Planetary Magnetic Fields
Planetary Interiors and Habitability
Final Report

June 2016
Study Report prepared for the Keck Institute for Space Studies (KISS)
Part I: August 12–16, 2013
Part II: February 18–20, 2014
Study Co-Leads: Joseph Lazio, Evgenya Shkolnik, Gregg Hallinan

Acknowledgements

The production of this report benefited from the contributions of multiple people. In particular, significant contributions were made by Hao Cao, Peter Driscoll, William Farrell, Kevin France, Jean-Mathias Griessmeier, Renyu Hu, Thomas Kuiper, Leslie Rogers, Dave Stevenson, Aline Vidotto, Jackie Villadsen, Michael Werner, Wesley Traub, and Philippe Zarka.

The entire Study program would not have been possible without the generous support of the W. M. Keck Foundation. We thank Michele Judd, Tom Prince, and the staff of the W. M. Keck Institute for Space Studies for their hospitality and attention to detail, such that the Study participants could turn their attention to focused discussions and innovative ideas. We also thank Charles ("Chuck") Carter of Eagre Games, Inc., for his assistance with graphics.

E. S. acknowledges supported by NASA Origins of the Solar System grant NNX13AH79G. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under a contract with the National Aeronautics and Space Administration.

Editing and Formatting: Meg Rosenberg

Cover Image: Chuck Carter/Keck Institute for Space Studies (KISS)

Header Images: NASA’s Scientific Visualization Studio

© June 2016. All rights reserved.
4.4  Multi-wavelength Observing Mission ........................................ 106

5  Roadmap and Future Steps ......................................................... 107

5.1  Theoretical and Modeling Efforts ............................................ 107

5.1.1  Planetary Structure & Evolution ......................................... 108

5.1.2  Planetary Dynamo Modeling ................................................ 110

5.2  Observational Approaches ..................................................... 111

5.2.1  Solar System Observations .................................................. 111

5.2.2  Magnetospheric Emissions ................................................... 113

5.2.3  Star-Planet Interactions ....................................................... 113

Bibliography ............................................................................. 115
List of Figures

2.1 Strengths and Structures of Planetary Magnetic Fields in the Solar System .......................... 20
2.2 Magnetic Dynamo Illustration ......................................................................................... 27
2.3 Planet Mass-Radius Diagram ......................................................................................... 36
2.4 Primary Spectra of Inner Solar System Charged Particles .................................................. 42
2.5 Variation in Cosmic Ray Flux .......................................................................................... 43
2.6 Charged Particle Trajectories in a Dipole Field ................................................................. 45
2.7 Magnetic Disconnection & Plasma Detachment at Mars .................................................... 48
2.8 Energetic Particle Dose Rates .......................................................................................... 50
2.9 UV Albedoes of Solar System Giant Planets ..................................................................... 54
2.10 Geographical Correlation Between UV-Dark Aerosols and Auroral Zones on Jupiter .... 55
2.11 Ion Chemistry Driven Reactions in Auroral Regions ......................................................... 56
2.12 Sensitivity of WFIRST Coronagraph for Imaging Planets around Nearby Stars ............ 57
2.13 Magnetic Field of the Planet-Hosting Star τ Boo ............................................................... 60
2.14 Jovian Hectometric Radiation Detected by Cassini and Galileo ..................................... 61

3.1 Refractive Index Near Electron Cyclotron Resonance ....................................................... 64
3.2 Solar System Planetary Radio Spectra ............................................................................... 67
3.3 Radio Emission Limits for Extrasolar Planets .................................................................. 71
3.4 UV Auroral Emission ...................................................................................................... 74
3.5 HD 209458b UV Auroral Spectrum .................................................................................. 76
3.6 UV Fluorescent Emission Lines from M Dwarfs ............................................................... 77
3.7 “UV Bode’s Law” ......................................................................................................... 78
3.8 HST Archival Auoral Emission Limits ............................................................................. 79
3.9 H$_3^+$ emission from Saturn .......................................................................................... 81
| 3.10 Illustration of Early Ingress Light Curves | 84 |
| 3.11 Planet-Induced Chromospheric Ca II K Emission for the Solar-Type Star HD 179949 | 87 |
| 3.12 MHD Simulation of the HD 179949 System | 88 |
| 3.13 Planetary Magnetic Moment-Stellar Activity Correlation | 89 |
| 3.14 Zeeman Doppler Imaging Magnetic Map of HD 179949 | 89 |
| 4.1 Auroral Kilometric Radiation Detected by RAE-2 | 92 |
| 4.2 Sensitivity Requirements for Radio Telescopes | 94 |
| 4.3 Array Constellation | 97 |
| 4.4 Jupiter & a Large Number, Small Diameter Array | 98 |
| 4.5 INSPIRE Spacecraft | 99 |
| 4.6 Large Diameter Space Aperture | 101 |
| 4.7 Wire Arrangement for a Large Reflector | 103 |
| 4.8 Wire Resistance & Diameter | 104 |
| 4.9 Large UV Space Observatory Concept Illustration | 105 |
Study Participants

Part I: August 12–16, 2013

Payam Banazadeh  
JPL, Caltech¹

Renyu Hu  
Caltech

Evgenya Shkolnik  
Lowell Observatory

Stephen Bourke  
Caltech²

Mary E. Knapp  
MIT

David J. Stevenson  
Caltech

Jean-Michel Desert  
Caltech

Thomas B.H. Kuiper  
JPL, Caltech

Margaret A. Thompson  
Princeton University

William M. Farrell  
NASA/Goddard

Joseph Lazio  
JPL, Caltech

Wesley A. Traub  
JPL, Caltech

Kevin France  
University of Colorado at Boulder

Steven M. Levin  
JPL, Caltech

Aline Vidotto  
Univ. of St. Andrews³

Gregg Hallinan  
Caltech

Paulett C. Liewer  
JPL, Caltech

Jacqueline R. Villadsen  
Caltech

Jacob M. Hartman  
JPL, Caltech

Walid A. Majid  
JPL, Caltech

Daniel Winterhalter  
JPL, Caltech

Leslie A. Rogers  
Caltech

¹Currently of Stanford University
²Currently of Onsala Space Observatory
³Currently of Trinity College Dublin
Part II: February 18–20, 2014

Payam Banazadeh  
JPL, Caltech\textsuperscript{4}  
Joe Kirschvink  
Caltech  
Leslie A. Rogers  
Caltech

Stephen Bourke  
Caltech\textsuperscript{5}  
Margaret Kivelson  
UCLA  
Gerald Schubert  
UCLA

Hao Cao  
UCLA  
Mary Knapp  
MIT  
Evgenya Shkolnik  
Lowell Observatory

Peter Driscoll  
University of Washington  
Thomas B.H. Kuiper  
JPL, Caltech  
David J. Stevenson  
Caltech

William M. Farrell  
NASA/Goddard  
Joseph Lazio  
JPL, Caltech  
Wesley A. Traub  
JPL, Caltech

Jean-Mathias Griessmeier  
Universite d’Orleans  
Steven M. Levin  
JPL, Caltech  
Aline Vidotto  
Univ. of St. Andrews\textsuperscript{6}

Gregg Hallinan  
Caltech  
Paulett C. Liewer  
JPL, Caltech  
Jacqueline R. Villadsen  
Caltech

Renyu Hu  
Caltech  
Walid A. Majid  
JPL, Caltech  
Daniel Winterhalter  
JPL, Caltech

Jonathan Nichols  
University of Leicester

\textsuperscript{4}Currently of Stanford University  
\textsuperscript{5}Currently of Onsala Space Observatory  
\textsuperscript{6}Currently of Trinity College Dublin
Soon after its detection, radio emission from Jupiter was quickly identified as a product of its planetary-scale magnetic field. Subsequent spacecraft investigations have revealed that many of the planets—and even some moons—either currently have or have had in the past a planetary-scale magnetic field. Generated by dynamo processes within the planet, planetary-scale magnetic fields provide a means of constraining the properties of a planet’s interior through remote sensing, and it may even be possible to measure the magnetic fields of extrasolar planets. If so, they will offer one of the few means available of understanding the potential diversity of planetary interiors. In the case of our own planet, the presence of Earth’s magnetic field has long been suspected to be partially responsible for its habitability. Thus, knowledge of the magnetic field of an extrasolar planet may be a valuable component to assess its habitability, or to understand an absence of life on an otherwise potentially habitable planet.

This report summarizes the investigations and conclusions from a William M. Keck Institute for Space Studies on planetary magnetic fields. It is organized as follows:

Chapter 2 focuses on the generation and implications of planetary magnetic fields. Among the topics considered during the Study were the state of knowledge of planetary dynamos and the generation of planetary-scale magnetic fields, using the solar system planets as a guide; what the implications are for a planet’s atmosphere and biology, depending upon whether the planet has a magnetic field; and how a planet’s magnetic field interacts with the stellar wind from its host star. Key conclusions from this chapter include:

- The detection of even a single extrasolar planetary magnetic field could provide essential information to extend our knowledge of planetary dynamos. One of the limiting factors
in understanding planetary dynamos is the relatively limited sample that we have in the solar system. Just as the first discoveries of “hot Jupiters” yielded crucial insights to the larger diversity of planets and even to the formation of our solar system, the discovery of extrasolar magnetic fields will likely improve our knowledge of magnetic dynamos, including those found in our solar system.

- Rotation does not appear to be a significant factor in the planetary dynamo mechanism, our relatively limited sample in the solar system notwithstanding. Planetary rotation periods, even in the case of Venus, are rapid with respect to those for generating planetary dynamos, a conclusion often not widely appreciated in the larger astronomical community.

- Further study is needed to understand the extent and mechanisms by which magnetic fields might shield a planet’s atmosphere and surface life. The Study participants did not reach a consensus on the importance of this factor. Considerations included comparisons of the relative thickness of the atmospheres of Earth and Mars, even though both seem to suffer similar mass loss rates.

Chapter 3 focuses on the detection of magnetic fields, particularly by remote sensing means suitable over interstellar distances. Five mechanisms were identified:

**Radio Emission** A proof-of-concept of this method already exists, in the detection of Jupiter’s radio emission, though there have been no detections to date of radio emissions from any extrasolar planets and the predicted flux densities are low.

**Far-Ultraviolet Auroral Emission** A proof-of-concept of this method exists, in the detection of Jovian aurorae by the Hubble telescope. There have been no detections to date by this method, though the number of planets targeted has also been small. More concerning, even if far-UV emission is detected, it may not be a unique indicator that it was generated near magnetic polar regions. In the absence of a magnetic field, the stellar wind would impinge directly on the planet’s atmosphere, potentially producing far-UV emission.

**Infrared H$_3^+$ Auroral Emission** A proof-of-concept of this method exists, in the detection of Jovian, Saturian, and Uranian aurorae, including from ground-based telescopes. Like far-UV emission, it is not clear that the presence of H$_3^+$ emission can only be generated near magnetic polar regions. Further, it is also not clear that even the James Webb Space Telescope will have sufficient sensitivity to detect this emission over interstellar distances.

**Transit Light Curve Variations** Two extrasolar planets, WASP-12b and HD 189733b, both hot Jupiters, show asymmetric transits consistent with the presence of a magnetospheric bow shock. Due to the orbital velocity of a hot Jupiter, a magnetospheric bow shock would not appear along the star-planet line but somewhat ahead of the planet in its orbit, thereby producing an early transit ingress. This method appears promising, given that there are
now two detections, but it is also limited to hot Jupiters because of the required geometry and it may not be possible to separate uniquely the strength of the planetary magnetic field and the stellar wind.

**Star-Planet Interactions** Multiple stars have shown line strength variations phased to the orbital period of a hot Jupiter. The line strength variations appear more likely to originate from magnetic field interactions than from tidal effects. However, these measurements do not provide a unique measure of the strength of the planetary magnetic field.

**Chapter 4** describes the mission concepts that were developed during the course of the Study. Given the concerns about the strength of H$_3^+$ emission relative to the sensitivity of JWST, attention was focused on just two concepts: a radio telescope and a UV-optimized space telescope. Neither concept is new, though the detection and measurement of planetary magnetic fields has not traditionally been recognized as a potential key component of such telescopes. Further, both a future UV-optimized telescope and large space-based radio telescope have been identified in planning documents for NASA. For the UV-optimized telescope, there is already considerable design work underway. For a radio telescope, the Study identified two possibilities: either a small number of large diameter antennas or a large number of small antennas. The potentially new technology relevant for a future radio telescope would determine the extent to which small spacecraft (potentially "cubesats") could be used to realize either concept. Further, while this KISS Study focused on planetary magnetic fields, a radio telescope capable of detecting and studying extrasolar planetary magnetic fields would also likely be capable of detecting the highly redshifted neutral hydrogen hyperfine line (at 21 cm) from the intergalactic medium during and before the formation of the first stars, an epoch identified as "Cosmic Dawn."

**Chapter 5** concludes the report with an identification of future steps to be taken. Considerable theoretical work could yet be done for modeling planetary structure and dynamos. Some insights into a planet’s interior can be gleaned from mass-radius relationships, which in turn can be used to develop or test dynamo models. Further, laboratory experiments or other developments in understanding the phase diagram of materials within planetary interiors will inform dynamo models. There are observational campaigns that can yet be undertaken, both for solar system planets and extrasolar planets. Illustrating this continued potential, the discovery of the asymmetric transit of HD 189733b occurred during the writing of this report, and efforts continue to detect the radio emission of extrasolar planets. In addition, there has been a campaign to explore the Saturnian magnetosphere with Cassini and Juno will explore the magnetic polar regions of Jupiter. There are significant technical advances that could be made for realizing future mission concepts, such as a UV-optimized telescope or a space-based radio telescope. Fortunately, in many cases, this technical work is proceeding, driven by a combination of funding from NASA, other Governmental agencies, and even private sources.
1. Study Overview

The magnetic field of the Earth\(^1\) has been known since antiquity and attempts to explain its origin date back over a century. Even early attempts tie the Earth’s magnetic field to its interior structure. Soon after the discovery of decametric wavelength radio emission from Jupiter (Burke & Franklin, 1955; Franklin & Burke, 1956), it was determined that this radiation was linked to Jupiter’s magnetic field (Carr & Gulkis, 1969, and references within), but it took some time for the existence of the Jovian magnetic field to be tied to the interior structure of the planet. As spacecraft began visiting other planets in the solar system, other planetary-scale magnetic fields began to be detected. Today, from a variety of both remote sensing and in situ spacecraft measurements, it is clear that the Earth, Mercury, Ganymede, and the giant planets of the solar system all contain internal dynamo currents that generate planetary-scale magnetic fields. Further, other objects in the solar system, notably Mars and the Moon, show residual magnetism indicative of past dynamo processes.

These internal dynamo currents arise from differential rotation, convection, compositional dynamics, or a combination of these processes in a planet’s interior. Extrapolated to extrasolar planets, knowledge of a planet’s magnetic field places constraints on the thermal state, composition, and dynamics of its interior—all of which will be difficult to determine by other means—as well as potentially crucial information about the extent to which the surface of a terrestrial planet is shielded from cosmic rays and hence potentially habitable.

\(^1\)Header image: Artist’s impression of the Earth’s magnetosphere, also illustrating its interaction with the solar wind emitted by the Sun. Within the solar system—in addition to the Earth—Mercury, Ganymede, and all of the giant planets generate magnetospheres. These magnetospheres result from internal currents in these bodies and therefore provide information about their interior structures. [The figure is not to scale. Image credit: NASA]
Numerous observational manifestations of planetary magnetic fields have been suggested, based in large part upon experience from the detection and study of planetary magnetic fields within the solar system. Electron cyclotron maser emission, resulting from an interaction between the planetary magnetosphere and the solar wind in the planetary magnetic polar regions, has been detected from all of the gas giants and the Earth in the solar system. In addition to electron cyclotron maser emission, planetary auroral regions produce ultraviolet emission, which may also be detectable over interstellar distances. There have been suggestions in the literature that the inflated radii of some “hot Jupiter” extrasolar planets may be due, in part, to Ohmic dissipation within the planet as the planetary magnetosphere moves through the magnetosphere of its host star. Finally, the magnetospheres of “hot Jupiters” have been predicted to produce a bow shock that would be detectable as an asymmetric transit light curve.

The W. M. Keck Institute for Space Studies Program on Planetary Magnetic Fields brought together an international team of experts to assess the current state of knowledge about planetary magnetic fields, for both solar system and extrasolar planets; to track the progress of the new ground-based instruments, particularly those in the radio (e.g., the Low Frequency Array [LOFAR], Long Wavelength Array at Owens Valley Radio Observatory [LWA-OVRO]), as they acquire their first data on extrasolar planetary systems; to develop complementary observing strategies from existing optical/UV telescopes; and to feed that forward to form an observational strategy for current and future radio and UV telescopes from both the ground and space.

This report summarizes the findings and conclusions of the Study. Further, following the conclusion of this Study, NASA’s Astrobiology Program undertook a strategic roadmap exercise. The Astrobiology Strategy Roadmap (Hays et al., 2015) recognizes the potential importance of a planet’s magnetic field in determining its habitability. Reflecting the likely lack of current measurements in an extrasolar context, however, the Astrobiology Strategy Roadmap contains little discussion of approaches to measuring extrasolar planetary magnetic fields. This Study report can be seen as a first effort to understand how it might be possible in the context of the current (and future) Astrobiology Strategy Roadmaps.
2. Searching for & Characterizing Extrasolar Planets

Though the concept of worlds outside of the solar system extends back at least a millennium, the (confirmed) detection of extrasolar planets resulted only in the last decade of the 20th Century. At this writing, nearly 2,000 extrasolar planets are known, with over half having been discovered by the Kepler mission (Borucki et al., 2010). Numerous comprehensive reviews describe the techniques by which extrasolar planets can be both detected and confirmed to exist, and the focus of work has broadened to include not only the detection of extrasolar planets but their characterization as well. Many instruments, telescopes, and missions are being developed to address this dual objective.

From the perspective of this workshop, notable space-based missions with extrasolar planets as either a specific focus or included in their science case include the following: The Transiting Exoplanet Survey Satellite (TESS), the James Webb Space Telescope (JWST), the Wide-Field Infrared Survey Telescope (WFIRST), and potentially the Exoplanet Characterization Observatory (EChO).

Most of the current community focus, however, has not included the potential role of magnetic fields as a means of detecting or characterizing extrasolar planets. This chapter provides the motivation for such searches and the use of magnetic fields for characterization of extrasolar planets. The interested reader may also wish to consult Characterizing Stellar and Exoplanetary

---

1 http://exoplanet.eu/catalog/
2 http://kepler.nasa.gov/
3 http://space.mit.edu/TESS/TESS/
4 http://www.jwst.nasa.gov/
5 http://wfirst.gsfc.nasa.gov/
6 http://echo-spacemission.com/
Chapter 2. Searching for & Characterizing Extrasolar Planets

Environments (Lammer & Khodachenko, 2015), which discusses interactions between planets and their host stars and includes, but is not limited to, the role of magnetic fields.

2.1 Searching for Planets

We defer a detailed discussion of the physics by which a planetary-scale magnetic field can generate an observational signature (or signatures) until Chapter 3. A key aspect of many of those observational signatures, however, is that they are generated by an interaction between the stellar wind and the planet’s magnetic field, such that a stronger stellar wind is more likely to generate a larger magnetically-generated observational signature.

In contrast, for many of the current methods of searching for extrasolar planets, the precision to which a planetary signal can be extracted from the data often depends, in part, on the properties of the host star. For instance, in radial velocity surveys, one of the limiting factors in the velocity precision is intrinsic stellar “jitter,” caused by starspots or other surface inhomogeneities. Such stellar jitter is well known to be correlated with stellar activity, the level of which declines with age (Butler et al., 1996; Saar & Donahue, 1997). Radial velocity surveys tend to select stars that are chromospherically quiet (Saar et al., 1998; Cumming et al., 2008), which is likely to introduce a bias toward older stars. Furthermore, the link between chromospheric activity and age means that distinguishing planetary transits from stellar surface features will probably be easier for older, less active stars (yet, see also Jenkins, 2002). Consequently, there is likely to be a selection bias against planets around stars with ages (significantly) less than that of the Sun.

Through measurements of the size of the astropause (i.e., the boundary between the stellar wind and the local interstellar medium), Wood et al. (2002, 2005) find the mass loss rate as a function of age, \( \dot{M} \propto t^x \), with \( x \approx -2 \), a dependence probably linked to the decrease in surface magnetic activity with stellar age. Thus, the stellar wind around a 1-Gyr-old star may be 25 times as intense as the current solar wind, suggesting a concomitant increase in the amplitude of the magnetically-generated observational signature from any planet in orbit about such a star.

Thus, an approach that complements standard searches for planets (i.e., radial velocity surveys and transits) would be to search for magnetically-generated emission from stars, particularly from younger stars or more active stars that might not be targeted for radial velocity or transit surveys. Lazio et al. (2010a) illustrate this approach for the specific case of looking for electron cyclotron masers from Jovian-mass planets around nearby stars.
2.2 Planetary Interiors: Magnetic Field Determination & Interior Structure

Planetary-scale magnetic fields offer a window into planetary interiors and are the only remote probes of planetary deep interior dynamics. While gravity field measurements provide information about the structure, a planet’s magnetic field comes from a dynamic process and reveals unique information that cannot be gleaned any other way.

We have a basic understanding that planetary magnetic fields come from dynamos. There is a richness to planetary interior structures, and many aspects of a planet’s interior structure factor into determining whether there might be a dynamo. Some of the most challenging problems today in planetary physics relate to the presence and absence of magnetic fields and the crustal magnetization signatures of extinct dynamos.

In this section, we summarize the magnetic field characteristics of the solar system planets (§2.2.1), distill guiding principles of planetary dynamo magnetic fields (§2.2.2), review planetary thermal evolution (§2.2.3), and then attempt to extend the discussion to the diverse array of extrasolar planets detected to date and what new insights magnetic fields measurements could offer into the interior structure of these distant worlds (§2.2.4). The subsequent section explores the connection to planetary rotation and formation (§2.3) and is related. We also refer the reader to Section 5.1 for a discussion of priorities for future theoretical and modeling efforts.

2.2.1 Solar System Planetary Magnetic Fields

Magnetic fields are ubiquitous in the solar system. All of the planets, with the possible exception of Venus, have or have had magnetic fields. Even some moons and small bodies have or have had magnetic fields (e.g., Ganymede, Vesta, Gaspra). Solar system planetary magnetic fields are measured in situ by magnetometers on flyby trajectories and orbiting spacecraft, by electron and proton reflectometers on low-orbiting spacecraft, and remotely by their radio emission, in the case of the giant planets (§3.1). The solar system planets show a wide diversity of planetary magnetic field strengths, structures, and histories (Table 2.1; Figure 2.1). Because the characteristics of magnetic fields in the solar system guide our approach toward magnetic fields in extrasolar planets, we review the properties of these more local magnetic fields, starting from the inside out. Detailed reviews can also be found in Stevenson (2010) and Schubert & Soderlund (2011).

Mercury

Mercury’s primarily iron, partially-molten core accounts for ~70% of the planet’s mass (Hauck et al., 2013). Mercury’s magnetic field was first measured by Mariner 10 in 1974 and 1975 (Ness et al., 1975, 1976), and recently measured in greater detail by the MESSENGER spacecraft (Anderson et al., 2008, 2011, 2012). Mercury’s magnetic field has three striking features.
Figure 2.1: Radial component of the intrinsic magnetic fields at the surfaces of those solar system planets having planetary-scale magnetic fields. For each map, the center is at 0° longitude, and the color scales differ from planet to planet. Figure constructed based on the magnetic field models from Anderson et al. (2012, Mercury), Finlay et al. (2010, Earth: IGRF10), Yu et al. (2010, Jupiter), Cao et al. (2012, Saturn), and Holme & Bloxham (1996, Uranus and Neptune).
### Table 2.1: Magnetic field characteristics of Solar System bodies (adapted from Schubert & Soderlund, 2011)

<table>
<thead>
<tr>
<th>Planet</th>
<th>Mass ((10^{24} \text{ kg}))</th>
<th>Radius ((\text{km}))</th>
<th>Surface Radial Magnetic Field ((\mu\text{T}))</th>
<th>Dipolarity</th>
<th>Dipole Tilt ((\degree))</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mercury</td>
<td>0.33</td>
<td>2440</td>
<td>0.3</td>
<td>0.71</td>
<td>3</td>
</tr>
<tr>
<td>Venus</td>
<td>4.87</td>
<td>6052</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Earth</td>
<td>5.97</td>
<td>6371</td>
<td>38</td>
<td>0.61</td>
<td>10</td>
</tr>
<tr>
<td>Moon</td>
<td>0.07</td>
<td>1738</td>
<td>(\lesssim 100)</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Mars</td>
<td>0.64</td>
<td>3390</td>
<td>(\lesssim 0.1)</td>
<td>0.61</td>
<td>10</td>
</tr>
<tr>
<td>Jupiter</td>
<td>1900</td>
<td>69,911</td>
<td>550</td>
<td>0.61</td>
<td>9</td>
</tr>
<tr>
<td>Io</td>
<td>0.09</td>
<td>1821</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Europa</td>
<td>0.05</td>
<td>1565</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Ganymede</td>
<td>0.15</td>
<td>2634</td>
<td>0.91</td>
<td>0.95</td>
<td>4</td>
</tr>
<tr>
<td>Callisto</td>
<td>0.11</td>
<td>2403</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Saturn</td>
<td>570</td>
<td>58,232</td>
<td>28</td>
<td>0.85</td>
<td>(&lt; 0.5)</td>
</tr>
<tr>
<td>Titan</td>
<td>0.13</td>
<td>2575</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
<tr>
<td>Uranus</td>
<td>87</td>
<td>25,362</td>
<td>32</td>
<td>0.42</td>
<td>59</td>
</tr>
<tr>
<td>Neptune</td>
<td>100</td>
<td>24,624</td>
<td>27</td>
<td>0.31</td>
<td>45</td>
</tr>
</tbody>
</table>

First, the field is axisymmetric about Mercury’s rotation axis, with relatively little variation with longitude. Second, the field is much stronger near the North Pole than at the South Pole (by a factor of 3.4). Lastly, the magnitude of Mercury’s magnetic field is two orders of magnitude smaller than that of the Earth, far less than predicted by dynamo theory (§2.2.2).

The high axisymmetry of Mercury’s magnetic field is likely a property of the dynamo and the host planet. A dynamo-generated magnetic field cannot be perfectly axisymmetric, as stated by one of the oldest and most renowned fundamental theorems of dynamo theory, Cowling’s theorem: *a constantly axisymmetric magnetic field cannot be maintained via dynamo action* (Cowling, 1933).

The near-axisymmetry may be a testament to the nature of the flows in Mercury’s interior, which must be controlled by the properties of Mercury, e.g., rotation period, buoyancy flux, core-mantle boundary heterogeneity. This near-axisymmetry could also be evidence of a stable layer at the top of the core (e.g., Christensen, 2006).

The significant north-south asymmetry of Mercury’s magnetic field could be a property of the dynamo or the planet itself, or a combination of both. Such symmetry breaking could be a natural result of a corresponding north-south asymmetry in the Mercury’s interior, for example, if the temperature at the base of the mantle is different between the northern and southern hemispheres. However, any such asymmetries in Mercury’s mantle and core properties have yet
to be discovered. Alternatively, dynamo models with north-south symmetric boundaries are also capable of explaining this type of symmetry breaking. Such dynamo models are examples of spontaneous symmetry breaking, and have been constructed for Mercury (Cao et al., 2014). In either scenario, the north-south asymmetry of Mercury’s magnetic field provides constraints on Mercury’s interior.

