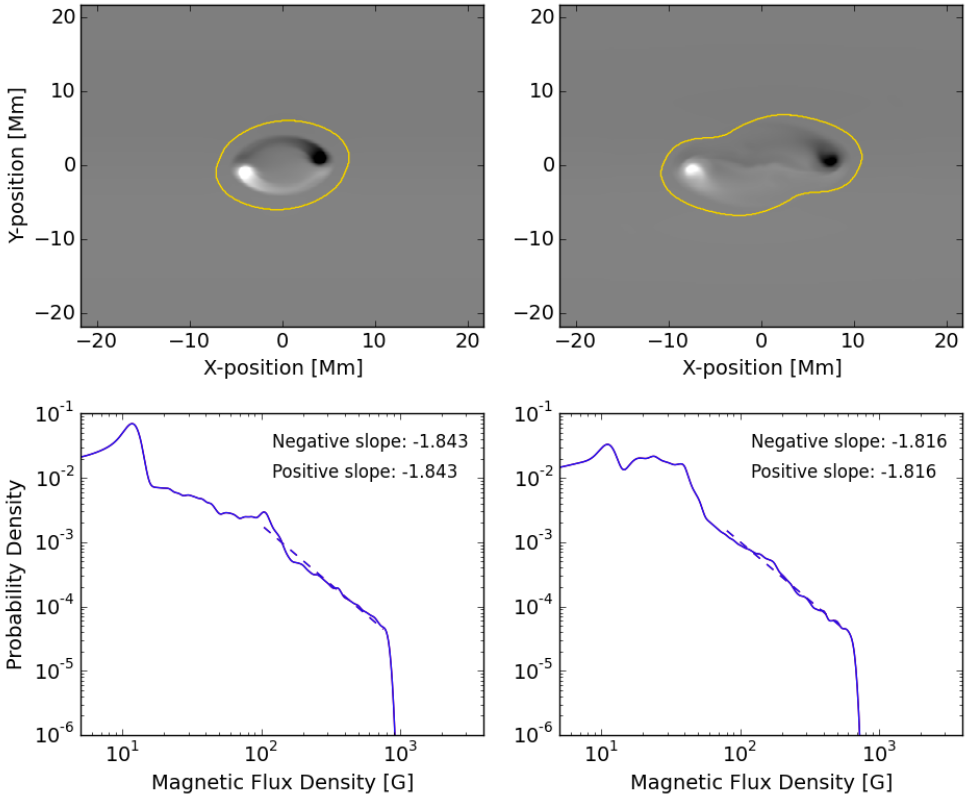

Figure 4.4: Magnetograms from one simulation run (James Leake and Mark Linton) at the third (left) and final (right) time steps and their corresponding KDEs.

decreasing (steepening) again as the resolution was further decreased.

It is possible that the slightly higher (less steep) values of the slopes associated with this simulation compared to the observations are related to the size of the emerging AR. The simulated region had a peak flux of 1.8×10^{22} Mx, about a factor of 10 larger than the ARs studied in Sections 3.2.2, 3.3 and 3.4. An observational study that looks at larger ARs and the dependence of the slope values on the flux as well as a simulation run with a smaller flux tube would be needed to confirm this.

4.3.2 Simulations without convection

For comparison to the simulation distributions found in Section 4.3.1, KDEs were also created for the simulated magnetograms of James Leake and Mark Linton (Figure 4.4). Their simulations do not include convection. Different features are seen in the distributions, particularly for low B_z values. There is no clear first knee, but a jagged distribution. On the other hand, at larger B_z values there is a roughly straight line portion and a second knee, similar to the observed distributions. The straight line portion increases as the emergence progresses and the distributions become more similar to the observed ones. The slope values were calculated for the straight line portion. Here, the first knee was defined as the B_z value greater than 80 G associated with the highest probability density. The second knee was calculated in a similar way to that described in Section 3.3.2 and a straight line was fitted between the two knees. The slope values are very similar to the observed slopes, taking values in the range $[-1.9, -1.4]$ (excluding the first time step, where the straight line part was very small) and peaking just before the time of peak flux (Figure 4.5).

