Günther Rüdiger and Rainer Hollerbach

The Magnetic Universe

Geophysical and Astrophysical Dynamo Theory

WILEY-VCH Verlag GmbH & Co. KGaA

Günther Rüdiger and Rainer Hollerbach

The Magnetic Universe

Geophysical and Astrophysical Dynamo Theory

Günther Rüdiger and Rainer Hollerbach

The Magnetic Universe

Geophysical and Astrophysical Dynamo Theory

WILEY-VCH Verlag GmbH & Co. KGaA

Authors

Günther Rüdiger
Astrophysical Institute Potsdam
gruediger@aip.de

Rainer Hollerbach
Dept. of Mathematics, University of Glasgow
rh@maths.gla.ac.uk

Cover picture
Total radio emission and magnetic field vectors of M51, obtained with the Very Large Array and the Effelsberg 100-m telescope (λ=6.2 cm, see Beck 2000). With kind permission of Rainer Beck, Max-Planck-Institut für Radioastronomie, Bonn.

This book was carefully produced. Nevertheless, authors, and publisher do not warrant the information contained therein to be free of errors. Readers are advised to keep in mind that statements, data, illustrations, procedural details or other items may inadvertently be inaccurate.

Library of Congress Card No.: applied for
British Library Cataloging-in-Publication Data:
A catalogue record for this book is available from the British Library

Bibliographic information published by
Die Deutsche Bibliothek
Die Deutsche Bibliothek lists this publication in the Deutsche Nationalbibliografie; detailed bibliographic data is available in the Internet at <http://dnb.ddb.de>.

© 2004 WILEY-VCH Verlag GmbH & Co. KGaA,
Weinheim

All rights reserved (including those of translation into other languages). No part of this book may be reproduced in any form – nor transmitted or translated into machine language without written permission from the publishers. Registered names, trademarks, etc. used in this book, even when not specifically marked as such, are not to be considered unprotected by law.

Printed in the Federal Republic of Germany
Printed on acid-free paper

Printing Strauss GmbH, Mörlenbach
Bookbinding Litges & Dopf GmbH ,
Heppenheim
ISBN 3-527-40409-0

Contents

Preface XI

1 Introduction 1

2 Earth and Planets 3
- 2.1 Observational Overview 3
  - 2.1.1 Reversals 4
  - 2.1.2 Other Time-Variability 6
- 2.2 Basic Equations and Parameters 6
  - 2.2.1 Anelastic and Boussinesq Equations 7
  - 2.2.2 Nondimensionalization 9
- 2.3 Magnetoconvection 12
  - 2.3.1 Rotation or Magnetism Alone 14
  - 2.3.2 Rotation and Magnetism Together 15
  - 2.3.3 Weak versus Strong Fields 16
  - 2.3.4 Oscillatory Convection Modes 18
- 2.4 Taylor's Constraint 18
  - 2.4.1 Taylor's Original Analysis 19
  - 2.4.2 Relaxation of Ro = E = 0 21
  - 2.4.3 Taylor States versus Ekman States 22
  - 2.4.4 From Ekman States to Taylor States 24
  - 2.4.5 Torsional Oscillations 28
  - 2.4.6 $\alpha\Omega$-Dynamos 29
  - 2.4.7 Taylor's Constraint in the Anelastic Approximation 30
- 2.5 Hydromagnetic Waves 30
- 2.6 The Inner Core 32
  - 2.6.1 Stewartson Layers on $\mathcal{C}$ 33
  - 2.6.2 Nonaxisymmetric Shear Layers on $\mathcal{C}$ 33
  - 2.6.3 Finite Conductivity of the Inner Core 36
  - 2.6.4 Rotation of the Inner Core 37
- 2.7 Numerical Simulations 38
- 2.8 Magnetic Instabilities 40
- 2.9 Other Planets 42
  - 2.9.1 Mercury, Venus and Mars 42

2.9.2 Jupiter's Moons . . . 44
2.9.3 Jupiter and Saturn . . . 45
2.9.4 Uranus and Neptune . . . 46

3 Differential Rotation Theory 47
3.1 The Solar Rotation . . . 47
3.1.1 Torsional Oscillations . . . 51
3.1.2 Meridional Flow . . . 52
3.1.3 Ward's Correlation . . . 53
3.1.4 Stellar Observations . . . 55
3.2 Angular Momentum Transport in Convection Zones . . . 57
3.2.1 The Taylor Number Puzzle . . . 63
3.2.2 The Λ-Effect . . . 64
3.2.3 The Eddy Viscosity Tensor . . . 72
3.2.4 Mean-Field Thermodynamics . . . 74
3.3 Differential Rotation and Meridional Circulation for Solar-Type Stars . . . 77
3.4 Kinetic Helicity and the DIV-CURL-Correlation . . . 81
3.5 Overshoot Region and the Tachocline . . . 84
3.5.1 The NIRVANA Code . . . 85
3.5.2 Penetration into the Stable Layer . . . 86
3.5.3 A Magnetic Theory of the Solar Tachocline . . . 89

4 The Stellar Dynamo 95
4.1 The Solar-Stellar Connection . . . 95
4.1.1 The Phase Relation . . . 96
4.1.2 The Nonlinear Cycle . . . 97
4.1.3 Parity . . . 99
4.1.4 Dynamo-related Stellar Observations . . . 101
4.1.5 The Flip-Flop Phenomenon . . . 104
4.1.6 More Cyclicities . . . 105
4.2 The α-Tensor . . . 111
4.2.1 The Magnetic-Field Advection . . . 112
4.2.2 The Highly Anisotropic α-Effect . . . 116
4.2.3 The Magnetic Quenching of the α-Effect . . . 122
4.2.4 Weak-Compressible Turbulence . . . 125
4.3 Magnetic-Diffusivity Tensor and η-Quenching . . . 129
4.3.1 The Eddy Diffusivity Tensor . . . 129
4.3.2 Sunspot Decay . . . 133
4.4 Mean-Field Stellar Dynamo Models . . . 135
4.4.1 The α^2-Dynamo . . . 137
4.4.2 The $\alpha\Omega$-Dynamo for Slow Rotation . . . 142
4.4.3 Meridional Flow Influence . . . 146
4.5 The Solar Dynamo . . . 146
4.5.1 The Overshoot Dynamo . . . 146
4.5.2 The Advection-Dominated Dynamo . . . 149

4.6 Dynamos with Random α 152
  4.6.1 A Turbulence Model 155
  4.6.2 Dynamo Models with Fluctuating α-Effect 155
4.7 Nonlinear Dynamo Models 158
  4.7.1 Malkus-Proctor Mechanism 159
  4.7.2 α-Quenching 160
  4.7.3 Magnetic Saturation by Turbulent Pumping 162
  4.7.4 η-Quenching 163
4.8 Λ-Quenching and Maunder Minimum 163

5 The Magnetorotational Instability (MRI) 167

5.1 Star Formation 167
  5.1.1 Molecular Clouds 167
  5.1.2 The Angular Momentum Problem 171
  5.1.3 Turbulence and Planet Formation 174
5.2 Stability of Differential Rotation in Hydrodynamics 174
  5.2.1 Combined Stability Conditions 176
  5.2.2 Sufficient Condition for Stability 178
  5.2.3 Numerical Simulations 179
  5.2.4 Vertical Shear 179
5.3 Stability of Differential Rotation in Hydromagnetics 181
  5.3.1 Ideal MHD 182
  5.3.2 Baroclinic Instability 183
5.4 Stability of Differential Rotation with Strong Hall Effect 184
  5.4.1 Criteria of Instability of Protostellar Disks 184
  5.4.2 Growth Rates 186
5.5 Global Models 187
  5.5.1 A Spherical Model with Shear 187
  5.5.2 A Global Disk Model 192
5.6 MRI of Differential Stellar Rotation 194
  5.6.1 T Tauri Stars (TTS) 194
  5.6.2 The Ap-Star Magnetism 195
  5.6.3 Decay of Differential Rotation 198
5.7 Circulation-Driven Stellar Dynamos 199
  5.7.1 The Gailitis Dynamo 200
  5.7.2 Meridional Circulation plus Shear 201
5.8 MRI in Kepler Disks 201
  5.8.1 The Shearing Box Model 202
  5.8.2 A Global Disk Dynamo? 205
5.9 Accretion-Disk Dynamo and Jet-Launching Theory 207
  5.9.1 Accretion-Disk Dynamo Models 207
  5.9.2 Jet-Launching 209
  5.9.3 Accretion-Disk Outflows 212
  5.9.4 Disk-Dynamo Interaction 213