The relatively weak strength of Mercury’s magnetic field is a puzzle, and several potential explanations have been suggested (e.g., Heimpel et al., 2005; Stanley et al., 2005; Takahashi & Matsushima, 2006; Christensen, 2006; Christensen & Wicht, 2008; Manglik et al., 2010; Vilim et al., 2010). The suggested explanations for Mercury’s small surface field strength can be divided into two general groups. In the first group, the magnetic field within Mercury’s core is indeed rather weak. Some support for this notion emerges from recent observations of the magnetic fields of low mass stars, which show several examples of bi-stability (i.e., stars with similar properties possessing magnetic fields with very different strengths). In the second group, the magnetic field within Mercury’s core is comparable to that within Earth’s core; however, secondary effects, such as electromagnetic filtering, only allow a small fraction of the field to be present outside the core. These explanations could have more direct implications about planetary interiors.

Venus & Earth

While Venus and Earth have many similar bulk properties, a noticeable difference is Venus’ absence of a magnetic field. While Earth has a dynamo-generated global magnetic field, Venus has no observed global field, and there is not even a remanent crustal magnetic field, which could serve as an indication of a dynamo-generated global magnetic field in Venus’ past. However, there is a simple explanation for the absence of even a crustal magnetic field on Venus: it may be a consequence of surface temperatures so high that the crust cannot retain a magnetic memory.

The two main suggestions for why Venus does not currently have a magnetic field are (i) that the planet does not have an inner core, or (ii) that, by virtue of being in a stagnant lid regime, the core cannot cool at a sufficient rate to sustain a dynamo. We emphasize that the lack of a magnetic field on Venus cannot readily be attributed to its slow rotation (§2.2.2)—Venus is in fact a fast rotator from the perspective of dynamo flows. We will investigate possible reasons for the differences between Earth’s and Venus’ magnetic fields in greater detail in Section 2.2.3.

Earth’s magnetic field is currently dipole-dominant, and the surface field strength is about 0.5 G. The dipole axis of the Earth’s magnetic field currently points south, and deviates 10° from the spin axis (Finlay et al., 2010). Paleomagnetic evidence indicates that Earth has had a global magnetic field for at least 3.5 billion years, and its polarity has undergone several reversals. While ground-based and low altitude satellite field measurements have informed us about Earth’s surface magnetic field to very fine scales, the finite amplitude of the crustal magnetic field limits our knowledge of the geomagnetic field at the core surface to spherical harmonic degrees less
than 14. Even within this limitation, though, it is clear that there is structure in the field at the
core-mantle boundary: there are intense flux patches near the equator, reversed flux patches in
both hemispheres, and field minimum regions near the poles. Earth’s magnetic field is currently
changing at a rate of about 0.1% yr$^{-1}$. If we attribute the field changes to advection of magnetic
field by core surface flows under the assumption of flux frozen-in, then zonal flows at the scale of
20 km yr$^{-1}$ near the equator and anticyclonic polar vortices are implied.

Moon

The Moon does not currently have a global magnetic field but does have small-scale, isolated
magnetic anomalies. The anomalies have been measured by nearby orbiting spacecraft (e.g., Dyal
et al., 1970; Purucker, 2008; Tsunakawa et al., 2010). The study of lunar rocks has also offered
insights into the primitive magnetization of the surface material (e.g., Runcorn et al., 1970; Garrick-
Bethell et al., 2009). Some of the largest magnetic anomalies are antipodal to major impact
basins. However, rather than necessarily indicating the presence of a previous global magnetic
field, these magnetic anomalies may be the result of seismically modified terrain or antipodal
impact ejecta acquiring a magnetization in the presence of a lunar magnetic field amplified at
the antipode by impact-generated plasma compression (Hood et al., 2001), or a combination of
both effects. The anomalies may even reflect not a previous magnetization of the Moon, but the
presence of highly magnetic material delivered by the impactors (Wieczorek et al., 2012).

It is difficult to understand the paleomagnetic field of the Moon without invoking a lunar dynamo
during an early period of lunar history (Garrick-Bethell et al., 2009; Hood, 2011). Evidence for
the magnetic field is seen in rocks spanning a 640 Myr period of time (Shea et al., 2012; Suavet
et al., 2013). The magnetic anomalies are also relatively strong, despite the weak magnetism of
lunar rocks. The paleointensity approximately 4.2 Gyr ago is estimated to be of the order 10–100
mG (Garrick-Bethell et al., 2009). Accounting for the radial $r^{-3}$ drop off in the magnetic field
between the lunar surface and the small lunar core, the Moon likely had a field that at its core
was as large as (if not larger than) that of the Earth at its core.

The main issue with the Moon, and possibly the greatest puzzle among planetary bodies, is how
to make the dynamo operate for hundreds of millions of years. A thermal convection-driven lunar
dynamo is not expected to last long before the core cooling rate (predicted by thermal evolution
models, e.g., Konrad & Spohn, 1997) declines to the extent that the Moon’s liquid outer core is
no longer convecting. It could be that the Moon is special and that a different dynamo mechanism
was operative for the Moon than for other solar system bodies. One possibility for generating the
lunar dynamo is precession. The lunar core does not follow the precession of the mantle, which
occurs on an 18.6-year timescale. The resulting differential motion between the core and mantle
could set up flows (similar to those set up by the mechanical stirring in a washing machine)
that would generate a magnetic field (Dwyer et al., 2011). A second possibility is to transiently
generate a magnetic field in the lunar core by impact-induced changes in the Moon’s rotation
rate (Le Bars et al., 2011). Following a large impact, the Moon would gradually spin down back to its current libration state (with one face toward the Earth), and large-scale flows excited in the core might generate a magnetic field, although it is difficult to understand how this mechanism could operate for 640 Myr. Alternatively, recent models have suggested that a compositionally driven dynamo in the Moon may be able to explain both the magnitude and duration of the dynamo-generated magnetic field on the moon (Scheinberg, Soderlund, & Schubert, 2015).

**Mars**

Mars, like the Moon, does not currently have a global magnetic field of internal origin. Mars does have a strong crustal magnetization that exceeds the remanent magnetization in Earth’s crust by an order of magnitude. Orbiting within 200 km of the martian surface near periareion, Mars Global Surveyor found that the crustal magnetization is frequently organized in linear features (magnetic lineations) and features a strong north-south asymmetry (Connerney et al., 2005). The strong crustal magnetic field implies that Mars had an active dynamo early in its history. There is no consensus regarding when and for how long the dynamo operated (e.g., Acuña et al., 1999, 2001; Schubert et al., 2000; Lillis et al., 2006; Milbury et al., 2012).

**Jupiter System**

Of the icy bodies in the solar system, currently only Jupiter’s moon Ganymede is known to have a dynamo magnetic field. Ganymede’s magnetosphere was mapped by the Galileo spacecraft (Kivelson et al., 1996). Ganymede’s field is anti-parallel to its rotation axis, directed opposite to the Jovian magnetic field, and dipole-dominated with a mean surface strength of about 10 mG. Ganymede itself is half ice, half rock. Though larger than Mercury, it has a smaller iron core (Schubert et al., 2004), which nonetheless appears to be more efficient or effective in generating a substantial magnetic field by dynamo action relative to Mercury (Schubert et al., 1996). Ganymede is the best example known to date bracketing the lower bound on the size of a core that can sustain a dynamo.

Why can Ganymede sustain a dynamo, while Mars (a much larger body) cannot? An important factor may be that Ganymede is volatile rich, and probably has considerably more sulfur in its core than does Mars. As a result, the phase diagram relevant to Ganymede’s core allows for the possibility of compositionally-driven convection. The sinking of Fe snow formed below the core-mantle boundary, or the upward flotation of solid FeS formed in the deep core, can lead to compositional buoyancy (Hauck et al., 2006). Dynamo calculations by Zhan & Schubert (2012) have shown that multipole-dominant magnetic fields are generated by Fe snow, while dipole-dominant dynamos are produced by FeS flotation and inner core growth. Ganymede’s present dipole-dominant magnetic field suggests that the Fe snow process does not play a primary role in driving Ganymede’s core convection.
It is unclear why Ganymede has a dynamo, whereas other large icy satellites such as Titan and Callisto do not (Jia et al., 2010). From gravity data, it is known that Ganymede is differentiated (Anderson et al., 1996), whereas the extent of differentiation of the other satellites is less clear (e.g., Showman et al., 1999; Iess et al., 2010). If Titan and Callisto are indeed not fully differentiated, that could suggest a different formation scenario for these two satellites as compared to that for Ganymede. The timing and duration of the accretion process, as well as the amount of later radiogenic heating, all affect the degree of differentiation.

Jupiter

Jupiter has the largest magnetic field of all the solar system planets, with a mean surface strength of 5.5 G. The magnetic field is produced by convectively-driven dynamo action in the highly electrically conducting metallic hydrogen layer, with possible contributions from the less electrically conducting region near the base of the molecular envelope. The Juno spacecraft, which will arrive at Jupiter in 2016, will determine the magnetic field of Jupiter to better precision than we have for the internal field of the Earth, mapping the field structure of Jupiter up to spherical harmonic degree 14. The tilt of Jupiter’s dipole is about 10° relative to its rotation axis (similar to the Earth). One consequence of this tilt is that Jupiter’s satellites experience a time-varying magnetic field, which induces currents in conducting regions in their interiors. The discovery of this electromagnetic induction revealed that Europa and Callisto, as well as potentially Ganymede, have salt water oceans (Khurana et al., 1998; Kivelson et al., 1999, 2000).

Saturn

Saturn’s magnetic field at the surface is dipole-dominant, weaker than those of Jupiter and Earth (0.3 G at 1 bar), and extremely axisymmetric. The dipole tilt of Saturn’s magnetic field must be less than 0.06°, as constrained by Cassini measurements (Cao et al., 2011). Such an extremely axisymmetric magnetic field is challenging for planetary dynamo theory (see the discussion above for Mercury). Furthermore, Cao et al. (2011) place a tight upper bound on the secular variation of the field. The secular variation of Saturn’s magnetic field, if it exists, is at least an order of magnitude smaller than its terrestrial counterpart. Such a slow secular variation does not have a proper explanation yet, and could possibly be linked to the extreme axisymmetry. Measurements obtained by Cassini also reveal properties of the Saturnian magnetic field other than axisymmetry and extremely slow secular variation, e.g., magnetic flux inside Saturn is expelled away from the equator and strongly concentrated near the spin poles (Cao et al., 2012). This behavior differs from the field properties at the core surface of the Earth, where the field near the spin poles is at a relative minimum compared to the field at mid-latitudes. The equatorial flux expulsion could be a result of strong equatorial zonal winds near the surface of the dynamo region, while the lack of polar field minima could be a result of a relatively small “rocky” core inside Saturn. The Cassini proximal orbits, which will measure Saturn’s magnetic field at an unprecedentedly close
distance, will reveal considerably more detail about Saturn’s magnetic field, which in turn can tell us more about Saturn’s interior structure, dynamics, and rotation period (which is currently only known to within 10 minutes).

**Uranus & Neptune**

The ice giants’ magnetic fields are fundamentally different from other known (dipole-dominated) planetary fields, in the sense that they have significant quadrupole components. In addition to having more complex surface fields, the dipole components of the fields for both Uranus and Neptune have very large tilts (~50°). Such complex field geometries, when expressed in spherical harmonics, are multipolar-dominant. The ice giant planets are not thought to have metallic hydrogen in their interiors. Rather, the conducting material in the dynamo generation region is likely a water-dominated material under high pressure. Shock-wave experiments and ab initio molecular dynamics simulations have shown that high pressure water is electrically conductive due to the mobility of protons (e.g., Mitchell & Nellis, 1982; Nellis et al., 1997; Cavazzoni et al., 1999; Lee et al., 2006; French et al., 2009). The presence of a magnetic field on Uranus and Neptune is thus evidence that a very conductive metal is not necessary for a dynamo; an ionic conductor is sufficient.

### 2.2.2 General Guiding Principles of Planetary Dynamos

Among the bodies in our solar system, there are five categories of planetary magnetic fields: (i) Large (~1 G) predominantly dipolar fields (e.g., Earth, Ganymede, Jupiter, Saturn); (ii) Weak and strongly north-south asymmetric magnetic fields (e.g., Mercury); (iii) Large predominantly non-dipolar fields (e.g., Uranus, Neptune); (iv) Small fields arising from crustal magnetism indicating possible past dynamos (e.g., Moon, Mars, some asteroids, and possibly Venus); and (v) Small fields arising from induction in the presence of a time-varying external field (e.g., Io, Europa, Callisto). Of relevance to extrasolar planets, only strong, long-lived, planetary-scale magnetic fields could be potentially observable. It is thought that the generation of large magnetic fields (in planets, the Sun and stars, and the Galaxy and other galaxies) requires a dynamo process.

A dynamo is a process through which kinetic energy is converted into magnetic energy. This process in astrophysical bodies is usually realized through fast (magnetic Reynolds number greater than 10) and complex (non-zero helicity) flows in electrically conducting fluids which maintain magnetic fields over astronomical timescales. The system illustrated in Figure 2.2 is a mechanical analog to a magnetic dynamo, and illustrates the basic principles. In the mechanical disk dynamo, there is a shaft connected to a spinning electrically conducting disk, and the whole system is threaded by a seed magnetic field. From Faraday’s law, the motion of the disk induces an electromotive force (EMF) between the shaft \( B \) and the sliding contact \( A \). The EMF drives a
current through the circuit. Because the wire is threaded in the right-hand sense, from Ampère’s law, the current produces a magnetic field directed in the same sense as the seed magnetic field. As work is done to spin the disk and offset the dissipation of the electrical current in the wire, the initial seed field is amplified.

Planets do not actually look like a mechanical disk dynamo, but the principles expressed in this mechanical dynamo analogy have a one-to-one correspondence with various aspects of generating a magnetic field in a planet or a star. First, a crucial attribute of the mechanical dynamo is differential motion. The disk is spinning relative to the wire; if the system were rigid, it would not generate a field. In a planet, differential motion comes from fluid motions. Second, an electrical conductor is crucial to allow current to flow. Third, the topology is an interesting and tricky aspect of the system. If the wire is wrapped in the opposite (left-hand) sense, the induced magnetic field opposes the seed field and an anti-dynamo is obtained. Planets with dynamo magnetic fields have managed to break symmetry in the fluid motions. Dynamos are intrinsically three-dimensional and symmetry breaking. Three spatial dimensions are required for a dynamo; if we lived in a two-dimensional universe, there would be no magnetic field generation.

The relevant differential equation for the magnetic field $B$ in a fluid dynamo is

$$
\frac{\partial B}{\partial t} = \nabla \times (\lambda \nabla \times B) + \nabla \times (\mathbf{v} \times B),
$$

where $\lambda = (\mu_0 \sigma)^{-1}$ is the magnetic diffusivity, $\sigma$ is the conductivity, and $\mathbf{v}$ is the fluid motion.
Chapter 2. Searching for & Characterizing Extrasolar Planets

relative to rigid rotation (e.g., Braginsky & Roberts, 1995). Assuming that the magnetic diffusivity is constant, a simplified form of the induction equation is obtained

\[
\frac{\partial \mathbf{B}}{\partial t} = \lambda \nabla^2 \mathbf{B} + \nabla \times (\mathbf{v} \times \mathbf{B}). \tag{2.2}
\]

In the absence of fluid motion, the magnetic field will decay diffusively on a timescale

\[
\tau \sim \frac{L^2}{\pi^2 \lambda}. \tag{2.3}
\]

Fluid motion is important (and can offset magnetic diffusion) provided that the magnetic Reynolds number is \( R_m = vL/\lambda \gtrsim 10 \), where \( L \) is the length scale of the motions and magnetic field structures. Then, perhaps a field is sustained. The “perhaps” qualification is important, because the particular flow-topology (e.g., non-zero helicity) is not guaranteed in all fluid flows. Thus, \( R_m > 10 \) is a necessary, but not sufficient, criterion for a dynamo.

A planetary dynamo originates from fluid motions in an electrical conductor, in the presence of a preexisting magnetic seed field. Large-scale vertical motions are needed to obtain a significant external magnetic field (Busse, 2000), and large-scale vertical motions suggest convection. Convection can arise through compositional buoyancy, resulting from the presence of material less or more dense than the other materials in the system, or through thermal buoyancy, resulting from the interior of the planet being at a higher temperature than the surface. Though there are many papers devoted to other dynamo generation mechanisms in the literature (e.g., by waves or other motions), all of the alternate mechanisms are far less efficient than convection at generating a magnetic field.

Cooling is usually the driver for convection in planet interiors. Cooling can provide thermal buoyancy (as in ordinary thermal convection), and can also promote phase transitions that lead to compositional buoyancy (e.g., the growth of the solid inner core of the Earth with a different composition than that of the liquid outer core). Indeed, the crucial reason that the Earth and other planets have a magnetic field is because they are cooling; if the Earth were to stop cooling, the growth of its solid inner core and the generation of its magnetic field would cease. In the case of the Earth, its core is cooling because the mantle is sufficiently cooler than the core (§2.2.3). More generally, whether the interior of a planet can cool requires understanding the planet in totality—including the effect of the atmosphere, if there is a thick one—to assess whether or not its interior is cooling and might potentially harbor a convective electrically conductive fluid dynamo-generation region.

Counterintuitively, a high electrical conductivity can be a disadvantage to driving a dynamo. By the Wiedemann-Franz relation, a metal’s electrical and thermal conductivities are proportional,
since electrons in the metal are responsible for carrying both electrical current and heat. If the electrical conductivity of a planet’s interior is too high, the heat flux will be carried by conduction, the interior will not convect, and there will be no dynamo action. In a homogeneous body, thermal convection requires a heat flow $F_{\text{total}}$ in excess of that which can be transported by thermal conduction along the adiabat $F_{\text{cond,ad}}$.

$$F_{\text{total}} > k \left( \frac{\alpha T g}{C_p} \right) = F_{\text{cond,ad}}.$$  \hspace{1cm} (2.4)

In this expression, $k$ is the thermal conductivity, $\alpha$ is the coefficient of thermal expansion, $T$ is the local temperature, $g$ is the local gravity, and $C_p$ is the heat capacity at constant pressure. In a heterogeneous body, the convective heat transport can be outward or inward because of a larger compositional buoyancy, e.g., inner solid iron core growth or snow/rain up or down depending on the phase diagram. A high thermal conductivity $k$, inadequate cooling, or both can impede convection and turn off a dynamo. The Earth’s energy budget is not too far from this threshold for convection; current estimates place the core heat flow by conduction along the adiabat at about 15 TW, within a factor of 3 of the total heat flow through Earth’s surface, about 44 TW.

All planets are fast rotators from the perspective of dynamo flows. The Rossby number $R_0$,

$$R_0 \equiv \frac{v}{2\Omega L},$$  \hspace{1cm} (2.5)

is a measure of the effect of rotation; small values of $R_0$ imply a large influence of rotation on fluid dynamics. For rotation rates of $\Omega \sim 10^{-4} \text{ s}^{-1}$ (Earth) and $\Omega \sim 10^{-6} \text{ s}^{-1}$ (Venus), the Rossby number is very much less than 1 ($R_0 \sim 10^{-4}$ to $10^{-6}$). Dynamo action requires that the characteristic timescale for fluid motions is short compared to the magnetic diffusion time, the timescale over which the planet’s magnetic field would decay away if the dynamo is suddenly extinguished. The magnetic diffusion timescale for Earth-like planets is

$$\tau \sim 3,000 \text{ yr} \left( \frac{1 \text{m}^2 \text{s}^{-1}}{\lambda} \right) \left( \frac{L}{1,000 \text{ km}} \right)^2.$$  \hspace{1cm} (2.6)

Magnetic diffusion is slow compared to rotation timescales, even for very slowly rotating bodies like Venus. It can be argued that Venus does not have a magnetic field not because of its rotation rate but rather because it is not cooling efficiently. Nonetheless, it is thought that rotation plays an important role in breaking the symmetry of the interior fluid motions and in allowing the dynamo to happen.
There does not currently exist an exact criterion for the presence of a planetary dynamo (e.g., Busse, 2000; Gubbins, 2001). Given a planet, with precise information on its composition, evolution, and heat flow, we cannot say for certain whether or not the planet will have a dynamo. It is suspected that the criterion for a planetary dynamo is close to the criterion for convection, i.e., if the buoyancy flux or heat flow is even mildly in excess of that needed for convection, the resulting convective motions will have the requisite magnetic Reynolds number and complexity. This supposition has yet to be quantified properly, however. Developing a predictive theory of planetary dynamos remains an unsolved problem.

When a dynamo does exist, what determines the field stability (i.e., whether the field is AC, as for the Sun, or DC with reversals, as for the Earth) and magnetic field geometry? While the timescale for magnetic reversals on the Sun is nearly constant and is related to a characteristic dynamical timescale, geomagnetic polarity intervals vary by several orders of magnitude (from $10^4$ yr to $10^7$ yr) and are much longer than the roughly 100-yr timescale for convective overturn in the Earth’s core (e.g., Merrill et al., 1996). For the Earth, it is thought that reversals result from large fluctuations in the structure of the mantle (e.g., Vogt, 1975; Jones, 1977; Glatzmaier et al., 1999; Driscoll & Olson, 2011b). Chaotic reversals also occur in numerical dynamos with stationary parameters, with reversal frequency correlated with core buoyancy flux (Driscoll & Olson, 2009a). The thickness of the electrically conducting convective layer, presence of shearing motions, locations of buoyancy sources, and the nature of boundary conditions can all have an effect on the field geometry and stability. In dynamo models, both dipole-dominant field geometry and multipolar field geometry can be reproduced (e.g., Christensen et al., 1999; Stanley & Bloxham, 2006). Regime boundaries for these two types of field geometries have been suggested based on numerical dynamo models (Stanley & Bloxham, 2006). However, we do not really know how to extrapolate the regime boundaries from numerical models to the low-Ekman number parameter regime of real planets.

Finally, we turn to the question of the amplitude of a planet’s magnetic field. It is obvious from the diversity of the solar system planets that any theory of field magnitude scaling must have a huge variance. Jupiter and Saturn have surprisingly different field strengths, Mercury has a surprisingly weak field, and the paleolunar field is surprisingly large. All of these surprises indicate that the current understanding of planetary magnetic fields is incomplete. There are two kinds of ideas for what determines the field magnitude: force balance (or convective efficiency arguments) and energy scaling (see e.g., Christensen, 2010, for a review). Both types of scaling laws are roughly satisfied in the solar system, but they each have different implications.
The force balance-field magnitude scaling relation is based on the assumption of magnetostrophic balance, in which the magnitude of the Coriolis and Lorentz forces are comparable. Balancing the Coriolis and Lorentz forces leads to an Elsasser number, $\Lambda$,

$$\Lambda \equiv \frac{B^2 \sigma}{2 \rho \Omega},$$

near unity, which in turn implies

$$B \sim \left(\frac{2 \rho \Omega}{\sigma}\right)^{1/2}. \quad (2.8)$$

In this expression, $B$ is a characteristic rms magnetic field strength inside the dynamo generation region, $\rho$ is the fluid density, $\sigma$ is the electrical conductivity, and $\Omega$ is the angular frequency of the planet’s rotation. Thus, this force balance criterion depends both on the electrical conductivity and on rotation. The solar system planets satisfy the force balance criterion, at best, only approximately well. Stevenson (2003) estimated that $\Lambda_{\text{Earth}} \sim \Lambda_{\text{Jupiter}} \sim \Lambda_{\text{Ganymede}} \sim \Lambda_{\text{Saturn}} \sim 0.3$ and $\Lambda_{\text{Neptune}} \sim \Lambda_{\text{Uranus}} \sim 0.01$ at the top of the field generation region, while Schubert & Soderlund (2011) estimated that $\Lambda_{\text{Earth}} \sim 0.1$, $\Lambda_{\text{Jupiter}} \sim 1$, $\Lambda_{\text{Saturn}} \sim 0.01$, $\Lambda_{\text{Ganymede}} \sim 10^{-3}$, $\Lambda_{\text{Neptune}} \sim \Lambda_{\text{Uranus}} \sim 10^{-4}$, and $\Lambda_{\text{Mercury}} \sim 10^{-5}$. The low estimated values of $\Lambda$ for Mercury, Ganymede, Uranus, and Neptune may be indicative of some tension with the assumption of magnetostrophic balance. These estimates of $\Lambda$ differ because of, and are complicated by, the necessity of extrapolating the measured surface magnetic fields down to the field generation layer, and should be considered lower bounds due to contributions from both toroidal field components inside the dynamo region and unresolved multipole components in the poloidal surface fields.

The energy scaling relation suggests an equipartition between the energy density in the magnetic field and the kinetic energy density of the convective fluid motions,

$$\frac{B^2}{2\mu_0} \sim \frac{1}{2} \rho v^2. \quad (2.9)$$

The convective velocity $v$ is related to the convective heat flux $F_{\text{conv}} = F_{\text{total}} - F_{\text{cond,ad}}$ (or the buoyancy equivalent) by mixing length theory,

$$v \sim \left(\frac{1 F_{\text{conv}}}{\rho H_T}\right)^{1/3}, \quad (2.10)$$
where $l$ is the mixing length, and $H_T = C_p/ (\alpha g)$ is the temperature scale height. This energy scaling predicts that the field strength depends on heat flux, but does not depend on electrical conductivity nor the rotation rate. Christensen et al. (2009) have shown that the energy scaling can roughly explain the observed field strengths of planets and rapidly rotating low-mass stars.

Future direct measurements in the mass gap between stars and planets, as well as measurements for objects of different ages, will help to constrain whether the Christensen et al. (2009) energy scaling relation is broadly applicable. For instance, the energy scaling relation suggests that younger brown dwarfs should have stronger magnetic field strengths than older brown dwarfs at the same mass (Reiners & Christensen, 2010).

### 2.2.3 Planetary Thermal Evolution

In the previous section, the importance of cooling (and its consequences) was developed as a crucial criterion for sustaining a planetary dynamo. We now turn to the thermal evolution of jovian, rocky, and envelope-shrouded planets in more depth.

In summary, though obtaining the electrically-conducting fluid convection necessary for dynamo action is “easy” in giant planets, it can be “difficult” in terrestrial planets, because it is easy to imagine conditions under which terrestrial planet dynamos will turn off. Extrasolar planets exhibit a wide diversity of compositions (§2.2.4), and a more expansive study of planetary thermal evolution is warranted to assess which planets might be conducive to dynamos.

#### Gas Giants

Gas giant planets are not static equilibrium bodies, but instead have been gradually cooling and contracting since their initial formation. To leading order, the solar system gas giants are convective throughout their bulk interior, with adiabatic temperature profiles. Microphysical energy transport (radiative diffusion, conduction, molecular diffusion) is insufficient to transport the planet’s intrinsic luminosity (which is dominated by primordial heat, trapped since the planet’s formation) along a sub-adiabatic temperature gradient. The specific entropy along the planet’s adiabat gradually decreases over time as the planet cools. Near the outer boundary of the planet at low pressures, the planet’s atmosphere becomes optically thin, and energy is transported by radiation. It is this atmosphere of the planet that controls the rate at which the planet cools. The planet’s effective temperature (approximately the temperature at the optical depth $\tau \sim 1$ surface) gradually decreases over time.

There are, of course, several important corrections to this simplified picture of adiabatic giant planet evolution. There may be a radiative zone in the planet’s interior between 1,000 K and 3,000 K where molecular hydrogen is rather transparent (Guillot et al., 1994). Compositional gradients can also have an effect in impeding convection and slowing planetary cooling (Leconte
& Chabrier, 2012). In general, homogeneous Kelvin contraction models for the solar system giant planets fare pretty well for Jupiter, but under-predict the intrinsic luminosity of Saturn, and over predict the intrinsic luminosities of Uranus and Neptune (e.g., Fortney et al., 2011).