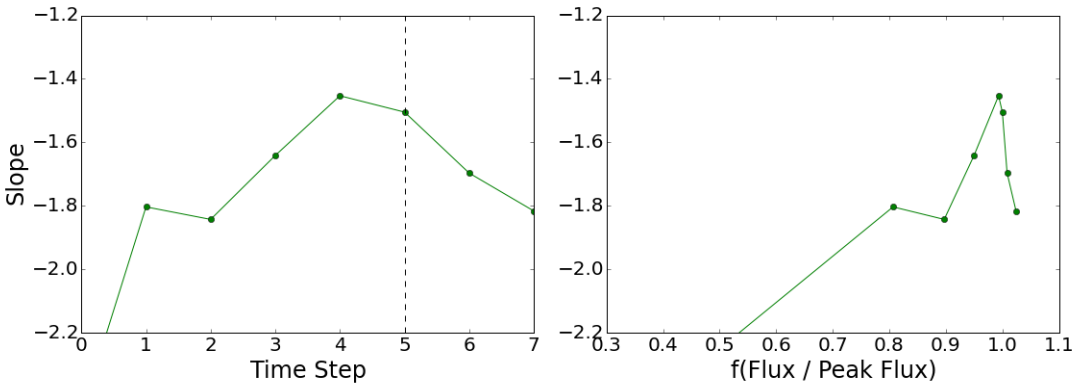

Figure 4.5: The evolution of the slope of the field distribution of a simulated AR from James Leake and Mark Linton. Both the positive and negative distributions were the same at all time steps. The slope values have been plotted against time step (left) with a vertical dashed line indicating the time of peak flux and against the normalised flux (right).

Plotting the slope against the normalised flux does not produce such convincing similarities, possibly due to the contribution of the azimuthal field of the strongly twisted simulated region to the total flux. The azimuthal component will also have an influence on the distributions particularly at early time steps and may be a reason why the distributions differ from the observations.

Again, a few small alterations needed to be made to the method. In this case an area selection procedure was used and this was very similar to that described in Section 3.2.2 for the observed ARs. The simulation uses an irregular Cartesian grid, such that the central region has a higher spatial resolution than the edges of the domain. This optimises the computational time and the quality of the interesting part of the simulation, where the flux emerges. The variable pixel size associated with the irregular grid was taken into account when producing the KDEs, with larger pixels making a correspondingly larger contribution to the distribution. The main difference to the observational procedure and that used for the Rempel & Cheung (2014) data is in the definition of the first knee. Here, the first knee was consistently at a larger value, but this was necessary due to the difference in the shape of the distribution at low B_z values.

Because the simulations of James Leake and Mark Linton do not include convection, they are less computationally expensive to run than those of Rempel & Cheung (2014), which would be advantageous for a parameter study. A region with a larger peak flux (1.2×10^{21} Mx, compared to 5.1×10^{19} Mx of the region discussed above) was also studied. The distributions have a similar shape to those of the previous region and the slope values show the same evolutionary trend (excluding the first time step). However, the slopes are steeper, with values more negative than those observed. It would be interesting to study a larger set of simulation runs with a selection of peak flux values to get a clearer idea of the effect of the peak flux on the distribution. The azimuthal field component is expected to have a large influence on the distribution at early times, so it would be good to analyse data from runs with a range of twists.

4.3.3 Comparison of the simulations

Both of the simulations studied (Sections 4.3.1 and 4.3.2) show distributions with some similarities to the observations. The slope values of the small flux region from James Leake and Mark Linton (Figures 4.4 and 4.5) are the closest to the observed values. However, the general shape of the distributions from

the simulations including convection (Rempel & Cheung 2014) is closer to that observed, particularly at low B_z values and during the early stages of emergence. This suggests that convection may be important for emergence and for the distribution at low magnetic field values during the whole evolution. From these simulations it is difficult to make inferences about the role of convection in the decay phase, as the simulations without convection (James Leake and Mark Linton) do not include this phase. This study has the potential to be extended much further by looking at how changes in the simulation parameters affect the distribution.

