A Comprehensive Study of Magnetic Map-making Techniques for Navigation and Localization

J. de Jong

TU Delft

A Comprehensive Study of Magnetic Map-making Techniques for Navigation and Localization

by

J. de Jong

to obtain the degree of Master of Science
at the Delft University of Technology,
to be defended publicly on Friday June 2, 2023 at 02:00 PM.

Student number:	4830164	
Project duration:	September 1, 2022 – June 02, 2023	
Thesis committee:	Prof. dr. ir. M. Verlaan	TU Delft
	Prof. dr. ir. M.B. van Gijzen	TU Delft
	Dr. ir. A.R.P.J. Vijn	TU Delft, TNO, supervisor
	R.G. Tan, MSc.	TNO

TU Delft

Preface

During my first year of pursuing a master's degree in Applied Mathematics, I enrolled in a course on Advanced Modelling which proved to be an enlightening experience. One of the projects assigned to us focused on Indoor Magnetic Navigation utilizing Magnetic Beacons, a subject that was not initially my top choice but had captured the interest of the double-degree students in our group, combining Applied Mathematics and Applied Physics. To my surprise, this project exposed me to a whole new realm of practical applications, which I had not encountered during my bachelor's and master's programs. I reached out to the supervisor of the project, Aad Vijn, to continue working on this topic as my thesis.

For the past nine months, I have delved into various subjects, including magnetism, navigation techniques, Kalman filters, particle filters, and the practical aspects of working with magnetic data, such as magnetometers and compensation. Through this comprehensive study, I have come to appreciate the significance of magnetic maps in achieving GNSS-independent navigation. The availability of accurate maps directly contributes to enhanced positioning accuracy.

Moreover, it is intriguing to observe that magnetic maps serve not only the purpose of precise positioning. It appeared to me that researchers are dedicating their efforts to exploring and solving captivating hide-and-seek games: the search for hidden treasures, valuable minerals, or ferromagnetic objects. This further highlights the multifaceted nature of magnetic mapping.

I would like to express my gratitude to Aad Vijn for his guidance and support throughout this project. I am also thankful to my colleagues at TNO, Mia Jukic, Eugene Lepelaars, and Reinier Tan, for their brainstorming sessions and feedback on my work and for providing me access to their measurement equipment. I would also like to thank Martin Verlaan for his valuable feedback and suggestions. A special thanks to Arian Joyandeh for our running sessions, which were a great source of motivation and in-depth discussion. Finally, I would like to thank my friends and family for their support and encouragement.

J. de Jong
Delft, June 2023

Contents

1 Introduction

1.1. Introduction and Motivation

Centuries ago, navigators used compasses to traverse the vast oceans, and even today, compasses remain a vital component of modern navigation systems. Despite the widespread adoption of satellite technology and GNSS, which has significantly reduced the direct use of compasses, they continue to exist as part of an Inertial Navigation System (INS). However, as humans seek to reduce their reliance on GNSS and explore alternative, GNSS-independent methods for achieving the same level of position accuracy, a resurgence of interest in magnetic field-based navigation appears to be emerging. Accurate mapping of the Earth's magnetic field could pave the way for a revolutionary shift in navigation, where we combine the best of traditional sea navigation and modern technologies to navigate the skies more effectively using aeromagnetic navigation.

The importance of GNSS-independent navigation is broader than one might think. Especially with the rise of autonomous vehicles, like cars, robots and drones, it is important to have a backup navigation system. Besides that, GNSS is not always available, for example underground, inside buildings, in a long tunnel or in a conflict zone. The latter is a very important reason to have a backup navigation system [14]. In a conflict zone, GNSS can be jammed or spoofed, and therefore it is not reliable.

In addition to aiming for GNSS-independent navigation, there are many other applications for magnetic maps. Magnetic maps have been used for oil- and gas exploration [65], finding metals and minerals [44, 40], for archeological reasons [21], in geophysical studies [44, 54] and for the detection of ferromagnetic objects [63]. A better understanding of magnetic maps can significantly improve these applications.

The specific interest in magnetic maps arises from the fact that the Earth's magnetic field has strong contributions from the Earth's core and crust. A major advantage over other maps, such as height maps and topographic maps, is that the magnetic field is everywhere. It is almost not affected by weather conditions, and human activity has little effect on the magnetic field on large scales since the intensity of a magnetic source is inversely proportional to the cube of the distance from the source. Altogether, this makes the magnetic field a very reliable field for navigation, exploration and other applications.

1.2. Background of Magnetic Map-Making

Mapping the magnetic field is not a new idea and has been extensively studied over the years. One of the most commonly studied types of magnetic maps is the global magnetic map [34]. These maps are created by measuring the magnetic field that originates from the Earth's core and are widely available [18, 1]. To represent these maps, spherical harmonics are commonly used (Chapter 6). The coefficients of the spherical harmonic equations are often estimated by magnetic models that incorporate the changes in the magnetic field over time.

On a regional, more localized, scale, we encounter limitations of the spherical harmonic equations. These equations are designed to approximate the field as a multipole expansion stemming from the core. However, the magnetic field at the surface does not solely stem from the core but has contributions from the crust and near-surface objects as well. While, in theory, an infinite number of spherical harmonic terms is perfectly capable of representing the magnetic field, this is not feasible in practice.

The common approach for accurately representing the magnetic field on a regional scale appears to be through local extensions of global magnetic maps. The widely available magnetic models can be utilized to

find a mean value of the magnetic field in a specific region. This value can be used as a reference value for the magnetic field. A local coordinate system can be defined around the region of interest and the magnetic field can be measured in this coordinate system. The measured magnetic field can then be compared to the reference value and the difference can be used to create a local magnetic anomaly map. This way, a local enhancement of the magnetic field can be achieved, which allows for a more accurate, high-resolution, representation of the magnetic field.

There are several ways to extend the magnetic field to a local scale. The methods range from taking measurements and interpolating the magnetic field using splines, to more advanced techniques that incorporate the physical properties of the magnetic field by making dipole approximations. One of these techniques is Equivalent Layer. Equivalent Layer is a linear optimization problem that tries to find the best fit between the measured magnetic field and a pre-defined layer of magnetic dipoles that are placed under the surface. Equivalent Layer is proven to be an effective method for representing the magnetic field, but it is not without its limitations. The main limitation is that Equivalent Layer requires a priori information about where to place the dipoles.

Mapping techniques are mainly focused on making a two-dimensional map of the magnetic field. However, if one wants to navigate an aircraft, it is important to have information about the magnetic field in three dimensions. A cornerstone of magnetic maps for aeromagnetic navigation is Upward Continuation. Upward Continuation is a technique that uses the magnetic field at a certain height to estimate the magnetic field at a higher altitude. It does so by leveraging the harmonicity properties of the magnetic field and applying Green's third identity. Therefore, under many circumstances, it is sufficient to measure and map the magnetic field at a single height to be able to estimate the magnetic field at higher altitudes.

1.3. Contributions

This thesis aims to advance the knowledge of magnetic map-making. The main contributions of this thesis are:

1. Providing an overview of the whole pipeline of magnetic map-making. This includes the theory of magnetism (Chapter 2), the physical properties of the magnetic field (Chapter 3), the different techniques for mapping the magnetic field (Part II), and the applications of magnetic maps (Chapter 13). Literature on magnetic map-making is often focused on a single aspect of the pipeline.

2. A review of the different magnetic models that are available and their quality (Chapter 6). This includes understanding the differences between the models and their limitations, and how the spatial resolution relates to the applicability of the models. Together with the overview of the whole pipeline, this provides a better understanding of the applicability of the different magnetic models.

3. Filling in missing parts of the literature. In the description of the Equivalent Layer, the derivation of the linear system is often missing (Section 8.1). This thesis provides a derivation of the Equivalent Layer problem, starting from a single dipole with a single measurement. This derivation is then extended to multiple dipoles and multiple measurements. In the literature, no detailed derivation of the Equivalent Layer problem is found.

4. Presenting a new method for reconstructing a magnetic field from a set of measurements. This method is based on the Anderson functions and does not require a priori information about the location of the magnetic sources (Chapter 9). An extension of the Anderson functions to two dimensions is presented, which allows for the reconstruction of a magnetic field from a set of measurements in a plane. These two-dimensional Anderson functions are used as a starting point for an orthonormalized wavelet family that can be used to reconstruct a magnetic field from a set of measurements in a plane. It also provides a way to detect and make position estimates of the magnetic sources based on a measurement signal.

5. A benchmark of the different map-making techniques on simulated data (Chapter 11). A Python package `MagMap` is developed to facilitate this benchmarking and all other simulations in this thesis. The benchmarking is done on a simulated magnetic field which is designed to mimic a regional magnetic field. The benchmarking is done by comparing the reconstructed magnetic field to the original magnetic field after applying interpolation and extrapolation techniques. These types of comparisons are lacking in the literature. The `MagMap` provides a way to compare different map-making techniques on a simulated magnetic field and can easily be extended to include other map-making techniques.

6. Validation of the map-making techniques on real-world data (Chapter 12). We share our experiences with the application of map-making techniques on real-world data. We discuss the challenges we encountered and how we overcame them.

1.4. Research Questions and Objectives

This research is focused on exploring the techniques for magnetic map-making and how magnetic maps can be applied. The general research question we try to answer is: What are the fundamental principles and concepts underlying magnetic map-making? This question is divided into the following sub-questions:

1. What is the definition and purpose of a magnetic (anomaly) map?
2. What are the established methodologies and techniques utilized in the creation of magnetic maps?
3. How can the depth of dipole sources be accurately determined from the measured signal?
4. What is the empirical performance of current interpolation and extrapolation techniques for magnetic maps?
5. What is the impact of environmental and platform noise on measurement data in the process of map-making, and what strategies can be employed to minimize this influence?
6. In what ways can magnetic (anomaly) maps be utilized for navigation, localization, and exploration?

The first objective is to give a clear definition of a magnetic map. The second objective is to explore the current techniques for making these magnetic maps. Magnetic maps have been made for a long time and are based on measurements of the magnetic field. The construction of physically accurate maps is a challenging task, especially when the measurements are sparse. Different techniques are used, depending on the size of the mapped area. Enlisting the techniques summarizes a part of the literature study.

It followed from the literature review that many of the techniques use the same basic assumption. In the context of magnetic maps, but also for the estimation of magnetic fields in general, the assumption is that the magnetic field can be approximated by a sum of dipoles. The dipole sources are assumed to be located below the surface of the Earth. The accuracy of the map-making techniques seems to depend heavily on how the dipole sources are positioned. The third objective is to explore a method for determining the depth of the dipole sources from a measured signal.

Another objective of this research is to compare map-making techniques. Before a comparison can be made, assessment criteria must be established. The quality of a magnetic map is a complex issue. It is not only a matter of how well the map represents the magnetic field but also to what extent the map can be used for the intended purpose. A central question in this research is how to define measurable criteria for the quality of a magnetic map. Once the criteria are set, map-making techniques can be compared. An extensive comparison is made using numerical simulations to understand the performance and behavior of each method.

This research not only focuses on the theoretical aspects of magnetic map-making but also on the practical aspects. The fifth objective is to investigate the impact of environmental and platform noise on the quality of measurements. When attaching a magnetometer to a ferromagnetic platform, the magnetic field measured by the magnetometer is distorted. Especially in the context of an aircraft, this is something that needs to be compensated for.

Finally, the possible applications of magnetic maps are explored. Some experiments with navigation and localization methods on simulated magnetic maps are conducted to demonstrate the possible use case. The purpose of which is to convince the readers that magnetic maps are a good candidate to extend or replace the current navigation and localization methods that are based on GNSS. As a final step, a method for finding an optimal path in a magnetic map is developed.

1.5. Scope

Magnetic maps have a broad range of applications. They can be used for navigation, localization, detection and exploration, under almost any conditions and on any scale. Highly accurate magnetic measurements help trace the state of a molecule and can therefore be used for magnetic resonance imaging (MRI) [71], while at the same time, large-scale magnetic maps can be used for studying other planets or exploring outer space [45].

Limiting ourselves to just the use of magnetic maps for human navigation on Earth seems too broad an area even of research. For example, the differences between in-building navigation and aeromagnetic navigation on continental scales are too large. However, it is worth noting that navigation on a continental scale has been resolved using spherical harmonics.

In contrast, in-building navigation requires different techniques due to the type of sources involved. Especially the presence of human-made sources such as metal objects and electrical devices and human interactions that might affect the position of the sources makes this type of navigation more challenging. This area of research is also actively studied, with many techniques being developed to improve indoor navigation [32], such as the use of beacons to overpower the effects of human-made sources (Section 14.4).

However, there is currently a gap in research when it comes to utilizing the magnetic field for regional-scale outdoor navigation, such as aeromagnetic navigation on the scale of a city or a country. This gap exists because the limits of spherical harmonics have been reached on that scale, while in-building techniques seem less practical for outdoor use. Aeromagnetic navigation on a regional scale is an area that is explored in this research, aiming to develop methods that can bridge the gap and provide reliable map-making techniques.

It is worth noting that the techniques and methods developed and evaluated in this research can also be applied on smaller and larger scales. Therefore, while the scope of the research is focused on regional-scale navigation, the findings may still be valuable and interesting for various purposes in different contexts.

1.6. Outline

In this research, we present a full overview of the whole pipeline of magnetic map-making. The research starts with three introductory chapters. Chapter 2 provides an overview of the theory of magnetism and the physical properties of the magnetic field, Chapter 3 describes the structure of the Earth and what types of magnetic sources are present in the ground and Chapter 4 gives a brief introduction to magnetic sensing. The rest of the report is divided into four parts.