Hot Jupiters (Jupiter-mass planets on orbits with periods $\lesssim 5$ days) evolve in the presence of a far stronger radiative forcing than the eponymous Jupiter. While the rate at which the solar system planets absorb energy from the Sun is comparable to their intrinsic luminosities, the stellar radiation received by hot Jupiters can be 3–4 orders of magnitude higher. Consequently, the surface radiative zone in irradiated hot Jupiters extends to far greater depths before the increase in opacity with pressure leads the radiative diffusion temperature gradient to become convectively unstable. For a hot Jupiter, its effective temperature at the $\tau \sim 1$ level is set by the equilibrium temperature in the stellar radiation field and does not vary significantly in time as the planet evolves (in contrast to its cooler solar system cousins). Hot Jupiters still cool, nonetheless; the specific entropy of their interior adiabats may decrease in time as the radiative-convective transition evolves to higher pressures. It is thus expected that both hot and cold Jupiters orbiting other stars will likely have the conditions necessary to generate a magnetic field, since they are convective at depth.

**Rocky Planets**

In considering the thermal evolution of rocky planets, namely a planet consisting of an iron core surrounded by a silicate mantle, we take the Earth as the prototypical example.

Both thermal and compositional convection in the iron core of a rocky planet require cooling of the core by the overlying mantle. The core itself does not have any significant heat sources but does have a reservoir of heat left over from planet formation. The dominant heat sources are in the mantle, since radioactive nuclei are lithophile elements, i.e., they have a strong affinity for oxygen and form compounds that do not sink into the iron core as the planet is differentiating. The decay of $^{238}$U, $^{235}$U, $^{232}$Th, and $^{40}$K produces a significant amount of radiogenic heating in Earth’s mantle, about 13 TW (Jaupart et al., 2007; Arevalo et al., 2009). Alternative sources of internal mantle heating include tidal dissipation associated with gravitational tides (e.g., Greenberg, 2009; Heller et al., 2011), gravitational energy released due to ongoing differentiation of the mantle, and latent heat released at phase boundaries in the mantle, including the possibility of a deep magma ocean (Labrosse et al., 2007). In the present-day Earth, these alternative mantle heat sources are thought to be minor compared to radiogenic heat production, although this may not have been the case in the early Earth.

It is generally, although not unanimously, believed that the temperature of the Earth’s mantle has cooled over time because the net heat loss has exceeded the net heat production. (e.g., from the decline of radioactivity in the mantle). Rocky planets with stronger mantle heat sources (e.g., due to higher abundances of radioactive nuclei or tidal heating) may have cores in which cooling
Chapter 2. Searching for & Characterizing Extrasolar Planets

is insufficient to sustain convection or a dynamo. Jupiter’s satellite Io may be the best example of a terrestrial body for which tidal heating of the mantle is likely preventing the core from cooling and suppressing dynamo action. Short orbital period extrasolar planets orbiting close to their host star are also likely to experience intense tidal heating driven by the gravitational torque from the star (Barnes et al., 2013). Tidal forces may also drive core flows by differential motion between the core and solid mantle (Tilgner, 2007; Le Bars et al., 2011; Dwyer et al., 2011), a so-called “precessional dynamo.” However, in a thermally stratified core, tidally driven core flows would be suppressed, so that the threshold for normal thermal convection must already be maintained.

In addition to internal heating, the efficiency of heat transport through the mantle can limit the core cooling rate. The plate tectonics style of mantle convection (“mobile lid”) is thought to be the most efficient long-term mode of mantle heat transport, as cold plates are subducted and hot mantle is exposed to the Earth’s surface temperature at mid-ocean ridges. However, Earth is unique in being the only known planet currently undergoing plate tectonics. All other terrestrial mantles can be categorized as either stagnant lid (Mercury, Venus, Moon, Mars, Ganymede, Callisto, Enceladus, and Titan) or as a heat pipe (Io) style of mantle convection. In a stagnant lid (“single plate”) planet, the mantle may or may not convect under a thick conductive lid, but the upwelling mantle does not reach the surface and heat loss is limited by conduction through the lid (Solomatov & Moresi, 2000). Alternatively, Io’s heat pipe mechanism removes heat by hot spot volcanism surrounded by slow lithospheric subsidence (or drips) (O’Reilly & Davies, 1981; Moore, 2003; Breuer & Moore, 2007). It is widely expected that heat loss through a stagnant lid is less than that through a mobile lid, and that this factor alone can prevent a core dynamo from operating (Nimmo, 2002; Stevenson, 2010), although the additional heat loss associated with large scale mantle melt eruption (e.g., heat pipe) is less clear. It has been proposed that Venus operates in a sporadic heat pipe mode (Turcotte, 1989) in which mantle heat is released in massive but rare volcanic episodes. Maintaining a heat pipe style of mantle cooling may require a large internal mantle heat source (e.g., tidal heating of Io), so that a slowly cooling terrestrial planet heated by radiogenic decay alone may favor heat conduction through a mobile or stagnant lid, with only occasional volcanic heat loss. Thus, plate tectonics may aid magnetic field generation in terrestrial planets by enabling efficient cooling, but is not a required condition for magnetic field generation (as the examples of Mercury and Ganymede illustrate).

Gas-Laden Super-Earth Planets

Planets are now being discovered around distant stars that bridge the gap between the gas giant and rocky planet categories found in the solar system. A new class of low-density ($\lesssim 1$ g cm$^{-3}$), low mass (a few M$_{\oplus}$) planets have been identified by Kepler on orbits interior to 0.1 au. One possible scenario for these planets is that they formed in situ in a massive protoplanetary disk, wherein rocky planet embryos assembled on approximately 1-Myr timescales managed to accrete
H and He before the gas disk dispersed (Hansen & Murray, 2012; Chiang & Laughlin, 2013). In this case, the planets would consist of rocky cores surrounded by H/He-dominated envelopes.

How would a strongly irradiated optically thick H/He envelope surrounding an Earth-like rocky-composition interior affect the planet’s thermal evolution and potential to sustain a dynamo magnetic field? In a rough sense, this is a hybrid scenario between the jovian and rocky planet thermal evolution histories discussed above. If the H/He envelope is sufficiently thick, the temperature and pressure conditions at its base may be such that the underlying silicates are in the liquid phase (magma ocean). In this scenario, the entropy of the magma ocean adiabat is directly linked to the entropy of the H/He layer, through the boundary conditions at the interface. This situation is very different from the case of the Earth, where the mantle is undergoing solid convection along an adiabat determined by the rheological properties of the silicates—Earth’s mantle is just hot enough so that it can flow to effectively eliminate the Earth’s interior heat flux. On the Earth, the decline of radiogenic heat sources over time allows the mantle to cool, which in turn allows the core to cool. In contrast, the mantle and core cooling rates of H/He-shrouded planets are controlled by the atmosphere. Strongly-irradiated H/He-shrouded super-Earths may evolve in a manner similar to hot Jupiters, with a deepening of the surface radiative zone over time enabling cooling even though the surface temperature at the $\tau \sim 1$ level is fixed by the stellar irradiation. Quantitative evolution studies of this scenario need to be pursued, to assess whether or not the H/He envelope will allow sufficient cooling to drive convection (and perhaps a dynamo) in the planet’s iron core.

### 2.2.4 Extrasolar Planets & Their Magnetic Fields

The extrasolar planets discovered to date indicate a tremendous degree of diversity, and their large number\(^7\) has opened the possibility of studying the magnetic fields and interior structures of planetary bodies under a wide range of conditions.

The initial discoveries favored hot Jupiters—Jovian-size planets on close-in orbits of a few days—as they are easiest to detect. Ground-based radial velocity surveys (Howard et al., 2010), planet microlensing surveys (Sumi et al., 2010), and the space-based Kepler transit survey (Fressin et al., 2013; Petigura et al., 2013) have subsequently revealed that the planet occurrence rate increases toward smaller planet masses and radii. One of the great surprises revealed by the Kepler mission was the high abundance of low-density sub-Neptune-size planets (e.g., Kepler-11b,c,d,e,f,g, Lissauer et al., 2011). It is an outstanding question as to whether these planets formed in situ near where they are currently observed, or whether they formed beyond the snow line and later migrated in. At the other density extreme, close-in rocky planets on orbital periods less than one day have also been discovered (e.g., Kepler-10b, CoRoT-7b, KOI1843.03, KIC 8435766).

---

\(^7\) More than 5,000 planets and planet candidates have been discovered (http://exoplanets.org/; Extrasolar Planet Encyclopaedia).
Figure 2.3: Planetary mass-radius relation for planets having independently determined masses (abscissa) and radii (ordinate). Black points represent extrasolar planets, red stars indicate the position of the solar system planets. The colored curves are theoretical mass-radius relations for constant planet compositions from Seager et al., (2007): hydrogen (cyan), pure water ice (blue), Earth-like composition (32% Fe, 68% silicate, brown), and pure iron (grey).

With sub-stellar point temperatures on the order of 2,000 K, these scorched worlds may have atmospheres of silicate vapor (Castan & Menou, 2011) and large-scale surface magma oceans (Léger et al., 2011). It is clear that the planets in the solar system represent only a subset of the possible planetary structures.

The main material building blocks from which planets form can be divided into three categories based on their condensation temperatures in a protoplanetary disk: rock (refractory material), astrophysical ices (such as H₂O, NH₃, and CO, which condense in the outer regions of the disk at least a few au from the star), and H/He gas. Low density condensibles (water and other astrophysical ices) are a tracer of a planet’s formation location: planets formed beyond the snow line are expected to initially contain an ice mass fraction comparable to the rock mass fraction, while planets formed on the close-in orbits ($P < 50$ days) are expected to only have trace amounts of astrophysical ices. On the other hand, the presence of H/He gas is a tracer of the timing of a planet’s formation; to accrete H/He from the protoplanetary disk, planets must form before the gas disk dissipates (within a few million years).

Planets with constraints on both the mass and radius are a valuable subsample for constraining planetary interior structure (Figure 2.2). The planet’s radius (measured from the transit depth) and the planet’s mass (measured from the radial velocity amplitude or transit timing variations)
together give the planet’s density and some handle on the planet’s composition. Inferring planetary compositions from measured planet masses and radii is a challenging under-constrained inversion problem, however, because planets with disparate interior compositions can have identical masses and radii (e.g., Valencia et al., 2007; Rogers & Seager, 2010a). The transiting super-Earth GJ 1214b (Charbonneau et al., 2009) provides a striking example of these degeneracies. Based on its measured mass and radius alone (6.5 \( M_\oplus \), 2.7 \( R_\oplus \)), GJ 1214b could have very different interior properties: it could be a miniature Neptune with an interior of ice and rock surrounded by a primordial H/He envelope containing 0.01%–5% of the planet’s mass; or it could be a water planet (composed of at least 47% \( H_2O \) by mass) shrouded in a layer of vapor from sublimated ices; or a terrestrial super-Earth harboring a hydrogen-rich outgassed atmosphere (Rogers & Seager, 2010b).

For planets with favorable transit depths orbiting the brightest stars, transit transmission spectroscopy provides insights into the composition of the planet’s upper atmosphere, which in turn can help inform models of planetary composition, if not break all degeneracies. Measuring the change in eclipse depth across spectral lines \( \Delta D \),

\[
\Delta D \propto \frac{H R_p}{R_* R_p},
\]

(2.11)
gives a measure of the planet atmospheric pressure scale height \( H \),

\[
H = \frac{kT}{\mu m g},
\]

(2.12)

which depends upon the planet’s surface gravity \( g \), its atmospheric temperature \( T \), and the mean molecular weight of the atmosphere \( \mu_m \). Low mean molecular weight (e.g., H/He-dominated) planet envelopes may have deeper features in their absorption spectra, while high mean molecular weight atmospheres (e.g., \( H_2O \)- or \( CO_2 \)-dominated) will have flatter spectra. High-level clouds and hazes can obfuscate the transmission spectrum constraints on the planet’s atmospheric composition, however. Kreidberg et al. (2014) observed 12 transits of GJ 1214b with HST-WFC3, and measured a flat transmission spectrum, despite having the precision to detect spectral features in cloud-free atmospheres dominated by \( H_2O \), \( CH_4 \), \( N_2 \), \( CO \), or \( CO_2 \). GJ 1214b’s atmosphere must contain clouds at the 0.001 to 1 mbar pressure level.

With the advent of the next generation space-based transit surveys (TESS, CHEOPS, and Plato) the number of transiting planets orbiting bright targets that are amenable to atmospheric spectroscopy and ground-based radial velocity follow-up will increase. In the future era of large aperture space telescopes (§4.3), the masses of transiting planets may even be constrained from
high quality transit spectroscopy alone (in the absence of radial velocity measurements) in cases where the atmospheric scale height, temperature, radius, and mean molecular weight of the planet’s atmosphere are all measured (de Wit & Seager, 2013). Degeneracies will remain, however, in the interior structure, insights into which cannot be inferred from a planet’s mass, radius, and atmospheric composition alone.

Measurement of extrasolar planet magnetic fields will add an extra dimension to the observables available to characterize planets. The planet mass-radius-incident flux diagram will be extended to a planet mass-radius-incident flux-magnetic field strength diagram. How can a measurement of magnetic field constrain planet structure and evolution, given the diversity of possible extrasolar planet compositions and our incomplete understanding of planetary magnetic field generation? The forward problem of relating known planetary interior structure to expected magnetic field is itself a challenge, let alone the inverse problem.

Most robustly, the detection of a magnetic field on an extrasolar planet indicates the presence of a convecting, electrically conducting fluid in the planet. Depending on planet mass and composition, different materials and pressure depths within the planet can contribute to the electrically conducting dynamo generation region.

Rocky planets (e.g., Mercury and Earth, §2.2.1) can have electrically conducting liquid iron cores. The presence of a liquid iron core on a rocky planet is not guaranteed, however. If the oxygen fugacity is sufficiently high during the rocky planet’s assembly, the planet’s iron complement may be oxidized and incorporated into the silicate mantle, instead of being differentiated into a central core (Elkins-Tanton & Seager, 2008). Partial solidification of the iron core may also limit the range of planet masses with sufficiently large liquid cores to sustain a dynamo. The extent to which an iron core solidifies is sensitive to the presence of volatiles or other minor constituents mixed into the iron core that have a substantial impact on the iron phase diagram. In addition, high-pressure high-temperature silicate melts may also be sufficiently electrically conducting (Tyburczy & Fisler, 1995) to support a dynamo in a magma ocean on young/hot, vigorously convecting rocky planets (Stevenson, 2001; Ziegler & Stegman, 2013).

In icy planets like Neptune and Uranus, water is electrically conducting above a few thousand Kelvin, in the ionic, plasma, and super-ionic solid phases (Cavazzoni et al., 1999; Lee et al., 2006; French et al., 2009). Finally, in jovian planets with massive H/He envelopes, hydrogen becomes metallic above ~25 GPa (Wigner & Huntington, 1935), and can support a dynamo.

In addition to the existence of an electrically conducting fluid region, the detection of a planetary magnetic field further indicates that the electrically conductive fluid must be convecting. This presence of convection might be most informative for small planets. As described in Section 2.2.3, while jovian planets are expected to generally be convective at depth, the energy budget for convection in Earth’s core is marginal. Higher equilibrium temperature (\(\gtrsim 1,500\) K), stronger tidal heating of the mantle, higher concentrations of radioactive nuclei in the mantle, the presence
of a thick H/He envelope, or a stagnant lid tectonic regime could turn off convection (and hence a dynamo) in the core of an otherwise Earth-like planet. The inference of convection in a planet via a magnetic field measurement would give an important new insight into the planet’s thermal evolution and energy budget, and may also serve as an indirect indication of plate tectonics.

Looking beyond a simple detection of an extrasolar planetary magnetic field, what could be learned from a measurement of the magnetic field strength? The field strength measured at the planet’s surface reflects both the field at the dynamo source region and its modification by the over-lying planet layers. The range of possible fields emanating from a dynamo region is uncertain. Even if the dynamo is strong (i.e., the field is strong in the dynamo generation region), the measured surface field can be small due to the radial fall-off of the magnetic field (proportional to $r^{-3}$ for a dipole). Measurements of planetary magnetic field strengths may provide the strongest insights into planetary interior structure when the magnetic field strengths are extreme. For example, a 10 G extrasolar planetary magnetic field measurement may indicate that the dynamo source region (e.g., iron core in rocky planets) is close to the surface. At the other extreme, if a weak magnetic field strength is measured, it might be argued that the radius of the convecting region is small compared to the radius of the planet.

We turn now to the topic of what could be learned if the geometry (multipolar structure) of an extrasolar planet’s magnetic field could be measured. If field geometry is somehow accessible to future measurements, it may yield insights into the location and geometry of, as well as ongoing phase transitions in, the dynamo generation region. Since higher multipole components of the field decay more quickly with distance, the measured ratio of quadruple to dipole may be related to the location of the field generation region relative to the planet surface (with dipole-dominated fields providing tentative evidence for a deeper dynamo). It has also been proposed that the geometry of the dynamo generation region (the thickness-to-radius aspect ratio) is an important factor in determining the intrinsic multipolar structure produced by the dynamo (Stanley & Bloxham, 2006); quadropole-dominated fields may provide evidence of a thin-shell geometry for the planet dynamo generation region. Finally, the multipolar structure of a planet’s magnetic field may also be a signature of the source of the buoyancy flux driving convection. Models of Ganymede (Zhan & Schubert, 2012) have found that Fe snow may lead to a multipolar field due to the buoyancy release close to the surface, while upward flotation of FeS could lead to a dipole-dominated field.

In addition to opening a new window into the interior structure and evolution of individual planets, a statistical sample of extrasolar planetary magnetic field measurements will also open the possibility of constraining dynamo theory. Measuring the magnetic field strengths of a wide diversity of compositions and structures may help to distinguish between the force balance and energy scaling relations for planetary magnetic fields (§2.2.2), each of which work decently well for the planets in the solar system. If there is an observational technique that will yield insights into the multipolar structure of extrasolar planet magnetic fields, determining whether planetary
magnetic fields are primarily dipolar versus quadrupolar will provide useful empirical information moving forward with dynamo theory. Furthermore, correlating measurements of magnetic field strength (and possibly geometry) with other planet properties (such as planet mass, density, age, rotation rate, orbital eccentricity, orbital period, and equilibrium temperature) will provide useful tests of the plausibility arguments developed to explain the diversity of solar system planet magnetic fields. The study of extrasolar planet magnetic fields will likely be a phenomenologically driven science at first. Accumulating a body of statistics on extrasolar planet magnetic fields will help to improve our understanding of our own solar system.

2.3 Rotation & Planetary Formation

If a planet has a magnetic field, and that field is not aligned with the rotation axis (or if it has a multipolar component), then it is possible to determine the planet’s rotation period from time-variation of the radio emissions. This information can constrain the formation and evolution history of the planet, at least for planets at large orbital radii.

Planets close to the star will tidally despin to synchronicity (orbital period = spin period) in many cases. However, even planets that have Venus-like or Mercury-like orbits will not necessarily despin to this state. Mercury is in a 3:2 spin-orbit resonance, and Venus has competing atmospheric and solid body tides that have led to a rotation period that is not simply related to the orbital period. Collecting examples of planets that have undergone tidal evolution of their spin but are not synchronous can inform us about their structure or orbital evolution.

Planets that have not undergone substantial tidal despinning have rotation periods that constrain theories of planetary formation, perhaps even the existence of moons. Earth could have a rotation period as short as 5 hours, but it rotates more slowly due to the presence of a large moon, such that 80% of the angular momentum of the Earth-Moon system is in the orbital motion of the Moon. Jupiter and Saturn are not close to rotational break-up even though the likely angular momentum budget at their time of formation was more than sufficient to create this state. It is likely that their slower rotation is the result of these bodies having formed disks (“miniature solar systems”) that were the sites of large satellite formation as well as reservoirs of angular momentum. These disk are affected by the orbital location (more distant planets can form larger disks because of a larger Hill radius).

At present, the theories for planetary spin admit many interpretations and possibilities. The accumulation of rotation information will not by itself solve the puzzles of planet formation and the possible existence of moons, but will be an important part of solving these puzzles.
2.4 Planetary Habitability: Magnetic Fields & Radiation Protection

Before considering the extent to which magnetic fields are important for planetary habitability, the concept of "habitability" must be defined. Discussions of habitability in an extrasolar context tend to have much narrower criteria than for the solar system. The extrasolar search for habitable planets can only consider signatures that can be detected via remote sensing; in situ examination of extrasolar planets is beyond the horizon of this report. In practice, this search focuses on planets that might support liquid water on their surfaces; the "habitable zone" defined by Kasting et al. (1993) refers to the range of orbital distances at which a planet with an Earth-like atmosphere will have surface temperatures that allow liquid water.

In contrast, within the solar system, more information is available through the combination of relatively local remote sensing and in situ measurements, revealing that diverse environments—such as Europa and Enceladus with their sub-surface oceans—are potentially habitable. Motivated by solar system discoveries, Lammer et al. (2009) divide habitable planets into four classes. Their Class I, defined as "bodies on which stellar and geophysical conditions allow Earth-analog planets to evolve so that complex multi-cellular life forms may originate," is generally most relevant for extrasolar planets. (Implicit in this definition is the criterion that the planetary surface conditions allow for liquid water over geological time scales.) Also of potential relevance to magnetic field detections are Class II planets, defined as "planets where, in the beginning, life may evolve because these planets start out on an evolution path similar to Class I types, but due to different stellar and geophysical conditions the planetary environments and life evolve differently than on Earth."

The space environments around the terrestrial planets in the solar system are harsh, and similar conditions are expected around extrasolar planets, due to the presence of high energy charged particle radiation in the form of the ever-present Galactic cosmic rays and solar (or stellar) energetic particle events. If directly incident on a surface, such radiation is expected to be destructive (harmful) to Earth-like biological tissue. In light of the diversity of potentially habitable environments within the solar system, we cannot state that planetary magnetic fields are a requirement for habitability, but the arguments presented below suggest that magnetic fields may facilitate favorable conditions for the emergence of life by protecting the surface from cosmic rays, shielding the atmosphere from violent chemical changes due to solar energetic protons, and potentially helping to retain water in the atmosphere.

The absence of a planetary-scale magnetic field may be an important piece of evidence in understanding or classifying a planet as Class I or potentially Class II. However, the KISS Study team also found that the effectiveness of magnetic fields for atmospheric retention remains ambiguous and merits further study.
2.4.1 High Energy Charged Particle Environment

In the solar system, there are two sources of high energy charged particles. Figure 2.3 shows the spectrum of these two primary sources. The first source is Galactic cosmic rays (GCRs) that consist of electrons, protons, and more massive ions at energies 10–10 000 MeV. The source of these charged particles is still under study, but they are thought mostly to originate in Galactic supernovae.\(^8\) The GCRs in the solar system represent an integrated effect from all Galactic sources. By analogy, other stellar systems should also be infused with GCRs, though the flux into a stellar system may vary with position within the Galaxy.

The GCR flux at Earth is modulated on decadal scales by the solar cycle. During solar maximum, the inner heliospheric magnetic field is at its largest time-averaged value in the cycle, which can reduce the GCR penetration to the inner solar system by as much as 50%. In contrast, at solar minimum, the magnetic field of the inner solar system is at its lowest time-averaged value,

\(^8\)Even higher energy particles, presumed to be extragalactic, are detected, but their flux is sufficiently low that they are unimportant for this discussion.
allowing relatively greater GCR flux to penetrate deeper into the inner heliosphere. Figure 2.4 shows the GCR proton flux (> 60 MeV), representing the quasi-constant background level, which exhibits modulation on timescales of a solar cycle.

Figure 2.5: Variation of energetic protons over 20 years of IMP-8 observations (from Stone et al., 1998). Both quasi-continuous GCR and impulsive SEP effects are captured. The former has a minimum in flux near solar maximum (1981, 1992, ...). The impulsive SEP events also tend to be clustered near solar maximum.

The second source of energetic particles is from the Sun/star itself. Such events in the solar system are called “solar energetic particles” (SEPs), which are characterized by 2–3 day episodes of enhanced emission of electrons and protons in the energy range of 10s to 1,000s of MeVs. Figure 2.4 shows the measured proton spectrum during the intense 1972 August 4 event. SEP events are found in association with solar storms, and specifically in the reconnection-related release of coronal mass ejections (CMEs). As such, episodes of these temporally-impulsive events tend to have a peak in occurrence during solar maximum, when CMEs tend to be most frequent (Figure 2.5).

The exact source of the SEPs and their relationship to the CME is currently a point of debate. Relativistic energetic particles are often detected simultaneously at the launch of a CME from the Sun (a possible magnetic reconnection). However, SEP fluxes can, in some cases, remain high for
3–5 days as the CME’s preceding shock propagates outward to 1 au. In some, but not all, cases, as the CME passes an observing spacecraft, a local increase in energetic particles called energetic storm particles (ESP) can be detected in association with the CME shock. These observations suggest that there may be multiple sources of energization associated with the SEPs: at the time of CME launch in association with magnetic disconnection and in the CME-driven shock itself.

For an exposed rocky surface with no atmosphere or large-scale magnetic field, both GCRs and SEPs are directly incident on the surface. The incoming charged particle nuclei interact with existing surface nuclei and can create secondary neutrons. As a consequence, the surface itself becomes an emitter of energetic neutrons ([Feldman et al.], 1998), which can themselves be extremely harmful since such neutrons interact most easily with hydrogen-bearing compounds (such as water and hydrocarbons). This secondary emission represents a third potent element to the radiation environment.

### 2.4.2 Planetary Dipole Magnetic Fields

Large scale planetary magnetic fields created by dynamos effectively protect a planetary surface, and potentially the atmosphere, from a large portion of high energy charged particle radiation. The field can shield GCRs and SEPs by deflecting the particles via the Lorentz $\mathbf{v} \times \mathbf{B}$ forces. In contrast to a planet without a magnetic field, the charged particle penetration to the planetary surface becomes a function of particle entry, dipole strength, and (magnetic) latitude.

Charged particle-planetary magnetic dipole interactions were first studied by in the early 1900s by C. Stormer. The Hamiltonian is derived for a particle under the influence of a dipole $\mathbf{B}$, and the magnetic potential and azimuthal particle momentum are combined to form an “effective potential” $\psi$. Accessible regions of space (allowed and forbidden trajectories) are dependent upon the particle’s total energy $W$, azimuthal momentum $p_\phi$, and dipole strength $M$. Figure 2.5 shows a generalized view of the allowed and forbidden regions for a planetary magnetic dipole.

As illustrated, there exists a set of particle energies and dipole strengths that allow charged particles access to the atmosphere, especially in polar regions. The boundary of the accessible region is ([Shepherd & Kress], 2007)

$$\rho = C_{st} \frac{\cos^2 \lambda}{1 + \sqrt{1 + \cos^2 \lambda}},$$  \hspace{1cm} (2.13)

where $\lambda$ is the magnetic latitude and an “effective blocking length” at the equator for a particle of mass $m$, charge $q$, and velocity $v$ is given by $C_{st} = \sqrt{M\mu_0 q/4\pi\gamma mv}$. In practice, the actual demarcation of the forbidden zone is about $0.4C_{st}$ and forms a toroid ([Shepherd & Kress], 2007). As the particle energy is increased, the forbidden region is reduced in extent both radially and
Figure 2.6: Forbidden and allowed charged particle trajectories in a dipole field, in a frame in which planetary magnetic moment $M$ defines the $z$ axis and the radial coordinate is $\rho$. Critical surfaces defined by the particle’s energy $W$ being equal to the “effective potential” $\psi$ are shown, for three different values of the particle’s azimuthal momentum $p_\phi$. (right) For $p_\phi < 0$, the particle is reflected and does not enter the magnetosphere. (middle) For modest, positive momenta, the particle can enter the magnetic polar regions. (left) For large, positive momenta, the allowed trajectories split into two domains, one trapped within the magnetosphere, which accounts for planetary radiation belts. [Figure from Gurnett & Bhattacharjee (2005).]

in latitude, allowing greater particle access to the upper atmosphere usually at higher magnetic latitudes (Figure 2.5, middle panel). Thus, a strong planetary dipole magnetic field can block a large portion of the incident energetic particle flux, with the protection being most effective at the magnetic equator.