Comparisons to other simulations would also be advantageous, in particular to those with a deeper convection zone (*e.g.*, Stein et al. 2011, 2012). These simulations can be considered a more accurate representation of flux emergence, as the magnetic field inserted into the computational domain is only minimally structured, unlike in most current models (including those studied in this Chapter). Until now, deep convection zone models have reproduced the emergence of small scale loops, but have been unable to self consistently produce an AR. Hopefully this will be achieved in the near future and studies of the simulated magnetic field distribution could be used to test the success of these models.

Chapter 5

Summary

A trend was found in the changes of the distribution (probability density) of the vertical component of the photospheric magnetic field of ARs as they emerge, evolve and eventually disperse into the quiet Sun. On a log-log plot, the distribution can be characterised by two turning points and an approximately linear part between them. The slope of the best fit straight line to this central portion was found to be steep and negative at the start of emergence, increasing to a maximum value ~ -1.5 just before the time of peak flux and then decreasing towards the quiet Sun value ~ -3. This differed to the result expected from simple classical diffusion, which predicts a slope of -1, suggesting that other processes may be important in the dispersion of flux. One of these processes could be magnetic flux reprocessing by (super)granular convective cells, in agreement with the turbulent diffusion model (*e.g.*, Petrovay & Moreno-Insertis 1997).

To further investigate the importance of convective processes, it would be instructive to analyse data produced from simulations with and without convection. An initial study of one simulation run with convection (Rempel & Cheung 2014) and a comparison with a simulation without convection (developed by James Leake and Mark Linton) suggests that convective processes are important, particularly at the start of emergence and for low magnetic field values. Further research using simulated magnetograms would help to confirm this. Additionally, further observational studies could be used to determine other factors which may influence the distributions, such as peak flux, twist, or phase of the solar cycle. Data from MDI could be used for this, with a few small adjustments to the methods to take into account the lower resolution, and the results could be directly compared to those from HMI.

Appendix A

Magnetohydrodynamics (MHD)

Magnetohydrodynamics (MHD) consists of a set of equations to describe the macroscopic behaviour of plasma in a magnetic field. As such it is frequently used to describe the behaviour of the Sun's plasma and magnetic field. The MHD equations can be derived by applying certain assumptions to equations describing the motion of a fluid, consisting of electrons and ions, subject to electric and magnetic fields. The assumptions applied in MHD are that electron inertia is negligible, the plasma is quasi-neutral, electromagnetic waves can be neglected (the displacement current is much smaller than the plasma current) and that the phenomena being studied occur on length scales larger than the Larmor radius (the radius of the circular motion of an ion in a uniform magnetic field). External forces, such as gravity, are often ignored, as is the case here. A further assumption is that changes in the fluid's pressure and volume occur adiabatically, *i.e.* there is no heating. Defining mass density as ρ_{M}, pressure as P, plasma velocity as $\mathbf{v}$, magnetic field as $\mathbf{B}$, electric field as $\mathbf{E}$, current density as $\mathbf{J}$, and with γ the ratio of specific heats, μ_0 the permeability of free space and η the magnetic diffusivity, the resulting equations are given as follows.

Mass conservation:

$$\frac{\partial \rho_{\mathrm{M}}}{\partial t} + \frac{\partial}{\partial \mathbf{r}} \cdot (\rho_{\mathrm{M}} \mathbf{v}) = 0 \tag{A.1}$$

Newton's second law of motion for fluids:

$$\rho_{\mathrm{M}} \frac{\mathrm{d}\mathbf{v}}{\mathrm{d}t} = \mathbf{J} \times \mathbf{B} - \frac{\partial \mathrm{P}}{\partial \mathbf{r}} \tag{A.2}$$

Energy conservation:

$$\frac{\mathrm{d}}{\mathrm{d}t}(\mathrm{P}\rho_{\mathrm{M}}^{-\gamma}) = 0 \tag{A.3}$$