Part II describes the different techniques for mapping the magnetic field. This part starts with the definition of a magnetic map in Chapter 5. A clean signal is required to make a magnetic map from a set of measurements. Chapter 6 describes the global magnetic models and how spherical harmonics can be used for their representation. The process of Aeromagnetic Compensation, to clean up a signal from platform noise in aeromagnetic surveys, is described in Chapter 7. After that, Equivalent Layer is studied in Chapter 8 as a local extension of the global magnetic models. Chapter 9 contains a description of the Anderson Functions which are then used as a wavelet basis for the reconstruction of the magnetic field and dipole depth estimation. Finally, Chapter 10 discusses under what circumstances the magnetic field at a certain height can be used to estimate the magnetic field at a higher altitude.

Part III is dedicated to simulating and validating the different techniques for magnetic map-making that are described in Part II. Chapter 11 describes the benchmarking of the different techniques on simulated data. In Chapter 12, the techniques are validated on real-world data. Part IV contains several examples of how magnetic maps can be applied. Finally, Part V concludes the research and gives recommendations for future work.

I

Theory of Magnetism

2

An introduction into Magnetic Fields

2.1. Maxwell Equations

The Maxwell Equations, named after James Clerk Maxwell, are a set of partial differential equations that describe how electric charges and currents create electric and magnetic fields, and how these fields interact with each other. The four Maxwell equations [29, p. 16] are:

$$\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0} \tag{2.1}$$

$$\nabla \cdot \mathbf{B} = \mathbf{0} \tag{2.2}$$

$$\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t} \tag{2.3}$$

$$\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}, \tag{2.4}$$

where $\mathbf{E}$ [$\mathrm{V m^{-1}}$] is the electric field, $\mathbf{B}$ [T] is the magnetic induction field, ρ [$\mathrm{C m^{-3}}$] is the charge density, $\mathbf{J}$ [$\mathrm{A m^{-2}}$] is the current density, ϵ_0 [$\mathrm{F m^{-1}}$] is the permittivity of free space, $\mu_0 \approx 1.2566 \times 10^{-6}$ [$\mathrm{H m^{-1}}$] is the magnetic permeability of free space, and t [s] is time. The four equations describe the following laws:

1. **Gauss's law for electricity**: $\nabla \cdot \mathbf{E} = \frac{\rho}{\epsilon_0}$.
   The first Maxwell Equation (2.1) is also known as Gauss's law for electricity. It states that the divergence of the electric field is equal to the charge density divided by the permittivity of free space. The left-hand side of the equation is the divergence of the electric field, a positive quantity indicates that the electric field is spreading out, while a negative quantity indicates convergence. The right-hand side describes the density of charge at that point in space. The equation, therefore, relates the electric field to the charge density of the system. A large positive charge results in a strong diverging electric field, while a negative charge leads to convergence.

2. **Gauss's law for magnetism**: $\nabla \cdot \mathbf{B} = 0$.
   The second Maxwell Equation (2.2) is also known as Gauss's law for magnetism. It states that the divergence of the magnetic field is zero, which means that there are no sources nor sinks in the magnetic field. Instead, a divergence of zero explains why magnetic field lines are closed loops, as they cannot diverge or converge. Besides that, it shows that the magnetic flux is conserved, which means that the magnetic flux through a closed surface is constant. Furthermore, this equation refutes the existence of magnetic monopoles, as fields are due solely to electric currents and that magnetic "charges" do not exist. [38, p. 208]

3. **Faraday's law**: $\nabla \times \mathbf{E} = -\frac{\partial \mathbf{B}}{\partial t}$.
   Faraday's law, third Maxwell Equation (2.3), implies that a time-varying magnetic field induces an electric field. There are many practical applications of this law since electrical induction is the underlying concept of generators, transformers and motors [29, p. 326].

4. **Ampère's law**: $\nabla \times \mathbf{B} = \mu_0 \mathbf{J} + \mu_0 \epsilon_0 \frac{\partial \mathbf{E}}{\partial t}$.
   The last Maxwell Equation (2.4), generally known as Ampère's law, relates the curl of the magnetic field to the currents and the rate of change in the electric field. The curl on the left-hand side of the equation indicates the circulation of the magnetic field around a closed surface. The right-hand side has two terms, the first term indicates the currents passing through the surface, while the second term describes the rate of change of the electric field. The equation, therefore, shows that a current **J** passing through a surface induces a magnetic induction field **B** and that a time-varying electric field **E** induces a magnetic induction field **B**.

Gauss's law for magnetism and Ampère's law are the two most important equations when it comes to the magnetic field. In the following sections, we discuss the field **B** in more detail.

2.1.1. Magnetic Induction Field B

The magnetic **B**-field is a vector field that measures the strength of the magnetic induction field at a particular point in space, expressed in teslas (T). This field describes the forces acting on a moving charge, and it typically has two components: one arising from the magnetic field **H**, and another arising from the magnetization field **M** [29, p. 40]. The relation between the field **B** and the fields **H** and **M** is given by:

$$\mathbf{B} = \mu_0 (\mathbf{H} + \mathbf{M}) , \tag{2.5}$$

where μ_0 [$\mathrm{H\,m^{-1}}$] is the permeability of free space. This expression is always true, regardless of the context, and relates the vector fields **B**, **H**, and **M** to each other. This relationship is necessary for solving the Maxwell equations, which are fundamental to the understanding of electromagnetism. The fields **H** and **M** are explained in the following sections.

2.1.2. Magnetic Field H

The magnetic **H**-field is a measure of magnetic field strength, and is measured in amperes per meter ($\mathrm{A\,m^{-1}}$). The field can be interpreted as an intensity field that describes the magnetic flux density per unit area, i.e. the density of magnetic field lines passing through a given area. This field is generated by the electrical currents outside the material, either by a solenoid, an electromagnet, or from a permanent magnet.

2.1.3. Magnetization Field M

There exists a relationship between the properties of a material and the magnetic field it experiences. The field **M**, the magnetization vector, expresses the magnetic moment per unit volume in amperes per meter ($\mathrm{A\,m^{-1}}$) [38, p. 40]. This relationship is given by the following equation:

$$\mathbf{M} = \frac{\mathbf{m}}{V} , \tag{2.6}$$

where **m** is the magnetic moment and V is the volume of the material. It describes the contribution to the induction from the magnetization of a material. The movement of electrons in a material can cause a magnetic moment. For example, a strongly magnetizable material, such as iron, becomes magnetized when it is placed in a magnetic field and produces a magnetic field on its own. The direction and magnitude of the magnetization vector are determined by the direction and magnitude of the applied magnetic field and the structure of the material. The sum of all the magnetic moments is the magnetization of the material. When a material is in a steady state, there are no free currents and the magnetization is zero. This is often the case outside of a ferromagnetic material, e.g. in air, water or vacuum.

Magnetic Moment

The magnetic moment is the underlying quantity that describes how a material responds to a magnetic field. The magnetic moment is expressed in amperes times meters squared ($\mathrm{A\,m^2}$). The general definition of the magnetic moment is given by an integral of the magnetization **M** over the volume Ω of the material:

$$\mathbf{m} = \int_{\Omega} \mathbf{M}\, dV . \tag{2.7}$$

In the case of a bar-magnet, the magnetic moment is given by:

$$\mathbf{m} = p l , \tag{2.8}$$

where p is the pole strength [A m] and l is the length of the magnet [m].

2.1.4. The current density J

Another variable in the Maxwell equations is the current density **J**. The current density **J** is the density of the current passing through a given area. This quantity consists of several contributions. The first contribution is the current density $\mathbf{J}_f$ due to the motion of free charges. Several examples [38, p. 454] of these types of currents are *conduction currents* and *electrolytic currents* and *motion of macroscopic charged bodides*. The second contribution, $\partial\mathbf{D}/\partial t$, is the current density due to displacements. This is a time-varying component that depends on the *polarization current* density $\partial\mathbf{P}/\partial t$ and the rate of change of the electric displacement field $\epsilon_0\partial\mathbf{E}/\partial t$. The third contribution $\nabla\times\mathbf{M}$ is the current density due to magnetization. The magnetization current and the polarization current form together the bounded current density $\mathbf{J}_b = \nabla\times\mathbf{M} + \partial\mathbf{P}/\partial t$. The total current is therefore given by:

$$\mathbf{J}_t = \mathbf{J}_f + \frac{\partial\mathbf{P}}{\partial t} + \epsilon_0\frac{\partial\mathbf{E}}{\partial t} + \nabla\times\mathbf{M}, \tag{2.9}$$

which is often separated into $\epsilon_0\partial\mathbf{E}/\partial t$ and **J**, where **J** consists of the free current density $\mathbf{J}_f$ and the bounded current density $\mathbf{J}_b$:

$$\mathbf{J} = \mathbf{J}_f + \frac{\partial\mathbf{P}}{\partial t} + \nabla\times\mathbf{M} = \mathbf{J}_f + \mathbf{J}_b\,. \tag{2.10}$$

2.2. Magnetostatics

The Maxwell equations are a general description of the electromagnetic field. However, in many cases, the Maxwell equations can be simplified. One such simplification is the magnetostatics approximation. It assumes that there are no time dependencies in the magnetic field and that the electric charges are stationary or moving very slowly. The magnetostatic model is a good approximation when the time-varying components of the electromagnetic field are negligible when compared to the time-invariant components. The so-called quasi-magnetostatics Maxwell Equations can be derived from the Maxwell Equations (2.1-2.4) by setting the time derivatives to zero. This breaks the equations down into two sets of equations, two equations for the electric field and two equations for the magnetic induction field. The magnetic induction field equations are [38, p. 455]:

$$\nabla\cdot\mathbf{B} = 0 \tag{2.11}$$

$$\nabla\times\mathbf{B} = \mu_0\mathbf{J} = \mu_0(\mathbf{J}_f + \nabla\times\mathbf{M})\,, \tag{2.12}$$

where **M** is the magnetization of the material and the last equation holds since $\partial\mathbf{P}/\partial t = \mathbf{0}$ in magnetostatics.

2.2.1. Magnetic vector Potential A

The magnetostatic equations (2.11-2.12) allow the magnetic induction field to be expressed in terms of a magnetic vector potential **A** [Tm] since every divergence-free vector field is the curl of some other vector field:

$$\nabla\cdot\mathbf{B} \implies \mathbf{B} = \nabla\times\mathbf{A}. \tag{2.13}$$

Until now, **A** is only restricted to be a curl-free vector field. It is possible to impose additional constraints on **A**, which lesad to a unique solution, this is called gauge fixing. A common gauge fixing condition is the Coulomb gauge, which requires that **A** is solenoidal: $\nabla\cdot\mathbf{A} = 0$. Under this assumption, the magnetic induction field can be expressed as:

$$\nabla\times\mathbf{B} = \nabla\times(\nabla\times\mathbf{A}) = \nabla(\nabla\cdot\mathbf{A}) - \nabla^2\mathbf{A} = -\nabla^2\mathbf{A} = \mu_0\mathbf{J}\,. \tag{2.14}$$

Therefore, in the absence of the magnetization term ($\mathbf{M} = \mathbf{0}$), the magnetic vector potential can be defined by the following Poisson equation:

$$\nabla^2\mathbf{A} = -\mu_0\mathbf{J}\,. \tag{2.15}$$

The solution [28, p. 180] of this system is given by:

$$\mathbf{A}(\mathbf{r}) = \frac{\mu_0}{4\pi}\int_\Omega \frac{\mathbf{J}(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}', \tag{2.16}$$

where $\mathbf{r}'$ is a point in space where the current density source is located and $\mathbf{r}$ is the point at which the magnetic vector potential is evaluated.

2.2.2. Magnetic Scalar Potential ϕ

For magnetostatic conditions, the magnetic field can be expressed in terms of a magnetic scalar potential ϕ [A]. Combining Equation 2.5 and Equation 2.12 yields:

$$\nabla\times\mathbf{H} = \nabla\times\left(\frac{1}{\mu_0}\mathbf{B}-\mathbf{M}\right) = \mathbf{M}+\mathbf{J}_f-\mathbf{M} = \mathbf{J}_f. \tag{2.17}$$

In a current free region, $\mathbf{J}_f = \mathbf{0}$, so the curl of the magnetic field strength is zero:

$$\nabla\times\mathbf{H} = \mathbf{0}. \tag{2.18}$$

Therefore, the magnetic field strength can be expressed as the gradient of a scalar potential ϕ:

$$\mathbf{H} = -\nabla\phi. \tag{2.19}$$

In the absence of magnetization ($\mathbf{M} = \mathbf{0}$), this magnetic scalar potential satisfies the Laplace equation because the magnetic field strength is divergence-free:

$$\nabla^2\phi = -\nabla\cdot\mathbf{H} = 0. \tag{2.20}$$

Therefore, the following equation holds:

$$\mathbf{B} = \mu_0\mathbf{H} = \mu_0(-\nabla\phi). \tag{2.21}$$

Note that the operator ∇ is the gradient operator, which is distributive under scalar multiplication. This means that there also exists a potential U that already incorporates the permeability of the material. This potential is given by:

$$\mathbf{B} = \mu_0(-\nabla\phi) = \nabla(-\mu\phi) = -\nabla U, \tag{2.22}$$

where the $U = \mu\phi$ represents the energy of the magnetic induction field $\mathbf{B}$.