6 The Galactic Dynamo **215**
6.1 Magnetic Fields of Galaxies . . . 215
6.1.1 Field Strength . . . 218
6.1.2 Pitch Angles . . . 218
6.1.3 Axisymmetry . . . 220
6.1.4 Two Exceptions: Magnetic Torus and Vertical Halo Fields . . . 221
6.1.5 The Disk Geometry . . . 223
6.2 Nonlinear Winding and the Seed Field Problem . . . 224
6.2.1 Uniform Initial Field . . . 224
6.2.2 Seed Field Amplitude and Geometry . . . 226
6.3 Interstellar Turbulence . . . 228
6.3.1 The Advection Problem . . . 228
6.3.2 Hydrostatic Equilibrium and Interstellar Turbulence . . . 229
6.4 From Spheres to Disks . . . 232
6.4.1 1D Dynamo Waves . . . 233
6.4.2 Oscillatory vs. Steady Solutions . . . 235
6.5 Linear 3D Models . . . 236
6.6 The Nonlinear Galactic Dynamo with Uniform Density . . . 238
6.6.1 The Model . . . 238
6.6.2 The Influences of Geometry and Turbulence Field . . . 240
6.7 Density Wave Theory and Swing Excitation . . . 242
6.7.1 Density Wave Theory . . . 242
6.7.2 The Short-Wave Approximation . . . 243
6.7.3 Swing Excitation in Magnetic Spirals . . . 244
6.7.4 Nonlocal Density Wave Theory in Kepler Disks . . . 248
6.8 Mean-Field Dynamos with Strong Halo Turbulence . . . 251
6.8.1 Nonlinear 2D Dynamo Model with Magnetic Supported Vertical Stratification . . . 252
6.8.2 Nonlinear 3D Dynamo Models for Spiral Galaxies . . . 253
6.9 New Simulations: Macroscale and Microscale . . . 255
6.9.1 Particle-Hydrodynamics for the Macroscale . . . 256
6.9.2 MHD for the Microscale . . . 258
6.10 MRI in Galaxies . . . 261

7 Neutron Star Magnetism **265**
7.1 Introduction . . . 265
7.2 Equations . . . 266
7.3 Without Stratification . . . 270
7.4 With Stratification . . . 271
7.5 Magnetic-Dominated Heat Transport . . . 276
7.6 White Dwarfs . . . 278

8 The Magnetic Taylor–Couette Flow **281**
8.1 History . . . 281
8.2 The Equations . . . 284

- 8.3 Results without Hall Effect 286
  - 8.3.1 Subcritical Excitation for Large Pm 286
  - 8.3.2 The Rayleigh Line ($a = 0$) and Beyond 286
  - 8.3.3 Excitation of Nonaxisymmetric or Oscillatory Modes 290
  - 8.3.4 Wave Number and Drift Frequencies 291
- 8.4 Results with Hall Effect 292
  - 8.4.1 Hall Effect with Positive Shear 293
  - 8.4.2 Hall Effect with Negative Shear 294
  - 8.4.3 A Hall-Driven Disk-Dynamo? 295
- 8.5 Endplate effects 297
- 8.6 Water Experiments 298
- 8.7 Taylor–Couette Flow as Kinematic Dynamo 299

9 Bibliography **301**

Index **327**

Preface

It is now 85 years since Sir Joseph Larmor first proposed that electromagnetic induction might be the origin of the Sun's magnetic field (Larmor 1919). Today this so-called dynamo effect is believed to generate the magnetic fields of not only the Sun and other stars, but also the Earth and other planets, and even entire galaxies. Indeed, most of the objects in the Universe have associated magnetic fields, and most of these are believed to be due to dynamo action. Quite an impressive record for a paper that is only two pages long, and was written before galaxies other than the Milky Way were even known!

However, despite this impressive list of objects to which Larmor's idea has now been applied, in no case can we say that we fully understand all the details. Enormous progress has undoubtedly been made, particularly with the huge increase in computational resources available in recent decades, but considerable progress remains to be made before we can say that we understand the magnetic fields even just of the Sun or the Earth, let alone some of the more exotic objects to which dynamo theory has been applied.

Our goal in writing this book was therefore to present an overview of these various applications of dynamo theory, and in each case discuss what is known so far, but also what is still unknown. We specifically include both geophysical and astrophysical applications. There is an unfortunate tendency in the literature to regard stellar and planetary magnetic fields as somehow quite distinct. How this state of affairs came about is not clear, although it is most likely simply due to the fact that geophysics and astrophysics are traditionally separate departments. Regardless of its cause, it is certainly regrettable. We believe the two have enough in common that researchers in either field would benefit from a certain familiarity with the other area as well. It is our hope therefore that this book will not only be of interest to workers in both fields, but that they will find new ideas on the 'other side of the fence' to stimulate further developments on their side (and maybe thereby help tear down the fence entirely).

Much of the final writing was done in the 2nd half of 2003. Without the technical support of Mrs. A. Trettin and M. Schultz from the Astrophysical Institute Potsdam it would not have been possible to finish the work in time. We gratefully acknowledge their kind and constant help. Many thanks also go to Axel Brandenburg, Detlef Elstner, and Manfred Schüssler – to name only three of the vast dynamo community – for their indispensable suggestions and never-ending discussions.

Potsdam and Glasgow, 2004

1 Introduction

Magnetism is one of the most pervasive features of the Universe, with planets, stars and entire galaxies all having associated magnetic fields. All of these fields are generated by the motion of electrically conducting fluids, via the so-called dynamo effect. The basics of this effect are almost trivial to explain: moving an electrical conductor through a magnetic field induces an emf (Faraday's law), which generates electric currents (Ohm's law), which have associated magnetic fields (Ampere's law). The hope is then that with the right combination of flows and fields the induced field will reinforce the original field, leading to (exponential) field amplification.

Of course, the details are rather more complicated than that. The basic physical principles may date back to the 19^{th} century, but it was not until the middle of the 20^{th} century that Backus (1958) and Herzenberg (1958) rigorously proved that this process can actually work, that is, that it is possible to find 'the right combination of flows and fields.' And even then their flows were carefully chosen to make the problem mathematically tractable, rather than physically realistic. For most of these magnetized objects mentioned above it is thus only now, at the start of the 21^{st} century, that we are beginning to unravel the details of how their fields are generated.

The purpose of this book is to examine some of this work. We will not discuss the basics of dynamo theory as such; for that we refer to the books by Roberts (1967), Moffatt (1978) and Krause & Rädler (1980), which are still highly relevant today. Instead, we wish to focus on some of the details specific to each particular application, and explore some of the similarities and differences.

For example, what is the mechanism that drives the fluid flow in the first place, and hence ultimately supplies the energy for the field? In planets and stars it turns out to be convection, whereas in accretion disks it is the differential rotation in the underlying Keplerian motion. In galaxies it could be either the differential rotation, or supernova-induced turbulence, or some combination of the two.

Next, what is the mechanism that ultimately equilibrates the field, and at what amplitude? The basic physics is again quite straightforward; what equilibrates the field is the Lorentz force in the momentum equation, which alters the flow, at least just enough to stop it amplifying the field any further. But again, the details are considerably more complicated, and again differ widely between different objects.

Another interesting question to ask concerns the nature of the initial field. In particular, do we need to worry about this at all, or can we always count on some more or less arbitrarily small stray field to start this dynamo process off? And yet again, the answer is very different for different objects. For planets we do not need to consider the initial field, since both the advective and diffusive timescales are so short compared with the age that any memory of the precise initial conditions is lost very quickly. In contrast, in stars the advective timescale is still short, but the diffusive timescale is long, so so-called fossil fields may play a role in certain aspects of stellar magnetism. And finally, in galaxies even the advective timescale is relatively long compared with the age, so there we do need to consider the initial field.