Maxwell equations[1]:

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{A.4}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} \tag{A.5}$$

Ohm's law, describing the evolution of the current density:

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = \mu_0 \eta \mathbf{J} \tag{A.6}$$

[1] In this form, there are four Maxwell equations, however only two are used in MHD. $\nabla \cdot \mathbf{E} = \frac{\rho}{\varepsilon_0}$ (where ρ is the charge density and ε_0 is the permittivity of free space) is not allowed in MHD, where it is assumed that the plasma is quasi-neutral. $\nabla \cdot \mathbf{B} = 0$ is an initial condition *i.e.* it does not provide any information about the evolution of the system.

In ideal MHD, electrical resistivity is ignored, $\mu_0\eta = 0$, and a further simplification can be made to Ohm's law:

$$\mathbf{E} + \mathbf{v} \times \mathbf{B} = 0 \tag{A.7}$$

The induction equation describes the relationship between the magnetic and the velocity fields. It can be derived from equations (A.4), (A.5) and (A.6), as follows. Taking the curl of equation (A.6) provides

$$\nabla \times \mathbf{E} = \nabla \times (-\mathbf{v} \times \mathbf{B}) + \mu_0\eta(\nabla \times \mathbf{J}). \tag{A.8}$$

Substituting equation (A.4) into equation (A.8) gives

$$-\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (-\mathbf{v} \times \mathbf{B}) + \mu_0\eta(\nabla \times \mathbf{J}). \tag{A.9}$$

To find $\nabla \times \mathbf{J}$ in terms of $\mathbf{B}$, the curl of equation (A.5) is taken, providing

$$\begin{aligned} \mu_0\eta(\nabla \times \mathbf{J}) &= \eta(\nabla \times (\nabla \times \mathbf{B})) \\ &= \eta(\nabla(\nabla \cdot \mathbf{B}) - \nabla \cdot (\nabla \mathbf{B})) \\ &= -\eta\nabla^2\mathbf{B}, \end{aligned} \tag{A.10}$$

using $\nabla \cdot \mathbf{B} = 0$. Substituting equation (A.10) into equation (A.9) gives the induction equation:

$$\frac{\partial \mathbf{B}}{\partial t} = \eta\nabla^2\mathbf{B} + \nabla \times (\mathbf{v} \times \mathbf{B}), \tag{A.11}$$

which for ideal MHD reduces to:

$$\frac{\partial \mathbf{B}}{\partial t} = \nabla \times (\mathbf{v} \times \mathbf{B}). \tag{A.12}$$

An important result arising from the ideal induction equation (A.12) is Alfvén's frozen-in-flux theorem, which states that plasma is unable to move across magnetic field lines. This can be derived as follows. The magnetic flux passing through a surface S tied to the fluid can be defined as

$$\phi_S = \int_S \mathbf{B} \cdot \mathrm{d}S. \tag{A.13}$$

Two processes might affect how ϕ_S changes with time, namely changes in $\mathbf{B}$ with time and movement of the surface S with the plasma to a position with a different $\mathbf{B}$. The first contribution is

$$\frac{\partial \phi_S}{\partial t} = \int_S \frac{\partial \mathbf{B}}{\partial t} \cdot \mathrm{d}S. \tag{A.14}$$

The second contribution can be evaluated as follows. A small vector $\mathrm{d}l$ along the circumference of S sweeps out an area $\mathrm{d}A = \mathbf{v}\mathrm{d}t \times \mathrm{d}l$ in time $\mathrm{d}t$. The flux passing through this area $\mathrm{d}\phi = \mathbf{B} \cdot \mathrm{d}A$. The total change in flux due to the movement of the surface can be found by integrating along S:

$$\frac{\mathrm{d}\phi}{\mathrm{d}t} = \oint \mathbf{B} \cdot (\mathbf{v} \times \mathrm{d}l), \tag{A.15}$$

which using standard vector identities and Stokes' theorem can be rewritten as

$$\begin{aligned} \frac{\mathrm{d}\phi}{\mathrm{d}t} &= \oint (\mathbf{B} \times \mathbf{v}) \cdot \mathrm{d}l \\ &= \int_S \nabla \times (\mathbf{B} \times \mathbf{v}) \cdot \mathrm{d}S. \end{aligned} \tag{A.16}$$