2.4.3 Planetary Atmosphere Loss

Orbital distance, typically used to define the habitable zone, is not the only condition for liquid surface water; a planet must also have an atmosphere with sufficient surface pressure and water content. Magnetic fields may play a role in retention of planetary atmospheres, thereby improving the likelihood of liquid surface water. The KISS Study found that the effectiveness of magnetic fields for atmospheric retention remains ambiguous and merits further study.

As discussed in §2.4.2 and §2.6, a planet’s magnetic field can shield its atmosphere by diverting cosmic rays, the incident stellar wind, and CMEs so that they mainly affect the planet’s atmosphere at the magnetic poles. Stormer theory predicts that external low energy ($\ll 100$ keV) solar and stellar wind will be deflected by a dipolar magnetic field. Spacecraft observations indeed confirm that the solar wind at the bow or nose of a magnetosphere forms a collisionless shock that slows/stagnates the incoming solar wind flow. The bulk of the plasma is then deflected around the magnetospheric cavity. As such, it seems plausible that a global magnetic dipole field reduces a planet’s rate of atmospheric mass loss, in particular helping to retain the hydrogen and
oxygen ions that make up water. Recent, dramatic evidence in favor of atmospheric shielding by a planetary magnetic field is provided by observations with the Martian Atmosphere and Volatile Evolution (MAVEN) spacecraft of the response of the Martian atmosphere as it is impacted by a coronal mass ejection (Jakosky et al., 2015a).

At first glance, a comparison of Earth to Mars and Venus supports the argument that magnetic fields can prevent loss of water from planetary atmospheres (Lundin et al., 2007; see also §2.2.1). Earth, with its strong dipole field, has a rich atmosphere that allows liquid surface water and sustains life. Mars, which at present lacks a strong global dipole field, has an atmospheric pressure less than 1% that of the Earth, but surface magnetization observed by the Mars Global Surveyor and surface morphology such as river and lake beds provide evidence that both a strong global magnetic field and surface liquid water existed on Mars about 4 billion years ago. The Venusian atmosphere, unprotected by a global magnetic field, has a surface pressure 90 times that of Earth, but with much less water content.

The apparent contradiction that Venus currently has a thick atmosphere but has suffered significant atmospheric (water) loss can be resolved by noting that hydrogen is the species preferentially lost because of its small mass and large scale height. Early water on Venus could have been disassociated, with the H lost to space and the O absorbed into crustal rocks. Without substantial amounts of water, CO$_2$ emitted into the Venusian atmosphere (e.g., by vulcanism) could not be dissolved and re-incorporated into rocks. Indeed, Venus and Earth have comparable carbon inventories, and Earth would have a substantially CO$_2$-dominated atmosphere were it not for the effect of water (in Earth’s oceans) in removing CO$_2$ from the atmosphere. Therefore, it is the escape of H that has left Venus dry, and Driscoll & Bercovici (2013) have proposed that Venus could have retained more H (and water) with a strong planetary magnetic field. However, if H escaped early in Venus’ history, then it is difficult to infer whether a global magnetic field, if one existed, slowed this process.

In order to consider whether magnetic fields would have affected the loss of water from Mars and Venus, it is useful to think of atmospheric escape in terms of the escape limiting process (i.e., the bottleneck). Often, there are several physical escape mechanisms occurring simultaneously, but one of them is a limiting factor. The two common limiting factors for escape are (i) the diffusion limit, at which escape is limited by the rate at which the escaping species can diffuse upwards to the escaping region (Hunten & Donahue, 1976), and (ii) the hydrodynamic limit, at which escape is limited by the rate at which the atmosphere can expand and blow off (Watson et al., 1981). Hydrodynamic processes depend on the escape velocity, which is significantly lower for Mars than for Venus and Earth, which may explain the lower net atmospheric pressure of Mars compared to the other two terrestrial planets. In addition to the two standard limits to escape, a planetary magnetic field could limit the rate of escape. Driscoll & Bercovici (2013) demonstrated that, if magnetic-limited escape is controlled by the density and erosion rate of H at the magnetopause, then the escape rate decreases with increasing planetary magnetic moment. This calculation
assumes that Kelvin-Helmholtz instabilities along the magnetopause are responsible for the H transport (Wolff et al., 1980; Brain et al., 2010), although other sweeping mechanisms could enhance this effect (Strangeway et al., 2005).

Recent, unpublished work (R. Strangeway, 2014, private communication) suggests that Venus, Earth, and Mars all have similar present-day O$^+$ mass loss rates, between $10^{24}$ and $10^{26}$ s$^{-1}$, implying that the net oxygen (and hence water) loss rate does not depend strongly on the presence of a magnetic field. Studies of the terrestrial polar ionospheric regions, where the solar wind is directly accessible via connecting magnetic field lines, indicate substantial and enhanced atmospheric (O$^+$) loss via plasma heating and ambipolar processes (Moore & Khazanov, 2010). Contrary to initial expectations, it may be that Earth’s strong magnetic dipole field enhances the atmospheric loss rate, rather than reducing it, through the process of magnetic reconnection and by increasing our planet’s cross-section for interaction with the solar wind. Reconnection at the dayside magnetopause is an effective mechanism for heating the polar atmosphere, thereby producing an outflowing polar wind that is a sink for atmospheric oxygen.

On the other hand, magnetospheric recycling is expected to occur to reduce the net loss rate. Seki et al. (2001) found that the outflow rate of O$^+$ ions from Earth’s ionosphere into the magnetosphere is an order of magnitude higher than the escape rate from the magnetosphere into interplanetary space, suggesting that a reservoir of atmospheric ions is stored in the lobe regions of the geomagnetic tail that convects inward to repopulate the plasmasphere. Atmospheric losses are then associated with magnetospheric transport and loss processes. However, Seki et al. (2001) caution that they have only examined four methods of escape from the magnetosphere into interplanetary space, and that they only studied ions with energies greater than 50 eV; lower-energy ions may constitute a significant element for balancing the escape budget.

Comparison of the present day escape rates among the terrestrial planets is complicated by the fact that they have such different atmospheres. For example, the escape rates may have been very different for most of their evolution, but are similar today because there is little H left on Venus and Mars, and Earth has most H locked up in the oceans. The influence of a planetary magnetic field on atmospheric escape is also muddied by the fact that the terrestrial atmospheres are so different today, and the influence of the magnetic field cannot be easily disentangled from other escape processes. Ideally, we could compare two planets with identical atmospheres, but with very different magnetic fields. Perhaps exoplanet studies in the future could provide such a comparison.

The efficiency of polar plasma losses, geotail storage, and magnetospheric recycling is a likely function of the specific exoplanet magnetospheric configuration. Effects are also likely to be time variable: atmospheric loss could be expected to increase during a geomagnetic storm when the storing tail magnetically disconnects from the magnetosphere. Thus, fundamental work remains to be performed to determine atmospheric loss processes, including the steady loss rates and
impulsive disconnection rates for magnetic and non-magnetic planets. The erosion and escape processes at the magnetic outer planets have not been thoroughly investigated. The MAVEN mission\(^9\) has the specific objective to investigate atmospheric loss in regions with direct access to the solar wind and in regions that are magnetically connected by the strong Martian remanent magnetic field (Jakosky et al., 2015b). As Figure 2.6 illustrates, such magnetic regions are susceptible to local impulsive magnetic disconnection and atmospheric escape (Crider et al., 2005; Brain et al., 2010).

Figure 2.7: Schematic representation of local magnetic disconnection and plasma detachment at Mars. The trajectory of the Mars Global Surveyor spacecraft is the dotted line, and the Sun is to the left in both panels. [Figure taken from Brain et al. (2010), copyright (2010) by the American Geophysical Union and reproduced with permission.] (a) Magnetic field lines are stretched by the solar wind but are still anchored to the crust. (b) Magnetic loops have detached from the crust, carrying with them ionospheric plasma.

Finally, whether or not magnetic fields have a net effect on the geological stability of a planetary atmosphere, moving the locus of O\(^+\) escape to the magnetic poles and away from the general atmosphere may have profound effects on planetary chemistry. For example, on Mars the present UV photodissociation of water (followed by loss of hydrogen via Jeans escape) leaves no net change in redox state because of a 2:1 feedback in H:O escape and sputtering (McElroy & Yung, 1976). Moving the atmospheric “escape hatch” to the poles will force the buildup of a net reservoir of O\(^+\) in the atmosphere, leading to the buildup of ozone screens (O\(_3\)) and perhaps even some amounts of free O\(_2\). This condition may have happened on ancient Mars around 4 billion years ago, when Mars had a strong planetary dynamo (Kirschvink et al., 1997; Acuña et al., 1998; Weiss et al., 2000, 2002) during the time of aqueous deposition of the carbonate blebs in the cracks of the orthopyroxene breccia (Kirschvink et al., 1997; Valley et al., 1997; Thomas-Keprta et al., 2009; Halevy et al., 2011). In these same carbonates, Farquhar et al. (1998) detected the telltale signature of mass-independent isotope fractionation produced by the cycling of oxygen through an

\(^9\)http://www.nasa.gov/mission_pages/maven/
atmospheric ozone stage, implying that early Mars had environmental redox gasses that ranged from pure H\(_2\) through molecular O\(_2\). As the bulk of Tharsis was in existence during this time (Phillips et al., 2001), and its 21-km height extends well above the approximate 11-km scale height of the Martian atmosphere, this ozone layer impinged directly on the Martian surface in the Tharsis area. A situation of this sort would provide a full cascade of redox-active compounds which might help promote the origin of the electron-transport steps of metabolism.

### 2.4.4 Biological Consequences

Penetrating high energy radiation can destroy biological tissue. The interaction can be direct, with the radiation’s incident energy altering DNA molecules and other supporting cell molecules and creating abnormal cell structures and mutations. However, the interaction is predominantly indirect, with radiation-created free radicals formed in intercellular water reacting with the key proteins, DNA, and RNA to induce cell death. There is great uncertainty in modeling the human reaction to space radiation because of insufficient data to estimate the body’s reaction to the protracted low dose rates found in the space environment. Uncertainties also lie in individual variations in biological repair (Hu et al., 2009) and the effects of heavy ions (Cucinotta et al., 2010).

Generally, the constant GCR environment is considered energetic and thus difficult to shield. However, it contributes a relatively low dose. During the Mars Science Laboratory/Curiosity cruise phase, the GCR dose rate was found to be 1.8 mSv d\(^{-1}\) (Zeitlin et al., 2013), comparable to a daily diagnostic CT scan (Bushberg, 2016). While not giving rise to acute radiation sickness (ARS), the dose rate when integrated over a 1-yr duration Mars mission is close to the astronaut career exposure limit based on an increased risk of later-life cancers (Cucinotta et al., 2010; Zeitlin et al., 2013).

In contrast, SEPs are of lower energy and thus more easily shielded by metal layering. However, the SEP flux is intense and impulsive, and dose rates can be many orders of magnitude higher than for GCRs (Figure 2.7). The variations in flux levels and energy spectra between the impulsive events (e.g., Figure 2.4) make reliable prediction difficult. In extreme cases, the dose can be harmful to exposed humans. For example, modeling of the extreme 1972 August 4 SEP event (Figure 2.3) suggests that an exposed astronaut (e.g., during an extravehicular activity or EVA) would have experienced recoverable but performance-degrading ARS with a mortality rate of less than 0.1% (Hu et al., 2009; Cucinotta et al., 2010). However, an event twice this intensity (like the energetic particle event in the 1859 Carrington solar storm (McCracken et al., 2001; Cliver & Svalgaard, 2004) would generate severe ARS for an exposed human, having a mortality probability near 4% (Hu et al., 2009). Shielding is effective for stopping SEPs, with 5 g cm\(^{-2}\) of Al reducing dose rates by about a factor of 3–5 (Hu et al., 2009; Cucinotta et al., 2010).

Drawing an analogy to stellar systems, the penetrating GCR flux is expected to have near-constant low dose similar to that of the solar system. However, stellar energetic particle activity from
stellar storms should be expected to be highly variable, and possibly even more extreme than in the solar system, creating acute lethal events.

Could a sufficiently thick atmosphere provide enough shielding, even in the absence of a planetary-scale magnetic field? Considering the 1972 August 4 event (Reagan et al., 1983), SEPs appeared capable of producing chemical and ionization effects down to altitudes of about 20 km (pressure ~10 mb) in the polar regions ([magnetic] invariant latitudes above about 60°). Thus, to shield a surface from energetic particle penetration associated with similar magnitude events, the atmosphere would have to be at least 10 mb at the surface. In the case of Mars (surface atmospheric pressure ≈6 mb), the 1972 August 4 event likely had an effect over most of the planet’s surface. Moreover, the 1972 August 4 is not the strongest known energetic particle event (e.g., the 1859 Carrington event). A stellar energetic particle event that was 10× more energetic, and with 10× the flux, might be able to produce effects even at the surface of a planet with an atmosphere as thick as the Earth’s.
2.4.5 Extensions to Extrasolar Planets

High Energy Particle Environment

In considering the propagation of energetic particles into extrasolar planetary magnetospheres, not only must the characteristics of the planets be considered but also the age and spectral type of the star, as they affect the strength of the stellar wind, which in turn affects how well higher energy particles from the Galaxy can be shielded (Grießmeier et al., 2005). Further considerations include the influence of the planet-star distance (Grießmeier et al., 2004; 2009), the presence or absence of tidal locking via its influence on the planetary magnetic field (Griessmeier et al., 2005a, 2009), and the effect of the planetary size and type, via the estimated magnetic field (Grießmeier et al., 2009).

More recently, Grießmeier et al. (2015) and Grießmeier et al. (2016) take a slightly different approach and systematically study the influence of the planetary magnetic field. Instead of applying a model to estimate the planetary magnetic moment, they show how magnetic protection varies as a function of the planetary magnetic dipole moment. They evaluate the efficiency of magnetospheric shielding as a function of the particle energy (in the range 16 MeV–500 GeV) and of the planetary magnetic field strength (in the range 0–10 times the Earth’s magnetic moment). They find that, in the absence of a strong magnetic field, the flux of GCRs to the planetary atmosphere can be greatly enhanced (up to two orders of magnitude for particles energies $< 256$ MeV).

Grenfell et al. (2007) and Grießmeier et al. (2016) discuss the modification of atmospheric chemistry by such an enhanced GCR flux, and the associated destruction of atmospheric biosignature or biomarker molecules. In particular, they find that ozone may be depleted by up to 20% above 40 km altitude, which is considerable, but probably not sufficient to be detectable in spectroscopic observations. Atri et al. (2013) and Grießmeier et al. (2016) present a quantitative treatment of the potential biological implications for a planetary surface as a function of the cosmic ray and UV flux. They find that the surface biological dose rate may increase by a factor of two for a vanishingly small planetary magnetic field. In comparison, the efficiency of an Earth-like magnetosphere (a factor of two between an Earth-like magnetosphere and an unmagnetized planet) was found to be much less than the shielding efficiency of an Earth-like atmosphere (a factor of 500 difference between atmospheres with atmospheric depth of 1000 and 100, respectively). Further, Segura et al. (2010), Grenfell et al. (2012), and Tabataba-Vakili et al. (2016) have considered the case of stellar energetic particles for the case of an unmagnetized planet. All find strong removal of atmospheric ozone (over 90%) in the case of strong stellar flares.
Planetary Atmosphere Loss in Young/Active Stellar Systems

For two reasons, planetary systems around young/active stars (including the early solar system) may enter a different regime in which planetary magnetic fields play an important role in retaining primordial planetary atmospheres. First, a young star likely has much higher X-ray/ extreme UV (EUV) activity than the present-day Sun, which heats and thus expands planetary atmospheres (Lammer et al., 2003; Grießmeier et al., 2004; Ribas et al., 2005). The exobase height rises and can reach altitudes of up to several planetary radii. Second, the stellar wind of a young star is much denser and faster, compressing the magnetopause (Grießmeier et al., 2004, 2007a, 2010). Both effects combined lead to exosphere altitudes which are much closer to the magnetopause than for present-day Earth. If, as a third ingredient, the planetary magnetic field is weak, the planetary atmosphere may be exposed to direct interaction with the stellar wind, enabling strong atmospheric erosion. Lundin et al. (2007) suggest that in contrast to Mars and Venus, Earth’s global dipole field may have enabled it to retain volatiles, especially water, from its primordial atmosphere during approximately the first 1 Gyr of the solar system’s history, when the solar wind may have been as much as 100 times stronger than it is at present (Wood, 2006).

Grießmeier et al. (2010) explore the region in parameter space in which the stellar wind and CMEs can compress the magnetopause down to a certain altitude level. They examine cases in which the stellar winds can compress the magnetosphere down to levels of approximately 1.15 times the Earth radius, i.e., an altitude of order 1,000 km above the planetary surface. This limit corresponds to the exobase altitude at a time when the solar EUV flux was 70 times higher than today, and can be considered a typical limit for strong atmospheric erosion (Khodachenko et al., 2007; Lammer et al., 2007). Under these conditions, Grießmeier et al. (2010) found strong atmospheric erosion for weakly magnetized planets, especially for stellar ages less than about 0.7 Gyr.

The majority of M dwarfs remain highly active (with detectable Hα emission in their spectra) for a billion years or more (West et al., 2008), much longer than most solar-mass stars. They can also have much higher flare rates than the Sun. Khodachenko et al. (2007) scale up the correlation observed between solar flares and CMEs to estimate the rate and energetics of CMEs from active M dwarfs. They find that CMEs can cause significant magnetospheric compression for planets in the habitable zone of such stars, exposing the dayside atmosphere to the stellar wind and CMEs. In a companion paper, Lammer et al. (2007) argued that the high EUV flux in the habitable zone of active M dwarfs heats and inflates exoplanet atmospheres, so that CME-induced ion pickup can effectively erode atmospheres of up to hundreds of bars.

To date, the signature of habitability for an extrasolar planet has focused on its location in the “habitable zone” of its host star and its atmospheric composition. However, the existence of a protective planetary magnetic field should be included in any inventory of habitability characteristics for a given extrasolar planet. Given an estimate of the planetary magnetic field
strength, one can assess the effectiveness of shielding against galactic and stellar high-energy particles and potential habitability of an extrasolar planet. Moreover, even if a planet itself is not habitable (e.g., a gas giant), its magnetic field may also be important in shielding moons such that they are habitable (Heller & Zuluaga, 2013).

Finally, much of this discussion has been based on the radiation environment produced by the Sun. The mass loss rate of stars as a function of age (estimated from measurements of the size of the astropause [Wood et al., (2002, 2005)]) goes as \( \dot{M} \propto t^x \), with \( x \approx -2 \), a dependence probably linked to the decrease in surface magnetic activity with stellar age (Butler et al., 1996; Saar & Donahue, 1997). Thus, the stellar wind around a 1-Gyr-old star may be 25 times as intense as the current solar wind (from a 4.5-Gyr-old star).

2.5 Atmospheric Chemistry: Magnetic Fields, Aerosols, & Planet Color

Color is a primary characteristic of a planet. If the giant planets in our solar system were observed as extrasolar planets, namely as point sources, Jupiter and Saturn would look red to yellowish, while Uranus and Neptune would look blue to greenish. Measurements of the geometric albedo of these giant planets would make these qualitative descriptions quantitative: for wavelengths shorter than 600 nm, the geometric albedoes of Jupiter and Saturn decrease towards shorter wavelengths while the geometric albedoes of Uranus and Neptune flatten towards shorter wavelengths (Karkoschka, 1994; Figure 2.9). Traub (2003) discusses the color of extrasolar planets similar to the \( R = 5 \) spectra in Figure 2.9.

These general trends are consistent with expectations from atmospheric structure models. A non-absorbing deep Rayleigh and Raman scattering atmosphere is expected to have a geometric albedo of 0.62 at 265 nm (Hansen & Travis, 1974). This value is fairly close to the measured albedoes of Uranus and Neptune, implying that their atmospheres are fully reflective at this wavelength. In contrast, the UV albedoes of Jupiter and Saturn are lower than 0.3, suggesting that they have materials in their atmospheres that absorb blue and UV light.

The UV-dark materials in the atmospheres of Jupiter and Saturn are likely formed by auroral processes (§3.2), as established by their geographical correlation. Figure 2.10 illustrates the typical observations of aerosols on Jupiter (West, Strobel, & Tomasko, 1986; Pryor & Hord, 1991; Zhang et al., 2013), and provides the following two lines of evidence. First, the concentration of the UV-dark aerosols reaches the highest levels towards the polar region, indicating that the sources of these aerosols are in the polar region. Second, the aerosol layer in the northern hemisphere is darker, thicker, and extends to a lower latitude than the aerosol layer in the southern hemisphere, which correlates with the fact that the northern auroral zone is greater than the southern auroral zone. Similar evidence has been found for Saturn (e.g., Pryor & Hord, 1991).
Chapter 2. Searching for & Characterizing Extrasolar Planets

Figure 2.9: Low-resolution albedo spectra of the giant planets in the solar system. Original high-resolution data are from Karkoschka (1994), and these panels are reprinted from Cahoy, Marley, & Fortney (2010). (Left) Spectral resolution of $R = 5$, comparable to what might be obtained with initial spectra of extrasolar planetary atmospheres. (Right) Spectral resolution of $R = 15$.

A likely candidate for the aurora-derived dark aerosols in the atmospheres of Jupiter and Saturn is polycyclic aromatic hydrocarbons (PAHs). This interpretation is motivated by the detection of benzene by a series of spacecraft observations (Kim et al., 1985; Niemann et al., 1998) and by a high abundance of methane in these atmospheres that is known to lead to production of complex hydrocarbons (Gladstone et al., 1996). Detailed modeling of chemical processes in the polar region of Jupiter has shown that neutral chemistry is unable to produce the required amount of dark aerosols (Wong et al., 2000; Friedson et al., 2002), whereas ion chemistry initiated by energetic electrons is able to form sufficient amounts of aerosols (Wong et al., 2003).

The sequence of reactions is initiated by the reaction between $\text{H}_2$ and energetic electrons that eventually produces $\text{H}_3^+$, which then attaches to $\text{CH}_4$ to form $\text{CH}_5^+$ (Yung & Strobel, 1980). The major reaction pathway to form PAHs is successive ion-neutral reactions of ions with $\text{C}_2\text{H}_2$ or $\text{H}_2$, producing $\text{C}_2\text{H}_3^+$, $\text{C}_4\text{H}_3^+$, and $\text{c-C}_6\text{H}_7^+$ (Figure 2.11). This ion chemistry pathway is approximately 100 times more efficient than chemical pathways that only involve neutral species (Wong et al., 2003). Furthermore, it has also been shown that for increasing energy flux of auroral electrons, the production and mixing ratio of the dark aerosols increases. Similar processes are thought to operate on Saturn. On Uranus, the auroral electrons cannot reach the homopause\(^\text{(10)}\).

\(^{10}\)The homopause is defined as the altitude at which the atmosphere can be approximated as well mixed, with a composition that can be represented by a single species, having the average mass of the constituent species and a single scale height.

of methane, while on Neptune, the auroral energy deposit is too small for significant aerosol production.

The current mainstream technique to characterize extrasolar planet atmospheres is via the planet transiting its host star. For measuring the reflection spectrum, one could expect the magnitude of the signal to scale as $A_g(R_p/D)$, where $A_g$ is the geometric albedo, $R_p$ is the planet’s radius, and $D$ is the semi-major axis of the planet’s orbit. Based on this scaling, the transit technique will work the best for giant planets at close-in orbits (i.e., hot Jupiters). Most hot Jupiters are dark at visible wavelengths, perhaps due to a lack of clouds and broad absorption from alkali metal vapor in their atmospheres (*Sudarsky et al.*, 2000). However, a few hot Jupiters have recently been found to have a large mean albedo at visible wavelengths (*Evans et al.*, 2013; *Demory et al.*, 2013). These broad observations show that hot Jupiters may have a large baseline albedo due to clouds, providing a starting point to approach the detection of the dark hazes. A low-resolution spectrum, as a step forward from the current broadband spectrum, will provide valuable information on the planet’s atmospheric constituents and may shed light on the existence of a magnetic field.
Figure 2.11: Important reaction pathways of benzene and PAH formation dominated by ion chemistry. \( A_1 \) represents benzene, \( A_2 \) represents naphthalene, \( A_1 \) and \( A_2 \) represent one- and two-ring radicals, \( A_1C_2H_2 \) represents benzene with a hydrogen atom replaced by a \( C_2H_2 \) group, and PAHs represent all ring compounds larger than \( A_2 \). [Reprint from Wong et al. (2003).]

A future space-based telescope equipped with a coronagraph could directly image extrasolar giant planets and measure their albedo spectra, and by extension constrain their magnetic fields. Recent definition of the WFIRST mission concept involves a 2.4 m Hubble-quality aperture with the capability of directly imaging extrasolar planets at contrast levels of \( 10^{-9} \) and an inner working angle of less than \( 0.2'' \) (WFIRST\(^{11}\)). If achieved, such a telescope would be able to measure reflection spectra of nearby Jupiter-sized extrasolar planets at 1–8 au from nearby Sun-like stars (Figure 2.12). These prospective planets will have suitable temperatures in their atmospheres for the formation of PAH aerosols, similar to those in Jupiter’s atmospheres; they will also likely have water clouds in their atmospheres to provide a large baseline albedo (Cahoy, Marley, & Fortney, 2010). Thus, if the slope of reflection spectra of these planets can be measured (Traub et al., 2016), dark hazes may be detected, indicating the presence of a magnetic field.

2.6 Stellar Winds, Flares, and Coronal Mass Ejections

Due to our privileged position immersed in the solar wind, there is a considerable quantity of data that allow a detailed understanding of the physics operating in the Sun. Recent sophisticated observations, e.g., SOHO and Ulysses (McComas et al., 1995; Suess & Smith, 1996; Wilhelm, 2006, among others) have shown that the solar atmosphere is a highly complex system. It consists of long-lived features, like the fast and slow solar wind, streamers, and coronal holes, as well as transient features, like CMEs, solar flares, and solar prominences.

\(^{11}\)http://wfirst.gsfc.nasa.gov/
Figure 2.12: Sensitivity of the WFIRST coronagraph for imaging planets around nearby stars. Solid lines mark the baseline technical goal of 1 ppb contrast and 0.2$''$ inner working angle, while dotted lines show the more aggressive goals of 0.1 ppb and 0.1$''$. Colored circles show a snapshot from a simulation of model planets, ranging in size from Mars-like to several times the radius of Jupiter, in orbit around approximately 200 stars within 30 pc. The model assumes roughly four planets per star with a mixture of gas giants, ice giants, and rocky planets, and size and radius distributions consistent with Kepler results and extrapolated to larger semi-major axes and lower masses. Color indicates planet mass, and size indicates planet radius. Crosses represent known radial velocity planets. WFIRST would be able to image a number of known radial velocity planets, and it would detect a number of new gas and ice giants. [Reprinted from Spergel et al. (2015).]

Similar to the Sun, other stars also experience mass loss in the form of winds and CMEs. These outflows permeate the interplanetary space, interacting with any body encountered on their way. The interaction between the stellar wind and a magnetized planet can give rise to observable signatures, which could then be used to probe planetary magnetic fields (§3.4).