2.3. The relationship between ϕ and M

The magnetic scalar potential can thus be used to express the strength of a current-free magnetic field. Outside of a material, where no magnetization is present, there exists a direct relationship (Equation 2.22) between the magnetic field and the magnetic scalar potential. However, inside a material, the magnetization term $\mathbf{M}$ is present, which complicates the relationship between the magnetic field and the magnetic scalar potential. Combining the definition of the magnetic induction field $\mathbf{B}$ (Equation 2.5) and Gauss's law of magnetism (Equation 2.2) gives the following relationship between the field $\mathbf{H}$ and $\mathbf{M}$:

$$\nabla\cdot\mathbf{H} = -\nabla\cdot\mathbf{M}. \tag{2.23}$$

A relationship between the magnetic scalar potential ϕ and the magnetization $\mathbf{M}$ can now be derived by substituting Equation 2.19 into Equation 2.23:

$$\Delta\phi = -\nabla\cdot\mathbf{M}. \tag{2.24}$$

This is a special case of the Poisson equation:

$$\Delta\phi = f, \tag{2.25}$$

with the source term $f = -\nabla\cdot\mathbf{M}$. The solution of the 3D Poisson equation can be found using Green's functions [28, pp. 125–127]. The Green's function for the three-dimensional Laplace operator is given by:

$$G(\mathbf{r},\mathbf{r}') = \frac{-1}{4\pi|\mathbf{r}-\mathbf{r}'|}. \tag{2.26}$$

Therefore, the solution of Equation 2.25 is given by:

$$\phi(\mathbf{r}) = \int_\Omega -G(\mathbf{r},\mathbf{r}')f(\mathbf{r}')d^3\mathbf{r}' = \frac{1}{4\pi}\int_\Omega \frac{f(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'. \tag{2.27}$$

Now, we consider the special case of the Poisson equation (Equation 2.24) with a source in the form of the divergence of a vector field. According to Equation 2.27, the solution of this equation is given by the convolution of the source term with the Green's function:

$$\phi(\mathbf{r}) = \int_\Omega G(\mathbf{r},\mathbf{r}')\nabla\cdot\mathbf{M}(\mathbf{r}')d^3\mathbf{r}' = -\int_\Omega \frac{\nabla\cdot\mathbf{M}(\mathbf{r}')}{4\pi|\mathbf{r}-\mathbf{r}'|}d^3\mathbf{r}'. \tag{2.28}$$

This shows that the magnetic scalar potential ϕ can be expressed as an integral of the magnetization $\mathbf{M}$ over the volume of the material, scaled by the distance from the point of interest to the magnetization source.

2.4. Harmonicity

Harmonicity is a property of a field that describes the field's regularity and smoothness. Harmonic vector fields have no abrupt change in direction or magnitude over a given region. A harmonic field is also continuous and does not have any singularities.

Definition 2.4.1 (Harmonic Vector Field) *A vector field* $\mathbf{F}$ *is said to be harmonic if it satisfies the Laplace equation:*

$$\nabla^2\mathbf{F} = 0. \tag{2.29}$$

A magnetic field is harmonic at every point except at the poles of the magnet.

2.4.1. Green's Third Identity

Special properties hold for harmonic functions when it comes to integration over a closed surface. Green's third identity defines a relationship between the boundary of a volume and the volume itself [30, p. 361]. It follows from Green's second identity:

$$\int_U (\psi\nabla^2\phi - \phi\nabla^2\psi)dV = \oint_{\partial U}(\psi\nabla\phi - \phi\nabla\psi)\cdot d\mathbf{S}. \tag{2.30}$$

The third identity can be derived by substituting $\phi = G$, where G is the Green's function for the Laplace operator (Equation 2.26). For any twice continuously differentiable function ψ on U it follows that:

$$\int_U [G(\mathbf{y},\eta)\Delta\psi\acute{C}(\mathbf{y})]dV_\mathbf{y} - \psi(\eta) = \oint_{\partial U}\left[G(\mathbf{y},\eta)\frac{\partial\psi(\mathbf{y})}{\partial n_\mathbf{y}} - \psi(\eta)\frac{\partial G(\mathbf{y},\eta)}{\partial n_\mathbf{y}}\right]dS_\mathbf{y}. \tag{2.31}$$

Specifically for harmonic functions ψ it follows that:

$$\psi(\eta) = \oint_{\partial U}\left[\psi(\mathbf{y})\frac{\partial G(\mathbf{y},\eta)}{\partial n_\mathbf{y}} - G(\mathbf{y},\eta)\frac{\partial\psi(\mathbf{y})}{\partial n_\mathbf{y}}\right]dS_\mathbf{y}. \tag{2.32}$$

A similar expression is obtained in [5, p. 314] to express the potential as a surface integral over the boundary of the volume. Green's third identity theorem has some important consequences. For example, if $\mathbf{F}$ is a harmonic vector field, and $\mathbf{F}$ is known on the boundary of a closed surface, then $\mathbf{F}$ is also known inside the surface. For the magnetic field, this means that if the magnetic field is known on the boundary of a closed surface, and the surface does not contain any poles, then the magnetic field inside the surface is completely defined by the boundary conditions. This property is the underlying principle of Upward Continuation (Chapter 10).

2.4.2. Harmonicity of the vector Magnetic Induction Field B

A commonly used property of the potential U is that it is harmonic. This allows us to use for example Green's third identity, which is the underlying principle of Upward Continuation. If we can show that the magnetic induction field $\mathbf{B}$ is also harmonic, then we can use the same principles to upward continue the magnetic induction field.

The magnetic induction field is related to the gradient of the magnetic scalar potential ϕ (Equation 2.22). If the magnetic scalar potential is harmonic, then its gradient is also harmonic. However, the magnetic induction field is not just the gradient of the magnetic scalar potential; it also contains the magnetization field $\mathbf{M}$. The magnetic induction field $\mathbf{B}$ is harmonic if the following equation is true:

$$\nabla^2 \mathbf{B} = 0\,, \tag{2.33}$$

from Equation 2.5 it follows that

$$\nabla^2 \mathbf{B} = \mu_0 \nabla^2 \mathbf{H} + \mu_0 \nabla^2 \mathbf{M} = 0\,. \tag{2.34}$$

Here, $\nabla^2 \mathbf{H} = 0$ is always true, since the $\mathbf{H}$-field is defined as the gradient of the (harmonic) magnetic scalar potential ϕ. If the magnetization field is not harmonic, then the magnetic field is not harmonic either. Luckily, the magnetization field can be assumed to be harmonic in most cases. Especially when we are further away from the source. This is because the magnetization field is caused by the magnetic moments of the atoms in the material.

2.5. Multipole Expansion of the Magnetic Scalar Potential

A magnetic scalar potential can be expressed in terms of a multipole expansion [11]. Multipoles are a series of terms that describe the magnetic field in terms of a monopole, a dipole, a quadrupole, and multipoles of higher order. The expansion is a series expansion, which approximates the magnetic field better as more terms are added to the series.

This expansion is a series expansion, similar to the Taylor series. The expansion thus starts with a monopole, followed by a dipole, a quadrupole, and multipoles of higher order. Physically, there does not exist a magnetic monopole, but theoretically, it is possible to express the magnetic field in terms of a monopole. The multipole expansion of the magnetic scalar potential ϕ is given by:

$$\phi(\theta, \phi') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} C_l^m Y_l^m(\theta, \phi')\,, \tag{2.35}$$

where C_l^m are the coefficients of the expansion and Y_l^m are the spherical harmonics, discussed in Appendix A. The potential energy of a magnetic pole decreases with the distance from the pole. Depending on the order of the multipole, the potential energy of the pole decreases with a different power of the distance. The potential energy of a monopole is inversely proportional to the distance $(1/r)$ and the potential energy of a dipole is inversely proportional to the distance squared $(1/r^2)$. The potential energy of a multipole of order l is inversely proportional to the distance to the power of l.

The magnetic field strength can be derived from the magnetic scalar potential by taking the gradient of the potential. The strength of a dipole is therefore proportional to $(1/r^3)$, and for multipoles of order l it holds that the field strength is proportional to $(1/r^{l+1})$.

2.5.1. Dipole Approximation

The dipole approximation is a simplification of the multipole expansion. The energy of a multipole is proportional to the distance and drops off faster for higher-order multipoles. If the distance to a magnetic source is sufficiently large, then the energy is dominated by the (monopole and) dipole terms. The magnetic induction field $\mathbf{B}$ of a dipole is given by

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{m} \cdot (\mathbf{r} - \mathbf{r_0}))(\mathbf{r} - \mathbf{r_0})}{|\mathbf{r} - \mathbf{r_0}|^5} - \frac{\mathbf{m}}{|\mathbf{r} - \mathbf{r_0}|^3} \right), \tag{2.36}$$

where $\mathbf{m}$ is the dipole moment, $\mathbf{r}$ is the position vector and $\mathbf{r_0}$ is the position of the dipole. A dipole field can also be written in terms of the magnetic scalar potential ϕ:

$$\phi(\mathbf{r}) = \frac{\mathbf{m} \cdot \mathbf{r}}{4\pi |\mathbf{r}|^3}\,, \tag{2.37}$$

where $\mathbf{B}(\mathbf{r}) = -\mu_0 \nabla \phi(\mathbf{r})$.

3

Structure of the Earth's Magnetic Field

3.1. Geomagnetic Fields

Earth's magnetic field is a complex system, and it is not possible to model it with a single field. The Earth magnetic field is the superposition of field contributions from internal and external sources. The internal sources are the main (core), crustal (lithosphere) and induced (mantle) fields. The external sources are located in the ionosphere and magnetosphere. The internal and external fields are superimposed onto each other, interact by induction and form together the Earth's magnetic field. The internal fields are the dominant contribution to the total Earth's magnetic field. A schematic of Earth's structure is shown in Figure 3.1. This chapter discusses the internal and external fields in more detail: the source that causes the field is identified, depth profiles are given and the fields are discussed in terms of their spatial and temporal variations.

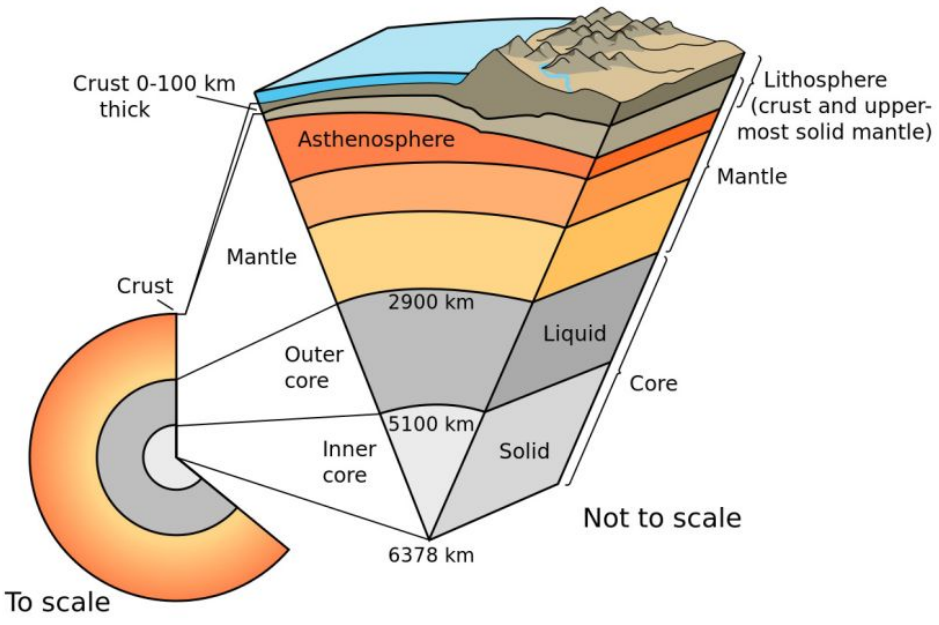


Figure 3.1: Cutaway view of the Earth's magnetic field. [19]

3.2. Main Earth Field

The main field is that part of the geomagnetic field that is generated in the outer core. It is the dominant contribution to the magnetic field at the surface and contributes approximately 95-99% to the Earth's magnetic field strength at the surface [26]. The main field is generated by the motion of electrically conducting fluids in the outer core. It is, by approximation, a dipole field with the dipole axis pointing (almost) toward the Earth's rotational axis. The intensity of the field is in line with this orientation and varies at the surface of the Earth from 30-40 μT at the equator to 60-70 μT at the poles. The field undergoes secular variation, which is the slow change in time of the main magnetic field on a time scale of at least a year. The liquid outer core is positioned between the solid inner core and the mantle, at a depth ranging from 2900 to 5100 km. This depth results in large spatial wavelengths of the main field, which are on the order of 10^3-10^4 km.

3.3. Induced Field

The induced field is that part of the geomagnetic field that is generated in the mantle. The currents in the conducting mantle are induced by external sources. Induced magnetization occurs when magnetic minerals align themselves with the surrounding magnetic field. This alignment happens spontaneously and is proportional to the strength of the ambient magnetic field, and it also varies across different locations [66]. These fields are difficult to model since they are dependent on time-varying external sources.

3.4. Crustal Field

The crustal field is that part of the geomagnetic field that is generated in the crust and upper mantle (lithosphere). Around 1% up to 5% of the total Earth's magnetic field strength at the surface is contributed by the crustal field [35], which has at the surface a magnitude in the order of 10^2 nT. The field is caused by the remanent magnetization of rocks. Remanent magnetization is the magnetic field that is left in a rock after the external field that caused it has been removed. Some rocks in the lithosphere, such as magnetite, are naturally ferromagnetic and may exhibit remanent magnetization. Other rocks exhibit thermoremanent magnetization, which is the magnetization that is left in a rock after it has been heated. The depth of the crustal field ranges from 0-100 km, resulting in wavelengths on the order of 10^2-10^3 km at Earth's surface. Variations are caused by changes in the lithosphere, such as the movement of tectonic plates. These changes occur on geological time scales. This, in combination with the relatively short wavelengths, makes the crustal field a good candidate for magnetic navigation.