Combining the two contributions gives

$$\frac{\mathrm{d}\phi_S}{\mathrm{d}t} = \int_S \left(\frac{\partial \mathbf{B}}{\partial t} - \nabla \times (\mathbf{v} \times \mathbf{B}) \right) \cdot \mathrm{d}S \,. \tag{A.17}$$

Comparing the integrand of equation (A.17) to the ideal induction equation (A.12), shows that

$$\frac{\mathrm{d}\phi_S}{\mathrm{d}t} = 0 \,, \tag{A.18}$$

the magnetic flux through a surface moving with the fluid does not change with time.

To determine whether the dynamics of an ideal MHD system are dominated by the magnetic field or by plasma motions, the ratio of the plasma pressure to the magnetic pressure is calculated. This important parameter is known as plasma beta:

$$\beta = \frac{2\mu_0 \, \mathrm{P}}{\mathrm{B}^2} \,. \tag{A.19}$$

Appendix B

Magnetic Reconnection

While ideal MHD provides a good description of the coronal magnetic field, it does not explain energetic events, such as flares, or changes in the coronal magnetic field structure. Magnetic reconnection changes the magnetic field topology and converts magnetic energy to kinetic and thermal energies, often observed on the Sun as brightenings in certain spectral lines. The classic model of reconnection proposed by Sweet (1958) and Parker (1957) is a two dimensional model, which assumes resistive MHD. In this approach, shown in Figure B.1, oppositely oriented field lines are brought towards each other, probably due to plasma flows in a high β regime, or forces resulting from pressure imbalance in a low β regime *e.g.*,when magnetic flux emerges into or is ejected away from the corona. Between the approaching oppositely oriented field lines, the magnetic field gradient increases, until a thin current sheet is formed between them. The plasma is incompressible and flows out of the current sheet (in a direction perpendicular to the inflow) at high speed. Magnetic flux is conserved (the magnetic field strength of the outflowing field lines divided by the width of the current sheet is equal to the magnetic field strength of the inflowing lines divided by the length of the current sheet), so the magnetic energy of the outflowing field lines is very small. The magnetic energy of the inflowing field lines is converted to kinetic energy of the outflowing plasma and thermal energy. However, the timescales over which reconnection occurs in this model are much slower than those observed. One way of dealing with this problem is to assume a different geometry, such as in the Petschek model, where the reconnection region is smaller and more square. Another possibility is that the resistivity in the current sheet takes a much higher value, which could be caused by turbulence in this region. MHD may not be

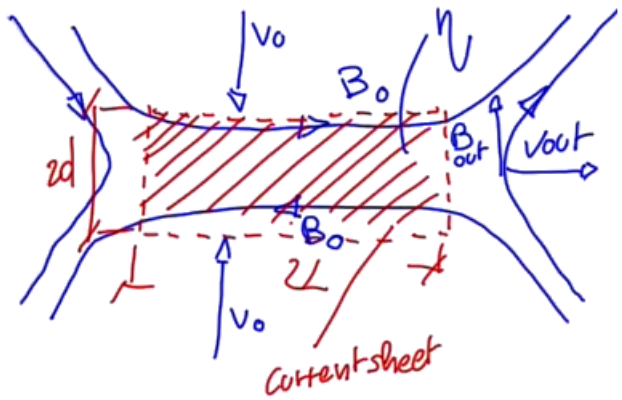

Figure B.1: The Sweet-Parker model of magnetic reconnection. Image courtesy of Ivo Furno from the École Polytechnique Fédérale de Lausanne (EPFL) Massive Open Online Course (MOOC).

suitable for modelling reconnection, and some more recent simulations (*e.g.*, Daughton et al. 2011) have used fully kinetic models (where each particle is modelled individually, rather than as a single fluid).