Accretion disks provide another interesting twist to this question of whether we need to consider the initial condition. The issue here is not whether the dynamo acts on a timescale short or long compared with the age, but whether it can act at all if the field is too weak. In particular, this Keplerian differential rotation by itself cannot act as a dynamo, so something must be perturbing it. It is believed that this perturbation is due to the Lorentz force itself, via the so-called magnetorotational instability. In other words, the dynamo can only operate at finite field strengths, but cannot amplify an infinitesimal seed field. One must therefore consider whether sufficiently strong seed fields are available in these systems.

Accretion disks also illustrate the effect that an object's magnetic field may have on its entire structure and evolution. As we saw above, the magnetic field always affects the flow, and hence the internal structure, in some way, but in accretion disks the effect is particularly dramatic. It turns out that the transport of angular momentum outward – which of course determines the rate at which mass moves inward – is dominated by the Lorentz force. Something as fundamental as the collapse of a gas cloud into a proto-stellar disk and ultimately into a star is thus magnetically controlled. That is, magnetism is not only pervasive throughout the Universe, it is also a crucial ingredient in forming stars, the most common objects found within it.

We hope therefore that this book will be of interest not just to geophysicists and astrophysicists, but to general physicists as well. The general outline is as follows: Chapter 2 presents the theory of planetary dynamos. Chapters 3 and 4 deal with stellar dynamos. We consider only those aspects of stellar hydrodynamics and magnetohydrodynamics that are relevant to the basic dynamo process; see for example Mestel (1999) for other aspects such as magnetic braking. Chapter 5 discusses this magnetorotational instability in Keplerian disks. Chapter 6 considers galaxies, in which the magnetorotational instability may also play a role. Chapter 7, concerning neutron stars, is slightly different from the others. In particular, whereas the other chapters deal with the origin of the particular body's magnetic field, in Chapt. 7 we take the neutron star's initial field as given, and consider the details of its subsequent decay. We consider only the field in the neutron star itself though; see Mestel (1999) for the physics of pulsar magnetospheres. Lastly, Chapt. 8 discusses the magnetorotational instability in cylindrical Couette flow. This geometry is not only particularly amenable to theoretical analysis, it is also the basis of a planned experiment. However, we also point out some of the difficulties one would have to overcome in any real cylinder, which would necessarily be bounded in z.

Where relevant, individual chapters of course refer to one another, to point out the various similarities and differences. However, most chapters can also be read more or less independently of the others. Most chapters also present both numerical as well as analytic/asymptotic results, and as much as possible we try to connect the two, showing how they mutually support each other. Finally, we discuss fields occurring on lengthscales from kilometers to megaparsecs, and ranging from 10^{-20} to 10^{15} G – truly the magnetic Universe.

2 Earth and Planets

2.1 Observational Overview

We begin with a brief overview of the field as it is today, as well as how it has varied in the past. See also Merrill, McElhinny & McFadden (1998) or Dormy, Valet & Courtillot (2000) for considerably more detailed accounts of the observational data, or Hollerbach (2003) for a discussion of the theoretical origin of some of the timescales on which the field varies.

Figure 2.1 shows the Earth's magnetic field as it exists today. The two most prominent features, are (i) that it is predominantly dipolar, and (ii) that this dipole is quite closely aligned with the rotation axis, with a tilt of only 11°. We would expect a successful geodynamo theory to be able to explain both of these features, as well as others, of course, such as why the field has the particular amplitude that it does.

Turning to the dipole dominance first, we begin by noting that much of this is an artifact of where we have chosen to observe the field, namely at the surface of the Earth. As we will see later, the field is actually created deep within the Earth, in the molten iron core, with the overlying mantle playing no direct role. Because the mantle (consisting of rock) is largely insulating, we can project the field back down to the core-mantle boundary (CMB). All components of the field are amplified when we do this, but the nondipole components are also amplified relative to the dipole, since they drop off faster with increasing radius, and hence increase faster when projected back inward again. Figure 2.1 also shows the resulting field at the CMB, which we note is indeed considerably less dipole dominated.

Figure 2.2 shows the corresponding power spectra, both at the surface and the CMB. The enhancement of the higher harmonics at the CMB is clearly visible. The other important point to note is that whereas the surface spectrum has been plotted to spherical harmonic degree $l = 25$, only the modes up to $l = 12$ have been projected inward to obtain the CMB spectrum. The reason for this is the sharp break observed in the surface spectrum at $l \approx 13$, with the power dropping off quite steeply up to there, but not at all thereafter. The generally accepted interpretation of this phenomenon is that this power in the $l > 12$ modes is due to crustal magnetism. These modes cannot therefore be projected back down to the CMB to obtain the spectrum there. Figure 2.1 (bottom) is thus not the true field at the CMB, but merely a filtered version of it, with all of the smallest scales having been filtered out. That is, the true field could very well exhibit highly localized features like sunspots, but this crustal contamination prevents us from ever observing them.

Turning next to the alignment of the dipole with the rotation axis, the probability that two vectors chosen at random would be aligned to within 11° or better is less than 2%. It seems more plausible therefore that this degree of alignment is not a coincidence, but instead reflects

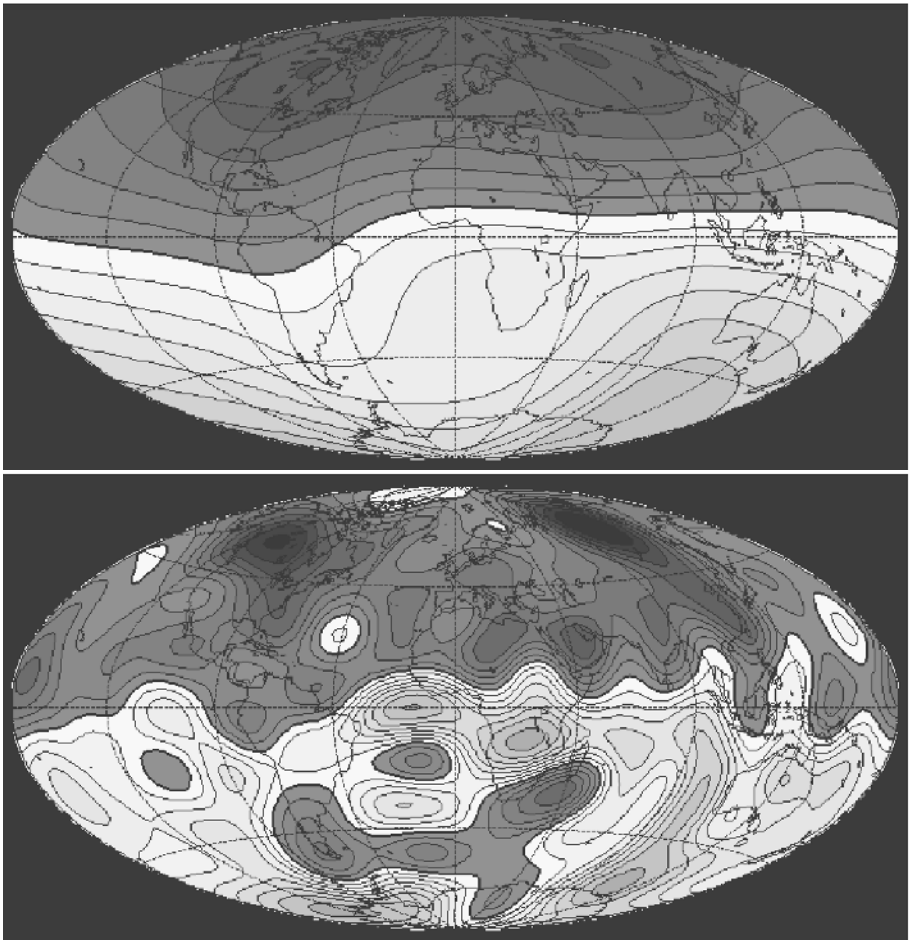

Figure 2.1: The radial component of the Earth's field at the surface (*top*), and projected down to the core-mantle boundary (*bottom*). Courtesy A. Jackson.

some controlling influence of rotation on the geodynamo. And indeed, we will see below that rotation exerts powerful constraints on the field (although it is not immediately obvious why this influence should lead to an alignment of the field with the rotation axis).