Most of the extrasolar planets found to date are orbiting cool, low-mass stars ($\lesssim 1.5 \, M_\odot$). In this Section, we present an overview of winds from these stars.
2.6.1 Stellar Winds

Similar to the Sun, stars experience mass loss in the form of winds during their lives. Along with the evolution of the stars, their winds change characteristics over time. For some stars, at certain evolutionary phases, the amount of mass lost through stellar winds can amount to a significant portion of their own mass. However, for cool dwarf stars during the main sequence phase, their winds carry only a small amount of mass compared to the total mass of the star ($\sim 10^{-12} \ M_\odot$ yr$^{-1}$ for the Sun). Even though winds of cool dwarf stars are quite rarefied, because they are magnetized they carry a significant amount of angular momentum, which affects the stellar rotational evolution (e.g., Bouvier et al., 1997). In turn, the loss of stellar angular momentum alters the star’s own magnetism through dynamo processes that are still not fully understood.

Direct measurements of tenuous coronal winds for stars other than the Sun have proven difficult to obtain. P-Cygni profiles, the traditional mass-loss signatures observed in the denser winds of giant and supergiant stars, are not formed in the rarefied winds of cool dwarf stars. To detect these winds, other indirect methods have been proposed (e.g., Lim & White, 1996; Wood et al., 2001; Wargelin & Drake, 2002). Unfortunately, most of the time these methods could only constrain upper limits. So far, the method developed by Wood et al. (2001) has been the most successful, enabling estimates of mass-loss rates for about a dozen cool dwarf stars.

The high coronal temperatures observed in low-mass stars (including the Sun) are thought to be due to the release of magnetic energy in the stars’ atmospheres, although the detailed coronal heating mechanism remains unknown. The solar inner corona has a temperature of about $\sim 1$ MK. Observations indicate that the coronae of solar-analogs can be much hotter than that, with temperatures of the X-ray emitting coronae exceeding 10 MK (Güdel, 2004). If the temperature of the X-ray emitting (closed) corona is related to the temperature of the stellar wind (flowing along open field lines), as one would naively expect, then we may expect stellar winds of cool dwarf stars to have temperatures that could be much larger than the solar wind temperature.

Through indicators of magnetic activity, such as surface spot coverage, emissions from the chromosphere, transition region, or corona, it has been recognized that these stars are magnetized. However, only recently has it become possible to reconstruct the large-scale surface magnetic fields of other stars. In particular, the Zeeman-Doppler Imaging (ZDI) technique has now been successfully used to investigate the magnetic topology of stars of different spectral types (see Donati & Landstreet [2009] for a review). The ZDI method is a tomographic imaging technique that allows us to reconstruct the large-scale magnetic field (intensity and orientation) at the surface of the star from a series of circular polarization spectra (Donati & Brown, 1997). This method has revealed fascinating differences between the magnetic fields of different stars. For example, solar-type stars that rotate about two times faster than our Sun show a substantial toroidal component of magnetic field, a component that is almost non-existent in the solar...
magnetic field (Petit et al., 2008). The magnetic topology of low-mass (<0.5 \( M_\odot \)) very active stars seems to be dictated by interior structure changes: while partially convective stars possess a weak non-axisymmetric field with a significant toroidal component, fully convective ones exhibit strong poloidal axisymmetric dipole-like topologies (Morin et al., 2008; Donati et al., 2008a).

The magnetic field observations of planet-hosting stars (Catala et al., 2007; Donati et al., 2008b; Fares et al., 2009, 2010, 2012, 2013) are of particular relevance for characterizing the interaction between winds and planets.

Figure 2.13a shows an example of the large-scale field reconstructed from the observations. The surface magnetic map shown is that of the hot-Jupiter hosting star, \( \tau \) Boo. Maps like this one are now being incorporated into simulations of stellar winds (Vidotto et al., 2011b, 2012; Jardine et al., 2013; Llama et al., 2013), where they are used as boundary conditions for stellar wind simulations. A potential field extrapolation is adopted at the initial state of the simulation (Fig. 2.13b). As the simulation evolves in time, stellar wind particles and magnetic field lines are allowed to interact with each other. Figure 2.13c shows the self-consistent solution found for the magnetic field lines for the case of \( \tau \) Boo (Vidotto et al., 2012). Note that the magnetic field lines become stressed, wrapping around the rotational axis of the star (pointing towards positive \( z \)). Color-coded are the reconstructed large-scale surface field of \( \tau \) Boo in 2006 June (Fig. 2.13a,b) and the wind velocity at the equatorial plane of the star (Fig. 2.13c).

Data-driven simulations like these are able to provide important constraints of the local condition on the stellar wind surrounding the planet. These conditions are required, for instance, to predict signatures arising from the interaction between winds and planets (§3.4 and §2.4).

### 2.6.2 Flares & Coronal Mass Ejections

In addition to the steady solar wind, the Sun also generates explosive transient events in the form of CMEs and solar energetic particles. The most energetic solar particles are associated with strong solar flares. Solar flares are caused by an explosive release of magnetic energy stored in active region magnetic fields. They release up to \( 10^{30} \) ergs on a time scales of hours. Similarly, CMEs are explosive releases of magnetized coronal plasma injected into the solar wind, with velocities in the range 200–3,000 km s\(^{-1}\), average masses of \( 1.6 \times 10^{12} \) kg, and kinetic energies in the range \( 10^{28}–10^{30} \) ergs. The energy source for CMEs is also the release of magnetic energy stored in coronal magnetic fields via magnetic reconnection. The CME rate is much higher and the speeds much faster during the maximum phase of solar activity. The most energetic CMEs generally come from the most energetic active regions and are generally accompanied by a large solar flare.

There are no confirmed detections yet of CMEs from other stars, but since flares have been detected from other stars, CMEs are expected as well. A possible spectroscopic detection of a
stellar CME was reported by Fuhrmeister \& Schmitt (2004), who observed blue-shifted Balmer lines during a flare on an M9 dwarf. The shift indicated upward mass motion at about 100 km s\(^{-1}\); they suggest that this might be the start of a CME. The energy of flares and CMEs is expected to scale as the available energy stored in the magnetic active region. More active stars or stars with larger magnetic fields and starspots might also have much more energetic flares and CMEs. Surprisingly, results from Kepler (Maehara et al., 2012) have shown that even some G-type stars have "super flares"—flares with energies larger than 10\(^{33}\) ergs. The durations of the detected superflares were typically a few hours, and their amplitudes were generally of order 0.1–1% of the stellar luminosity. Comparing different types of stars, Maehara et al. (2012) found that while the rate of flares was correlated with the stellar rotation rate, the flare amplitude was not. They also found that younger, faster-rotating G-type stars had a higher frequency of superflares, but even slow rotators like the Sun could produce superflares (viz. §2.4). Quasi-periodic brightness

Figure 2.13: (a) Observationally reconstructed large-scale field of the planet-hosting star \(\tau\) Boo (Catala et al., 2007). (b) Potential field extrapolation used as an initial condition for the stellar wind simulations. The distribution of the observationally reconstructed surface magnetic field (used as a boundary condition) is shown color-coded. (c) Self-consistent solution of the magnetized stellar wind. The wind velocity is shown at the equatorial plane. [Figure adapted from Vidotto (2013).]
modulations observed in the solar-type stars with superflares suggest, however, that they have much larger starspots than does the Sun. Since the energy source for flares and CMEs is the same, stars with superflares could potentially eject CMEs with greater than 1,000 times more energy than those from the Sun.

CMEs, and the shocks that they drive, can interact strongly with planetary magnetic fields and cause (significant) increases in radio emissions. Thus, magnetized planets can serve as probes of the space environment around their host stars. Increases in radio emission due to CMEs and their shocks have been observed at Earth, Jupiter, and Saturn (Figure 2.13). The energy from CMEs and shocks is injected into the planet’s magnetosphere both by compression, due to the ram pressure of the CME and its shock, and by magnetic reconnection between the magnetic fields of the planet and the CME. The reconnection accelerates electrons along field lines that cause radio emissions via the electron cyclotron maser instability (§3.1).
Detecting magnetic fields in extrasolar planets over interstellar distances requires that measurement techniques focus on the relatively strong fields generated by planetary-scale dynamos (§2.2.2). Immersed in the stellar winds of the host stars, these magnetic fields produce large-scale magnetospheres that can interact with and extract energy from the stellar winds.

In this chapter, we review the suite of measurement techniques that was the focus of the KISS Study as means for determining the presence of and measuring the strength of extrasolar planetary magnetic fields. Our focus is on the underlying physics responsible for producing a detectable signal and what current and near-future measurements capabilities are. In addition to the methods described here, it may be possible to use measurements of energetic neutral atoms (ENAs) to constrain the properties of extrasolar planetary magnetospheres (Ekenbäck et al., 2010; Lammer et al., 2011; Kislyakova et al., 2014).

3.1 Radio Emission

All of the “radio-active” planets in the solar system (Earth, Jupiter, Saturn, Uranus, and Neptune) produce radio emission via the electron cyclotron maser instability (ECMI). This section summarizes the conditions necessary to create the ECMI and the implications for its detection from extrasolar planets. (Consult Treumann [2006] for a comprehensive review of the ECMI.)
3.1.1 Theoretical Background

In order to create this intense, coherent emission, two key conditions have to be met in the magnetosphere of a planet:

- The relativistic cyclotron resonance must couple to an existing extraordinary mode (X-mode) escaping plasma mode in the region, and
- At this coupling point, energy flows from supra-thermal electrons to the waves.

Given these two conditions, the associated plasma conditions to create an ECMI and electron energies associated with wave activity can be estimated.

Figure 3.1 illustrates the index of refraction for an X-mode escaping wave in low plasma density and high plasma density conditions (assuming emission is emitted nearly perpendicular to the local auroral magnetic field). For frequencies $\omega$ much greater than the electron cyclotron frequency $\Omega_{ce}$, the index of refraction for the X-mode lies near unity. This mode is effectively a freely-propagating radio wave.

However, in the source region, as $\omega$ progressively decreases, and approaches $\Omega_{ce}$, the index of refraction for the mode decreases, going to zero (mode ceases to exist) at the "$R = 0$" cutoff frequency, which lies slightly above the cyclotron frequency. Specifically,

$$\omega_{R=0} = \Omega_{ce} + \omega_{pe}/\Omega_{ce}, \quad (3.1)$$

with $\omega_{pe}$ being the electron plasma frequency. As such, as the local plasma density and associated $\omega_{pe}$ decreases, the frequency difference between the cutoff frequency and the electron cyclotron...
3.1 Radio Emission

frequency also decreases. In a low density plasma, the two frequencies, \( \omega_{R=0} \) and \( \Omega_{ce} \), should lie very close together.

The frequency separation \( \epsilon \) between the cyclotron resonance and the mode creates a “stop band,” disallowing electron cyclotron resonance with the wave in a stationary frame. However, the relativistic cyclotron resonance condition has a Doppler shift term \( \omega = \Omega_{ce}/\gamma + k||\nu|| \), thereby allowing electron cyclotron motion to interact directly with a circularly polarized wave. The index of refraction defined by the resonance condition is

\[
n_{\text{res}} \sim A \left[ 1 - \frac{\Omega_{ce}}{\omega} \left( 1 - \frac{1}{2} \beta^2 \right) \right],
\]

where \( A = c/v|| \cos(\theta) \) and \( \beta = v/c \), with \( v \) being the velocity of the electrons and \( \theta \) being the wave normal angle between \( k \) and the magnetic field \( B \) (Farrell, 2001; Farrell et al., 2004). Figure 3.1 illustrates \( n_{\text{res}} \) for large \( \theta \). The gyrating electrons will be in resonance with the wave mode at locations where \( n_{\text{res}} = n_{RX} \), thus allowing energy to be exchanged between the electrons and wave.

The direction of the energy flow (wave → particle or particles → waves) at resonance depends upon the suprathermal electron component in resonance with the X-mode wave. If slightly more electrons are gyrating slightly faster than the spinning wave potential, the potential acts to slow this faster population by reducing their perpendicular velocity \( v_\perp \). In doing so, the deceleration will increase the wave potential, creating wave growth. In contrast, if the large number of trapped electrons have slightly more electrons gyrating slightly slower than the potential, the potential will "push" these straggling electrons, giving them added perpendicular velocity \( v_\perp \). These electrons will gain energy at the expense of wave energy (i.e., wave damping).

Electromagnetic wave growth is then related to the distribution of electrons, \( f(v||, v_\perp) \) in velocity space. A distribution containing more electrons at higher perpendicular speeds has a positive slope in the distribution function, \( df/dv_\perp > 0 \), thereby providing wave growth near resonance. In an analogy to laser terminology, such an electron distribution is often referred to as an "energy-inverted population," since there is a population of electrons at relatively higher energy states compared to a thermal Maxwellian distribution in the region. In auroral regions, such distributions with inverted populations that generate X-mode emission include loss cones (Wu & Lee, 1979) and horseshoe distributions (Mutel et al., 2010). Some separate physical action usually occurs to place electrons at higher relative energies. In the case of the horseshoe distribution, it is the passage of the electrons through the parallel potential along an auroral field line. In the case of the loss cone, higher energy electrons exist due to the near-complete loss of lower energy electrons via magnetic containment at the magnetic field line-atmosphere footprint. In both of these cases, the electron distribution is acted upon by some other direct force to create an energy-inverted population.
While the resonance condition is independent of $\omega_{pe}$ (or electron density), the mode’s index of refraction profile in frequency is dependent on $\omega_{pe}$. As illustrated in Figure 3.1, in a region with a low electron density, the X-mode index of refraction varies from 0 to 1 in a very small frequency band near the $R = 0$ cutoff, which itself has a small value of $\epsilon$. In this case, $dn/df$ for the mode is steep and will intercept the $n_{res}$ curve representing the resonance condition—thereby allowing emission. In contrast, in a region with high electron density, the mode varies from 0 to 1 over a broad band. As evident in Figure 3.1, $dn/df$ for this case is relatively gentle, and the mode does not intercept the $n_{res}$ curve representing the resonance condition. Emission will not occur in this high density case. By comparing the resonance condition to the mode, it can be demonstrated that the two will intercept when $\omega_{pe}/\omega_{ce} < 1$, though it is also dependent on the $v_{||}$ of the electrons, which affects the slope of the $n_{res}$ vs. $f$ profile.

As suggested by Figure 3.1, there are then three required conditions to obtain the ECMI:

1. The ratio $\omega_{pe}/\omega_{ce} \ll 1$ to bring the mode as close as possible to the electron cyclotron resonance;

2. The electrons must have a substantial parallel velocity component $v_{||}$ to allow the electron cyclotron resonance to be Doppler-shifted into resonance with the waves; and

3. The electrons near resonance must be distributed such that $df/dv_{\perp} > 0$ to allow waves to gain energy from the particles at resonance.

### 3.1.2 Observational Considerations

From an observational perspective, the planet will radiate up to a maximum (radio) frequency determined by the largest magnetic field strength within the region where the conditions for the ECMI can be sustained \citep{Farrell1999}. In practice, this region is typically near the magnetic polar regions, for which the maximum radiated frequency is then

$$
\nu_{max} = \frac{eMR_p^3}{2\pi m_e},
$$

(3.3)

\[ = \frac{eB_{pole}}{2\pi m_e}, \]

(3.4)

\[ \approx 2.8 \text{ MHz} \left( \frac{B_{pole}}{1 \text{ G}} \right), \]

(3.5)

where $e$ is the charge on the electron, $m_e$ is the mass of the electron, $R_p$ is the radius of the planet, $M$ is the magnetic moment of the planet at the surface or cloud tops (as distinct from the magnetic moment at the “surface” of the dynamo region, §2.2), and $B_{pole}$ is the magnetic
field strength at the surface of the planet or cloud tops in the magnetic polar regions, which are assumed to be the relevant regions for the ECMI radiation.

Figure 3.2 shows the radio spectra for the planets in the solar system that sustain the ECMI (Zarka, 1992). Jupiter, with a polar magnetic field strength at the cloud tops of about 14 G, is clearly the most intense emitter, and the only solar system planet detectable from the ground (Burke & Franklin, 1955; Franklin & Burke, 1956). With the other magnetic planets having much smaller magnetic moments—the Earth’s polar magnetic field strength is only about 1 G—their maximum emission frequencies are below the terrestrial ionospheric cutoff (∼ 10 MHz), which makes their emissions unobservable from the ground.

![AURORAL RADIO EMISSION SPECTRA](image)

**Figure 3.2:** Spectra of the radio emission from solar system planets, scaled to a common distance of 1 au. For Jupiter, three different components are shown: the hectometric emission (HOM), the decametric component linked to the moon Io (Io-DAM), and the decametric component not linked to Io (non-Io-DAM). For the other planets, there is one main contribution to the radio emission: the Saturnian kilometric radiation (SKR), the terrestrial kilometric radiation (TKR), the Uranian kilometric radiation (UKR), and the Neptunian kilometric radiation (NKR). The TKR is also often called the auroral kilometric radiation (AKR). The vertical solid line at $10^4$ kHz (= 10 MHz) indicates the approximate terrestrial ionospheric cutoff frequency below which ground-based observations are impractical. [Reprinted from Advances in Space Research, Vol. 12, P. Zarka, “The auroral radio emissions from planetary magnetospheres—What do we know, what don’t we know, what do we learn from them?” Pages No. 99–115, Copyright (1992), with permission from Elsevier.]
Another notable feature of Figure 3.2 is the intensity of terrestrial radio emissions relative to those from Uranus and Neptune. Both Uranus and Neptune have larger magnetic moments than the Earth. However, the solar wind loading onto their magnetospheres is much less, due to their greater distances from the Sun, and their radio luminosities (or radiated powers) are lower than that of Earth. Figure 3.2 thus also illustrates the basis for much of the extension to extrasolar planets (§3.1.3); a smaller star-planet distance should produce a higher radio luminosity.

3.1.3 Extensions to & Predictions for Extrasolar Planets

After the recognition that solar system planets could be radio emitters and the development of initial scaling laws (Desch & Kaiser, 1984), Fennelly & Matloff (1974), Yantis et al. (1977), Wingee et al. (1986), and Bastian et al. (2000) speculated about and conducted searches for analogous emission from extrasolar planets. If extrasolar planets host magnetic fields, it is reasonable to expect them to generate radio emission via the ECMI as well, though the challenges in detecting it are clear from simple considerations. At distances of at least $10^5$ times larger than for solar system planets, the flux densities of extrasolar planets should be lower by factors of at least $10^{10}$, though there may also be mechanisms that would lead to (much) enhanced flux densities relative to those that such simple considerations might predict.

The predicted flux density from an extrasolar planet depends upon the sources of available energy to the planetary magnetosphere. Five different input sources have been considered:

**Stellar Wind Kinetic Energy** The flux of protons from within the host star’s stellar wind and incident on the planet’s magnetosphere provides a power input proportional to $\rho v^2$, for a stellar wind of density $\rho$ and velocity $v$. This input energy source has been most frequently considered for extrasolar planets (Zarka et al., 1997; Farrell et al., 1999; Zarka et al., 2001; Lazio et al., 2004; Stevens, 2005; Grießmeier et al., 2005; Grießmeier, 2007; Grießmeier et al., 2007a; Grießmeier et al., 2007b).

**Stellar Wind Magnetic Energy** The flux of magnetic energy, or the electromagnetic Poynting flux, from the interplanetary magnetic field and incident on the planet's magnetosphere provides a power input proportional to $B^2v$, for an interplanetary magnetic field strength $B$ embedded in the stellar wind (Zarka et al., 2001; Zarka, 2007; Grießmeier et al., 2007b; Jardine & Collier Cameron, 2008).

**Stellar Coronal Mass Ejections** The kinetic energy of a stellar CME impacting a planetary magnetosphere provides power to the magnetosphere in a manner akin to the “Stellar Wind Kinetic Energy” described above. The distinction is that the kinetic energy of a CME is sufficiently large that the radio emission of the planet can be enhanced substantially relative
3.1 Radio Emission

to "normal" or quiet stellar conditions (Gallagher & D'Angelo, 1981; Grießmeier et al., 2005; Grießmeier, 2007; Grießmeier et al., 2007a; Grießmeier et al., 2007b).

Internal Magnetospheric Plasma Sources For Jupiter, the magnetic flux tube linking its satellite Io to its magnetic polar region produces the Io-decametric emission (Io-DAM, Figure 3.2). A rapidly-rotating extrasolar planet with a satellite could generate strong radio emission, even without significant stellar wind input to its magnetosphere (Nichols, 2011, 2012; Noyola et al., 2014, 2016).

Unipolar Interaction An analogy to the Jupiter-Io system could exist for an unmagnetized planet orbiting a magnetized star (Zarka et al., 2001; Zarka, 2006, 2007; Grießmeier et al., 2007a; Jardine & Collier Cameron, 2008). It is not clear that power levels, or radio luminosities, large enough to be detected over interstellar distances can be generated with this mechanism, and it is also not clear that this mechanism would provide much insight into the planet itself.

Generally, most of these models predict that close-in planets, especially "hot Jupiters," should have more intense emissions, due to the higher stellar wind loading of the magnetosphere.

Some caution is required when applying such a simple scaling law, however. The solar wind-auroral radio emission connection may not be direct in large co-rotating magnetospheres. Earth’s convection-driven magnetosphere is especially sensitive to solar wind pressure, and auroral kilometric radiation (AKR) can show a factor of 100 increase in power for a factor of about 2 increase in solar wind velocity (Gallagher & D’Angelo, 1981). However, Jovian aurorae are driven by currents that form in the co-rotating outer magnetosphere (Nichols, 2011), where the solar wind may only impose an indirect controlling influence. There is a correlation between Jovian radio power and solar wind, but it is not as evident as the terrestrial AKR case (Gurnett et al., 2002).

A notable exception to the concept of solar wind control of a planetary magnetosphere is the internal magnetospheric plasma source mechanism, by which a planet may be able to generate intense emission even if it is several astronomical units from its host star. Nichols (2011, 2012) showed that planetary co-rotation-dominated magnetospheres at stars with high X-ray emission (for enhanced magnetosphere-ionosphere coupling) but also in low stellar wind pressure environments (further from the star) may be strong radio sources. Further, at very close distances, within the closed magnetosphere of the host star, the ECMI mechanism may saturate rather than continue to increase (Jardine & Collier Cameron, 2008).

As the stellar wind parameters strongly depend on stellar age (Wood et al., 2001; Wood et al., 2002, 2005), for those mechanisms that depend upon energy input from the stellar wind, the age of the host star must also be incorporated into predictions (Stevens, 2005; Grießmeier et al., 2005; Lazio et al., 2010a). The radio flux of a planet around a young star may be orders of magnitude
higher than for a planet in an older system. Unfortunately, stellar ages are often poorly constrained, in turn often leading to a significant range in the potential planetary flux density.

By the same token, if a planet is in an eccentric orbit, the effective stellar wind density and velocity at the planet’s magnetosphere will vary over the course of the planet’s orbit, in turn modulating the planet’s emission (e.g., Grießmeier et al., 2007a). In the most dramatic cases, the resulting modulation of the planet’s radio emission might approach a factor of $10^3$ (Lazio et al., 2010b). However, if the planet’s orbit is sufficiently eccentric, it may be carried into a region where the stellar wind plasma density is high. If the resulting stellar wind plasma frequency is higher than the planetary ECMI frequency, the ECMI emission may be able to be generated within the planetary magnetosphere, but it would not be able escape from the local stellar environment (Grießmeier et al., 2007b; Hess et al., 2011).

Regardless of the energy source powering the radio emission, the same constraints of equation (3.5) apply for extrasolar planets as for solar system planets; namely, only those with sufficiently strong magnetic fields will generate radio emission at a high enough frequency to be detectable from the ground. Evaluating this frequency for an extrasolar planet requires an estimate of the planetary magnetic moment, which is often ill-constrained. Two main approaches have been adopted. Farrell et al. (1999) and Grießmeier et al. (2007b) assume the planetary magnetic moment can be calculated by a force balance, and find a planetary magnetic field which depends on the planetary rotation rate. In contrast, Reiners & Christensen (2010) assume the planetary magnetic moment to be primarily driven by the energy flux from the planetary core. Thus, they find no dependence on the planetary rotation rate; however, they obtain stronger magnetic fields and more favorable observing conditions for young planets. Driscoll & Olson (2011a) considered the specific case of terrestrial planets. They found that anomalously strong fields ($3 \times$ larger than the most optimistic prediction) are required for emission at frequencies above the Earth’s ionospheric cutoff; furthermore, the expected flux levels are very low.

Finally, while the focus of this KISS Study largely remained on planets orbiting main sequence stars, in the interest of completeness, we note that planets in more “exotic” environments have been considered as possible radio emitters. These environments include terrestrial planets around white dwarfs (Willes et al., 2005), planets around evolved cool stars (Ignace et al., 2010; Fujii et al., 2015), and planets around T Tauri stars (Vidotto et al., 2010a). Even interstellar “rogue planets,” i.e., planets not bound to a star, have been considered (Vanhamaki et al., 2011).

### 3.1.4 Observational Constraints for Extrasolar Planets

Figure 3.3 presents a graphical summary of most published limits on the radio emission from extrasolar planets (Zarka et al., 1997; Lazio et al., 2004; Lazio & Farrell, 2007; Lecavelier Des Etangs et al., 2009; Lazio et al., 2010a; Lecavelier Des Etangs et al., 2011; Lecavelier des Etangs et al., 2013; Hallinan et al., 2013; Murphy et al., 2014). Not shown are a few observations at
3.1 Radio Emission

frequencies above 1,000 MHz and a few observations at frequencies around 20 MHz—limits above 1,000 MHz are not shown as these are likely to be at too high of a frequency, and would require planetary magnetic field strengths larger than 500 G, while the published limits around 20 MHz are typically above the range of flux densities shown, lack adequate information to assess, or both. Based on a number of predictions that it is a promising target for detection, the most intensively studied planet to date is τ Boo b.

![Figure 3.3](image_url)

**Figure 3.3:** Published upper limits for radio emissions from extrasolar planets. Observations have been obtained the Very Large Array (blue), Giant Metrewave Radio Telescope (GMRT, orange), Ukrainian T-shaped Radio Telescope (cyan), Murchison Wide Field Array (pink), and Green Bank Telescope (red). The solid diamond at 74 MHz shows the limit on the average planetary radio emission from planets orbiting nearby solar-type stars (Lazio et al., 2010a). The solid star at 150 MHz is the tentative detection, on a single day, of radio emission from HAT-P-11b (Lecavelier des Etangs et al., 2013). For clarity, another 171 upper limits at 150 MHz from the GMRT (Sirothia et al., 2014) are not shown.

Several features of figure 3.3 deserve mention. First, the trend of upper limits becoming less constraining at lower frequencies is real and represents limits on radio telescope sensitivities at these frequencies. A primary factor determining the telescope sensitivity is $A_{\text{eff}}/T_{\text{sys}}$, the ratio between the effective area of the telescope and the system temperature. For a given telescope (e.g., VLA or GMRT), the effective area $A_{\text{eff}}$ is essentially fixed (by the number of antennas and the diameter of each antenna). At these frequencies, the dominant contribution to the system
temperature $T_{sys}$ is the sky temperature or the power contributed by the Milky Way Galaxy’s synchrotron radiation. This temperature increases dramatically at lower frequencies, scaling approximately with frequency as $\nu^{-2.6}$. Consequently, the limits become less constraining at lower frequencies. For dipole-based arrays, such LOFAR or the LWA, the effective area of the individual dipoles scales with frequency as approximately $\nu^{-2}$, so that any limits that they place should be much more constant with frequency.