Appendix C

Table of ARs

37 ARs that emerged on the solar disk between June 2010 and the end of 2014 were studied and are listed in the following table. These ARs emerged in relatively magnetic-field free regions prior to (or around) their central meridian passage. The first column contains the AR number and the second the time the AR passes the central meridian (CMP). The third and fourth columns contain the time difference with CMP of the first appearance and of the maximum unsigned flux, respectively. The fifth column contains the value of the maximum unsigned flux (the average of magnetic flux of both polarities).

NOAA	Time CMP [UT]	Δt_{start} [hours]	Δt_{max} [hours]	Maximum Flux [10^{21} Mx]
11199	2011-04-25 12:00	0	78	6.70
11211	2011-05-08 11:00	-5	31	0.73
11242	2011-06-28 20:00	-20	46	3.89
11273	2011-08-17 21:00	-33	9	1.42
11284	2011-08-26 19:00	-37	47	0.65
11398	2012-01-13 10:00	-10	68	4.64
11437	2012-03-19 06:00	-66	-36	0.60
11446	2012-03-24 00:00	-30	18	1.06
11480	2012-05-11 07:00	-43	-1	0.79
11551	2012-08-21 00:00	-30	12	1.76
11631	2012-12-12 14:00	-20	64	5.11
11645	2013-01-03 17:00	-23	37	2.45
11697	2013-03-13 09:00	3	45	1.95
11702	2013-03-20 10:00	-22	80	3.22
11706	2013-03-28 08:00	-38	-2	0.82
11707	2013-03-31 06:00	-66	24	2.26
11709	2013-03-28 10:00	-40	20	0.94
11750	2013-05-15 00:00	0	84	5.06
11765	2013-06-07 09:00	-63	75	6.40
11768	2013-06-11 15:00	-3	75	8.03
11776	2013-06-19 08:00	-20	70	6.32
11781	2013-06-28 17:00	-23	79	6.23
11807	2013-07-28 17:00	-11	73	2.09
11813	2013-08-07 12:00	-30	30	2.73
11821	2013-08-13 07:00	-1	35	1.09
11824	2013-08-17 04:00	-4	50	3.05
11825	2013-08-18 02:00	-8	76	4.61
11837	2013-09-01 17:00	-29	61	5.90
11886	2013-10-29 10:00	-34	26	2.34
11888	2013-10-29 17:00	-53	73	3.43
11919	2013-12-08 09:00	15	69	2.35
11922	2013-12-09 21:00	-33	69	5.26
11945	2014-01-04 09:00	-57	33	1.59
12039	2014-04-16 20:00	-26	40	1.17
12089	2014-06-13 06:00	-60	84	12.40
12098	2014-06-26 10:00	-64	2	2.41
12119	2014-07-20 04:00	-46	26	2.59

References

Abramenko, V. & Yurchyshyn, V. 2010, Astrophys. J., 722, 122

Arber, T. D., Haynes, M., & Leake, J. E. 2007, Astrophys. J., 666, 541

Archontis, V., Hood, a. W., Savcheva, A., Golub, L., & Deluca, E. 2009, Astrophys. J., 691, 1276

Archontis, V., Moreno-Insertis, F., Galsgaard, K., Hood, A., & O'Shea, E. 2004, Astron. Astrophys., 426, 1047

Borrero, J. M. & Solanki, S. K. 2008, Astrophys. J., 687, 10

Borrero, J. M., Tomczyk, S., Kubo, M., et al. 2011, Sol. Phys., 273, 267

Canfield, R. C., Hudson, H. S., & McKenzie, D. E. 1999, Geophys. Res. Lett., 26, 627

Chen, P. F. 2011, Living Rev. Sol. Phys., 8

Cheung, M. C. M. & Isobe, H. 2014, Living Reviews in Solar Physics, 11

Cheung, M. C. M., Rempel, M., Title, a. M., & Schüssler, M. 2010, Astrophys. J., 720, 233