2.1.1 Reversals

Figure 2.1 shows the field as it is today. The field is not static, however, varying instead on timescales as short as minutes or even seconds, and as long as tens or even hundreds of millions of years. Of all of these variations, the most dramatic are reversals, in which the entire

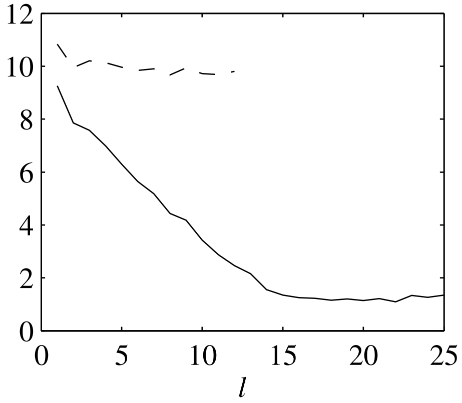

Figure 2.2: Power spectra of the Earth's field at the surface (solid) and the core-mantle boundary (dashed).

field switches polarity. See, for example, Gubbins (1994) or Merrill & McFadden (1999) for reviews devoted specifically to reversals.

Figure 2.3 shows the reversal record for the past 40 million years. The field is seen to reverse on the average every few hundred thousand years, but with considerable variation about that average. These relatively infrequent and irregular reversals of the Earth's field are thus very different from the comparatively regular, and much faster solar cycle.

Unlike the interval between reversals, the time it takes for the reversal itself seems to be a relatively constant 5–10 thousand years. During the reversal, the field is weaker, and considerably more complicated and less dipolar than in Fig. 2.1. Between reversals, however, it is generally similar to today's field, in terms of both field strength, dipole-dominated structure, and alignment with the rotation axis. This last point, of course, provides additional evidence that this alignment is not due to chance, but instead reflects the powerful influence of rotation.

Finally, the average interval between reversals itself varies on timescales of tens and hundreds of millions of years. For example, there were no reversals at all between 83 and 121 million years ago. Because these timescales are so much longer than any of the timescales 'naturally' present in the core, it is generally believed that this very long-term behavior is

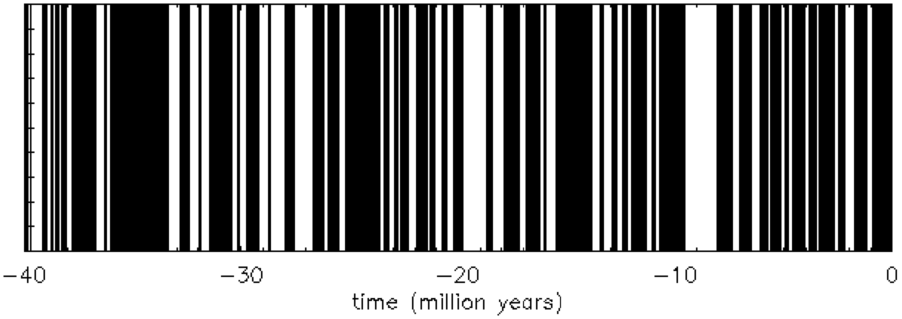

Figure 2.3: The reversal sequence for the past 40 million years. Courtesy A. Witt.

of external origin. In particular, the timescale of mantle convection is precisely tens to hundreds of millions of years (e.g. Schubert, Turcotte & Olson 2001), so the thermal boundary conditions that the mantle imposes on the core will also evolve on these timescales. See, for example, Glatzmaier et al. (1999) for a series of numerical simulations in which different thermal boundary conditions did indeed lead to different reversal rates.

2.1.2 Other Time-Variability

As noted above, reversals are only the most dramatic variation in time found in the field. Between reversals the field varies as well, again with a broad range of timescales and amplitudes. Most familiar is the so-called secular variation, in which some of the nondipolar features fluctuate on timescales of decades to centuries. See for example Bloxham, Gubbins & Jackson (1989) or Jackson, Jonkers & Walker (2000) for summaries of the secular variation observed in the historical record. Intermediate between secular variation and reversals are also excursions, in which the field varies by considerably more than the usual secular variation, but does not actually reverse either. Excursions are around ten times more numerous than reversals, but of similar duration.

At the other extreme, the shortest timescales that can be observed within the core are geomagnetic jerks, in which the usual secular variation changes abruptly – and over the whole Earth – within a single year. Around three or four such events have been recorded in the past century (LeHuy et al. 1998). Note also that these events may well occur even faster than the one-year timescale on which they are recorded at the surface; the mantle is not a perfect insulator, and its weak conductivity effectively screens out any variations in the core occurring on timescales faster than a year. (For this reason also the variations in the field occurring on timescales as short as minutes or seconds must be of external origin, i.e. magnetospheric or ionospheric.)

2.2 Basic Equations and Parameters

The Earth's interior consists of a series of concentric spherical shells nested rather like the layers of an onion. The most fundamental division is that between the core and the mantle. The core, consisting mostly of iron, extends from the center out to a radius of 3480 km; the mantle, consisting of rock, extends from there essentially to the Earth's surface at $R = 6370$ km. In fact, the top 30 km or so are sufficiently different in their material properties (brittle rather than plastic, due to the much lower pressures and temperatures) that they are further distinguished from the mantle, and referred to as the crust. However, as important as the distinction between crust and mantle may be for phenomena such as plate tectonics, volcanism, earthquakes, etc. (e.g. Schubert, Turcotte & Olson 2001), the fact that both consist largely of rock, which is a very poor electrical conductor, immediately suggests that we must seek the origin of the Earth's magnetic field elsewhere, namely in the core. From the point of view of geodynamo theory, the mantle and crust are merely 3000 km of 'inconvenience' blocking what we would really like to observe (see Sect. 2.9.1 though).

Turning to the core then, it is further divided into a solid inner core of radius $R_{\rm in} = 1220$ km, and a fluid outer core of radius $R_{\rm out} = 3480$ km. The inner core was first detected seismically in 1936. See for example Gubbins (1997) for a review devoted specifically to the inner core. Further seismic studies show it to be sufficiently rigid to sustain shear waves (although it may actually be a so-called mushy layer right to the center, see, for example, Fearn, Loper & Roberts 1981). In contrast, the outer core is as fluid as water, with a viscosity of around $10^{-2}\ \mathrm{cm}^2/\mathrm{s}$ (Poirier 1994, De Wijs et al. 1998).

Further seismic (and other) studies also indicate that the density of the outer core increases from around 9.9 g/cm^3 at R_{out} to 12.2 g/cm^3 at R_{in}, at which point there is an abrupt jump to 12.8 g/cm^3 in the inner core. This value for the inner core is consistent with the density of around 90% pure iron (at the corresponding pressures and temperatures). The 5% jump across the inner core boundary cannot be explained purely by the phase transition from solid to liquid though; the outer core must contain perhaps 15–20% lighter impurities (with S, Si and O being the most likely candidates, e.g. Alfè et al. 2002).

With this basic structure of the core in place, we can begin to understand the dynamics that ultimately lead to the emergence of the Earth's magnetic field. As the Earth slowly cooles over billions of years, the core gradually solidifies, that is, the inner core grows. (The reason it solidifies from the center, even though it is hottest there, is due to the influence of the pressure on the melting temperature.) As it freezes, most of the impurities get rejected back into the fluid (just as freezing salt water will reject most of the salt, leaving relatively fresh water in the ice). As Braginsky (1963) was the first to point out, there are then two sources of buoyancy at the inner core boundary, namely that due to these light impurities being rejected back into the fluid, and that due to the release of latent heat from the freezing process itself. Additionally, of course, there is the usual source of (negative) buoyancy at the outer core boundary, namely that due to the fluid there losing heat to the mantle and hence becoming denser. It is these various sources of buoyancy that drive the convection that ultimately generates the magnetic field.

Incidentally, note also that we can extrapolate this cooling process backward to estimate when the inner core first formed. Buffett et al. (1992, 1996) considered detailed models of the thermal evolution of the core, and concluded that the inner core started to solidify around two billion years ago, and also that at present thermal and compositional effects are of comparable importance in powering the geodynamo. The precise age of the inner core continues to be debated though; recent estimates vary between one and three billion years (Labrosse & Macouin 2003 and Gubbins et al. 2003, respectively). It is quite interesting then that there is paleomagnetic evidence for the existence of a field as long ago as 3.5 billion years (McElhinny & Senanayake 1980). That is, there was most likely a dynamo even before the inner core formed, and hence before these various buoyancy sources at the inner core boundary became available.