Second, the solid diamond at 74 MHz is the upper limit on the average planetary radio emission from planets orbiting nearby solar-type stars (Lazio et al., 2010a). It was constructed from a stacking analysis of the radio emission in the direction of stars within 40 pc. As such, it represents a limit on the combination of the average planetary radio luminosity and the fraction of solar-type stars hosting planets that radiate at 74 MHz.

Third, the solid star at 150 MHz is the tentative detection, on a single day, of radio emission from HAT-P-11b (Lecavelier des Etangs et al., 2013). If this measurement represents an actual detection, it implies that the magnetic field strength of HAT-P-11b is 50 G. However, equally sensitive observations on another day did not detect any radio emission. Thus, as the authors acknowledge, some caution is warranted in concluding that this measurement represents the first discovery of extrasolar planetary radio emission.

Finally, while not yet a consideration, any claim of detection of the radio emission from an extrasolar planet must address whether it is the extrasolar planet or the host star that has been detected. In general, extrasolar planetary radio emission is expected to exceed the emission of the planetary host star (Zarka et al., 1997; Grießmeier et al., 2005). Further strategies suggested to distinguish planetary from stellar radio emission have included checking whether the radio emission is modulated with the planet’s orbital period (if known) or with a time scale characteristic of solar system planetary rotational periods ($\approx 10$ hr). However, Fares et al. (2010) note that even a purely planetary signal may be partially modulated by the stellar rotation period, which could complicate the discrimination between a stellar and a planetary radio signal.

### 3.2 Far-Ultraviolet Auroral Emissions

The majority of studies investigating the atmospheric conditions of transiting extrasolar planets have employed absorption techniques. HD 209458b was the first extrasolar planet whose atmosphere was investigated using far-ultraviolet (far-UV) absorption line spectroscopy (Vidal-Madjar et al., 2003, 2004). The far-UV bandpass is unique in that it offers direct access to the strongest transitions of the atoms and molecules that constitute the majority of the mass in extrasolar planets (e.g., H, O, C, H$_2$, CO, etc.), as well as those that are involved in the chemical reactions that produce more complex molecules observed in gas giants (H$_2$O, CH$_4$, CO$_2$, TiO, VO; Swain et al., 2009; Sing et al., 2008; Désert, 2009). Some authors have noted that the
presence of ions in the inflated exospheres of hot Jupiters (C$^+$ and Si$_2^+$) can be used to set upper limits on dipolar fields of transiting planets (e.g., Koskinen et al., 2013, and references therein). Studies of extrasolar planetary thermal emission in the mid-IR have shown that these objects can be observed directly at wavelengths at which the relative planet/star contrast is favorable (Knutson et al., 2008, 2009). Analogously, both Jupiter and Saturn are strong far-UV emitters, well above any contribution from scattered sunlight at $\lambda < 1700$ Å (Feldman et al., 1993; Gustin et al., 2002, 2009).

The brightest emission lines in Earth’s upper atmosphere are those of atomic hydrogen (the Lyman series, Ly$\alpha$ 1216 Å, Ly$\beta$ 1025 Å, Ly$\gamma$ 973 Å, etc.) and oxygen (O i 1356 Å, 1304 Å, 1027 Å, 989 Å, etc.) with fainter emissions from atomic and molecular nitrogen (N i 1200 Å, 1135 Å; N$_2$ 958 Å) as well as other trace species (e.g., Feldman et al., 2001, and references therein). In practice, these lines will be very challenging to observe owing to their intrinsic faintness at distances of a few to tens of parsecs, as well as several logistical issues. First, the far-UV spectra of almost all cool stars are dominated by chromospheric Ly$\alpha$ emission (as well as strong lines of atomic oxygen), presenting a low planet/star contrast ratio at the wavelengths at which Earth-like planets might be expected at their peak brightness. Second, the interstellar medium is opaque to Lyman-series (and often O i) photons due to resonant scattering by neutral hydrogen, deuterium, oxygen, etc., in the local interstellar medium. Finally, the largest astronomical background for UV instruments in low Earth orbit (LEO) is geocoronal Ly$\alpha$. This background emission is brighter than Ly$\alpha$ emission from the chromospheres of most cool stars beyond about 5 pc. The first two challenges are set by the astrophysics of stars and the interstellar medium, and while larger telescopes may increase the overall number of photons detected, high-contrast angular resolution techniques at far-UV wavelengths will be required to separate Earth-like planetary emission from their host stars. Placing a future large (8–16-m) UVOIR space telescope beyond LEO will help mitigate the background levels at Ly$\alpha$ and other geocoronal emission wavelengths until the interplanetary medium is the dominant background source. High-resolution UV spectrographs may provide a means of separating short-period Habitable Zone planets (i.e., those orbiting M dwarfs) from their parent star’s relatively narrow Ly$\alpha$ emission line profile by observing near quadrature, where the velocity shift of the planetary emission may make it detectable.

Far-UV emission from Jupiter and Saturn is dominated by atomic H (Lyman series) and molecular hydrogen (H$_2$) lines, the primary constituents of gas giant planets (Sudarsky et al., 2003), and H$_2$ may be responsible for the Rayleigh scattering observed in optical spectra of HD 209458b’s atmosphere (Lecavelier des Etangs et al., 2008). Molecular hydrogen in the atmosphere of giant planets is excited by two primary mechanisms: electron bombardment and fluorescence pumped by stellar emission lines. Both processes excite the ambient molecules to higher-lying electronic states whose decay produces a highly structured spectrum in the 700–1650 Å bandpass. The majority of these lines can be attributed to fluorescence from the Lyman and Werner bands of H$_2$ (B$^1\Sigma_g^+$ and C$^1\Pi_u$ to $X^1\Sigma_g^+$, however higher electronic states contribute to the observed electron
impact spectrum). Gas giant aurorae produce the highest surface brightnesses (Clarke et al., 1998) as the electron impact spectrum is most intense where the magnetic field lines connect with the atmosphere (Figure 3.4). Dayglow (non-auroral illuminated disk) emission from the giant planets also shows strong features of H$_2$, with differing levels of molecular excitation and contributions from solar fluorescence (Feldman et al., 1993; Wolven & Feldman, 1998).

It follows that the far-UV could be an ideal wavelength regime in which to probe the atmospheres of hot Jupiters through their direct emission (Yelle, 2004). Most transiting extrasolar giant planets are at relatively small separations from their parent star ($a < 0.1$ au), where the line flux from the host star should be more intense than at Jupiter by a large factor ($> 10^3$) and bombardment from a stellar wind may be enhanced relative to the Jovian environment. Additionally, by virtue of their short-period orbits, these planets have large orbital velocities ($v \sim 100$ km s$^{-1}$). Hence, by observing the system near both quadrature positions (phase 0.25 and 0.75), even with modest spectral resolution, the observed velocity shift between the two epochs can provide confirmation that the signal is planetary in origin.

### 3.2.1 Previous Studies

France et al. (2010a) combined a Jovian aurora-based spectral model and early observations from the HST/Cosmic Origins Spectrograph to search for this electron-impact excited H$_2$ emission signature in the well-studied transiting hot Jupiter HD 209458b. Synthetic spectra of fluorescent

---

**Figure 3.4:** (Left) HST/STIS image of UV auroral ovals at the pole of Jupiter (image credit: NASA). (Right) Spectra of the Jovian aurorae. The upper (black) curve shows a section of the FUSE spectrum of the Jovian aurora taken at the time of the New Horizons Jupiter flyby (2007). The lower (pink) curve shows a synthetic model spectrum for electron-impact and photo-excited H$_2$ in the auroral regions of Jupiter that was adapted by France et al. (2010a) to calculate the expected spectrum of the transiting hot Jupiter HD 209458b. The two curves are offset for clarity.
emission from H$_2$ (Figure 3.4, right panel) can be made by computing the collisional and radiative excitation rates into the upper electronic states. This model (see also Wolven et al. [1997] and Liu & Dalgarno [1996]) assumes a ground electronic state population, then uses excitation cross-sections and an incident radiation field to calculate the population distribution in the rovibrational levels of the upper electronic state (predominantly B$^1\Sigma^+_g$ and C$^1\Pi_u$). The molecules will then return to the ground electronic state following the appropriate branching ratios, producing far-UV emission lines ($700 \, \AA \leq \lambda \leq 1650 \, \AA$), and leaving the molecules in excited rovibrational levels. Figure 3.4 shows a comparison of an example of the model created for HD 209458b with Far-Ultraviolet Spectroscopic Explorer observations of Jupiter.

France et al. (2010a) used the quadrature observations of HD 209458b to search for auroral emission from the planet. They used models of the well-studied H$_2$ emission spectrum of the Jovian aurorae (Wolven et al., 1997) to identify spectral regions within which the planet-star contrast was expected to be high. Figure 3.5 displays the two regions of the far-UV spectrum of HD 209458b with the auroral H$_2$ model, showing that no obvious emission from the planet is observed. We refer the reader to France et al. (2010a) for a detailed description of the derivation of flux upper limits that would be suitable for comparison with future models of extrasolar planetary aurorae. The non-detection of electron-impact excited H$_2$ emission suggests that surface magnetic fields do not play as significant a role on HD 209458b as they do on Jupiter, where they drive the observed far-UV spectrum.

An additional potential complication, realized first when considering auroral H$_3^+$ emission (§3.3), is that it may not be possible to directly infer the presence of a magnetic field from the detection of far-UV emission from an extrasolar planet. In the absence of a magnetic field, there would still be electron bombardment, though it would be distributed across the disk of the planet rather than concentrated into the auroral regions.

While not clearly related to magnetic activity, the prospects for detecting fluorescent dayglow emission from hot extrasolar planetary material appears promising. France et al. (2013) observe fluorescent H$_2$ emission lines in the M dwarf extrasolar planet host stars GJ 581, GJ 876, GJ 436, and GJ 832. These features are photoexcited (“pumped”) by Ly$\alpha$ photons out of the $v = 2$ level of the ground electronic state (Shull, 1978). Under the assumption of thermalized rovibrational populations, this process is a signpost for 2,000–4,000 K molecular gas. They detect between four and eight fluorescent emission lines in each of these targets, pumped by Ly$\alpha$ through the (1–2) R(6) $\lambda$1215.73 Å and (1–2) P(5) $\lambda$1216.07 Å transitions (Figure 3.6). Ly$\alpha$-pumped H$_2$ emission lines were first detected in sounding rocket spectra of sunspots (Jordan et al., 1977) and have been modeled assuming a thermalized H$_2$ population at a temperature of 3,200 K (Shull, 1978). These H$_2$ fluorescence lines are also powerful diagnostics of the molecular surface layers of protoplanetary disks over a similar range of temperatures (Herczeg et al., 2004; France et al., 2012a). These lines are not clearly detected in previous STIS observations of M dwarfs without planets (AD Leo,
Figure 3.5: Comparison of deep far-UV observations of HD 209458b and a model for electron-impact and fluorescently excited H$_2$ atmosphere, scaled from a model appropriate to the B field and illumination at Jupiter. No electron-impact H$_2$ emission is detected. [From France et al. (2010a).]

Proxima Cen, EV Lac, and AU Mic3), although these non-detections could be the result of much larger instrumental backgrounds associated with the STIS echelle modes.

All four of the stars showing strong H$_2$ fluorescence host planets of Neptune to super-Jupiter masses. The analysis of France et al. (2013) cannot rule out the possibility that the ubiquitous Ly$_\alpha$-pumped H$_2$ fluorescence observed towards M dwarf planet hosts is the result of photoexcited dayglow emission in planetary atmospheres. In light of the recent tentative detections of FUV H$_2$ fluorescent emission from the planetary systems around HD 209458 (France et al., 2010a) and brown dwarf 2MASS J12073346-3932539 (France et al., 2010b), it is tempting to attribute an extrasolar planetary origin to these lines. The authors caution that further study is required.

3.2.2 Estimates of the “UV Bode’s Law” & UV Observational Limits for Extrasolar Planetary Systems

In order to present a complementary relation to the radiometric Bode’s Law discussed above (§3.1), we have searched the literature to present a rough estimate of the equivalent relationship for UV auroral emission from planets within our solar system, the “UV Bode’s Law” (Figure 3.7). Taking the incident kinetic flow power from Zarka (1998) and the UV auroral emission strength
3.2 Far-Ultraviolet Auroral Emissions

Figure 3.6: Coadded H\textsubscript{2} emission spectra detected toward four M dwarf extrasolar planet host stars. Each velocity profile is the co-addition of four to six fluorescent emission lines. The dotted line is the intrinsic Gaussian line profile that has been convolved with the HST/COS line spread function (solid blue line) for comparison with the data (black histogram). The red bars are representative of the uncertainties on the normalized line profiles. [From France et al. (2013).]

from various literature sources (Brahdwaı̈ & Gladstone, 2000 [Jupiter and Saturn]; Herbert & Sandel, 1994 [Uranus]; Sandel et al., 1990 [Neptune]; and Clarke et al., 2005 and Gustin, private communication [Earth]), we have constructed the relation shown as the straight line in the log-log plot in Figure 3.7. This expression follows the relation $\log(L(\text{UV aurora})) = a_{\text{UV}} + b_{\text{UV}} \cdot \log(L(\text{incident}))$, where $[a_{\text{UV}}, b_{\text{UV}}] = [-7.0 \pm 4.3, 1.30 \pm 0.25]$.

Additionally, we searched the HST archive for all planet-hosting stars with far-UV observations from either the STIS or the COS instruments. In principle, all of these observations (save geometric alignments of transiting planets for which the planet is in secondary eclipse) would contain UV auroral emission from the orbiting planets if it existed. We made no clear detections of the two strongest planetary emission features from jovian planets in the solar system: (i) 1230–1300 Å emission from the strong Werner-band line complexes of electron-impact excited H\textsubscript{2}, and (ii) 1575–1608 Å “bunny ears” features that are a combination of discrete line emission from
Figure 3.7: Initial estimate for the “UV Bode’s Law.”

electron-impact processes, molecular dissociation continuum from Lyman band transitions to $v'' > 14$, and fluorescence of solar Lyβ, the strongest dayglow emission line on Jupiter.
3.2 Far-Ultraviolet Auroral Emissions

Figure 3.8: Flux limits on the 1230 and 1608 Å auroral emission from all extrasolar planets with HST/STIS and COS observations in the Archive. The masses are color-coded by Earth/super-Earth (green), Neptunian (blue), and Jovian (red). The objects with large upper limits are those observed in the high-resolution modes of STIS, for which instrumental backgrounds are a factor of about $100 \times$ those of COS.

In order to quantify these spectra for future researchers when physically-motivated models for UV auroral emission may exist, we have used these archival observations to obtain upper limits in two line complexes that would be spectrally resolved by the STIS and COS modes used to make these observations:

1. We chose the 1227.5–1232.5 Å spectral region to be representative of electron-impact emission, as this feature is one of the brightest from a Jovian-like spectrum and is at the peak of the HST/COS effective area curve ($A_{\text{eff}} \sim 2500 \text{ cm}^2$).

2. We honed in on the 1607.3–1607.7 Å region to be representative of the dayglow emissions from a Jupiter-like planet (keeping in mind that the dayglow on shorter period worlds will likely be dominated by Ly$\alpha$ fluorescence, as described above for the M dwarf extrasolar planet host stars).

Figure 3.8 shows upper limits for all available observations of extrasolar planet host stars orbiting cool stars (G–M dwarfs), ordered by planetary mass. We note that some of these planets are super-Earth mass objects in which Ly$\alpha$, as opposed to H$_2$, is likely the brightest auroral feature (see above). Airglow emission from Earth’s geocoronal and resonant scattering by interstellar hydrogen and deuterium prevents deep exposures of astrophysical objects in the Ly$\alpha$ line core, so we measured uniform wavelength regions for consistency.

While broader wavelength regions could have been integrated to refine these limits, we remind the reader that the far-UV spectra of cool stars are littered with chromospheric, transition
region, and coronal emission lines, and these narrow windows were specifically designed to avoid contamination from known features. Furthermore, in the absence of bulk kinematic effects (such as entrainment in a rotating disk), H$_2$ lines are always narrow. Expanding over wider spectral bandpasses will serve to add noise to the measurement without significantly increasing the potential signal detection.

3.3 H$_3^+$ Infrared Auroral Emissions

In our solar system, H$_3^+$ emission in the 2–4 µm region has been imaged in the polar regions of Jupiter, Saturn, and Uranus, and can be considered a tracer of auroral emission and magnetic fields. In these cases, it is thought that energetic solar wind particles are precipitated through the magnetic field into the polar caps and catalyze the formation of H$_3^+$ (via ionizing H$_2$ to create H$_2^+$, which reacts with a second H$_2$ molecule to form H$_3^+$). The energetic particles are also thought to be directly or indirectly responsible for exciting the H$_3^+$ emission. In addition, fainter H$_3^+$ emission is seen from Jupiter’s disk, although it is not clear if its origin is auroral as well.

The phenomenology of H$_3^+$ emission is rather complicated but may be summarized as follows: Many lines of H$_3^+$ have been detected in 2–4 µm spectra. The images, however, have only been made in the 3.4–4 µm region because the H$_3^+$ lines there (from the fundamental band) are the strongest and other radiation from Jupiter, Saturn, and Uranus is weak at these wavelengths (due to absorption of solar radiation by CH$_4$). The images have been made across entire planetary disks. On Jupiter, H$_3^+$ emission is by far the brightest in the polar regions but is seen across the entire planet (Figure 3.9). On Saturn, it is visible only in the polar regions. On Uranus, where the locations of the poles are now uncertain, it is more or less uniform across the entire disk and observations have not revealed a polar maximum. To date, all of the work on H$_3^+$ emission in the solar system has been done from large ground-based telescopes such as the IRTF and the Gemini Observatory.

Rough estimates based on the observed H$_3^+$ emission from our own gas giant planets suggested that it might be detectable from extrasolar planets by JWST, although there are some substantial observational problems related to flat fielding and photon noise from the star, which presumably would be in the spectrometer aperture. During initial discussions within the Study, it was thought that such emission would be diagnostic of an extrasolar planetary magnetic field. However, upon further thought, we realized that, in the absence of the magnetic field, the H$_3^+$ emission could still be seen, because the solar wind would then crash into and light up the entire planetary disk. It is the spatial distribution of the H$_3^+$ on the solar system planets which is indicative of the auroral origin and the relationship to planetary magnetic fields. Therefore, if H$_3^+$ emission were to be detected from an extrasolar planet, there would be a need for further studies to determine

---

1 The poles were identified when Voyager 2 passed by, but the rotation period of Uranus is sufficiently uncertain that the location is now unknown.
3.3 $\text{H}_3^+$ Infrared Auroral Emissions

Figure 3.9: Auroral $\text{H}_3^+$ emission from Saturn, as captured by the Visible and Infrared Mapping Spectrometer (VIMS) on board Cassini. In addition to the auroral $\text{H}_3^+$ ring (in green), this mosaic of images also shows reflected sunlight (blue) and thermal emission from beneath the clouds (red). [Image credit: NASA/JPL/University of Leicester/University of Arizona.]

the distribution of the emission across the planet’s surface to show an auroral origin and suggest the presence of an extrasolar planetary magnetic field.

Thus, detection of $\text{H}_3^+$ emission and association with an extrasolar planet magnetic field seems challenging. However, our current lack of knowledge regarding extrasolar planets and the continual surprises they present make it unwise to drop this possibility totally from the toolkit. We should be alert to the chance of a serendipitous detection of $\text{H}_3^+$ emission and be prepared to follow it up. A possible scenario would be that strong $\text{H}_3^+$ lines are detected unexpectedly in emission spectroscopy of a transiting planet in secondary eclipse; such observations are likely to encompass the 2–4 $\mu$m wavelength range within which $\text{H}_3^+$ emission is detected. If such a detection were to occur, repeated extrasolar planet spectroscopy using specialized techniques, or even measurement of a phase curve as the planet orbits, might show that the emission is concentrated at particular regions on the extrasolar planet and provide evidence for an auroral phenomenon mediated by an extrasolar planet magnetic field.

Consistent with our conclusions, Lenz et al. (2016) recently reported only an upper limit on the $\text{H}_3^+$ emission from the hot Jupiter HD 209458b. This search was conducted on the Very Large Telescope (VLT), and their observations suffer a significant noise penalty relative to future JWST observations. Nevertheless, the authors report that their upper limit could be only a factor of 10
Chapter 3. Magnetic Field Determinations

(although as much as a factor of 1000) above their predictions for the intensity of several H$_3^+$ lines around 4 µm.

3.4 Transit Light Curve Variations—Magnetosphere Standoff Distance

As a stellar wind flows from a star, it permeates the entire extrasolar system, interacting with any body that it encounters on its way. The interaction between stellar winds and extrasolar planets can lead to observable signatures, some of which are absent in our own solar system.

3.4.1 The Early Ingresses of WASP-12b and HD 189733b

A series of narrow-band near-UV spectroscopic observations with the Hubble Space Telescope has suggested that the transit light curves of the close-in giant planets WASP-12b and HD 189733b present early ingresses when compared to their optical transits (Fossati et al., 2010b; Ben-Jaffel & Ballester, 2013; Bourrier et al., 2013; Cauley et al., 2015), but see also Turner et al. (2016a), who do not find evidence for such asymmetric transits. In the case of WASP-12b, there are also indications of excess absorption during the transit. The asymmetric transits indicate the presence of an asymmetric distribution of material surrounding the planet.

Several mechanisms have been suggested to explain asymmetric transits. Close-in giant gas planets are often inflated and can develop exospheres that fill or even overflow their Roche lobes (Gu et al., 2003; Ibgui et al., 2010). The resulting mass transfer through a Lagrangian point to the star could cause an asymmetry in the appearance of the transiting planet-star system. Lai et al. (2010) suggest that this mechanism may explain the early near-UV ingress of WASP-12b (see also purely hydrodynamic simulations describing the pattern of the mass transfer by Bisikalo et al. [2013]). Asymmetries could be produced by cometary tails, although a radiation-driven cometary tail would produce a late egress of the planetary transit light curve, instead of an early ingress. Kislyakova et al. (2016) have suggested that mass loss from Trojan satellites could produce significant accumulations of matter ahead of a planet in its orbit.

Motivated by these transit observations, Vidotto et al. (2010b) suggested that the presence of bow shocks surrounding the magnetospheres of close-in planets might lead to transit asymmetries at certain wavelengths; Cauley et al. (2015) make a similar suggestion for HD 189733b. More recent analyses (Alexander et al., 2016; Turner et al., 2016b) highlight potential challenges in a magnetospheric explanation, but, consistent with the work in the KISS Study, we discuss how a planetary magnetosphere would be able to result in an asymmetric transit.

The main difference between bow shocks formed around extrasolar planets and the ones formed around planets in the solar system is the shock orientation, determined by the net velocity of the particles impacting on the planet’s magnetosphere. In the case of the Earth, the solar wind is
dominated by a radial component, which is much larger than the orbital velocity of the Earth. Because of that, the bow shock surrounding the Earth’s magnetosphere forms facing the Sun. However, for close-in extrasolar planets that possess high orbital velocities and are frequently located at regions where the host star’s wind velocity is comparatively much smaller, a shock may develop ahead of the planet. In general, we expect that shocks are formed at intermediate angles (see also Vidotto et al. [2011c] and Llama et al. [2013]).

In the case of WASP-12b, due to its extremely close proximity to the star, the flux of coronal particles impacting on the planet comes mainly from the azimuthal direction, as the planet moves at a Keplerian orbital velocity of $u_K = (GM_*/R_{\text{orb}})^{1/2} \sim 230$ km s$^{-1}$ around the star. Therefore, stellar coronal material is compressed ahead of the planetary orbital motion, possibly forming a bow shock ahead of the planet. If such compressed material is optically thin, the planetary transit light curve is symmetrical with respect to the ingress and the egress of the planet. Indeed, this is the case when the transit is observed at optical wavelengths (Hebb et al., 2009). However, if the shocked material ahead of the planet can absorb enough stellar radiation, the observer will note an early ingress of the planet in the stellar disk, but no difference will be seen at the time of egress, as the shocked material is present only ahead of the planetary motion. In addition, because of the extra absorption caused by bow shock material, the transit will also appear deeper in the spectral lines, as they trace the shock absorption with respect to the broadband optical ingress (Vidotto et al., 2010b). Figure 3.10 illustrates this idea. This suggestion was later verified by Llama et al. (2011), who performed Monte Carlo radiation transfer simulations of the near-UV transit of WASP-12b, confirming that the presence of a bow shock indeed breaks the symmetry of the transit light curve.

### 3.4.2 Using Bow Shock Signatures to Measure Planetary Magnetic Fields

The phases at which the transit begins in the optical ($\varphi_{\text{op}}$) and in the near-UV ($\varphi_{\text{UV}}$) can be used to trace the extension of the absorbing material (cf. Figure 3.10). If we take the stand-off distance of the absorbing material as the magnetosphere size $r_M$, we find

$$r_M = \frac{(1 - \varphi_{\text{UV}})}{(1 - \varphi_{\text{op}})}[(R_2^2 - b^2)^{1/2} + R_p] - (R_2^2 - b^2)^{1/2},$$

where $b$ is the impact parameter of the transit and $R_p$ and $R_*$ are the radii of the planet and star, respectively (Vidotto et al., 2011d). From the near-UV observations (Fossati et al., 2010b), one finds $r_M = 4.2 R_p$ for WASP-12b.

At the magnetopause, pressure balance between the coronal total pressure and the planet total pressure requires that

$$\rho_c \Delta u^2 + \frac{[B_c(a)]^2}{8\pi} + p_c = \frac{[B_p(r_M)]^2}{8\pi} + p_p,$$

where $B_c(a)$ and $B_p(r_M)$ are the magnetic field strengths in the coronal and planetary regions, respectively.
Figure 3.10: Sketch of the light curves obtained through observations in the optical and in certain near-UV lines, where the bow shock surrounding the planet’s magnetosphere is also able to absorb stellar radiation. The stand-off distance from the shock to the centre of the planet is assumed to trace the extent of the planetary magnetosphere $r_M$, which can be estimated by measuring the phases at which the near-UV ($\varphi_{\text{UV}}$) and the optical ($\varphi_{\text{op}}$) transits begin.

where $\rho_c$, $p_c$, and $B_c(a)$ are the local coronal mass density, thermal pressure, and magnetic field intensity at orbital radius $a$, and $p_p$ and $B_p(r_M)$ are the planetary thermal pressure and magnetic field intensity at $r_M$, respectively. The relative velocity between the material surrounding the planet and the planet itself is $\Delta u$. In the case of a magnetized planet with a magnetosphere of a few planetary radii, the planet total pressure is usually dominated by the contribution from the planetary magnetic pressure (i.e., $p_p \sim 0$). The local conditions surrounding the planet ($\rho_c$, $p_c$, $B_c$, and $\Delta u$) can be derived from models of stellar winds (Section 2.6; Vidotto et al., 2012; Llama et al., 2013). Vidotto et al. (2010b) showed that, in the case of WASP-12b, the dominant terms in equation (3.7) are the magnetic terms. Therefore, equation (3.7) can be approximated as $B_c(a) \simeq B_p(r_M)$. Further, assuming that stellar and planetary magnetic fields are dipolar, we have

$$B_p = B_\star \left( \frac{R_\star/a}{R_p/r_M} \right)^3,$$

where $B_\star$ and $B_p$ are the magnetic field intensities at the stellar and planetary surfaces, respectively.

Equation (3.8) shows that the planetary magnetic field intensity can be derived from observed quantities. For WASP-12, using the upper limit of $B_\star < 10$ G (Fossati et al., 2010a) and the stand-off distance obtained from the near-UV transit observation ($r_M = 4.2 R_p$), Vidotto et al. (2010b) predicted an upper limit for WASP-12b’s planetary magnetic field of $B_p < 24$ G.
Detecting Magnetic Fields in Other Hot Jupiters

The model developed for WASP-12b can also be used to detect magnetic fields in other transiting planets. If we assume that the bow shock absorption would occur at similar wavelengths as that observed for WASP-12b, then, in order to derive planetary magnetic field strengths, near-UV observations of other systems are required. Because acquisition of near-UV transit data requires the use of space-borne facilities, follow-ups and new target detections are rather expensive.