Choudhuri, A. R. 2015, Nature's Third Cycle: A Story of Sunspots (Oxford University Press), 281

Couvidat, S., Rajaguru, S. P., Wachter, R., et al. 2012, Sol. Phys., 278, 217

Couvidat, S., Schou, J., Hoeksema, J. T., et al. 2016, Sol. Phys., 291, 1887

Daughton, W., Roytershteyn, V., Karimabadi, H., et al. 2011, Nat. Phys., 7, 539

de Jager, O. C., Raubenheimer, B. C., & Swanepoel, J. W. H. 1986, Astron. Astrophys., 170, 187

Démoulin, P., Mandrini, C. H., Rovira, M. G., Henoux, J. C., & Machado, M. E. 1994, Sol. Phys., 150, 221

Domingo, V., Fleck, B., & Poland, A. I. 1995, Sol. Phys., 162, 1

Emonet, T. & Moreno-Insertis, F. 1998, Astrophys. J., 492, 804

Falco, M., Borrero, J. M., Guglielmino, S. L., et al. 2016, Sol. Phys., 291, 1939

Fan, Y. 2001, Astrophys. J., 554, L111

Fu, Y. & Welsch, B. T. 2016, Sol. Phys., 291, 383

Harvey, K. & Harvey, J. 1973, Sol. Phys., 28, 61

Harvey, K. L. 1993, PhD thesis, Univ. Utrecht

Hathaway, D. H. & Choudhary, D. P. 2008, Sol. Phys., 250, 269

Hoeksema, J. T., Liu, Y., Hayashi, K., et al. 2014, Sol. Phys., 289, 3483

Hood, a. W., Archontis, V., & MacTaggart, D. 2012, Sol. Phys., 278, 3

Howard, R. A., Moses, J. D., Vourlidas, A., et al. 2008, Space. Sci. Rev., 136, 67

Kaiser, M. L., Kucera, T. A., Davila, J. M., et al. 2008, Space. Sci. Rev., 136, 5

Krall, J., Chen, J., Santoro, R., et al. 1998, Astrophys. J., 500, 992

Kubo, M., Lites, B. W., Shimizu, T., & Ichimoto, K. 2008, Astrophys. J., 686, 1447

Liu, J. & Zhang, H. 2006, Sol. Phys., 234, 21

MacTaggart, D. & Hood, A. W. 2009, Astron. Astrophys., 507, 995

Magara, T. & Longcope, D. W. 2001, Astrophys. J. Lett., 559, L55

Manchester IV, W. 2001, Astrophys. J., 547, 503

Murray, M. J., Hood, a. W., Moreno-Insertis, F., Galsgaard, K., & Archontis, V. 2006, Astron. Astrophys., 460, 909

Ortiz, A., Bellot Rubio, L. R., & van der Voort, L. R. 2010, Astrophys. J., 713, 1282

Pariat, E., Aulanier, G., Schmieder, B., et al. 2004, Astrophys. J., 614, 1099

Parker, E. N. 1957, J. Geophys. Res., 62, 509

Pesnell, W. D., Thompson, B. J., & Chamberlin, P. C. 2012, Sol. Phys., 275, 3

Petrovay, K. & Moreno-Insertis, F. 1997, Astrophys. J., 485, 398

Petrovay, K. & van Driel-Gesztelyi, L. 1997, Sol. Phys., 176, 249

Poisson, M., Démoulin, P., López Fuentes, M., & Mandrini, C. H. 2016, Sol. Phys., 291, 1625

Poisson, M., Mandrini, C. H., Démoulin, P., & López Fuentes, M. 2015, Sol. Phys., 290, 727