2.2.1 Anelastic and Boussinesq Equations

Having discussed in qualitative terms the dynamics that lead to core convection and ultimately a magnetic field, our next task is to write down the specific equations. The most detailed analysis of these equations, and the various approximations one can make, is by Braginsky & Roberts (1995); here we merely summarize some of their findings. Linearizing the thermodynamics about an adiabatic reference state with density ρ_{a}, the momentum equation they ultimately end up with is

$$\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\nabla P + C\,\boldsymbol{g}_{\mathrm{a}} + \frac{1}{\mu_0 \rho_{\mathrm{a}}}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} + \nu \Delta \boldsymbol{u}. \tag{2.1}$$

The so-called co-density C is given by $C = -\alpha_S S - \alpha_\xi \xi$, where S and ξ are the entropy and composition perturbations, respectively, and

$$\alpha_S = -\frac{1}{\rho}\frac{\partial \rho}{\partial S}, \qquad \alpha_\xi = -\frac{1}{\rho}\frac{\partial \rho}{\partial \xi} \tag{2.2}$$

determine how variations in S and ξ translate into relative density variations (this means of course that we also need a suitable equation of state $\rho = \rho(P, S, \xi)$ to determine these coefficients). One other point worth stating explicitly is that the gravity $\boldsymbol{g}_a$ appearing in Eq. (2.1) is that due to the adiabatic reference state only (hence the subscript); Braginsky & Roberts show that the self-gravity induced by the convective density perturbations themselves can be incorporated into the reduced pressure P. This is obviously a considerable simplification, as $\boldsymbol{g}_a$ is then known (varying roughly as $-\boldsymbol{r}$), rather than having to be solved for at every timestep of the other equations.

The continuity equation associated with Eq. (2.1) is $\nabla \cdot (\rho_a \boldsymbol{u}) = 0$, that is, rather than considering the fully compressible continuity equation we have made the anelastic approximation, and thereby filtered out sound waves[1]. The timescale for sound waves to traverse the entire core is around ten minutes, which is so much faster than any of the other dynamics we will be interested in that filtering them out completely is a reasonable approximation. (Note that this is very different from many astrophysical situations, where the Alfvén speed is often comparable with or even greater than the sound speed.) Finally, with the usual advection-diffusion equations for S and ξ, and of course the induction equation for $\boldsymbol{B}$, we have a complete set of equations that we should be able to timestep for S, ξ, $\boldsymbol{u}$ and $\boldsymbol{B}$.

As we will see in the remainder of this chapter, making actual progress with these equations is a formidable undertaking, primarily because some of the nondimensional parameters take on such extreme values. Many models therefore simplify these equations further still, in a variety of ways. For example, even though we saw that compositional and thermal sources of buoyancy are both important, most models neglect compositional effects, and consider thermal convection only. Given how different thermal and compositional convection can be (e.g. Worster 2000), this probably does affect at least the details of the solutions; neglecting compositional effects certainly cannot be rigorously justified. The only 'justification' one can offer is that we cannot even get the details of thermal convection right, so there is little point in worrying about the precise differences between thermal and compositional convection. For example, the compositional diffusivity is several orders of magnitude smaller than the thermal (e.g. Roberts & Glatzmaier 2000), but even the thermal diffusivity is orders of magnitude smaller than anything that any numerical model can cope with. So if both have to be increased to artificially large values, much of the difference between the two effects is also likely to disappear (although there are other differences as well, such as very different boundary conditions).

Another common simplification is to make the Boussinesq approximation, in which density variations are neglected everywhere except in the buoyancy term itself. That is, we replace the adiabatic density profile ρ_a by a constant, ρ_0. The Boussinesq approximation also cannot be rigorously justified (once again, see Braginsky & Roberts 1995). In particular, the variations in ρ_a that are being neglected are orders of magnitude greater than the convective density perturbations that are being included (very much unlike laboratory convection). However, given that the density contrast across the outer core is only $\sim 20\%$ (as we saw above), it seems likely that Boussinesq and anelastic results also will not differ by too much. There certainly do not appear to be any fundamental differences between the two.

---

[1] see Lantz & Fan (1999) for a recent discussion of the anelastic approximation

We are therefore left with

$$\begin{aligned} &\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + 2\boldsymbol{\Omega} \times \boldsymbol{u} = -\nabla P - \alpha T\, \boldsymbol{g} + \frac{1}{\mu_0 \rho_0}(\nabla \times \boldsymbol{B}) \times \boldsymbol{B} + \nu \Delta \boldsymbol{u}, \\ &\frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{B}) + \eta \Delta \boldsymbol{B}, \qquad \left(\frac{\partial}{\partial t} + \boldsymbol{u} \cdot \nabla\right) T = \chi \Delta T, \end{aligned} \tag{2.3}$$

with $\nabla \cdot \boldsymbol{u} = 0$ and $\nabla \cdot \boldsymbol{B}$ as the simplest set of equations still 'reasonably' consistent with the original physics. (Note that when we neglect compositional effects, the entropy S can be replaced by the temperature T, with α then being the usual coefficient of thermal expansion.) These are the equations we will focus on, although in Sect. 2.4.7 we will return briefly to the original anelastic equation.

2.2.2 Nondimensionalization

Having settled on the equations, the next point we want to consider is how to nondimensionalize them, and what that might already tell us about the dynamics (that is, which terms are small or large, etc.). For a lengthscale, the obvious choice is the outer core radius $R_{\mathrm{out}} = 3480$ km (many numerical models actually take $R_{\mathrm{out}} - R_{\mathrm{in}}$, but such minor details need not concern us here). The timescale is not quite so obvious, but a natural choice is the magnetic diffusive timescale R_{out}^2/η. Using the value $\eta \approx 2 \cdot 10^4$ cm^2/s appropriate for molten iron (Poirier 1994), this comes out to around 200,000 yr. (Incidentally, we see therefore that the range of timescales observed in the field varies from much shorter to much longer than this diffusive timescale.)

The fluid flow is then scaled by $\mathrm{length/time} = \eta/R_{\mathrm{out}} = O(10^{-6})$ m/s, so the advective and diffusive terms in the induction equation are (formally) comparable. Note though that the actual magnitude of the flow can only emerge from a full solution of the problem, and may turn out to be different from this value. Indeed, if the time evolution of the field at the core-mantle boundary is used to estimate the flow, one obtains magnitudes on the order of $10^{-4} - 10^{-3}$ m/s (Bloxham & Jackson 1991). That is, we would expect $\boldsymbol{u}$ to equilibrate at 10^{2-3} rather than order 1. This value of a few hundred is then also the magnetic Reynolds number $\mathrm{Rm} = uR_{\mathrm{out}}/\eta$ in the core.

The magnetic field is scaled by $(\Omega \rho_0 \mu_0 \eta)^{1/2} \approx 10$ G, which ensures that the Coriolis and Lorentz forces in the momentum equation are formally comparable. This is believed to be the appropriate balance at which the field equilibrates, for reasons that will become clear later. It also compares rather well with the $\sim$ 3-G field observed at the CMB (particularly when we remember that the field deep within the core is likely to be at least somewhat stronger than right at the boundary). But once again, the actual magnitude of the field can only emerge from the complete solution. And as before with the magnitude of $\boldsymbol{u}$ giving us Rm, the magnitude of $\boldsymbol{B}$ (squared in this case) gives us the Elsasser number

$$\Lambda = \frac{B^2}{\varOmega \rho_0 \mu_0 \eta}. \tag{2.4}$$

We see therefore that Λ is 0.1 to perhaps 1 in the core.

Finally, the natural scale for the temperature is simply the temperature difference δT across the core. However, there is one very considerable difficulty with this, namely estimating what δT actually is. In particular, the dynamically relevant temperature difference is only what is left over after the adiabatic temperature difference has been subtracted out. This ends up being virtually everything though: of the more than 1000 K difference across the core, the super-adiabatic δT that actually drives convection amounts to a small fraction of 1 K. In other words, δT cannot be estimated by taking the known temperature difference and subtracting out the adiabat; the errors would overwhelm the signal. Instead, δT can only be inferred indirectly by energetic/thermodynamic considerations.