In order to optimize target selection, Vidotto et al. (2011a) presented a classification of the known transiting systems according to their potential for producing shocks that could cause observable light curve asymmetries. The main considered assumption was that, once the conditions for shock formation are met, for it to be detected, the shock must compress the local plasma to a density sufficiently high to cause an observable level of optical depth. This last hypothesis requires knowledge of the local ambient medium that surrounds the planet, which should be derived from models of stellar winds and coronae.

By adopting simplified hypotheses, namely that up to the planetary orbit the stellar corona can be treated as in hydrostatic equilibrium and isothermal, Vidotto et al. (2011a) predicted the characteristics of the ambient medium that surrounds the planet for a sample of 125 transiting systems, and discussed whether such characteristics present favorable conditions for the presence and detection of a bow shock. They found that the hot Jupiters WASP-19b, WASP-4b, and WASP-18b are among the extrasolar planets with the highest potential to present early-ingress in the near-UV.

The models derived in Vidotto et al. (2010b, 2011a) considered these shocks to be static. However, a second set of near-UV observations showed that the early ingress observed in the near-UV light curve of WASP-12b presents temporal variations. Haswell et al. (2012) found that the early ingress of WASP-12b was more delayed in the observations of 2010 March than in the first set of observations from 2009 September. This temporal variation may indicate that the stand-off distance between the shock and the planet is varying. It has been suggested that the variability of the stand-off distance is a direct response of the magnetosphere of the planet to variations in the stellar coronal material itself (Vidotto et al., 2011c). Such variation might be due to many potential factors including a non-axisymmetric stellar corona, planetary obliquity (allowing the planet to move through different regions of the host star’s corona), or intrinsic variations of the stellar magnetic field (resulting in stellar wind changes, CMEs, flares, magnetic cycles).

This suggestion has been verified recently in the simulated near-UV transit light curve of HD 189733b. Llama et al. (2013) performed a detailed modeling of the stellar wind, which incorporated the observationally reconstructed magnetic map of the host star. The three-dimensional stellar wind simulations were then used to determine the local stellar wind conditions.
throughout the orbital path of the planet. Llama et al. (2013) confirmed that, depending on the nature of the stellar magnetic field, and hence its wind, both the transit duration and ingress time can vary when compared to optical light curves. They also found that the signature of the bow-shock can become undetectable when the planet plunges through low-density stellar wind regions. As a result, consecutive near-UV transit light curves may vary significantly. Therefore, continuous monitoring of transits can not only serve as a way to detect planetary magnetic fields, but also could provide invaluable insights into the structure and evolution of the stellar wind.

3.5 Star-Planet Interactions

The approximately 20% of the extrasolar planetary systems containing hot Jupiters—planetary systems characterized by giant planets located a few stellar radii from their parent stars—provide a laboratory for the indirect detection of planetary magnetic fields as a result of star-planet interactions (SPI). Conceptually, the idea is the following. Zeeman Doppler imaging (ZDI) can be used to determine the line-of-sight magnetic field component, or even reconstruct the magnetic map of the star, thereby measuring or constraining the stellar magnetic field $B^\ast$. If an effect that depends upon the interaction of the stellar and planetary magnetic fields can be measured, then the planetary magnetic field $B_p$ can be constrained.

Over a dozen studies of hot Jupiter systems (e.g., Shkolnik et al., 2003, 2005, 2008; Walker et al., 2008; Pagano et al., 2009; Lanza, 2009, 2010; Pillitteri et al., 2010) have independently converged on the same scenario: a short-period planet can induce activity on the photosphere and upper atmosphere of its host star, making the star itself a probe of its planet. This class of interactions between a planet and its host star are known as star-planet interactions (SPI).

The first such monitoring campaign of chromospheric emission from hot Jupiter host stars revealed that stellar activity tracers vary with the planet’s orbital period rather than the star’s rotation for several systems (Figure 3.14; Shkolnik et al., 2003, 2005; Gurdemir et al., 2012; Pillitteri et al., 2011) also reported repeated coronal X-ray flares from one of the stars monitored, HD 189733, at the same orbital phase. These planet-phased phenomena are interpreted as evidence for magnetic reconnections in the stellar magnetosphere induced by the magnetized planet.

Magnetic SPI in hot Jupiter systems is detectable because the planets in general lie within the Alfvén radius of their parent stars ($< 10 R_\ast$). Within this distance, the Alfvén speed is higher than the stellar wind speed, thereby allowing direct magnetic interaction with the stellar surface. If a hot Jupiter is magnetized, mechanisms by which its magnetosphere can interact with the open coronal fields of its star throughout its orbital motion include magnetic reconnection (Lanza, 2008, 2009; Cohen et al., 2009), propagation of Alfvén waves within the stellar wind.
Figure 3.11: Planet-induced chromospheric Ca II K emission at 3933 Å for the solar-type star HD 179949 is phased at the planet’s orbital period of 3.1 days with an amplitude of 0.3% of the continuum. Shown are data from 2001 and 2002 (black) and 2005 (red). The slight phase offset between the 2001–2002 and 2005 data is not significant, but the phase offset of the peak from the subplanetary point ($\phi = 0$) is significant and indicates an interaction with spiraled stellar magnetic field lines and provides a measure of the twisting of the stellar field. For clarity, the data are repeated for two orbital cycles. [Figure adapted from Shkolnik et al. (2008).]

(Preusse et al., 2006; Kopp et al., 2011), or the generation of an electron beam which strikes the base of a stellar corona (Gu & Suzuki, 2009). Figure 3.12 shows the magnetic SPI from an MHD simulation of the HD 179949.

3.5.1 Planetary Effects on Stellar Angular Momentum Evolution

The magnetized stellar winds of main-sequence FGK stars act as brakes on the stellar rotation, thereby decreasing the global stellar activity. This magnetic braking produces age-rotation-activity relationships (e.g. Mamajek & Hillenbrand, 2008). However, if a hot Jupiter is affecting the star’s angular momentum, then the age-activity relation of these systems will systematically underestimate the star’s age, rendering “gyrochronology” (Barnes, 2007) inapplicable to such systems. A systematic underestimation of stellar ages would then affect conclusions about the evolution and migration of their planets.

Conversely, several stars with planetary companions display excess rotation compared to evolutionary models, an excess presumed to be due to (tidal) spin-up of the star by the planet (Pont, 2009; Brown et al., 2011; Schröter et al., 2011; Pillitteri et al., 2011). In two of the cases, no X-ray emission from known M dwarf companions is detected (CoRoT-2 and HD 189733). The lack of X-ray emission indicates that the systems are more than 2 Gyr old (West et al., 2008), but the activity-rotation age of the planet hosts are 100–300 Myr for CoRoT-2 and 600 Myr for HD 189733. These discrepancies would be resolved if the excess rotation and activity on the primaries were due to interactions with the planets, and not their proposed youth.
Figure 3.12: MHD simulation of the HD 179949 system. The coronal field lines that would be opened by the wind remain closed because of magnetic SPI, and the plasma in these loops does not escape. This effect reproduces three observed features: (1) enhancement of total X-ray flux, (2) appearance of coronal hot spots, and (3) phase shift of the hot spots from the subplanetary phase ($\phi = 0$) (Cohen et al., 2009).

Lanza (2010) showed that, in these systems, tides are too weak to spin-up the star. Rather he proposed that the excess rotation is due to interactions between the planetary field and stellar coronal field, which lead to a stellar magnetic field topology with predominantly closed field lines, thereby limiting the stellar wind flow and consequent angular momentum loss. Adopting an analytic linear force-free model, he computed the radial extension of the corona and its angular momentum loss rate. He found that stars with hot Jupiters experience angular momentum loss at a significantly slower rate than similar stars without close-in massive planets. This reduction in angular momentum loss due to the interaction between the stellar and planetary magnetic fields is confirmed in MHD calculations (Cohen et al., 2009, 2010; Vidotto et al., 2011b).

Figure 3.13 shows $M_p \sin i / P_{p,\text{rot}}$, a value potentially proportional to the planet’s magnetic moment, versus $\langle \text{MADK} \rangle$, the average of the integrated Mean Absolute Deviation of the Ca II K line variability for observing runs of stars potentially showing magnetic SPI. There is a potential correlation between a planet’s magnetic moment and the night-to-night chromospheric activity of its host star. A notable deviation from this potential correlation is presented by τ Boo, which has the most massive planet. Because of the near-zero relative motion of the planet relative to τ Boo’s magnetosphere, as a result of the tidal locking of both the star and the planet ($P_{*,\text{rot}} = P_{p,\text{rot}} = P_{\text{orb}}$), only weak Alfvén waves are generated within the stellar magnetosphere,
little excess energy is transported to the stellar surface along the magnetic field lines, and the magnetic SPI is minimal SPI (Gu et al., 2005; Lanza, 2009).

Figure 3.13: $M_p \sin i/P_{rot}$, a value proportional to a planet’s magnetic moment, versus the mean night-to-night chromospheric activity, assuming the planet is tidally locked, i.e., $P_{rot} = P_{orb}$. Filled-in circles are stars with detected planet-induced activity. The τ Boo star-planet system does not follow the trend and represents an exception that supports a proposed Alfvén wave model in which the near-zero relative motion of the planet through the star’s magnetosphere produces minimal magnetic SPI (Shkolnik et al., 2008).

Figure 3.14: Zeeman Doppler Imaging magnetic map of HD 179949. The star is shown in a flattened polar projection down to latitudes of -30°, with the equator depicted as a bold circle and parallels as dashed circles. The three components of the field are displayed in the color scale, with flux values labeled in Gauss. Ticks are the rotational phases of the 10 observations.

Donati et al. (2008b) and Fares et al. (2010, 2012) have used ZDI to determine the magnetic field strength and topology of 10 extrasolar planetary host stars and to look for SPI signatures in their spectra. Figure 3.14 illustrates some of this on-going work, with the ZDI magnetic map for HD 179949.
3.5.2 A Simple Scaling Law to Measure Relative Planetary B-Field Strengths

The existing models of magnetic SPI generally relate the dissipated power $P_d$ to the coronal field strength $B_\ast$, the strength of the planetary field $B_p$, and the relative velocity of the planet with respect to the coronal field lines $v$. Lanza (2009) finds that the dissipated power scales as

$$P_d \propto B_\ast^{4/3} B_p^{2/3} v.$$  \hspace{1cm} (3.9)

This equation is also valid, on dimensional grounds, for dissipation mechanisms more complex than a simple reconnection occurring at the boundary of the planetary magnetosphere, such as those extracting magnetic energy from extended coronal loops interconnecting the star with the planet as, e.g., in the numerical model by Cohen et al. (2011). For a planetary system in which the orbital velocity can be measured and the strength of the stellar field derived from spectropolarimetric measurements or estimated from spectroscopic activity/rotation diagnostics (e.g., Collier Cameron & Jianke, 1994), relative values of the planetary field strength can be inferred from observations of the excess power radiated by a chromospheric hot spot. For example, the magnetic field of the planet HD 179949b is inferred to be approximately 7 times stronger than that of HD 189733b.

Equation (3.9) provides relative, not absolute, magnetic field strengths because the functional dependence on $B_p$ is valid independently of the specific details of the interaction mechanism, provided that the planetary field is dipolar (Lanza, 2012). If combined with an absolute measurement of a planetary magnetic field, such as discussed previously in this chapter, this relation could be calibrated to provide estimates of the magnetic field strength on a number of other planets, even in the absence of a direct measurement.
A main focus of this Study was to identify mission concepts capable of detecting and measuring magnetic fields from extrasolar planets. Guided by experience gained from studying the solar system, we anticipate that the initial detection of magnetic field signatures may be possible from the ground, but that any such detections are likely to be special cases (e.g., massive planets, close-in planets, etc.). Full characterization, particularly with respect to the habitability of planets, will require space-based missions. In this section, we summarize the experience from solar system planets motivating this approach. For a discussion of a possible approach to detecting magnetic fields using the WFIRST mission planned as the next NASA Astrophysics flagship mission following the James Webb Space Telescope, see §2.5.

The radio emission emitted by the electron cyclotron maser in the Jovian magnetosphere is well known (Burke & Franklin, 1955; Franklin & Burke, 1956) and is easily detectable from the ground. The radio emission from other solar system planets requires space-based observations as it occurs at frequencies below the typical transmission frequency of the Earth’s ionosphere. Notably, the requirement to conduct space-based observations occurs even in the case of the Earth’s radio emissions (Figure 4.1)!

Terrestrial aurorae are beautiful representations of the terrestrial magnetic field. However, as discussed in §3.2, the strongest emission lines for solar system planets, and likely the highest contrast ratio with the host stars of nearby extrasolar planets, occurs shortward of about 1500 Å.

---

1 The Radio JOVE project enables students to build a simple decametric wavelength radio telescope capable of detecting the radio emissions from Jupiter, http://radiojove.gsfc.nasa.gov/.
Figure 4.1: Auroral kilometric radiation (AKR) from the electron cyclotron maser instability in the terrestrial magnetosphere as discovered by the Radio Astronomy Explorer-2 (RAE-2) mission. The abscissa shows Universal Time on 1973 December 12, as the orbit of RAE-2 took it behind the Moon, with the times of immersion and emersion of the Earth from behind the Moon labeled (approximately 15:00 to 15:32 UT). The top panel shows the dynamic spectrum (with the frequency labeled on the right hand side of the panel). The lower panels show the time series (light curve) at various frequencies (MHz), with the received power characterized by the antenna temperature. AKR is distinguished by the temperature increase when the Earth is in view relative to when it is occulted. The temperature increase prior to emersion illustrates that the AKR emission region is larger than the Earth itself. For reference, the Earth’s ionospheric cutoff frequency, at which the ionosphere becomes opaque, typically exceeds 10 MHz. [Figure from Alexander et al. (1975)]

Similarly, should $\text{H}_3^+$ emissions be able to be used to study planetary magnetic fields (§3.3), the emissions occur in the 2–4 µm band. While portions of this band are accessible from the ground, as Spitzer has demonstrated, full access requires a space-based mission.

4.1 Summary of Concepts

In the course of this Study, a number of future mission concepts were explored as being capable of studying the magnetic fields of extrasolar planets. After an initial wide-ranging exploration,
discussions of technical feasibility and scientific return led the Study participants to focus on four mission concepts, presented here in no particular priority order.

**Large Number, Small Diameter Radio Receptors (§4.2.2)** This mission concept involves a constellation of spacecraft, each spacecraft carrying one (or a few) radio receptors with individually very little gain, but combined as a phased array to obtain sufficient sensitivity to detect radio emission from extrasolar planets.

**Large Diameter, Small Number Radio Receptors (§4.2.3)** This mission concept involves one or a small-number constellation of radio receptors with significant gain. If multiple radio receptors are required to obtain sufficient sensitivity, they would also be operated as a phased array. Like the previous concept, this mission would aim to detect radio emission from extrasolar planets.

**UVIS-optimized Next Generation Space Telescope (§4.3)** This mission concept is a UV-optimized space telescope with a significantly larger aperture than the Hubble Space Telescope, consistent with that of the Large UV/O/IR Surveyor (*Astrophysics Roadmap*, 2014) and the High Definition Space Telescope (*Dalcanton et al.*, 2015). One of the science goals of such a mission could involve the detection of UV aurorae from extrasolar planets.

**Multi-wavelength Observing Mission (§4.4)** This program would involve the study of both extrasolar planets and their host stars and could involve a significant, or even exclusive, aspect of ground-based observations.

We now describe these concepts in somewhat more detail.

### 4.2 Space-based Radio Telescope

#### 4.2.1 General Considerations

Key design parameters for a telescope are sensitivity, angular resolution, and frequency range. Figure 4.2 illustrates the requirements for sensitivity and frequency range, using Jupiter as a guide. The other planets in the solar system do not emit above 1 MHz, and their flux densities are typically less than that of Jupiter.

An immediate conclusion is that a space-based radio telescope will be required, operating at frequencies around 10 MHz or lower. A space-based radio telescope is required because the Earth’s ionosphere generally becomes opaque at these frequencies. As the frequency range maps directly to the polar magnetic field strength of the planet (§3.1), it may be possible to detect giant planets with the strongest magnetic fields from the ground (e.g., *Zarka et al.*, 2015), but the full range of planetary magnetic field strengths will require a space-based telescope.
It is also apparent that substantial sensitivity or collecting area will be required ($> 10^6 \text{ m}^2$). Moreover, the sensitivity and angular resolution requirements are coupled because of confusion, or the integrated flux density of weak sources within a resolution element. For terrestrial telescopes, both gravity and cost limit the practical size of a single reflector. Based on existing radio telescopes, the maximum diameter for a fully steerable reflector appears to be about 100 m. At the wavelengths relevant for magnetospheric radio emissions ($\lambda \sim 30 \text{ m}$, $\nu \sim 10 \text{ MHz}$), such a single reflector has insufficient sensitivity and poor angular resolution ($\theta \gtrsim 4^\circ$). Such a telescope would be incapable of detecting extrasolar planets not only because its raw sensitivity would be insufficient, but also because its confusion flux density limit would be well above that reasonably expected for any extrasolar planet.

The confusion limit is a general property of telescopes, and the only mitigation is to improve the angular resolution. At radio wavelengths, the technique adopted has been interferometry or aperture synthesis (Thompson, Moran, & Swenson, 2007), in which the total aperture is split into a number of smaller apertures and the signals from those smaller apertures are combined coherently to synthesize an aperture with an effective diameter comparable to that of the largest separation between the smaller, component apertures. A classic problem in designing a radio array is determining the ideal number of receivers vs. antenna size. This problem has been much studied for ground-based interferometers (e.g., Schilizzi et al., 2007), and the solutions can be divided into two general classes: combining a small number of relatively large diameter apertures (SNLD) or combining a large number of relatively small diameter apertures (LNSD).
For a space-based telescope, significantly different considerations apply, and the solutions generally adopted on the ground must be revisited. For instance, while total mass remains a significant cost driver, gravity is no longer a factor in determining the maximum size of a mechanically stable antenna, and very little mass is required to hold the shape of the antenna. Thus, for a single aperture or an SNLD interferometer, much larger diameter apertures can be considered. For an LNSD concept, at the relevant frequencies, low-gain elements (e.g., dipoles) could naturally be used, but these elements will be sensitive to radiation from nearly the full 4\pi sr of sky in contrast to ground-based dipole-based interferometers that have a natural ground-plane and receive radiation from no more than 2\pi sr. Consequently, new signal processing algorithms are likely to be needed.

The KISS Study did not attempt to make a choice between these two potential solutions. Rather, illustrative concepts were developed, with the recognition that considerable refinement is needed, ideally guided by the detection of extrasolar planetary radio emission.

Finally, while the focus of the KISS Study was on extrasolar planets, it was also recognized that having such capability at radio wavelengths would be of significant benefit for cosmology. The New Worlds, New Horizons Decadal Survey also identified “Cosmic Dawn” as one of three key science objectives (Blandford et al., 2010). Following recombination (redshift $z \sim 1100$), the Universe entered a largely neutral state in which neutral hydrogen (H\textsubscript{I}) was the dominant baryonic component of the intergalactic medium (IGM). The highly redshifted 21 cm hyperfine transition of H\textsubscript{I} provides unique information about the state of the IGM and large-scale structures during the formation of the first stars and potentially can probe the IGM prior to their formation. Multiple epochs can be identified, but only poorly constrained by current observations (Furlanetto et al., 2006; Pritchard & Loeb, 2012).

In particular, there is a predicted redshifted H\textsubscript{I} spectral feature arising during the “Dark Ages,” at $100 \lesssim z \lesssim 35$, before the formation of the first stars. During this epoch, the H\textsubscript{I} gas was influenced only by collisions and absorption of photons from the cosmic microwave background (CMB). The H\textsubscript{I} spectral feature is recognized as potentially the only means of probing the Universe between recombination and the formation of the first stars. Moreover, because it is a spectral signature, the H\textsubscript{I} spectral feature potentially allows the evolution of the Universe to be tracked, in contrast to the CMB, which is a continuum measurement at essentially a single redshift.

Both extrasolar planetary magnetic fields and studies of the Dark Ages require observations in a common frequency range, creating a potential synergy. The redshifted H\textsubscript{I} spectral feature from the Dark Ages occurs at radio frequencies $10$ MHz $\lesssim \nu \lesssim 40$ MHz, frequencies that are potentially well matched to the range over which some planets might produce electron cyclotron maser emission. Moreover, while not a specific science target described for the Cosmic Dawn
Mapper (Astrophysics Roadmap, 2014), the capabilities of that instrument would likely have considerable applicability for extrasolar planetary radio emission.

### 4.2.2 Large Number, Small Diameter Interferometric Array

This architecture is the standard for ground-based interferometric arrays, and an extensive literature exists on the design considerations of such arrays (e.g., Thompson, Moran, & Swenson, 2007). Briefly, such arrays have a hierarchical structure consisting of a number of distributed receptors that couple to the incident radiation field, then transmit the signals to a central location for further processing, a central signal processing capability at which specified processing such as cross-correlation occurs, and additional sites at which higher level data products are produced. The hierarchy can have multiple levels, e.g., the signals from a group of receptors (often termed a “station” or “node”) may be combined before being transmitted to the central signal processing location.

There is a long history of experience with ground-based interferometric arrays at radio frequencies comparable to those relevant for the detection and study of the electron cyclotron maser emission from extrasolar planets. Indeed, the first detection of the radio emission from Jupiter occurred with the Carnegie Institution of Washington’s Department of Terrestrial Magnetism “Mills Cross” array, approximately 0.6 km in size and operating at 22 MHz (Burke & Franklin, 1955; Franklin & Burke, 1956), and the first attempt to detect the radio emission from extrasolar planets was conducted at the Clark Lake Radio Observatory, which had maximum dimensions of 3 km × 0.4 km and was operating at 26 MHz at the time (Erickson & Fisher, 1971; Yantis et al., 1977).

Existing large interferometric arrays, with architectures potentially similar to what could be realized for a space-based array, include the Giant Metrewave Radio Telescope (Swarup, 1990; Ananthakrishnan, 1995), the 74 MHz Very Large Array (Kassim et al., 2007), the Ukrainian T-shaped Radio Telescope (UTR-2, Braude et al., 1978), Low Frequency Array (van Haarlem et al., 2013), the Long Wavelength Array (Ellingson et al., 2013; Taylor et al., 2012), and the Murchison Widefield Array (Tingay et al., 2013). Figure 4.4 illustrates the concept.

The concept of a space-based radio interferometer is not new (e.g., French et al., 1967), and there have been initial demonstrations of the capability to conduct both radio astronomical and radio astronomical interferometric observations from space. The first Radio Astronomy Explorer (RAE-1) spacecraft was in an Earth orbit and made the first measurements of the Galaxy’s spectrum between 0.4 and 6.5 MHz (Alexander et al., 1969), while the second Radio Astronomy Explorer (RAE-2) spacecraft was in a lunar orbit and observed between 25 kHz and 13 MHz (Alexander et al., 1975). The Earth’s auroral kilometric radiation, generated by an electron cyclotron maser in the magnetic polar regions, has been studied by simple space-based arrays including a single-element interferometer consisting of the ISEE-1 and ISEE-2 spacecrafts.
(Baumberg et al., 1986) and a time-difference-of-arrival (TDOA) analysis with the Cluster spacecraft (Mutel et al., 2004).

There have also been initial descriptions and proposals of concepts for radio astronomy arrays of small spacecraft, notably including the Astronomical Low Frequency Array (ALFA) mission concept (Jones et al., 2000), and “cubesat”-based arrays (Banazadeh et al., 2013), for which the detection and study of extrasolar planets was either a part of the science mission or the prime science mission. In this regard, the notion of cubesats—small spacecraft with defined form factors and standards for many components—make the realization of an array of a large number of cubesats feasible.\(^2\) In this concept, each cubesat or small spacecraft would carry a single (dual polarization) antenna, with the array synthesized from the collection of spacecraft (Figure 4.3).

\[\text{Figure 4.3: Artist's illustration of how a large number of small spacecraft ("constellation") could be realized as an interferometric array. In this concept, each small spacecraft, which could be a cubesat, would host a radio antenna. Combining the signals coherently produces an effective aperture.}\]

\(^2\) Cubesats are defined in terms of “units,” with 1 unit (1U) defined to be a volume of 10 cm × 10 cm × 10 cm (1,000 cm\(^3\)) and a mass of no more than 1.33 kg per unit. Multiple unit cubesats are also possible, with 2U and 3U cubesats as common examples.
Figure 4.4: (Top) Long Wavelength Array at Owens Valley Radio Observatory (LWA-OVRO), an illustration of a large-number, small-diameter interferometric array. The signals from all of the individual dipole antennas are transmitted to a central location for processing, producing an effective aperture equivalent to the approximate maximum separation between the antennas ($\approx 300$ m for the LWA-OVRO). A similar concept applies for a space-based array. (In the background are other antennas at the Owens Valley Radio Observatory.) (Bottom) LWA-OVRO sky in the 30–43 MHz (left) and 47–78 MHz bands (right). Strong sources are labeled, notably including Jupiter and the Sun. The absence of Jupiter from the higher frequency image is consistent with the exceptionally strong cutoff of cyclotron maser emission when the local plasma frequency exceeds the local cyclotron frequency within the planet’s magnetosphere (§3.1.1). A future space-based array could search for and study extrasolar planets in a manner similar to how the LWA-OVRO can observe Jupiter. [LWA-OVRO sky images courtesy of M. Anderson]
Many of the relevant components have been or soon will be demonstrated in space, e.g., the U.S. Naval Academy’s Radar Fence Transponder (RAFT1)\(^3\) cubesat demonstrated a “high frequency” (HF) antenna for the relevant frequency range for extrasolar planetary radio emissions, and the Interplanetary NanoSpacecraft Pathfinder In a Relevant Environment (INSPIRE) pair of cubesats will demonstrate the ability of cubesats to operate in a deep space environment (Figure 4.5, Klesh et al., 2013).

While many of the components for a small spacecraft/cubesat-based array are feasible, there remain system-level aspects of the concept for which additional work is needed.

**Communications** By traditional ground-based standards, the data rate from an individual antenna is relatively modest. The incident radiation field is noise-like, so a relatively small number of bits for sampling can be used. For a notional 10 MHz bandwidth, with Nyquist sampling, the data rate from an individual antenna could easily be less than 40 Mbit s\(^{-1}\), well within the capability of a wireless router employing the 802.11 standard. However, for an array of antennas, the data rate can become quite large. Generally, the data rate will be

\[
\dot{D}_{\text{array}} = N_{\text{pol}}(2 \times \Delta \nu) N_{\text{bit}} N_{\text{ant}},
\]

where \(N_{\text{pol}}\) is the number of polarizations measured (and 2 would be the standard), \(\Delta \nu\) is the bandwidth (with the factor of 2 accounting for Nyquist sampling), \(N_{\text{bit}}\) is the number

\(^3\) \url{http://aprs.org/raft.html}
Chapter 4. Mission Concepts

of bits used in the analog-to-digital converter, and \(N_{\text{ant}}\) is the number of antennas in the array. A modest number of antennas (\(N_{\text{ant}} = 256\), comparable to that of ground-based arrays) can easily produce data rates well in excess of 10 Gbit s\(^{-1}\).

Such data rates are too large to transfer to Earth. There are multiple approaches to handling such data rates (Babuscia et al., 2013), including efficient methods for data compression and using a "mothership" for (a portion of the) initial processing to reduce the data rate before transmission to the Earth.