3.4.1. Secular Variation by Plate Tectonics

The crustal field is also the field that is used by paleomagnetists to study the history of Earth's magnetic field. They use the field to determine the orientation of the Earth's magnetic field at different times in the past. A remanent magnetized rock records the orientation of the Earth's magnetic field at the time when it was formed, so measuring the orientation gives information about the magnetic field at that time. On short time scales, the orientation of the Earth's magnetic field does not change much.

On large time scales, continental drift and plate tectonics cause the crust to move. This movement causes slight changes in the crustal field. On even larger time scales, reversals in the polarity of Earth's magnetic field occur. These reversals occur on the order of 10^6-10^7 years. The last reversal occurred 780-800 thousand years ago. Reversals in polarity can be recorded using a magnetometer at places where plates are diverging. Anomalies are observed when measuring in the direction the plates move in. In Figure 3.2 the total-field magnetic anomalies for some ridges in the southwest of Vancouver Island are shown. Clear stripes are observed, which correspond with the moments at which the polarity of the Earth's magnetic field reversed.

3.5. External Fields

The external fields are that part of the geomagnetic field that is generated in the ionosphere and magnetosphere. The ionosphere is the region of the Earth's atmosphere that is ionized by solar radiation. The magnetosphere is the region of the Earth's magnetic field that is generated by the solar wind. These anomalies have a short duration ranging from several minutes to several days. Polar and equatorial regions are most affected by these fields.

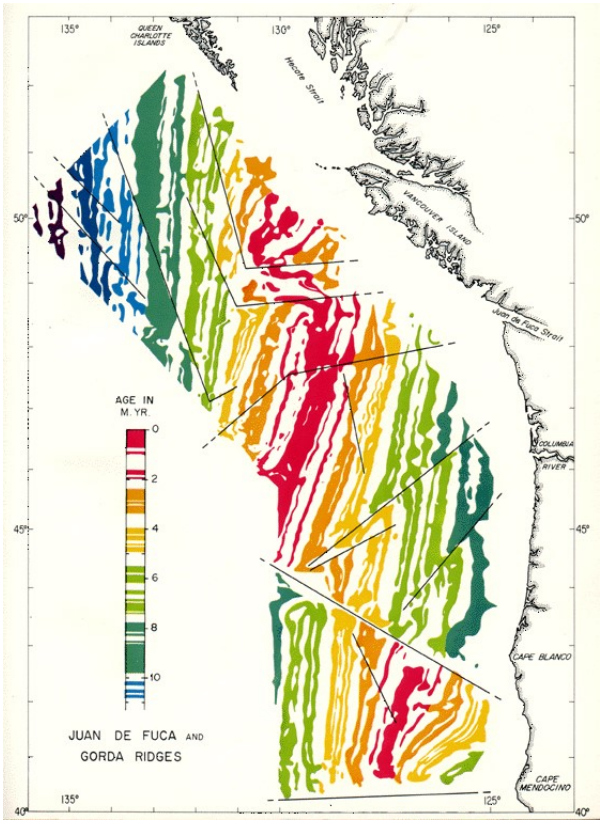


Figure 3.2: Total-field magnetic anomalies in the southwest of Vancouver Island. The colored areas indicate positive anomalies. The alternation between positive and negative anomalies corresponds with the reversal time scale.

4

Magnetic Sensing

The properties of magnetic fields are not always known to us, but what we can measure is the magnetic field itself. Magnetometers are used for this purpose, and they come in two categories: total field sensors and vector magnetometers. Measuring the strength, direction, and inclination of the magnetic field are some of the most crucial properties that can be determined. This chapter covers the various measurable characteristics of magnetic fields and the different types of magnetometers. It is important to distinguish between two types of magnetometers: total field sensors and vector magnetometers, as they measure different properties of the magnetic field. Map-making techniques might differ depending on the type of measurements that are taken.

4.1. Reference Frames

The Earth is often seen as a sphere, but this is not the case. The Earth is an oblate spheroid, which means that the Earth is flattened at the poles, this is caused by the rotation of the Earth. When working with magnetic fields on Earth it is thus important to know which reference frame is used. A reference frame is a coordinate system that is used to describe the position of an object. Two reference frames are used for magnetic fields on Earth: the Geodetic reference frame and the Geocentric reference frame.

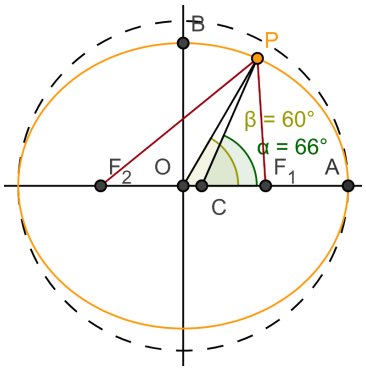


Figure 4.1: Geocentric and Geodetic reference frames in black and yellow respectively. This figure is taken from [62].

The Geocentric reference frame is a reference frame that is centered at the center of the Earth. The reference frame assumes that the Earth is a perfect sphere, the coordinate system is therefore identical to spherical coordinates with axis aligned with the Earth's axis. The Geodetic reference frame is a more accurate representation of the Earth. The Geodetic reference frame is therefore more commonly used in scientific research. A commonly used Geodetic reference frame is the World Geodetic System 1984 (WGS84).

Models of magnetic fields on Earth (Chapter 6) are often described with functions defined in a spherical coordinate system, e.g. Spherical Harmonics (Section 6.5). This requires a transformation from the Geocentric reference frame to the Geodetic reference frame. Geodetic coordinates (λ, ϕ, h) can be transformed to Geocentric coordinates (λ, ϕ', r) using the following equations [18]:

$$\begin{aligned} p &= (R_c + h)\cos\phi \\ z &= (R_c(1-e^2) + h)\sin\phi \\ r &= \sqrt{p^2 + z^2} \\ \phi' &= \arctan\frac{z}{p} \end{aligned} \quad , \tag{4.1}$$

where $p = \sqrt{(x^2 + y^2)}$, where x, y and z are the coordinates of a geocentric Cartesian coordinate system. The coefficients R_c and e for the WGS84 reference frame are given by:

$$\begin{aligned} A &= 6378137\text{m} \\ \frac{1}{f} &= 298.2572 \\ e^2 &= 2f - f^2 \\ R_c &= \frac{A}{\sqrt{1 - e^2\sin^2\phi}} \end{aligned} \quad . \tag{4.2}$$

While both reference frames serve their purpose, choosing the appropriate one for a given task is crucial. The Geocentric reference frame, considering the Earth as a perfect sphere, is generally simpler to work with and is often used in preliminary calculations or when precision is not the primary concern. However, for more precise measurements and particularly when dealing with surface-based applications, the Geodetic reference frame, which considers the Earth's actual oblate spheroid shape, is preferred. This is particularly true for GNSS systems, and surveying larger areas where the actual shape of the Earth is more relevant.

4.2. Measurable Components

4.2.1. Intensity

The intensity of a magnetic induction field **B** is defined as the magnitude of the magnetic field vector. This is given by the following equation:

$$I = |\mathbf{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2}. \tag{4.3}$$

4.2.2. Direction

The direction of a magnetic induction field **B** is given by the unit vector **u**, which points in the direction of the magnetic field. This unit vector is given by the following equation:

$$\mathbf{u} = \frac{\mathbf{B}}{|\mathbf{B}|}. \tag{4.4}$$

The combination of the intensity and the direction defines the magnetic field vector **B**.

4.2.3. Declination

The declination of a magnetic induction field **B** is the angle on the horizontal plane between the magnetic field vector **B** and the north direction [16]. On Earth, there is a slight difference between the magnetic north and the true north. Especially around the poles, the difference between the two is significant, this can be seen in figure 4.2. The declination is important for navigation because for example, a compass points to the magnetic north and not to the true north. In a North-East-Down (NED) coordinate system, the declination is the clockwise angle between the North direction $\mathbf{B_x}$ and the East direction $\mathbf{B_y}$. This is given by the following equation:

$$\delta = \arctan\left(\frac{B_y}{B_x}\right). \tag{4.5}$$

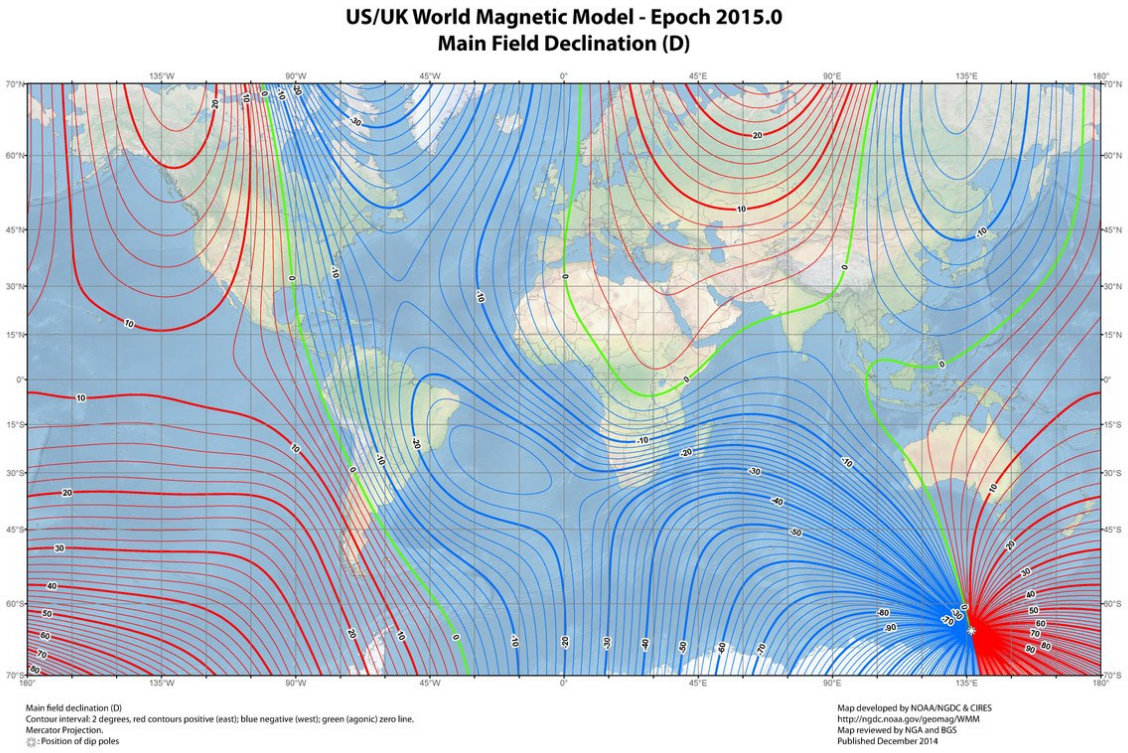


Figure 4.2: Declination of a magnetic induction field **B** based on the World Magnetic Model (WMM) for the year 2015. This figure is taken from [47].

4.2.4. Inclination

The inclination of a magnetic induction field **B**, also known as dip angle, is the angle which the magnetic field lines have with the horizontal plane. On Earth, the inclination is the angle at which the field lines go into the ground. In a North-East-Down (NED) coordinate system, the inclination is the angle between the vertical direction $\mathbf{B_z}$ and the horizontal plane[16]. This is given by the following equation:

$$\iota = \arctan\left(\frac{B_z}{\sqrt{B_x^2 + B_y^2}}\right). \tag{4.6}$$

4.3. Magnetometers

Magnetometers are sensors that measure the magnetic field. Magnetometers can be divided into two categories: scalar and vector magnetometers. Scalar magnetometers measure the intensity of the magnetic field, while vector magnetometers measure the direction and the intensity of the magnetic field. With the development of micro- and nanotechnology, magnetometers have become smaller and more accurate [43]. In this section, we give some examples of scalar and vector magnetometers.

4.3.1. Scalar Magnetometers

A scalar magnetometer measures the intensity of the magnetic field in tesla (T). They are known to be very accurate and reliable, but this comes at the cost of losing information about the direction of the magnetic field. The current scalar magnetometers have come so far that their accuracy in some cases depends on scientific constants. For example, Proton Magnetometers, which are based on the principle that protons in a magnetic field absorb energy and emit radiation at a specific frequency[27], are so sensitive because their precision is only based on scientific constants and the applied magnetic field [25]. Another popular and very sensitive type of magnetometer is the one that is based on alkali vapor cells.

Proton Magnetometers
A proton magnetometer employs the precession of atomic protons in a magnetic field to determine the field's strength (and orientation) [50]. When atoms are placed in a magnetic field, their protons start moving in a circular motion around the field's direction, which is called precession. The rate of this precession, so the frequency of the circular motion, is proportional to the strength of the magnetic field. Therefore, measuring the frequency at which the atoms move around gives a measurement of the magnetic field, this can be done by using a RF-pulse.