With these scalings, the nondimensionalized Boussinesq equations become

$$\begin{gathered}\mathrm{Ro}\,\frac{\mathrm{D}\boldsymbol{u}}{\mathrm{D}t} + 2\hat{\boldsymbol{e}}_z \times \boldsymbol{u} = -\nabla P + q\,\widehat{\mathrm{Ra}}\,T\,\boldsymbol{r} + (\nabla \times \boldsymbol{B}) \times \boldsymbol{B} + \mathrm{E}\Delta\boldsymbol{u}, \\ \frac{\partial \boldsymbol{B}}{\partial t} = \nabla \times (\boldsymbol{u} \times \boldsymbol{B}) + \Delta\boldsymbol{B}, \qquad \left(\frac{\partial}{\partial t} + \boldsymbol{u}\cdot\nabla\right)T = \mathrm{q}\Delta T.\end{gathered} \tag{2.5}$$

The nondimensional parameters appearing in these equations are, first, the (modified) Rayleigh number

$$\widehat{\mathrm{Ra}} = \frac{g_0 \alpha \delta T R_{\mathrm{out}}}{\varOmega \chi}, \tag{2.6}$$

where $g_0 = |\boldsymbol{g}(R_{\mathrm{out}})|$ (and by replacing $\boldsymbol{g}$ by $-\boldsymbol{r}$ in Eq. $(2.5)_1$ we are assuming for simplicity that gravity varies linearly with r). Note that this Rayleigh number measures the buoyancy force against the Coriolis force, rather than against the viscous force, as in classical Rayleigh–Benard convection. And once again, we remember that because of these uncertainties in δT, it is not clear just how large $\widehat{\mathrm{Ra}}$ is in the core. See, however, Gubbins (2001) for the latest estimates, and also Kono & Roberts (2001) for how $\widehat{\mathrm{Ra}}$ should even be defined when both thermal and compositional effects are important.

Next we have the Rossby number

$$\mathrm{Ro} = \frac{\eta}{\varOmega R_{\mathrm{out}}^2}, \tag{2.7}$$

measuring the ratio of the rotational timescale $\varOmega^{-1}$ (=$1/2\pi$ day) to the diffusive timescale R_{out}^2/η (=200,000 yr, as we saw above). That is, $\mathrm{Ro} = O(10^{-9})$. The Ekman number E (measuring viscous to Coriolis forces) and the Roberts number q

$$\mathrm{E} = \frac{\nu}{\varOmega R_{\mathrm{out}}^2}, \qquad \mathrm{q} = \frac{\chi}{\eta}, \tag{2.8}$$

(the latter measuring the ratio of thermal to magnetic duffusivity) come out to be $O(10^{-15})$ and $O(10^{-6})$, resp.

It is the extreme smallness of these three parameters that then makes the geodynamo equations so difficult. For example, if the advective term is at least as important as the diffusive term in Eq. $(2.5)_2$ (as we saw it is, and indeed must be to have any chance of achieving dynamo action), then in Eq. $(2.5)_3$ the advective term will dominate the diffusive term by many orders of magnitude, leading to extremely small lengthscales in T, which will certainly cause numerical difficulties, if nothing else. See also Christensen, Olson & Glatzmaier (1999) for further difficulties associated with the smallness of q.

These difficulties associated with q are usually 'solved' by invoking turbulent diffusivities, in which case all three diffusivities ν_{T}, η_{T} and χ_{T} will most likely be comparable, yielding $\mathrm{q_T} = O(1)$ – which is indeed the range used in virtually all numerical models. However, one has not really solved the problem thereby, merely deferred it to a proper investigation of this small-scale turbulence. See, for example, Braginsky & Meytlis (1990), St. Pierre (1996), Davidson & Siso-Nadal (2002) and Buffett (2003) for models that begin to explore the precise nature of such rotating MHD turbulence.

And finally, even if an appeal to turbulent diffusivities solves (or rather ignores) the difficulties associated with q, those associated with Ro and E remain. In particular, η_{T} (and hence also ν_{T}) cannot be increased much beyond 100 $\mathrm{m^2/s}$, otherwise the field would simply decay faster than it can be sustained. This means though that even $\mathrm{Ro_T}$ and $\mathrm{E_T}$ are at most 10^{-7} – which is still several orders of magnitude smaller than most numerical models can cope with. Much of the remainder of this chapter will be devoted to discovering just why small Ro and E should pose such problems.

But first, there is one more general feature of Eqs. (2.5) worth mentioning, namely the associated energy equation. If we add the dot products of Eq. $(2.5)_1$ with $\boldsymbol{u}$ and Eq. $(2.5)_2$ with $\boldsymbol{B}$, after a little algebra we obtain the global energy balance

$$\frac{\partial}{\partial t}\frac{1}{2}\int\left(|\boldsymbol{B}|^2 + \mathrm{Ro}\,|\boldsymbol{u}|^2\right)\mathrm{d}V$$

$$= \mathrm{q}\,\widehat{\mathrm{Ra}}\int u_r T r\,\mathrm{d}V - \int\left(|\boldsymbol{\nabla}\times\boldsymbol{B}|^2 + \mathrm{E}|\boldsymbol{\nabla}\times\boldsymbol{u}|^2\right)\mathrm{d}V. \tag{2.9}$$

The point we wish to focus on here is not so much the right-hand side (that is, how the energy changes), but rather the left, what the energy is in the first place. In particular, we recognize that if our nondimensionalization is correct, so that $\boldsymbol{u}$ and $\boldsymbol{B}$ do indeed equilibrate at roughly $O(1)$ values, then the magnetic energy will be several orders of magnitude *greater* than the kinetic. And because Ro is *so* small, this remains true even if $\boldsymbol{u}$ equilibrates at $O(10^3)$, as we saw above that it does. This is in sharp contrast to most astrophysical systems, where the magnetic energy is typically orders of magnitude smaller, or at best reaches equipartition.

Of course, if we included the energy stored in the Earth's rotation, we would be back in the astrophysically more familiar situation where the kinetic energy dominates by far. The rotational energy is not available though, since angular momentum must be conserved, so only deviations from solid-body rotation could be converted into magnetic (or other) forms of energy. And here again we see an enormous difference between the Earth and the Sun, for example; whereas in the Sun the differential rotation is a significant fraction of the overall rotation ($\sim$28%), in the Earth it is almost infinitesimal ($<$ 0.01%).

2.3 Magnetoconvection

Rotating, magnetic convection is a complicated process. Following Chandrasekhar (1961), let us therefore begin with classical Rayleigh–Benard convection, and first consider how rotation and magnetism separately alter the dynamics. Then we will explore how they act together, and finally what implications that might have for planetary dynamos, where the magnetic field is created by the convection itself, rather than being externally imposed.

Consider an infinite plane layer, heated from below and cooled from above. Additionally, there is an overall rotation $\Omega\hat{\boldsymbol{e}}_z$, and an externally imposed magnetic field $B_0\hat{\boldsymbol{e}}_z$. Linearizing about this basic state, the perturbation equations become

$$\begin{aligned}
&\frac{\partial \boldsymbol{u}}{\partial t} + 2\mathrm{E}^{-1}\hat{\boldsymbol{e}}_z \times \boldsymbol{u} = -\nabla P + \Delta\boldsymbol{u} + \mathrm{Ra}\,\mathrm{Pr}^{-1}\,T\,\hat{\boldsymbol{e}}_z + \mathrm{Ha}^2\,\mathrm{Pm}^{-1}\,(\nabla\times\boldsymbol{b})\times\hat{\boldsymbol{e}}_z \\
&\frac{\partial \boldsymbol{b}}{\partial t} = \nabla\times(\boldsymbol{u}\times\hat{\boldsymbol{e}}_z) + \mathrm{Pm}^{-1}\Delta\boldsymbol{b}, \qquad \frac{\partial T}{\partial t} - \boldsymbol{u}\cdot\hat{\boldsymbol{e}}_z = \mathrm{Pr}^{-1}\Delta T,
\end{aligned} \tag{2.10}$$

where length has been nondimensionalized by the layer thickness d, time by d^2/ν, $\boldsymbol{u}$ by ν/d, $\boldsymbol{b}$ by the imposed field B_0, and T by the imposed temperature difference δT. The nondimensional parameters are then the usual two Prandtl numbers $\mathrm{Pr} = \nu/\chi$ and $\mathrm{Pm} = \nu/\eta$, the Rayleigh number

$$\mathrm{Ra} = \frac{g\alpha\delta T d^3}{\nu\chi}, \tag{2.11}$$

measuring the thermal forcing, the (inverse) Ekman number

$$\mathrm{E}^{-1} = \frac{\Omega d^2}{\nu}, \tag{2.12}$$

measuring the rotation, and finally the Hartmann number

$$\mathrm{Ha} = \frac{B_0 d}{\sqrt{\mu_0\rho\nu\eta}}, \tag{2.13}$$

measuring the imposed magnetic field. Note also that the details of the nondimensionalization here – and hence the nondimensional parameters that arise – are different from those in Sect. 2.2.2. The reason for this is that here we want to start with classical Rayleigh–Benard convection, and only then add in rotation and magnetism, and study their effects. We must therefore also start with the classical nondimensionalization, so, for example, the usual Rayleigh number measuring buoyancy against viscosity, rather than against the Coriolis force, as in Eq. (2.6). Later on we will 'translate' the insight gained here into the geophysically more relevant parameters introduced in Sect. 2.2.2.