Rempel, M. & Cheung, M. C. M. 2014, Astrophys. J., 785, 90

Rimmele, T. R. 2004, Astrophys. J., 604, 906

Rust, D. M. 1994, Geophys. Res. Lett., 21, 241

Sainz Dalda, A. & Bellot Rubio, L. R. 2008, Astron. Astrophys., 481, L21

Scherrer, P. H., Bogart, R. S., Bush, R. I., et al. 1995, Sol. Phys., 162, 129

Scherrer, P. H., Schou, J., Bush, R. I., et al. 2012, Sol. Phys., 275, 207

Schou, J., Antia, H. M., Basu, S., et al. 1998, Astrophys. J., 505, 390

Schou, J., Scherrer, P. H., Bush, R. I., et al. 2012, Sol. Phys., 275, 229

Schrijver, C. J. & Harvey, K. L. 1994, Sol. Phys., 150, 1

Schrijver, C. J. & Title, A. M. 1999, Sol. Phys., 188, 331

Schulze-Hartung, T., Launhardt, R., & Henning, T. 2012, Astron. Astrophys., 545, A79

Schüssler, M. & Rempel, M. 2005, Astron. Astrophys., 441, 337

Schüssler, M. & Vögler, A. 2006, Astrophys. J., 641, L73

Silverman, B. W. 1986, Density estimation for statistics and data analysis, Vol. 26 (CRC press)

Sobotka, M. & Roudier, T. 2007, Astron. Astrophys., 472, 277

Socas-Navarro, H., Martínez Pillet, V., Sobotka, M., & Vázquez, M. 2004, Astrophys. J., 614, 448

Spruit, H. C. & van Ballegooijen, A. A. 1982, Astron. Astrophys., 106, 58

Stein, R. F., Lagerfjärd, A., Nordlund, ., & Georgobiani, D. 2012, in Astronomical Society of the Pacific Conference Series, Vol. 455, 4th Hinode Science Meeting: Unsolved Problems and Recent Insights, ed. L. Bellot Rubio, F. Reale, & M. Carlsson, 133

Stein, R. F., Lagerfjärd, A., Nordlund, Å., & Georgobiani, D. 2011, Sol. Phys., 268, 271

Strous, L. H., Scharmer, G., Tarbell, T. D., Title, A. M., & Zwaan, C. 1996, Astron. Astrophys., 306, 947

Sui, N., Li, M., & He, P. 2014, Mon. Not. R. Astron. Soc., 445, 4211

Švanda, M., Sobotka, M., & Bárta, T. 2014, Astrophys. J., 790, 135

Sweet, P. A. 1958, in Electromagn. Phenom. Cosm. Phys., 123–134

Toriumi, S. & Yokoyama, T. 2011, Astrophys. J., 735, 126

Tortosa-Andreu, A. & Moreno-Insertis, F. 2009, Astron. Astrophys., 507, 949

van Driel-Gesztelyi, L., Démoulin, P., Mandrini, C. H., Harra, L., & Klimchuk, J. A. 2003, Astrophys. J., 586, 579

van Driel-Gesztelyi, L. & Green, L. M. 2015, Living Rev. Sol. Phys., 12, 1

van Driel-Gesztelyi, L., Mandrini, C. H., Thompson, B., et al. 1999, in Astronomical Society of the Pacific Conference Series, Vol. 184, Third Advances in Solar Physics Euroconference: Magnetic Fields and Oscillations, ed. B. Schmieder, A. Hofmann, & J. Staude, 302–306

Vio, R., Fasano, G., Lazzarin, M., & Lessi, O. 1994, Astron. Astrophys., 289, 640

Wang, Y. M. & Sheeley, N. R. 1989, Sol. Phys., 124, 81

Wegman, E. J. 1972, Technometrics, 14, 533

Wilson, P. R., Burtonclay, D., & Li, Y. 1997, Astrophys. J., 489, 395

Wuelser, J.-P., Lemen, J. R., Tarbell, T. D., et al. 2004, in Proc. SPIE, Vol. 5171, Telescopes and Instrumentation for Solar Astrophysics, ed. S. Fineschi & M. A. Gummin, 111–122

Yardley, S. L., Green, L. M., Williams, D. R., et al. 2016, Astrophys. J., 827, 151

Zwaan, C. 1978, Sol. Phys., 60, 213