Vapor Magnetometers
Vapor magnetometers, or optically pumped magnetometers, are a very sensitive type of magnetometer that is based on the principle that alkali metals can exist in different levels of energy [7]. When a laser is shined on the alkali metal, the energy of the laser is absorbed by the metal and the metal excites to a higher energy level. The atoms in higher energy levels do not absorb the laser anymore, so the cell becomes transparent to the laser when all atoms are in the higher energy level, now the cell is said to be optically pumped. Applying a magnetic field to an optically pumped cell causes the atoms to move to a different energy level, which in its turn causes the cell to become less opaque. The change in opacity can be measured with a photodetector. The rate at which the cell becomes less opaque is proportional to the strength of the magnetic field, making a vapor magnetometer a scalar magnetometer.

4.3.2. Vector Magnetometers
A vector magnetometer, or tri-axial magnetometer, measures the flux density value in a specific direction. The intensity as well as the direction of the magnetic field are measured since the flux density is a vector. The resulting measurement can be expressed in a vector ($\mathbf{B} = (B_x, B_y, B_z)$) with three components of the magnetic field. The accuracy of a vector magnetometer is generally lower than that of a scalar magnetometer. This is because a vector magnetometer is more susceptible to errors, for example, perfect alignment of the axis within the sensor and the orientation of the sensor to the magnetic field is required.

It is not difficult to show that there cannot exist a vector magnetometer that is more accurate than any magnetometer that measures the total-field intensity. If such an instrument existed, it would also be able to measure the intensity of the magnetic field with the same, or better, accuracy. This is because the total-field intensity is the square root of the sum of the squares of the three components of the magnetic field (Equation 4.7). Therefore, a good scalar magnetometer can also be used to measure the total-field intensity by neglecting information about the direction of the magnetic field.

$$|\mathbf{B}| = \sqrt{B_x^2 + B_y^2 + B_z^2}. \tag{4.7}$$

Fluxgate Magnetometers
A fluxgate magnetometer is a commonly used type of vector magnetometer. It consists of two coils that are wrapped around a small magnetic core [51]. It utilizes the principle that a magnetic field induces a voltage in a coil. Faraday's law (Equation 2.3), one of the Maxwell equations, states that the produced voltage is proportional to the rate of change of the magnetic flux through the coil.

4.3.3. Sensitivity
The sensitivity of a magnetometer is the smallest change in the magnetic field that can be measured by the magnetometer. A common measure of the sensitivity of a magnetometer can be derived from the Amplitude Spectral Density (ASD) of the magnetometer. The ASD of a signal is the square root of the Power Spectral Density (PSD) (Equation 4.9) and indicates the amplitude of the signal at a specific frequency.

Definition 4.3.1 (Power Spectral Density) *Let $x(t)$ be a time-limited signal with Fourier transform $X(\omega;\eta,T)$, then the Power Spectral Density (PSD) [56] of the signal is defined as*

$$PSD = S_{xx}(\omega) = \lim_{T\to\infty} \frac{1}{T} S_{xx}(\omega;T) \tag{4.8}$$

which represents the power of the signal at a certain frequency ω.

The Amplitude Spectral Density (ASD) of a signal is the square root of the Power Spectral Density (PSD) of the signal:

$$ASD = \sqrt{PSD} = \sqrt{S_{xx}(\omega)} \tag{4.9}$$

which represents the amplitude of the signal at a certain frequency ω. The units of the ASD are the units of the signal per square root Hertz, so for a magnetometer, the unit is $\mathrm{T}/\sqrt{\mathrm{Hz}}$.

We were unable to find literature about the procedure for determining the sensitivity of a signal, but presumably, it is done by measuring the ASD with multiple sensors of the same type at the same location:

1. Place two sensors of the same type next to each other, e.g. at a distance of 1 meter.
2. Measure the magnetic field for a long period at both sensors, e.g. 1 hour.
3. Calculate the ASDs of both signals $x_1(t)$ and $x_2(t)$ by transforming the signals to the frequency domain and taking absolute values of the Fourier coefficients.
4. Subtract the ASDs of both signals from each other, this gives the error in the ASD for this type of sensor.

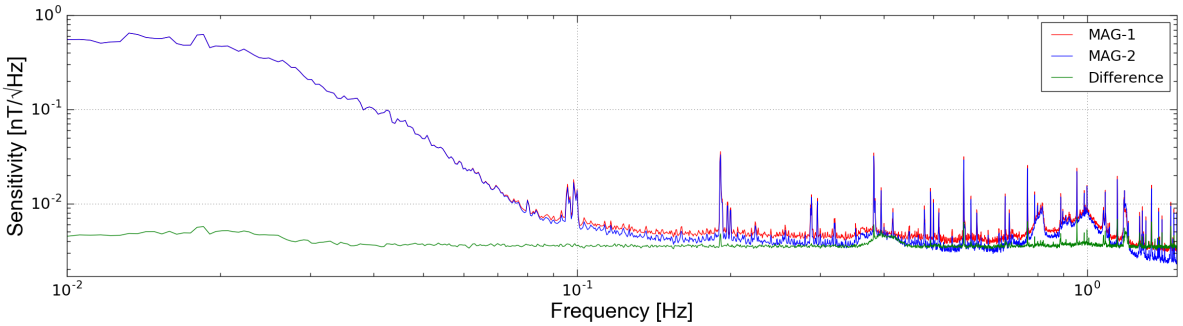


Figure 4.3: Two QTFMs separated by 1m: The subtracted signal gives the combined noise of both magnetometers plus any noise in the field gradient. This figure is taken from [53].

The sensitivity is often expressed by a certain frequency band, for example, the maximum error in the ASD for a frequency between 0.1 and 100 Hz. An example of an ASD for a QuSpin [53] Total-Field Magnetometer (QTFM) is illustrated in figure 4.3. This sensor has a field sensitivity of < 1 pT$/\sqrt{\mathrm{Hz}}$ in the $0.1 - 100$ Hz band.

There are some remarks to be made about this procedure. First, the sensors are not positioned in the same place, and therefore the measurements are slightly off. It is difficult to say how this affects the ASD, since the ASD is derived from two time signals and does not incorporate any spatial information. However, when the distance between the sensors is significantly smaller than the distance to the magnetic source, then the sensors are measuring approximately the same magnetic field. Furthermore, this procedure only uses two sensors, which might not be enough to accurately determine the sensitivity. Adding more sensors to the process allows us to calculate the mean and standard deviation of the sensitivity, which can give us a better idea of the sensitivity of the sensor. Finally, since the sensors are stationary, it is not guaranteed that all frequencies are present in the source signal, the ASD might be biased towards certain frequencies. Placing the two sensors in a simulated environment, where the magnetic field can be controlled, can help to determine the sensitivity of a specific frequency more accurately.

II

Magnetic Map-making Techniques

5

Magnetic Maps

5.1. Magnetic Maps

The Magnetic Field is not constant on Earth, therefore, a map of the magnetic field can be of use in many applications. The non-uniform distribution of magnetizable materials on Earth makes the values of the magnetic field location specific. Moreover, motions of the Earth's core, crust and human activities can cause the magnetic field to change over time. This makes the magnetic field time specific as well.

It has been stated in Chapter 3 that one of the main contributions to the magnetic field on Earth's surface comes from the outer core, but also magnetizable rocks in the crust or even magnetizable objects near the surface can have a significant contribution to the magnetic field. A premise that is made in this section is that sources at large distances contribute to low spatial frequencies, while sources closer to the surface are responsible for high spatial frequencies. This substantiates the choice to model the magnetic field in steps, starting with a model of the core field, then the crustal field, and then further local anomalies.

A map of the magnetic field exists in different variations. First, one can represent the magnetic field with vectorial (directional) or scalar (total-field intensity) measurements, or by the underlying scalar potential field. In addition, for some applications, the entire magnetic field is not required but it is sufficient to map only the variation. This is where magnetic anomaly maps come in.

5.2. Vector-Field and Total-Field Maps

A distinction between Vector-Field and Total-Field maps must be made. A Vector-Field map is a map of the direction and intensity of the magnetic field, it defines the three components of the magnetic field at each point. Any value in a Vector-Field map can be represented by a five-dimensional vector, which consists of the two position coordinates (x, y), the height (z) is fixed, and the three components of the magnetic field (B_x, B_y, B_z). A Total-Field map is a map of the intensity of the magnetic field and is often obtained by a scalar magnetometer (Chapter 4.3), but it can also be computed from a Vector-Field map by taking the L_2-norm. The values in a Total-Field map are three-dimensional vectors, two for the position and one for the intensity of the magnetic field $I = |\mathbf{B}|$.

5.3. Maps of the Scalar Potential

The magnetic field in a static state has a scalar potential field representation (Section 2.2.2). This allows for a compact representation of the magnetic field using scalar potential field maps. The use of a scalar potential field map is unfortunately somewhat less user-friendly, since the potential cannot be measured directly. Therefore, a derivative often has to be calculated to compare a measurement with the map. The scalar potential field representation can however have computational advantages [5, Chapter 4], especially when large computations have to be performed separately on the three components of the magnetic field. Transforming this vector field to a scalar potential and performing the computations directly on the potential field can speed up the computations. This research does not focus on the scalar potential field representation, but it is mentioned here for completeness.

5.4. Magnetic Anomaly Maps

The large models that represent the Earth's magnetic field on a global scale are sometimes not detailed enough for local applications. The limitations of these models, often represented by Spherical Harmonics, are discussed in Section 6.4. Magnetic anomaly maps however are a widely used solution for the representation of regional magnetic fields. The magnetic anomaly map indicates the deviation of the magnetic field from a pre-defined reference field with a lower resolution. They are very suitable for locally enhancing the magnetic field in a higher resolution when only a more global lower resolution model is available. Lower resolution (reference) fields that are often used are the core field (Section 3.2) or the crustal field (Section 3.4) which are already well mapped by the IGRF (Section 6.1.2) or the Enhanced Magnetic Model, a higher resolution model which is not discussed in this thesis.

The best that a low-resolution reference field can do for a higher-resolution representation is to provide a value close to the average magnetic field in the region of interest. The anomaly map is a map on top of this reference field, restricted by the domain of the reference field, but not restricted by its resolution. On the contrary, anomaly maps are there to enhance the resolution of the reference field by indicating the deviation from the mean value. This explains also why anomaly maps are often considered to be *mean-zero*.

The anomaly maps are also very useful for the visualization of the magnetic field. On Earth's surface, the magnetic field intensity is in the order of 30 to 70 μT (Section 3.2). This is quite large compared to the spatial variations coming from the crustal field (Section 3.4). The variations that originate from the crust are in the order of 10 to 100 nT, which is a factor of 10^3 smaller than the main field. These variations are often not visible on a map of the total magnetic field intensity and are therefore better represented by a magnetic anomaly map, which is a map of the deviation from the mean value.

5.4.1. Expression of the Magnetic Anomaly Field

There are multiple ways to express a magnetic anomaly field. We distinguish between two types of magnetic anomaly maps, the vector deviation field and the total-field anomaly map. A vector deviation field $\mathbf{B}_a$ is defined as the difference between the measured magnetic induction field $\mathbf{B}$ and the reference field $\mathbf{B}_r$:

$$\mathbf{B}_a := \mathbf{B} - \mathbf{B}_r \,. \tag{5.1}$$

This representation is useful when the magnetic field is measured with a vector magnetometer (Chapter 4.3). When the magnetic field is measured with a Total-Field magnetometer, the orientation of the vector $\mathbf{B}$ is unknown. Therefore, the vector deviation field $\mathbf{B}_a$ (Equation 5.1) cannot be directly computed. We can only say something about the differences in two Total-Field measurements of the magnetic field:

$$\Delta T := |\mathbf{B}| - |\mathbf{B}_r| \neq |\mathbf{B}_a| \,, \tag{5.2}$$

where $|\mathbf{B}|$ is the magnitude of the measured magnetic field and $|\mathbf{B}_r|$ is the magnitude of the reference field. It is important to notice that the magnetic anomaly ΔT (Equation 5.2) is not equal to the magnitude of $\mathbf{B}_a$ (Equation 5.1). It is only a first-order approximation of the magnitude of the vector deviation field $|\mathbf{B}_a|$ (Equation 5.1). Blakely [5, p. 179] derived the following approximation

$$\begin{aligned}
\Delta T &= |\mathbf{B}_r + \mathbf{B}_a| - |\mathbf{B}_r| && (5.3)\\
&\approx \sqrt{\mathbf{B}_r \cdot \mathbf{B}_r + 2(\mathbf{B}_r \cdot \mathbf{B}_a)} - |\mathbf{B}_r| && (5.4)\\
&\approx \sqrt{\mathbf{B}_r \cdot \mathbf{B}_r} + \sqrt{\mathbf{B}_r \cdot \mathbf{B}_r}(\mathbf{B}_r \cdot \mathbf{B}_a) - |\mathbf{B}_r| && (5.5)\\
&= \frac{\mathbf{B}_r \cdot \mathbf{B}_a}{|\mathbf{B}_r|} = \hat{\mathbf{B}}_r \cdot \mathbf{B}_a \,, && (5.6)
\end{aligned}$$

where $\hat{\mathbf{B}}_r = \frac{\mathbf{B}_r}{|\mathbf{B}_r|}$ is the unit vector pointing in the direction of the reference field. It was remarked that this approximation is valid if the anomaly field $\mathbf{B}_a$ is small compared to the reference field $\mathbf{B}_r$ so when $|\mathbf{B}_a| \ll |\mathbf{B}_r|$, or when $\mathbf{B}_a$ has the same orientation as $\mathbf{B}_r$. This assumption holds for most geophysical applications, where the reference field is often induced by the core and in the orders of microtesla and the anomaly field, e.g. a field originating from the crust, is in the orders of nanotesla. Rocks that are permanently magnetized retain their magnetization even in the absence of an external field, or when the external field changes direction. These rocks might be misaligned with the main field and can therefore cause an error in the approximation (Equation 5.3). The approximation is therefore also a good approximation when the magnitude of the permanently magnetized rocks is a lot smaller than the induced fields and the main field.