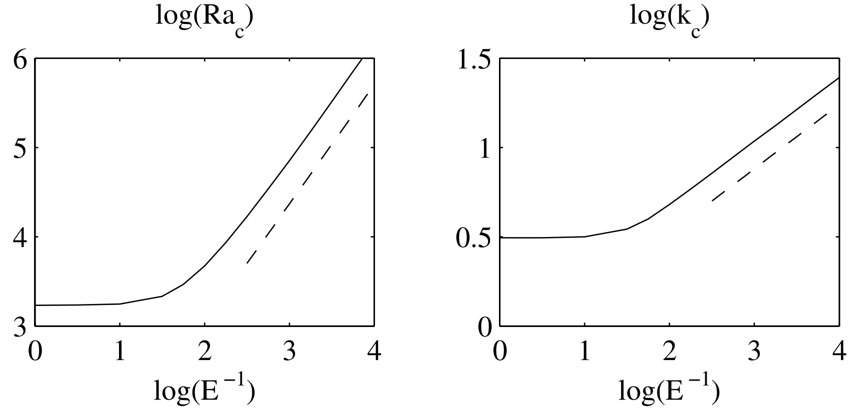

Figure 2.4: The influence of rotation without magnetism. *Left*: $\mathrm{Ra_c}$ as a function of E^{-1}. *Right*: k_c as a function of E^{-1}. The dashed lines have slopes 4/3 and 1/3, respectively, and indicate the scalings in the asymptotic limit.

Taking all quantities in Eq. (2.10) proportional to $\exp(\sigma t + \mathrm{i}k_x x + \mathrm{i}k_y y)$, we end up with the five equations

$$\begin{aligned}
&\sigma T = u_z + \mathrm{Pr}^{-1}\Delta T,\\
&\sigma \Delta u_z = -2\mathrm{E}^{-1}\omega_z' + \nabla^4 u_z - \mathrm{Ra}\,\mathrm{Pr}^{-1}\,k^2 T + \mathrm{Ha}^2\,\mathrm{Pm}^{-1}\,\Delta b_z',\\
&\sigma \omega_z = 2\mathrm{E}^{-1}u_z' + \Delta\omega_z + \mathrm{Ha}^2\,\mathrm{Pm}^{-1} j_z',\\
&\sigma b_z = u_z' + \mathrm{Pm}^{-1}\Delta b_z,\\
&\sigma j_z = \omega_z' + \mathrm{Pm}^{-1}\Delta j_z,
\end{aligned} \tag{2.14}$$

where u_z and b_z are the z-components of $\boldsymbol{u}$ and $\boldsymbol{b}$, ω_z and j_z the z-components of $\nabla \times \boldsymbol{u}$ and $\nabla \times \boldsymbol{b}$, the primes denote differentiation with respect to z, and $k^2 = k_x^2 + k_y^2$. Together with the boundary conditions

$$T = 0, \qquad u_z = u_z' = 0, \qquad \omega_z = 0, \qquad b_z = \pm b_z'/k, \qquad j_z = 0, \tag{2.15}$$

at $z = \pm d/2$, corresponding to rigid boundaries and electrically insulating exteriors, this system forms a well-defined eigenvalue problem that can be solved (numerically) for σ for any set of values for k, Ra, E^{-1} and Ha. Just as in Rayleigh–Benard convection, we are interested in the particular values $\mathrm{Ra_c}$ (and corresponding k_c) for which we first obtain exponentially growing solutions, that is, modes with $\Re(\sigma) > 0$. In the absence of rotation and magnetism, this critical Rayleigh number for the onset of convection is 1708, with associated wave number $k_\mathrm{c} = 3.12$. We would like to discover then what effect nonzero E^{-1} and Ha have on this value, that is, whether rotation and magnetism help or hinder the onset of convection, and most importantly, how they interact with one another.

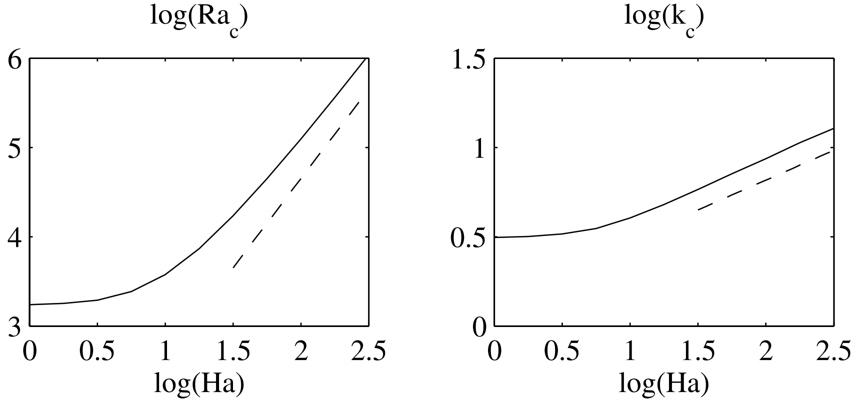

Figure 2.5: The influence of magnetism without rotation. *Left*: $\mathrm{Ra_c}$ as a function of Ha. *Right*: k_c as a function of Ha. The dashed lines have slopes 2 and 1/3, respectively, and indicate the scalings in the asymptotic limit.

2.3.1 Rotation or Magnetism Alone

Figure 2.4 (left) shows $\mathrm{Ra_c}$ as a function of E^{-1}, when $\mathrm{Ha} = 0$. We note that it increases monotonically, ultimately scaling as $\mathrm{E}^{-4/3}$ in the rapidly rotating limit. Rotation therefore suppresses convection. To see why, we turn to Eq. $(2.14)_3$, and note that for increasingly rapid rotation it becomes increasingly difficult to balance the term $2\mathrm{E}^{-1}u'_z$ against any of the others: the magnetic term is out, because we are taking $\mathrm{Ha} = 0$ here; the inertial term is also out, because these modes turn out to be steady, so $\sigma = 0$. If it were not for the viscous term, we would therefore have $u'_z = 0$ – which is of course just the familiar Taylor–Proudman theorem. Together with the boundary conditions, this would imply $u_z = 0$ though, eliminating the possibility of convective overturning. For convection to occur we must therefore break this Taylor–Proudman result, and as we just saw, the only way to achieve that is to balance the Coriolis term $2\mathrm{E}^{-1}u'_z$ against the viscous term $\Delta\omega_z$. This in turn implies that the convection must occur on very short horizontal lengthscales, since only then can the viscous term compete with this very large factor E^{-1} in the Coriolis term. Indeed, we see in Fig. 2.4 (right) that k_c also increases monotonically, ultimately scaling as $\mathrm{E}^{-1/3}$. Convection on ever shorter horizontal lengthscales is increasingly inefficient though, thereby explaining why $\mathrm{Ra_c}$ increases.