**Orbits and Propulsion** The required orbit is beyond low-Earth orbit (LEO), and likely beyond geosynchronous (GEO). An ideal location for a space-based array would be at 5–10 times more distant than the Moon, as that distance is sufficiently far that the terrestrial AKR is reduced to a comparable intensity as the Galactic emission (Figure 4.1). An acceptable distance is likely to be the lunar distance. The clear trade-off is that more distant locations require more fuel, longer transit times, and have more demanding communication requirements.

Families of orbits have been identified that require very little energy to access, the so-called “interplanetary superhighway” (Staehle et al., 2013). With the relatively small amounts of propulsion that an individual small spacecraft/cubesat can carry, accessing some of these orbits can require a long time. An alternate approach is to have a "mothership," with a larger propulsion capacity to carry the array to the desired location. A "mothership" would also likely to be able to provide more power to the telecommunications system, thereby reducing the telecommunications demands on the individual small spacecraft/cubesats.

### 4.2.3 Large Diameter, Small Number Interferometric Array

On the ground, stability concerns and gravitational stresses limit single apertures to diameters of typically about 100 m. Exceptions are those telescopes built by taking advantage of special geographical formations, such as the 300 m diameter William E. Gordon Telescope of the Arecibo Observatory and the 500 m diameter Five hundred meter Aperture Spherical Telescope (FAST). Even ignoring ionospheric reflection, these single aperture telescopes would be of limited utility for radio emissions from planetary magnetospheres. Taking the “rule of thumb” that an aperture must be approximately six wavelengths (6\(\lambda\)) in diameter to serve as an effective reflector, the Arecibo telescope would not function effectively at wavelengths longer than about 50 m (frequencies below 6 MHz).

Gravitational stresses are significantly reduced in space, allowing much larger apertures to be considered (e.g., Hedgepeth, 1970), which can function at the relevant wavelengths for
Figure 4.6: Artist’s illustration of how a large diameter single aperture could be realized in space. In this concept, small spacecraft (“cubesats”) would control the shape and steering of the telescope. Moreover, while the focus of this work is on the use of such a telescope for studying extrasolar planets, it might also be able to be used for studies of the highly redshifted neutral hydrogen hyperfine line from the Cosmic “Dark Ages.”

In a conductor, currents tend to flow along surfaces and edges because of the mutual repulsion of the conducting charges. For this reason, an acceptable Strehl ratio can be maintained even if the reflector is constructed from a wire mesh rather than a solid surface. Indeed, for approximately the first decade of its operation, the Arecibo telescope was constructed from a wire mesh, supported by cables.

\(^4\) http://www.kiss.caltech.edu/study/largeststructure/index.html
Because a shallow parabolic reflector has a better field-of-view and because in the absence of gravity essentially any shape is possible, we adopt an $f/D$ ratio (focal length over diameter) of 1.0. We illustrate the design of a wire mesh reflector, considering the size of the wire, the spacing between the wires and the number of points at which the reflector needs to be anchored to approximate a parabola. The results of these analyses can be re-scaled for different diameters and wavelengths.

We do not consider the feed system, but the design of a reflector is to focus the radiation to a point (or a plane). A single small spacecraft (or small number of them) could be equipped with feeds and receivers to sample the focused electric fields.

**Wire Spacing** Various theoretical calculations and numerical simulations suggest that a reflecting grid will perform reasonably well with wire spacings as large as a quarter of a wavelength. For an upper frequency of 10 MHz ($\lambda \sim 30$ m), the resulting wire spacing would be 7.5 m, providing a reflectivity of 90% ([Hill & Wait, 1976](#)). The effect of wire spacing on aperture efficiency is more critical. A continuous surface would have an efficiency of 0.8. With a spacing of a fifth of a wavelength, the efficiency would be 0.5 and, for a quarter wavelength, the efficiency would be 0.25. As the purpose of this study was not an optimized point design, we adopted 10 m spacing.

**Wire Arrangement** The simplest arrangement would be a series of radial wires connected by azimuthal wires. Such an arrangement has a greater concentration at the center than necessary, and so the central parts of some of the radial wires can be omitted while still staying within the wire separation constraint (Figure 4.7). This adjustment reduces the amount of wire by about 20% for a total wire length of about 95 km.

**Wire Thickness** The reflectivity of the mesh depends on its sheet resistance, which in turn depends upon the resistance of the wires. Because the current flows near the surface of the wire, we can reduce the wire size until the effective cross sectional area of the wire is affected. We required the reflectivity to be 95%, which translates into a requirement that the resistance of each 20-m wire segment be less than 10 $\Omega$, or a limit on the resistivity of 0.5 $\Omega$ m$^{-1}$. For copper wires, a diameter of 0.5 mm (AWG 24) meets this requirement (Figure 4.8), and, for aluminum wires with a 1.5$\times$ higher resistivity, a diameter of 0.75 mm (AWG 20) meets the requirement.

**Wire Mass** Given the length, diameter, and density of the material, the mass of the wire needed to construct a single 1 km diameter reflector can be estimated. For copper wires, the total mass of wire would be approximately 180 kg. The density of aluminum is 30% that of
copper, resulting in a total mass of only slightly more than 120 kg, but the volume would be $2.25 \times$ larger.

**Wire Anchoring** In order to maintain the shape of the reflector, a series of anchor points will be required (i.e., specific points at which station keeping would be required). These positions could be maintained by small spacecraft. We required the position of the wire to be maintained to within $0.08\lambda$. This requirement can be achieved with five anchors along a radial wire. We consider each cell of the mesh to be a plane and, with 24 of the radial wires so anchored, the radial center lines of these planes would not exceed the specified tolerance. Thus, the required shape can be maintained by 120 anchors.

![Diagram](image)

**Figure 4.7:** Arrangement of wires to produce a large reflector. The thickness of the wires is not to scale. For an actual 1-km-diameter reflector, the wires would not be visible.

4.3 UVIS-optimized Next Generation Space Telescope

In the near-term, the primary facility for UV studies of extrasolar planets will continue to be HST. With the repair of the Space Telescope Imaging Spectrograph (STIS) and the installation of the Cosmic Origins Spectrograph (COS), HST is at the peak of its capability for UV extrasolar planet
Figure 4.8: Wire resistance as a function of wire diameter, for copper wires. The various curves show the resistance at different frequencies.

studies. Examples of the current capability of HST include a growing number of transiting planets observed in Lyα and metal line absorption (e.g., Linsky et al., 2010; Ehrenreich et al., 2012; Lecavelier des Etangs et al., 2012; Ben-Jaffel & Ballester, 2013), as well as the emergence of programs that have combined the strengths of STIS and COS to provide the first quantitative measurements of the energetic radiation environments in the habitable zones around M dwarf extrasolar planet host stars and potential effects on the production of biomarker molecules (e.g., France et al., 2013; Tian et al., 2013).

Following the demise of HST (possibly in the 2018–2021 timeframe), the intermediate-term options through NASA for access to UV imaging and spectroscopy will likely be from (Small) Explorer ([SM]EX) missions and stand-alone missions of opportunity (SALMONs). Internationally, there will be the Russian/Spanish-led World Space Observatory-Ultraviolet mission (WSO-UV, Shustov et al., 2011). WSO-UV is a 1.7-m primary aperture space observatory that may be launched as early as 2018 and would include far- and near-UV spectroscopy as well as far-UV imaging. The spectral sensitivity in the near-UV is projected to be larger than HST at some wavelengths, though the far-UV spectroscopic channel will have STIS-like effective area while employing CCD detectors, which will almost certainly limit its effective sensitivity for faint targets. The far-UV imaging instrument on WSO-UV, ISSIS, will have HST/ACS- or HST/SBC-like imaging performance over approximately 4 times larger fields and may be a valuable instrument for directly imaging auroral emission from exo-gas giants around the nearest cool stars.
In the longer term, the astronomical community can look forward to a UVOIR flagship observatory that will be the successor to Hubble, and the recent NASA Astrophysics Roadmap describes a Large UV/O/IR (LUVOIR) Surveyor as one of its notional mission concepts (Astrophysics Roadmap, 2014). Realizations of LUVOIR include the ATLAST (Postman et al., 2009) and, more recently, the High Definition Space Telescope (Dalcanton et al., 2015), which focus on an 8–16-m primary aperture segmented mirror telescope that would unfold in space (Figure 4.9). In addition to wide-field, high angular-resolution imaging of extrasolar planetary systems, this instrument will likely be equipped with UV spectroscopic capabilities. The near- and far-UV have been shown to be the wavelength regions of largest effective radius for Earth-like planets orbiting solar-type stars (owing to the peak of the photoabsorption cross-sections of \( \text{O}_3 \) and \( \text{O}_2 \), respectively (Bétrémieux & Kaltenegger, 2013), and one of the prime motivations for the ATLAST concept is the direct detection of biomarker molecules in the atmospheres of habitable zone planets.

![Figure 4.9: Artist's impression of one concept for the ATLAST flagship observatory, optimized for observations in the ultraviolet to visible wavelength spectral range. This figure shows a 16-m-diameter aperture after having been deployed. [Credit: STScI.]](image)

A future NASA mission with a UV spectroscopic capability would provide fundamentally new insights into how extrasolar planetary systems form and the physics that governs their atmospheres. However, in order to achieve all of the goals of a long-term large UV-enabled mission, larger telescopes alone are likely not the solution within reasonable cost-envelopes for such a project. Because high-sensitivity and low background equivalent fluxes are a requirement for this mission, improvement on primary optical and detector subsystems (e.g., Sembach et al., 2009; COR Technology Report, 2012) will offer gains for the ultraviolet wavelength range comparable to the gains achieved by increasing the primary aperture diameter from 2.4 m to 8 m. Advances in component technology such as high-reflectivity UV coatings (factor of 3 improvement per optic
at $\lambda < 1100$ Å; Beasley et al., 2012) and low-noise borosilicate glass photon-counting detectors (Siegmund et al., 2011; factor of approximately 10 lower noise than HST/COS detectors) will provide many of the advantages of a large telescope for a fraction of the cost. We suggest that including both a high-resolution point source spectrograph and a multi-object spectrograph operating at lower resolution will provide the largest grasp of observatory discovery space for extrasolar planet and related research. Technology investment in a low-scatter echelle UV spectrographs (e.g., France et al., 2012b; factor of up to approximately 10 improvement in scattered light control at $R > 10^5$) would provide a means for achieving the order-of-magnitude gains necessary to carry out the science without the commensurate increase in telescope diameter.

### 4.4 Multi-wavelength Observing Mission

During the Study, the concept of a multi-wavelength telescope was identified as being potentially valuable for monitoring the space weather and star-planet interactions of nearby stars. This concept was not developed in great detail, but, for completeness, we record the essential elements of such a mission concept.

The scientific objectives of such a mission would be to serve both as a monitoring capability and an “alert” system. By monitoring nearby stars, the telescope could look for bow shocks associated with any transiting systems. By monitoring nearby stars, it could also detect any stellar flares. Such detections could serve several purposes, including characterizing the rates of such flares and producing alerts for other telescopes to search for specific signatures of star-planet interactions.

The implementation of such a mission would focus on photometry. There was some discussion that low resolution spectroscopy would be useful, but a future trade study would have to be conducted regarding the actual spectral resolution that might be obtainable. The wavelength coverage of such a mission would be at least simultaneous X-ray, UV, and visible wavelengths, with consideration also given to near-infrared coverage. Reasonably moderate apertures would likely suffice, e.g., a 50-cm-diameter aperture for the UV and visible wavelengths as that of the Galaxy Evolution Explorer (GALEX), and an X-ray aperture comparable to that of ROentgen SATellite (ROSAT).

Two possibilities for mission implementations were considered. One was an integrated set of apertures, mounted on the same platform. A possible approach would be a modified Evolved Expendable Launch Vehicle (EELV) Secondary Payload Adapter (ESPA). The alternate approach would be a constellation of small spacecraft conducting coordinated observations. Deployable apertures obtaining 20-cm or larger diameters now appear feasible from small spacecraft, including cubesats (e.g., Andersen, 2011; Dearborn et al., 2014).
5. Roadmap and Future Steps

5.1 Theoretical and Modeling Efforts

The KISS Study found that the effectiveness of magnetic fields for atmospheric retention remains ambiguous and merits further study.

In the next decade, we can take steps to advance our theoretical understanding of dynamos and of planet evolution, to better interpret eventual observations of extrasolar planet magnetic fields. Measurements of planetary magnetic fields will, in turn, help to further improve magnetic field strength scaling laws, and our understanding of the solar system planets.

Two kinds of modeling efforts are needed, addressing questions such as those listed below, which we then describe further in the subsequent sub-sections.

**Planetary Structure and Evolution**
What are the possible planetary structures and their evolutions?

- How are these related to accumulating data on the radii, masses, and atmospheres of extrasolar planets (if any)?
- Which of these structures support a convecting, electrically conducting region and, for those that do, for what age range or ranges?
- What are the relevant material properties? Addressing this question may require new experimental and (quantum mechanical) theoretical efforts, for example, determining
the behavior of silicate liquids under extreme pressures and temperatures, where they may become electrically conducting.

- How are planetary properties affected by formation location, tidal evolution, presence or absence of a moon, nature of the central star, and variability in radiogenic heat sources and composition (e.g., oxygen fugacity in the formation region)?

**Planetary Dynamos**

What is the expected field geometry and strength for a given planetary structure?

- What exactly is the criterion for a dynamo? Is it sufficient to have convection? Or is the criterion substantially more restrictive?

- Is there a scaling law for planetary magnetic fields? What are the most important parameters? Buoyancy flux? Rotation? Conductivity?

- Is there a possible connection between the field we measure and other potential observations (e.g., atmospheric winds)?

### 5.1.1 Planetary Structure & Evolution

Planetary structure is an area of broader interest than merely the existence or absence of a magnetic field, and there is much current work in this area, motivated by Kepler data in particular. However, much of this effort has not concerned itself with the nature of any electrically conducting regions or whether they are fluid and convecting in that region. There are many issues here that require attention. What are the thermodynamic properties of candidate materials and their mixtures? What are the possible outcomes of planetary accretion? How do these planets evolve (differentiate and cool) over time?

In general, cooling is likely to be crucial to understanding whether a planet can support a dynamo. In Earth’s core, for example, while there is no current consensus about the specific mechanism or mechanisms by which the Earth has maintained a magnetic field over billions of years, there is consensus that any reasonable model requires the core to cool, because even if compositional convection is involved, the buoyancy it provides is determined by cooling rate. In most cases planetary dynamos operate in a region that is overlain by a non-dynamo region (a mantle, an ocean, or an atmosphere). The properties of that non-dynamo region must also be understood as it influences, and often determines, whether the dynamo region is able to cool. For example, plate tectonics on Earth is a very efficient planetary cooling mechanism. It is possible that Earth, and other extrasolar terrestrial planets, would lack a magnetic field if it did not have plate tectonics.

Experimental studies of planetary materials have an important role to play, particularly at the pressures characteristic of the interiors of terrestrial mass planets (as it can be difficult to reach
5.1 Theoretical and Modeling Efforts

much higher pressures). For example, it is important to know whether silicates become electrically conducting at sufficiently high pressure or temperature. Transport properties are also important as they are needed to determine whether convection arises for a given heat flux. Theoretical calculations of material properties are needed, especially at higher pressures (e.g., for super-Earths or ice giants). Current capability in density functional theory makes these calculations possible, though it is still a major challenge to handle the most interesting case: liquids. Determining the structure (the local configurations of atoms, molecules, and ions) is part of the problem, usually approached through classical Monte Carlo or molecular dynamics simulations guided by quantum mechanics for the interaction energy for many different local configurations. The output of these kinds of calculations is then input for planetary structure and evolution calculations.

To date, there has not been an organized approach to modeling the thermal evolution for a wide range of possible planets. A useful approach would be a modeling effort that would quantify the range of heat flows each planet might have and explore whether or not each planet would convect in a region where there is a fluid conductor (be it iron, super-ionic water, or metallic hydrogen). Ultimately, thermal evolution models could be used to estimate—as a function of planet rotation—the non-dimensional numbers that would be helpful inputs to planet dynamo models (e.g., Ekman number).

There is also a considerable range of phase space for planetary structures that has not yet been explored with detailed thermal evolution models. Important parameters include planet mass, planet composition (e.g., rock/ice/gas ratios, which affect the boundary conditions on the dynamo generation region as well as thermal evolution), and the planet equilibrium temperature (which can also affect the thermal evolution in some cases). In terms of thermal evolution, the only planet scenario that is straightforward and has been modeled to date is that of hot Jupiters. Hot Jupiters do convect and cool as their outer radiative zones deepen over time, but complications arise in attempting to expand these models to consider conducting regions that are not composed of hydrogen—conduction becomes important, and the models are more involved and uncertain. An important goal would be to map out, in terms of planet mass, H/He envelope mass, and equilibrium temperature, planet configurations that allow electrically conducting convecting regions.

Initial studies could tackle subsets of this expansive parameter space. For instance, if Uranus or Neptune were scaled to a lower mass, but otherwise everything else were self-similar, would that planet have the opportunity to generate a dynamo in the water-rich part? Would the dynamo eventually switch over to the iron core at lower masses?
5.1.2 Planetary Dynamo Modeling

A more systematic and expansive approach to modeling planetary dynamos would be useful for extrapolating to extrasolar planets. So far only individual planet case studies have been approached, and there are opportunities to consider a wider range of cases.

No precise criterion for the existence of a dynamo exists. It is thought to be similar to the criterion for convection, but further simulations are required to elucidate the range of conditions for which this assumption is correct. One of the difficult issues in addressing this question is the possibility of multiple states—that is, a planet could exist with small or no field or in a strong field state depending on past evolution (path dependence). Usually dynamo simulations are done by imposing a field that is already dynamically important, rather than starting from an infinitesimal field. This initial condition is a quite different state because of the dynamic effect of the field on the characteristic size and pattern of the convective motions.

Our current understanding of dynamos offers only order of magnitude predictions for field strength at best. There is a difference between the field in the dynamo region (which is what matters for the MHD fluid dynamics) and the field that we measure. These differences arise in four ways:

- Radius of the dynamo region vs. radius of the planet: With an inverse cube effect for a dipole, the field that we measure depends upon the location and size of the dynamo region within the planet relative to the size of the planet itself;
- Field geometry: Dipolar vs. higher-order multipoles that decay even faster with distance from the dynamo region, such that the field at the surface may be significantly weaker than that in the dynamo region;
- Poloidal vs. toroidal fields: The field within the dynamo has both, but only the poloidal field is externally observed; and
- Time variability: It could be that any given dynamo varies by as much as an order of magnitude in field strength with time.

These factors can easily combine to give an order of magnitude uncertainty in predicted fields even if one has a reliable scaling law for the typical field in the dynamo region. More work is needed on these factors as well as the “generic” scaling law, assuming such a law even exists.

It is worth noting that brute force dynamo simulations are hard; for a detailed discussion, see Roberts & King (2013). Running a realistic, brute force dynamo simulation requires about 16 orders of magnitude more computer power than currently exists (i.e., 4 orders of magnitude higher resolution in 3 space + 1 time dimensions). Such a large improvement in computational power will not be available soon, if ever. Cleverness rather than brute force is needed. For instance, no modelers use realistic (very low) values of Ekman number (which measures the importance of
viscosity vs. rotation). The relative values of various transport parameters (e.g., ratio of viscosity to thermal diffusivity or magnetic diffusivity) may also matter. It has been argued that it is not necessary to obtain the actual Ekman numbers relevant for planets, as one should reach an asymptotic regime, perhaps one that has some very small dependence on Ekman number, such that extrapolation is possible. Some existing proposals for dynamo scaling laws also assume no dependence on the Ekman number. This assumption is plausible for turbulent convection, for which viscosity has no dynamical role (but is needed in the computer simulation to provide stability and convergence). However, it is also possible that we have not yet encountered the asymptotic regime in existing calculations. There are two complementary approaches. One is the move towards a multi-scale approach to the problem (parameterize sub-grid physics with smaller scale simulations). Another is the attempt to simulate ever smaller Ekman number (though still well short of the planetary value), as in the Japanese supercomputer efforts (Miyagoshi et al., 2011; Kageyama et al., 2008).

Many existing dynamo codes are either Boussinesq (no variation of density in the fluid other than the small differences arising from convection) or allow for only a small range of material property variation (e.g., one pressure scale height). Current codes may not be sufficient to understand bodies for which the material parameters change by many orders of magnitude in the region of interest. There is no published Jupiter dynamo model that reproduces the dipolar character of the observed field and is faithful to the huge conductivity variation in the outer region of likely dynamo generation.

Even assuming dynamo modeling progresses to the point at which dynamo behavior can be predicted for a given set of planetary variables, there will be a need to merge dynamo modeling and thermal history modeling to some extent (Driscoll & Olson, 2009b, 2011b). This goal does not mean running a dynamo model for the equivalent of billions of years but rather identifying “snapshots” of 1D evolution models that define the parameters for the dynamo. Doing so will require good communication between two scientific communities that are often disparate.

5.2 Observational Approaches

In this section we illustrate how progress can be made in the near-term, largely using existing or “in progress” telescopes and facilities that can be used to strengthen the approach toward the detection and study of extrasolar planetary magnetic fields.

5.2.1 Solar System Observations

There are three primary means by which observations of solar system objects can inform current and future searches for extrasolar planetary magnetic fields.
The first is observations of the Sun, with a focus on “the Sun as a star.” Full disk observations of the Sun present a unique opportunity to study a planet’s host and its radiation and particle effects on the atmospheres of the planets that orbit it. The Solar Dynamics Observatory (SDO) has been observing the Sun since 2010. On board is an imager, the Atmospheric Imaging Assembly (AIA), which monitors the full disk at 10 wavelengths every 10 seconds spanning the soft X-ray/EUV through to the optical with a spatial resolution of 1″. In addition, there is an EUV spectrometer, the Extreme Ultraviolet Variability Experiment (EVE), which measures the solar irradiance at these high energies with unprecedented spectral resolution, temporal cadence, and precision. Lastly, HMI (the Helioseismic and Magnetic Imager) extends the capabilities of the SOHO/MDI instrument with continual full-disk coverage at higher spatial resolution and new vector magnetogram capabilities.

Although SDO is monitoring just one planet host, these three instruments monitoring the Sun’s variability across much of the electromagnetic spectrum provide ample opportunities to apply such rich data sets to exoplanet studies. For example, Llama & Shkolnik (2015) showed the effects that solar-like variability would have on planetary transits and the subsequent measurements of planetary radii, and there has been much work in using the 2012 transit of Venus across the solar disk as an exoplanet analog (e.g., Chiavassa et al., 2015).

The second approach involves observations of planetary magnetospheres. The broad outlines of the interactions between a planet’s magnetosphere and the solar wind have been clear for some time (e.g., Gallagher & D’Angelo, 1981; Desch & Rucker, 1983; Desch & Barrow, 1984; Clarke et al., 2009), including dramatic examples in which the Saturnian radio emission ceases when Saturn passes into the Jovian magnetotail (Desch, 1983). Nonetheless, particularly with the discovery of extrasolar planets, the importance of comparative magnetospheric studies remains high and was highlighted as a topic for continued study in both Solar and Space Physics: A Science for a Technological Society (Heliophysics Decadal Survey, Baker et al., 2012) and Vision and Voyages for Planetary Science in the Decade 2013–2022 (Planetary Sciences Decadal Survey, Squyres et al., 2011).

The Saturn Aurora Campaign\(^1\) was conducted during 2013 as part of the Cassini Solstice Mission. This campaign was an effort to conduct a coordinated set of observations with both Cassini instruments and various ground-based telescopes to study the effects of the solar wind during solar maximum (“solar max”) on the Saturnian magnetosphere. At the time of writing of this report, the results of these observations are still being analyzed, but the current plan is that there will be a future special issue of the journal *Icarus* to present the results of these observations and analysis. A similar approach might be conducted once Juno reaches Jupiter (2016 July), though Juno’s prime science mission will occur after solar maximum.

The third approach requires observations of planetary atmosphere loss (§2.4.3). During the KISS Study, there was considerable discussion and no consensus about the extent to which planetary magnetic fields are a factor in shielding planetary atmospheres. While clear evidence has been observed of atmospheric loss from an unshielded planet (Mars), there are also indications that Earth’s present-day O$^+$ mass loss rate is similar to that of Venus and Mars. Continued comparisons of mass loss rates from the terrestrial planets are needed, particularly to ensure that mass loss rate estimates are not biased by observations of a limited number of atmospheric constituents or species that are relatively easy to detect.

5.2.2 Magnetospheric Emissions

Among the various approaches to detecting and studying magnetospheric emissions (Chapter 3), significant effort has already been devoted to searches in both the far-UV and the radio ranges. As described in §3.2, current searches for auroral UV emissions from hot Jupiters has not yet been successful. There is also the concern that the detection of far-UV emission may not provide actual constraints on the strength or properties of an extrasolar planetary magnetic field. Nonetheless, future HST observations may be warranted if extrasolar planets are found around extremely nearby stars (e.g., with TESS).

Section 3.1.4 summarizes current limits on magnetospheric radio emissions from extrasolar planets. As discussed there, most of the published limits are at frequencies above 100 MHz, potentially too high for likely planetary cyclotron maser emission (given that Jupiter’s emission ceases above 40 MHz). There are significant programs underway at a variety of ground-based radio telescopes, including LOFAR, the LWA, the LWA-OVRO, and the UTR-2. Many of these telescopes provide significant improvements on the sensitivity available at frequencies at which planetary cyclotron masers are likely to generate radio emission. Thus, within a few years, it is possible that there will be a detection of extrasolar planetary radio emission.

Section 3.3 discusses the potential of JWST observations of the H$_3^+$ emission from extrasolar planets. An initial analysis suggests that the detection of H$_3^+$ emission will be challenging, in addition to the concern that its emission may not be tied directly to the presence or strength of a magnetic field. Nonetheless, future discoveries may warrant revisiting whether a JWST search for H$_3^+$ emission from nearby extrasolar planets is called for.

5.2.3 Star-Planet Interactions

Building a larger sample of stars within which magnetic SPI has been detected is critical to understanding its underlying mechanisms and extracting planetary field information using the models described in §3.5. In addition, the presence of a chromospheric hot spot should be associated with a greater probability of radio emission from a hot Jupiter (Lanza, 2009;
Lazio et al., 2009; Vidotto et al., 2012). With a larger sample, and improved statistics on the chromospheric SPI in systems with hot Jupiters, there will be a natural symmetry between the radio and SPI observations. Quantitative predictions of the radio flux density can be produced for stars displaying magnetic SPI, while radio detections of planetary systems producing magnetic SPI will calibrate the relative field strengths to an absolute scale.


Brown, D. J. A., Collier Cameron, A., Hall, C., Hebb, L., & Smalley, B. 2011, "Are falling planets spinning up their host stars?" MNRAS, 415, 605


exoplanets.org Web site; http://www.exoplanets.org/


Jakosky, B. M., et al. 2015a, "MAVEN observations of the response of Mars to an interplanetary coronal mass ejection," Science, 350, 210


Bibliography

CME Impact on Expected Magnetospheres of Earth-Like Exoplanets in Close-In Habitable Zones,” Astrobiology, 7, 167


Koskinen, T. T., Yelle, R. V., Harris, M. J., & Lavvas, P. 2013, “The escape of heavy atoms from the ionosphere of HD209458b. II. Interpretation of the observations,” Icarus, 226, 1695


Swarup, G. 1990, “Giant metrewave radio telescope (GMRT) - Scientific objectives and design aspects,” Indian J. Radio Space, 19, 493


Vidotto, A. 2013, “Protecting planets from their stars,” Astron. Geophys., 54, 010001