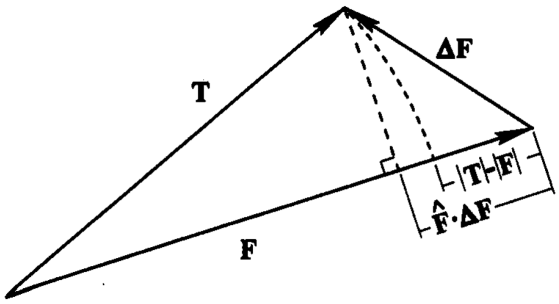


Figure 5.1: Vector representation of the Total-field anomaly (here $\Delta\mathbf{F}$). The length $|\mathbf{T}| - |\mathbf{F}|$ is a good approximation of the projection of the vector deviation field $\Delta\mathbf{F}$ onto the direction of the reference field $\hat{\mathbf{F}}$ whenever $|\Delta\mathbf{F}| \ll |\mathbf{F}|$. This figure is taken from [5].

A good illustration of the approximation is given in Figure 5.1. From this figure, it is clear that the Total-field anomaly ΔT is approximately equal to the projection of the vector deviation field $\mathbf{B}_a$ onto the direction of the reference field $\mathbf{B}_r$.

5.5. Three-dimensional Maps

In the description of maps, we have until now only considered two-dimensional maps. This is sufficient when one wants to navigate or find something on Earth's surface. However, when one strives to use magnetic maps for aeromagnetic navigation, it is necessary to have information about the magnetic field in a three-dimensional space, since aircrafts do not attain a constant altitude during their flight. This information, however, can under certain circumstances be obtained from two-dimensional maps. Especially in the context of Aeromagnetic Navigation, it is a good assumption that all the magnetic sources lie underneath your aircraft. This allows the application of Green's third identity (Section 2.4.1) to obtain any value of the magnetic field above a plane at a certain altitude. This principle is called Upward Continuation which is thoroughly described in Chapter 10.

5.6. Temporal Variations

The magnetic field on Earth is not constant over time, therefore, the validity of magnetic maps is limited in time. Several factors can cause the magnetic field to change and these changes happen on different timescales. We distinguish two types of temporal variations: short-term (or transient) variations and long-term (or secular) variations. This section enlists the most important factors that cause temporal variations in the magnetic field and provides some solutions to compensate for these variations.

5.6.1. Transient variations

The most transient (short-term) variations in the magnetic field are related to the weather. A geomagnetic storm, caused by solar wind, can cause a short-term variation in the magnetic field. In quiet times, the contribution of solar winds to the magnetic field at the equator ranges from -20 to 20 nT. When a geomagnetic storm is present, deviations up to hundreds of nanoteslas can occur [17]. The storms can last for several days and there are several ways to describe their intensity. Another short-term cause is lightning. When lightning strikes, a brief burst of electromagnetic radiation is emitted, but this radiation decays very quickly and is therefore not a problem when outliers are removed from the data. There exist many other short-term variations in the magnetic field, in general, all these variations are caused by external sources [64]. It is difficult to predict these variations and therefore, it is not easy to compensate for them in a magnetic map.

5.6.2. Secular variations

Secular variations are the variations that occur on timescales of at least a year. These variations are mostly caused by internal processes in the Earth. Most of these variations are caused by the Earth's core. On geological timescales, the movement of magnetized rocks by plate tectonics can also cause variations in the Earth's magnetic field (Section 3.4.1).

5.6.3. Ground Stations for transient variations

To most effective way to compensate for transient variations is to use ground stations. Ground stations are placed in regions where the magnetic field is relatively stable. They measure the magnetic field over a long period and use this data to correct for temporal variations in the external fields. For example, when a measurement campaign of several days is performed, the data from the ground station can be used to correct for variations that occur on a minutely, hourly or daily scale. The deviations measured at the ground station can be subtracted from the measurements of the mobile sensors.

5.6.4. Time-dependent differential equations for secular variations

Especially on a regional scale, the magnetic field is relatively stable and therefore, no variations occur on a short-term scale. One might however be interested in a map of the magnetic field that is valid for a long time. If one also demands that this map is nanotesla-accurate, then it is necessary to compensate for secular variations. The most common way to compensate for secular variations is to use a time-dependent differential equation. Several studies have been conducted to find the best way to model the (outer) core of the Earth, and the movement of Earth's plates. One can predict how the magnetic field is affected by these processes and use this information to keep the magnetic map up-to-date over time. We refer to these maps as magnetic models. In Chapter 6, we discuss the most important magnetic models in more detail. These models are often expressed by using a spherical harmonic expansion with time-dependent Gauss coefficients (Equation 6.1).

5.7. Examples of Magnetic Maps

A nice example of a magnetic anomaly map on a regional scale is a map of the Netherlands. The map was created by TNO for the Dutch Ministry of economic affairs to see where gas fields are located. The map is based on various Total-field measurements of the magnetic field in the Netherlands. The measurements on land were done at a set of locations that were evenly distributed over the country. The measurements at sea were taken by a ship that sailed up and down the coast and parts of the North. After the measurements were taken, data selection and data weighting (Section 6.2.2 - 6.2.3) were performed to make the map as accurate as possible. The resulting output, a map of the Total-field values of the Netherlands, was projected on a reference field to create a magnetic anomaly map. This map is shown in Figure 5.2. The reference field that is used for this map is DGRF85, a core-field model (Section 6.1) that is based on the IGRF model of 1985 (Section 6.1.2).

The map is not of a very high resolution, but it is detailed enough to notice some things: some large deviations are visible around the islands of Texel and Vlieland in the northwest of the Netherlands. This anomaly is caused by a dead volcano: the "Zuidwal volcano" has not been active for about 150 million years. It is estimated that this volcano lies at a depth of 2 kilometers.

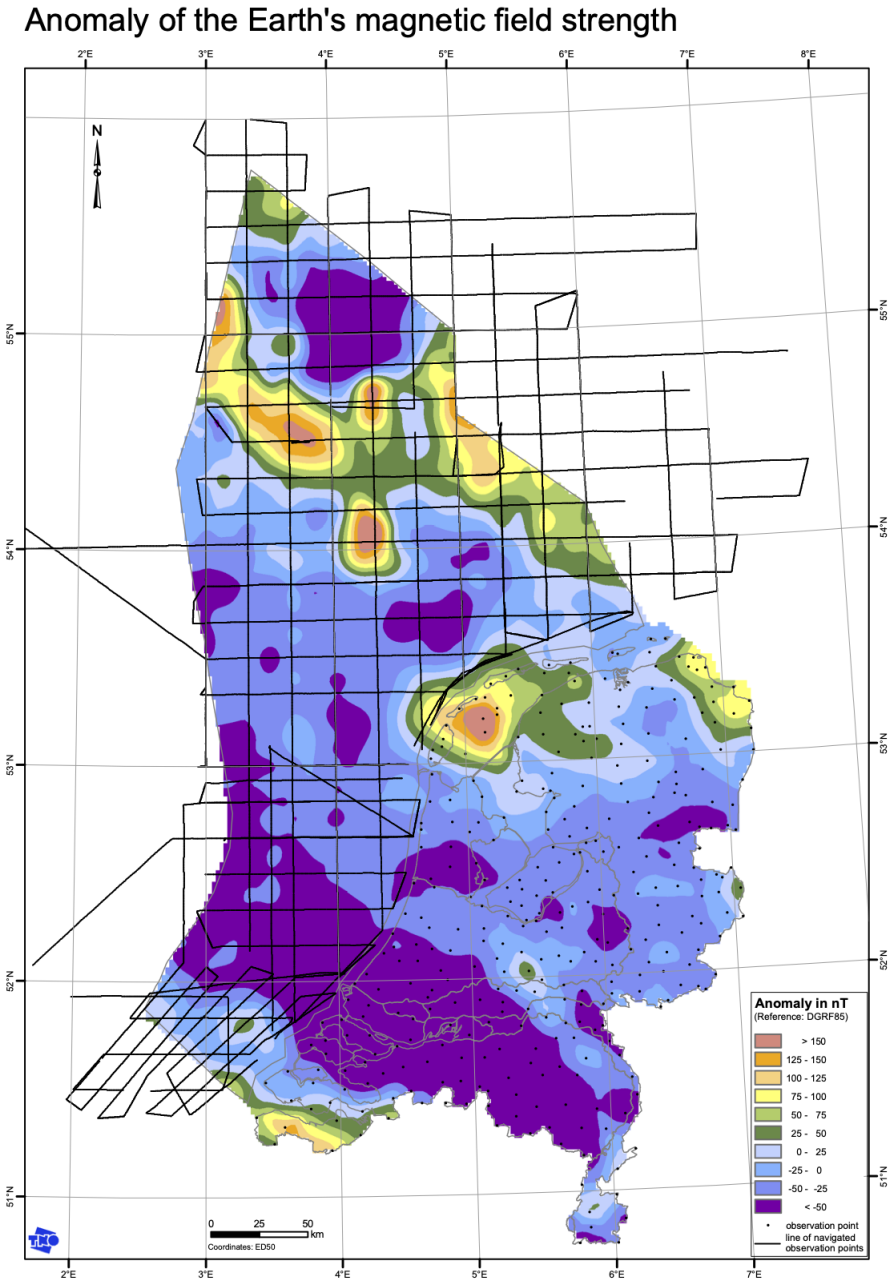


Figure 5.2: Magnetic anomaly map of the Netherlands. The map is based on measurements of the magnetic field in the Netherlands. The measurements were taken on land and at sea. The map is projected on a reference field to create a magnetic anomaly map. The reference field that is used for this map is DGRF85, a core-field model (Section 6.1) that is based on the IGRF model of 1985 (Section 6.1.2). This figure is taken from [46].

6

Modelling the Geomagnetic Field

The magnetic anomaly maps that are introduced in Chapter 5 represent often a deviation from a reference field. This reference field is usually a model of the Earth's magnetic field. The geomagnetic field can be represented in several ways. The choice of representation depends on the scale of the problem. On a large scale, e.g. the Earth's surface or continents, the curvature of the Earth can not be neglected. This gives rise to representing the harmonic field as a sum of spherical harmonics. On smaller scales, e.g. in a region of a few kilometers, the curvature of the Earth can be neglected. Here, plane waves in a cartesian coordinate system are often used. This chapter presents the techniques used to represent large-scale fields, such as the Earth's core or crustal field. Local enhancements of the field using a cartesian coordinate system are discussed in Chapter 8 and 9.

6.1. Core Field Models

The magnetic field originating from the Earth's core is the dominant field on Earth. A good model of the core field is therefore essential for applications such as geophysical exploration and navigation. Several models of the core field exist, but the most commonly used are the International Geomagnetic Reference Field (IGRF) and World Magnetic Model (WMM). This section gives a brief overview of these models and how they are derived using spherical harmonic analysis.

It is important to note that core field models like WMM and IGRF13 only model the main field, generated in the outer core. They do not model the geomagnetic fields that originate in the mantle, crust, ionosphere or magnetosphere. This means that, when using these models as a reference field, anomalies might occur when taking measurements with the magnetic sensor data from a compass or magnetometer [52]. Spatial anomalies on land can have several causes, such as the presence of ferromagnetic minerals in the soil, geological faults or human-made properties such as vehicles, power lines and railroad tracks. Anomalies in oceanic areas have more often a geological cause: ocean ridges, seamounts and submarine volcanoes are the most common causes.

6.1.1. World Magnetic Model (WMM)

The World Magnetic Model (WMM) is the model that is used by the U.S. Department of Defense, the U.K. Ministry of Defence, the North Atlantic Treaty Organization (NATO). The model is maintained by the United States National Oceanic and Atmospheric Administration (NOAA) and the United Kingdom's Defence Geographic Centre (DPC). It is based on historical data and current observations and tries to predict the behaviour of the field in the future. This results in a model that is accurate in the short term, but degrades with time. The current model WMM2020 is valid until 31 December 2024, after that, the model is prone to have significant errors. On its release date, the model is the most accurate and its accuracy degrades with time. Therefore, every 5 years a new model is released. The WMM is constructed using Spherical Harmonics (Section 6.5, Appendix A) of degree 12, spectral analysis of the main field shows that the degree of truncation is sufficient to model the main field (Section 6.3). The primary geomagnetic data set is generated using data from three Swarm satellites [18]. Figure 6.1 shows the Total-field intensity of the WMM2020 model.