Figure 2.5 shows $\mathrm{Ra_c}$ and k_c as functions of Ha, when $\mathrm{E}^{-1} = 0$. Both again increase monotonically, with $\mathrm{Ra_c}$ scaling as Ha^2 in the strongly magnetic limit, and k_c scaling as $\mathrm{Ha}^{1/3}$. The reason why $\mathrm{Ra_c}$ increases is therefore just as before, because the convection is again being forced to occur on ever shorter horizontal lengthscales. This in turn is also easy to understand; the magnetic field tends to suppress all motion perpendicular to it, forcing the flow into tall, thin convection cells. More mathematically, the difficulty this time is in balancing the term $\mathrm{Ha}^2\,\mathrm{Pm}^{-1}\,\Delta b'_z$ in Eq. $(2.14)_2$. If b'_z were zero though, Eq. $(2.14)_4$ would again yield the unacceptable result $u'_z = 0$.

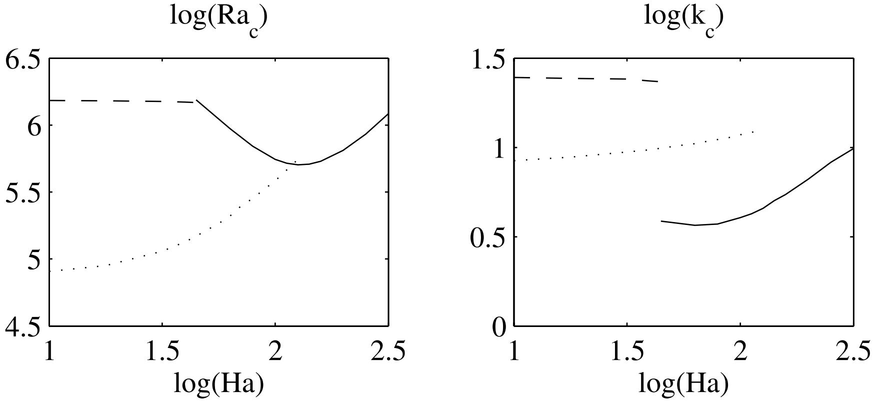

Figure 2.6: The effect of rotation and magnetism together. *Left*: $\mathrm{Ra_c}$ as a function of Ha. *Right*: k_c as a function of Ha. The dashed and solid lines denote the two different modes of convection discussed in the text. The dotted lines will be discussed in Sect. 2.3.4.

2.3.2 Rotation and Magnetism Together

We see therefore that acting alone, rotation and magnetism each suppress convection. When both act together though, the results could well be quite different. In particular, we note that then we can balance the Coriolis term $2\mathrm{E}^{-1}u_z'$ against the magnetic term $\mathrm{Ha}^2\,\mathrm{Pm}^{-1}j_z'$ in Eq. $(2.14)_3$, and similarly in Eq. $(2.14)_2$. That is, the mechanisms that forced the convection to adopt very short horizontal lengthscales in either of the previous two cases do not apply here. If convection can occur with $k_\mathrm{c} = O(1)$ though, $\mathrm{Ra_c}$ should also be much less than in either of the previous two cases.

Figure 2.6 shows $\mathrm{Ra_c}$ and k_c as functions of Ha, when $\mathrm{E}^{-1} = 10^4$, and validates this argument. We see that initially (the dashed line) increasing Ha has almost no effect, with the rapid rotation continuing to suppress the convection. However, once Ha reaches a critical value, a transition takes place to a completely different mode of convection (the solid line), which occurs with $k_\mathrm{c} = O(1)$, and correspondingly much lower $\mathrm{Ra_c}$, exactly as suggested above. Doing the asymptotic analysis (Chandrasekhar 1961), one finds that this transition takes place when $\mathrm{Ha} = O(\mathrm{E}^{-1/3})$. And once on this second branch, the minimum occurs when $\mathrm{Ha} = O(\mathrm{E}^{-1/2})$, at which point $\mathrm{Ra_c}$ is also $O(\mathrm{E}^{-1})$ (so the Coriolis, buoyancy and magnetic terms in Eq. $(2.14)_2$ are all comparable).

To summarize then, we have seen that while rotation and magnetism separately suppress convection, adding a magnetic field to a rotating system can facilitate convection again, reducing $\mathrm{Ra_c}$ from $O(\mathrm{E}^{-4/3})$ for $\mathrm{Ha} < O(\mathrm{E}^{-1/3})$ down to $O(\mathrm{E}^{-1})$ for $\mathrm{Ha} = O(\mathrm{E}^{-1/2})$. In the next section we will then (i) translate these results back into the geophysically more relevant parameters, and (ii) try to understand what implications they might have for planetary dynamos.

2.3.3 Weak versus Strong Fields

Doing the translation first, we note that the Ekman number is the same here and in Sect. 2.2.2, whereas the Rayleigh numbers are related by $\widehat{\mathrm{Ra}} = \mathrm{E}\,\mathrm{Ra}$. The Hartmann number is similarly related to the Elsasser number by $\Lambda = \mathrm{E}\,\mathrm{Ha}^2$. We therefore have that $\widehat{\mathrm{Ra}}_\mathrm{c} = O(\mathrm{E}^{-1/3})$ for $\Lambda < O(\mathrm{E}^{1/3})$, and $\widehat{\mathrm{Ra}}_\mathrm{c} = O(1)$ for $\Lambda = O(1)$ (having these last two quantities independent of E is, of course, what makes the nondimensionalization in Sect. 2.2.2 particularly convenient).

To assess what these results might imply for the geodynamo, we must consider the differences between our idealized Rayleigh–Benard problem and the real Earth. Most obviously, in the Earth we have a spherical shell rather than an infinite plane layer. This certainly makes the analysis considerably more complicated, and indeed adds various subtleties not present before. However, the main results are unchanged. Roberts (1968) and Busse (1970) considered rotating, nonmagnetic convection in spherical shells, and found that just as in the plane layer, it does not occur until $\widehat{\mathrm{Ra}} = O(\mathrm{E}^{-1/3})$. See also Jones, Soward & Mussa (2000) for the final(?) word on this problem. Similarly, Eltayeb & Kumar (1977), Fearn (1979) and Jones, Mussa & Worland (2003) considered rotating, magnetic convection, and found that there too the main results are as above.

Far more fundamental than this geometrical difference is the origin of the magnetic field; in this idealized problem it is externally imposed, whereas in the real Earth it is internally generated. That is, in the analysis above we could adjust Ha at will, but in the Earth we cannot adjust Λ; the amplitude of the field can only emerge as part of the full solution. Needless to say, this makes the problem considerably more difficult. Nevertheless, let us at least speculate about some of the implications that these results might have for internally generated rather than externally imposed fields.

In particular, imagine taking the Earth's core, and gradually increasing the Rayleigh number from zero. What sort of a sequence of bifurcations would we obtain? For $\widehat{\mathrm{Ra}} = 0$ we would clearly have $\boldsymbol{u} = 0$, and hence also $\boldsymbol{B} = 0$. The initial onset of convection therefore would be nonmagnetic, and would thus occur when $\widehat{\mathrm{Ra}} = O(\mathrm{E}^{-1/3})$. Increasing $\widehat{\mathrm{Ra}}$ further, the convection would presumably become more and more vigorous, until eventually a second critical value is reached where the flow acts as a dynamo. Immediately beyond this value, the field would most likely equilibrate as some very small value, but increasing $\widehat{\mathrm{Ra}}$ further still, both $\boldsymbol{u}$ and hence also $\boldsymbol{B}$ would presumably equilibrate at ever larger values.

In slowly rotating systems, this would presumably be all there is to it; the greater $\widehat{\mathrm{Ra}}$ is, the greater $\boldsymbol{u}$ and eventually $\boldsymbol{B}$ are, and that is it. If the system is rotating sufficiently rapidly though, the above analysis suggests that something quite dramatic could happen. Roberts (1978) conjectured that once the field exceeds $\Lambda = O(\mathrm{E}^{1/3})$, it would begin to facilitate the convection. A more vigorous flow would then yield a stronger field, which would further increase the flow, and so on. The resulting runaway growth would cease only when the field reaches $\Lambda = O(1)$, and the whole pattern of convection has switched from $O(\mathrm{E}^{1/3})$ to $O(1)$ lengthscales. Then once the system has switched to this new mode of convection, according to the results above it should also be possible to reduce $\widehat{\mathrm{Ra}}$ back down to some $O(1)$ value, and still maintain both the flow as well as the field. That is, the magnetic field facilitates convection to such an extent that one can have not only convection, but dynamo action, at a Rayleigh number lower than that for the initial onset of nonmagnetic convection. Indeed,