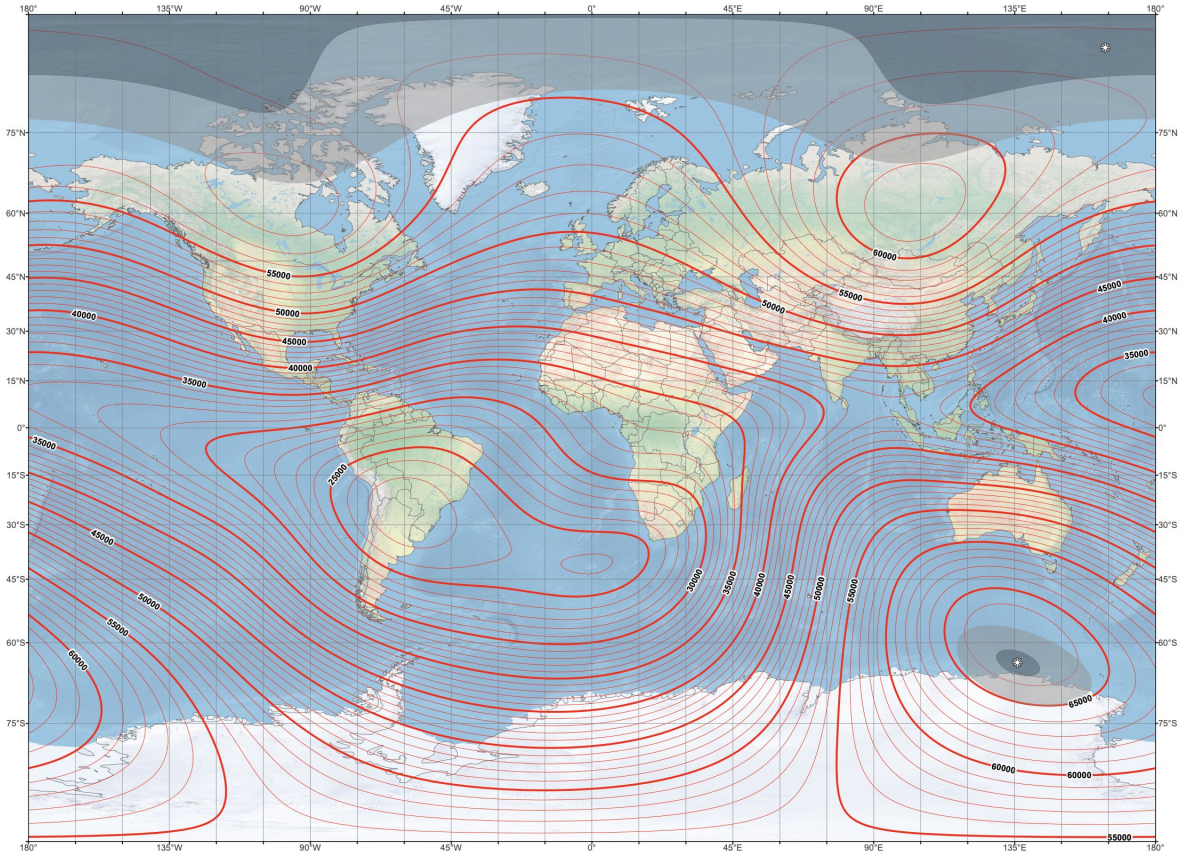


Figure 6.1: The World Magnetic Model (WMM) of 2020. The red contour lines indicate the Total-field intensity of Earth's main field in steps fo 1000 nT. The gray regions near the poles are blackout zones, where the model is less reliable. This figure is taken from [18].

6.1.2. International Geomagnetic Reference Field (IGRF)

The International Geomagnetic Reference Field (IGRF) is modeled and developed by the International Association of Geomagnetism and Aeronomy (IAGA). IAGA has a scientific community that contributes voluntarily to the development of the model. This model is used by the European Space Agency (ESA) and the European Union (EU). The model is very similar to the WMM model, but it is not maintained by two government agencies. The current model IGRF13 is valid until 2025 (Figure 6.2).

The IAGA keeps active track of the geomagnetic field. They compare the predicted IGRF model with current observations. The estimated coefficients are then updated to reflect the current state of the field. This leads to the Definitive Geomagnetic Reference Field (DGRF) model. The values of the DGRF model are more accurate than the IGRF model, but the DGRF model can only be evaluated at historical moments in time.

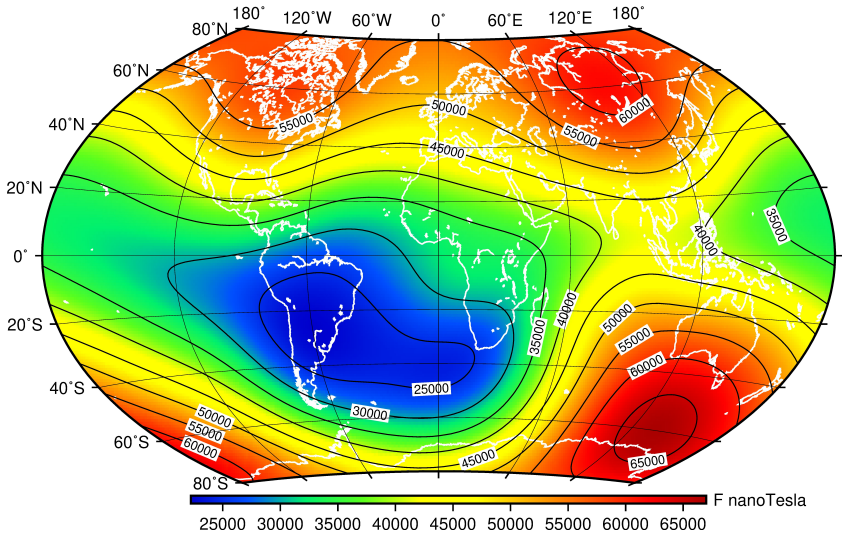


Figure 6.2: The Total-field intensities given by IGRF-13. The 13th generation of the International Geomagnetic Reference Field (IGRF) model is created in 2019 and is valid until 2025. This figure is taken from [1]

6.2. Selecting and Processing Data for Large-Scale Models

6.2.1. Data Sources

The data for most core field models is collected using satellites in combination with ground-based measurements. The following satellites are often used for this purpose:

- **CHAMP** (CHAllenging Minisatellite Payload) is a satellite that was launched in 2000. It was designed to study the Earth's gravity field and the Earth's magnetic field [55]. The mission was successful and finished in 2010.

- **Swarm** is a European Space Agency (ESA) mission that was launched in 2013. It consists of three satellites it is currently in a polar orbit and maintains an altitude of approximately 450 km. [49]

The advantage of using satellites is twofold. First, the satellites can be placed in an orbit, which allows them to collect data from a large area of the Earth. Second, the altitude of the satellites has a large impact on the accuracy of the measurements. A higher altitude results in a lower noise level, since noise is often of a higher frequency and high frequencies are attenuated at higher altitudes. A sufficiently high altitude reduces the power of sources located on the ground or in the crust. The technique of using altitude to reduce noise is described in-depth in Section 7.2.1.

Besides satellites, ground stations are also used to collect data. Ground stations have a fixed position on Earth's surface and are therefore spatially not that interesting. However, their fixed position makes them perfect for measuring the temporal variation of the geomagnetic field. Small variations that originate from external sources, such as the solar wind and the ionosphere, can be measured with high accuracy. The data from ground stations is therefore often used to correct the data from satellites.

6.2.2. Data Selection

The measured data is not always suitable for use in a model. The data is often contaminated with noise, some of the noise is significant and makes the data unusable. For this reason, data selection is an important step. Several criteria can be used to select data. Two common criteria are the K_p index and the R_c index.

K_p index

The K_p index is a measure of the geomagnetic activity on Earth. The p stands for planetary since the index is an average of local K indices on Earth. K indices are measured at ground stations. They range from 0 to 9. A K index of 0 indicates that the magnetic environment is quiet, while a K index of 9 indicates a significant magnetic storm (Section 5.6.1). A change in K_p index can have several causes, but the most common cause is solar activity such as Solar flares and Coronal Mass Ejections (CME). For the data selection of the IGRF, the K_p index was not allowed to exceed 2 [2].

R_c index

The R_c index describes the rate of change of the geomagnetic field. A ground station at a fixed position measures a near-constant geomagnetic field if noise is absent. The rate of change is therefore a good indication of the presence of noise. Data that exceeded a threshold of 3 $\mathrm{nT\,h^{-1}}$ was discarded in the selection process of the IGRF model [2].

6.2.3. Data Weighting

Data weighting is a technique that is used to correct for the bias in data that is retrieved from non-uniformly distributed measurements. Satellites are often placed in a polar orbit, which means that they take more measurements at the poles than at the equator. At the same time, ground stations are often placed on land, which means that there are fewer measurements available at sea. Both of these factors result in a bias in the data, if the data is not weighted properly.

Researchers of the IGRF model shared their data selection and weighting techniques in [2]. They used three different weighting factors:

- A spatial weighting factor that is inversely proportional to the distance to the nearest measurement point. This ensures uniformly distributed values over all latitudes and longitudes and ensures that sparsely sampled regions are not underrepresented.
- A weighting factor that indicates the quality of a dataset. This is often related to the variance of the data, but also to for example the noise levels based on Kp and Rc indices.
- A weighting factor that reduces the effect of outliers. This is done by using a robust estimator, such as the median.

6.3. Geomagnetic Field Spectrum

The spectrum of the Geomagnetic Field was extensively studied by Langel [33]. He noted a so-called "knee" in the spectrum of the geomagnetic field. After a degree of 15, corresponding to a wavelength of 2668km, the spectrum of the geomagnetic field is almost flat (Figure 6.3). Up to a degree of 15, the spectrum is dominated by the core field, while after a degree of 15, the spectrum is dominated by the crustal field. When plotting the Gauss coefficients as a function of the degree of the spherical harmonics, a clear break in the density function is observed between a degree of 12 and 15 [52]. This indicates that the to low-degree spherical harmonics are sufficient to model the main field. Contributions to higher degree spherical harmonics are caused by the crustal and induced fields.

6.4. Limitations of the Spherical Harmonics Model

The power spectrum of the geomagnetic field (Figure 6.3) gives good insights into the effectiveness of the spherical harmonics model. At a certain point, the power spectrum attains on the logarithmic scale almost a constant value of approximately 10^2 to 10^3 nT^2. Beyond this degree, adding more spherical harmonics still improves the model, but the improvement is small. At the same time, the computational and data requirements increase: the smaller the wavelength, the more measurement points are required to accurately model the field. Retrieving coarse measurements on Earth can be difficult for oceanic regions or regions that are not covered by an orbiting satellite and requires weighting techniques to correct for the bias in the data.

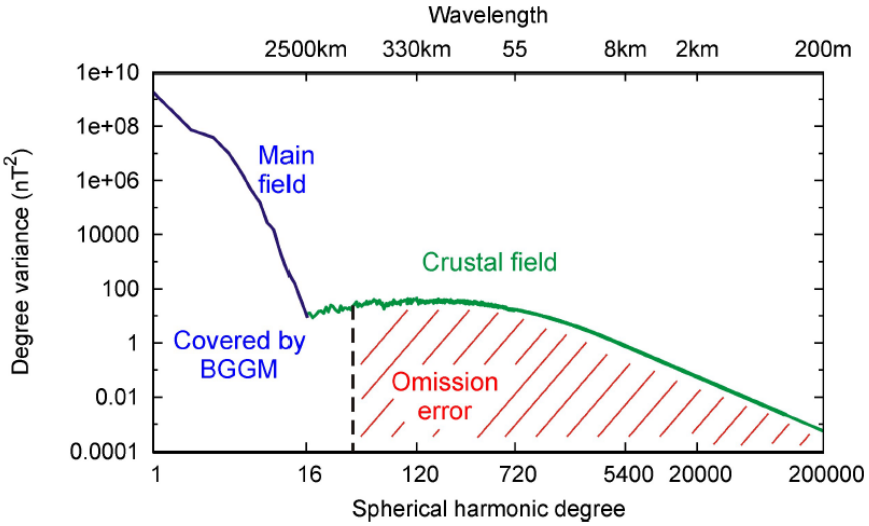


Figure 6.3: Power spectrum of the geomagnetic field at the Earth's surface. The "knee" is located at a degree of 15. This figure is taken from [42].

Higher resolution models using spherical harmonics with a degree up to 720 are available, but they require more frequent updates of the coefficients as the degree increases since they model the magnetic field that originates from smaller sources, closer to or at Earth's surface which are more subject to changes. Therefore, spherical harmonic models are not useful on regional scales with a resolution of for example 1km.

6.5. Representing the Geomagnetic Field

Spherical harmonics are an efficient way to approximate an arbitrary function on a sphere using harmonic functions, similar to how Fourier series are used to approximate an arbitrary function on a line using sinusoidal functions. As the name suggests, the basis consists of harmonic functions, as they are the solution of the Laplace equation on a sphere. This is useful to represent magnetic scalar potential, which is also harmonic (Section 2.4). The potential U, in this context a function of the latitude θ, longitude ϕ, radius r in the geocentric coordinate system and time t, can be represented as a sum of spherical harmonics:

$$U(r,\theta,\phi,t) = a + \sum_{n=1}^{N} \left(\frac{a}{r}\right)^{n+1} \sum_{m=0}^{n} \left(\tilde{g}_n^m(t)\cos(m\phi) + \tilde{h}_n^m(t)\sin(m\phi)\right) \check{P}_n^m(\cos\theta)\,, \tag{6.1}$$

where the number N is the degree of truncation of the series. The coefficients $\tilde{g}_n^m(t)$ and $\tilde{h}_n^m(t)$ are the time-dependent coefficients of the spherical harmonic series, the maintainers of the model update these coefficients to correct for secular variations in the field. These coefficients are called Gauss coefficients, named after Carl Friedrich Gauss who used spherical harmonics to create the first model of the Earth's magnetic field to show that the main contribution to the field comes from sources below Earth's surface [52].

The function $\check{P}_n^m(x)$ is the Legendre polynomial with a Schmidt normalization. The Schmidt normalization is the most common normalization used in magnetics and allows finding the most significant coefficients in the series by inspection. It is defined by:

$$\check{P}_n^m(x) = \sqrt{2\frac{(n-m)!}{(n+m)!}} P_n^m(x)\,, \tag{6.2}$$

where $P_n^m(x)$ is the Legendre polynomial of degree n and order m defined by Equation A.16. A general spherical harmonic basis is of the form:

$$Y_l^m(\lambda,\phi) = Ne^{im\phi} P_l^m(\cos\lambda)\,, \tag{6.3}$$

where N is a normalization constant. In Equation A.9 it is shown that the general solution to the Laplace equation on a sphere can be separated into radial parts with positive powers of r and radial parts with negative powers of r.

The radial parts with positive powers $R_l^m(r)$ can be interpreted as the influence of sources that lie outside the sphere and have no singularities inside the sphere. The negative powers $I_l^m(r)$ are useful to describe internal sources and have singularities at the origin. Depending on the application, either the positive or negative powers can be used. For studies of the magnetic spheres on a small sphere with no internal sources, only the positive powers are needed. Large-scale studies of for example the Earth's magnetic field assume often that all the sources lie below the surface (or inside the sphere). And thus for geomagnetic studies mainly the negative powers are used [58]. The coefficients described in Equation 6.1 are, as desired, describing the spherical harmonic expansion caused by internal sources.

Equation 6.1 can be derived from this general form by incorporating some physical properties of magnetic fields: there do not exist magnetic monopoles, resulting in $a = 0$ and the coefficients for odd l are zero. The potential of Earth's dipole field can also be expressed using spherical harmonics by a first-order approximation with only the first two terms of the series ($l = 0$ and $l = 1$). This implies also that the potential degrades with the square of the distance from the origin. Resulting in a r^{-3} term in the corresponding magnetic dipole field.

6.5.1. Wavelengths of Truncated Spherical Harmonics

Equation 6.1 expresses the potential of the geomagnetic field in spherical harmonics. Models of Earth's magnetic field are often truncated versions of this series. Truncation of the series reduces the computational complexity of the model at the cost of resolution. As shown in Section 6.3, a truncated series of spherical harmonics is a sufficient approximation. For a N degree truncated series of spherical harmonics describing the geomagnetic field, the minimum wavelength of the field is given by:

$$\lambda_{\text{min}} = \frac{2\pi R}{N}, \tag{6.4}$$

where $R = 6371$km is the Earth's radius. The minimum wavelengths for $N = 12$ and $N = 13$ are respectively 3336 km and 3079 km. A large wavelength corresponds to a low-resolution model. The WMM and IGRF13 models are thus low-resolution models.

6.5.2. Estimating Gauss Coefficients

Based on the scale of the problem, the number of measurements and the desired accuracy, it is difficult to estimate the Gauss coefficients analytically. Most models of Earth's magnetic field are based on a least squares fit of the Gauss coefficients to a set of measurements. The maintainers of IGRF used a least squares fit with a cost function that incorporates the weighting of the measurements and minimizes the difference between the model and the measurements [2].

Secular variations of the core field are also captured by the Gauss coefficients. It is assumed that the secular variations are slow and smooth. In most models, a linear approximation of the Gauss coefficients over a period of 5 years is used [1]:

$$\frac{dg_n^m(t)}{dt} = \frac{g_n^m(t + 5\text{ years}) - g_n^m(t)}{5\text{ years}}, \tag{6.5}$$

here $g_n^m(t)$ are the Gauss coefficients written in a compact form (see Equation 6.7).

Secular accelerations are not included in this linear model. Cubic B-splines of order 6 between uniform knots separated by half a year can be used to incorporate the secular accelerations $\frac{d^2}{dt^2} g_n^m(t)$, this is introduced in [2]:

$$g_n^m(t) = \sum_i g_{n,i}^m N_{i,k}(t), \tag{6.6}$$

where $N_{i,k}(t)$ is the kth order B-spline with knot t_i and $g_{n,i}^m$ are the coefficients of the B-spline.

6.5.3. Compact Representation of the Potential

To simplify further calculations, the potential can be expressed in a compact form. This form is taken from [2]. The Gauss coefficients $\tilde{g}_n^m(t)$ and $\tilde{h}_n^m(t)$ are replaced by a single coefficient $\tilde{g}_n^m(t)$ and a new function $S_n^m(\theta,\phi)$ is introduced. We define

$$g_n^m(t) = \begin{cases} \tilde{g}_n^m(t) & \text{if } m \geq 0 \\ \tilde{h}_n^m(t) & \text{if } m < 0, \end{cases} \tag{6.7}$$

$$S_n^m(\theta,\phi) = \begin{cases} \cos\big(m\phi\big)\check{P}_n^m(\cos\theta) & \text{if } m \geq 0 \\ \sin\big(|m|\phi\big)\check{P}_n^{|m|}(\cos\theta) & \text{if } m < 0, \end{cases} \tag{6.8}$$

$$U_n^m(r,\theta,\phi) = a\left(\frac{a}{r}\right)^{n+1} S_n^m(\theta,\phi), \tag{6.9}$$

so that the potential can be expressed as

$$U(r,\theta,\phi,t) = \sum_{n=1}^{N}\sum_{m=-n}^{n} g_n^m(t)U_n^m(r,\theta,\phi). \tag{6.10}$$

6.5.4. Compact Representation of the Magnetic Field

The compact representation of the magnetic field follows from differentiating Equation 6.10 with respect r, θ and ϕ: $\mathbf{B} = -\nabla U$. The magnetic field is then given by

$$\mathbf{B}(r,\theta,\phi,t) = \sum_{n=1}^{N}\sum_{m=-n}^{n} g_n^m(t)\mathbf{B}_n^m(r,\theta,\phi), \tag{6.11}$$

where $\mathbf{B}_n^m = -\nabla U_n^m$ is the vector containing the partial derivatives of the potential. The individual components can also be expressed in compact form:

$$\begin{pmatrix} \mathbf{B}_{n,r}^m \\ \mathbf{B}_{n,\theta}^m \\ \mathbf{B}_{n,\phi}^m \end{pmatrix} = \left(\frac{a}{r}\right)^{n+2} \begin{pmatrix} (n+1)S_n^m(\theta,\phi) \\ -\partial_\theta S_n^m(\theta,\phi) \\ -\partial_\phi S_n^m(\theta,\phi), \end{pmatrix} \tag{6.12}$$

where

$$\partial_\phi S_n^m(\theta,\phi) = \begin{cases} -m\sin\big(m\phi\big)\check{P}_n^m(\cos\theta) & \text{if } m \geq 0 \\ |m|\cos\big(m\phi\big)\check{P}_n^{|m|}(\cos\theta) & \text{if } m < 0, \end{cases} \tag{6.13}$$

$$\partial_\theta S_n^m(\theta,\phi) = \begin{cases} \cos\big(m\phi\big)\partial_\theta\check{P}_n^m(\cos\theta) & \text{if } m \geq 0 \\ \sin\big(m\phi\big)\partial_\theta\check{P}_n^{|m|}(\cos\theta) & \text{if } m < 0, \end{cases} \tag{6.14}$$

where $\partial_\theta\check{P}_n^m(\cos\theta)$ is the derivative of the Schmidt normalized associated Legendre polynomial with respect to ϕ. The Schmidt-normalization is not dependent on ϕ and thus acts as a constant. Differentiating the associated Legendre polynomial $\check{P}_n^m(x)$ with respect to x and applying the chain rule gives

$$\partial_x P_n^m(x) = \partial_x\left[(-1)^m(1-x^2)^{m/2}\frac{\partial^m}{\partial x^m}P_n(x)\right] \tag{6.15}$$

$$= (-1)^m\frac{m}{2}(-2x)(1-x^2)^{m/2-1}\frac{\partial^m}{\partial x^m}P_n(x) \tag{6.16}$$

$$+ (-1)^m(1-x^2)^{m/2}\frac{\partial^{m+1}}{\partial x^{m+1}}P_n(x) \tag{6.17}$$

$$= \frac{x^2-1}{mx}P_n^m(x) + (-1)^m(1-x^2)^{-1/2}P_n^{m+1}(x), \tag{6.18}$$

and substituting $x = \cos\theta$ gives

$$\partial_\theta P_n^m(\cos\theta) = -\left[\frac{\cos^2\theta - 1}{m\cos\theta}P_n^m(\cos\theta) + (-1)^m(1-\cos^2\theta)^{-1/2}P_n^{m+1}(\cos\theta)\right]\sin\theta \tag{6.19}$$

$$= \frac{\sin^3\theta}{m\cos\theta}P_n^m(\cos\theta) + (-1)^{m+1}\sin^2\theta P_n^{m+1}(\cos\theta). \tag{6.20}$$

6.6. Example of Spherical Harmonics using Gaussian Process Regression

In this section, we show how spherical harmonics can be used to model a magnetic field. We use the Gaussian Process Regression (GPR) model to fit a magnetic field model to a set of data points. Gaussian Process Regression is discussed in more detail in Appendix B. The example in this section demonstrates how spherical harmonics can represent a field on a sphere. The example is not meant to be a realistic model of the Earth's magnetic field.

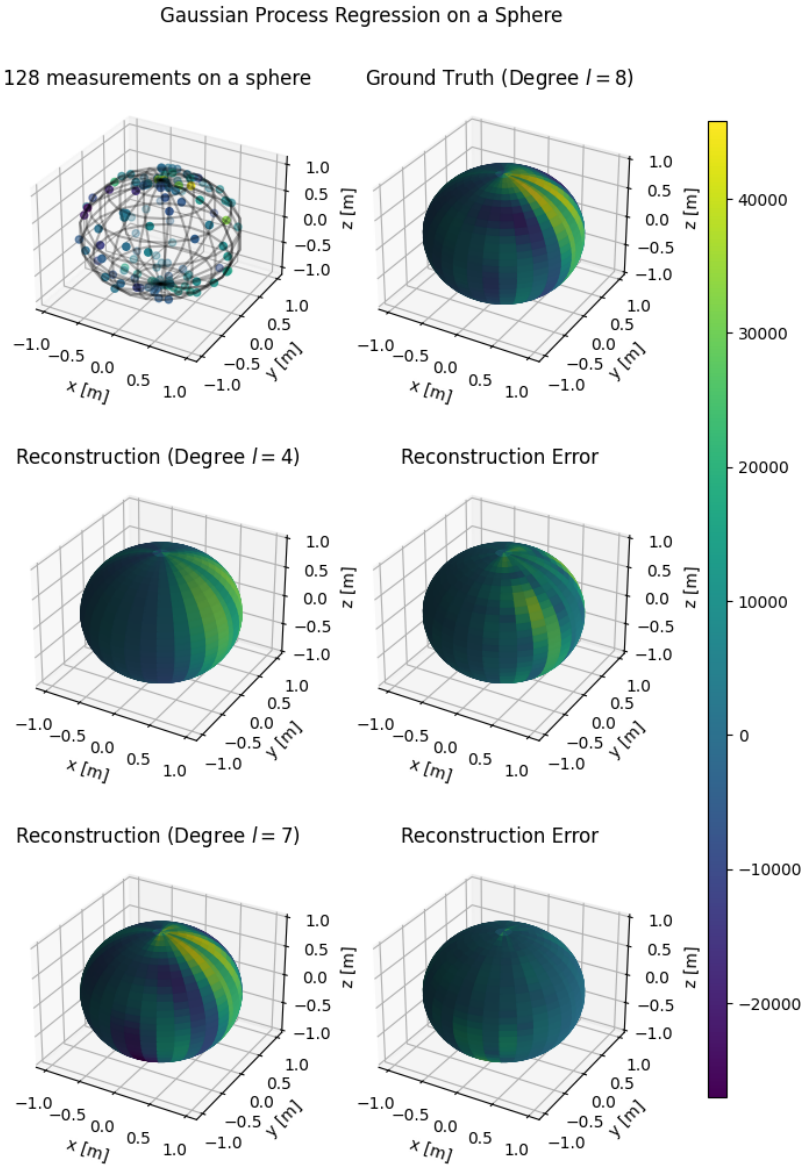


Figure 6.4: Example of a magnetic field model using spherical harmonics. Measurements are taken at 128 randomly distributed points on the sphere. The model is fitted to this data using Gaussian Process Regression. The regression model uses two bases: one with a degree of $l = 4$ and one with a degree of $l = 7$. The reconstruction and error for both models are shown.

Figure 6.4 demonstrates the principles of spherical harmonics in combination with Gaussian Process Regression. Measurements are taken at uniformly distributed points on the sphere. The results are visualized for two different degrees of spherical harmonics. The first model uses a degree of $l = 4$ and can therefore only represent the lower spatial frequencies on the sphere. The second reconstruction, using spherical harmonics up to a degree of $l = 7$, results in a more detailed reconstruction.

7

Aeromagnetic Compensation

The conventional method for surveying large areas of land or ocean involves the use of satellite measurements [2]. The World Magnetic Model (Section 6.1.1) and IGRF (Section 6.1.2) of the previous chapter are examples of models that are based on satellite measurements. Satellites offer, due to their stable platform, a low-noise environment for measuring the magnetic field. However, the inevitable trade-off is the lower spatial resolution of satellite-based measurements that comes from the large distance between the satellite and the magnetic sources, potentially compromising the detailed examination of smaller regions of interest.

To enhance resolution, airborne measurements become a viable option. These are typically taken using drones, airplanes or even helicopters, which can get closer to the Earth's surface, thus offering higher-resolution measurements of the magnetic field. Despite the resolution improvement, the more dynamic, less predictable environment these platforms operate in introduces more noise into the measurements. Vibrations from the platform's movement, interference from engines, and other electrical equipment contribute to this noise, necessitating proper compensation techniques [3, 22]. This section describes how magnetic sensor data can be compensated for the platform. We break down the magnetic field into several components and describe how calibration flights and compensation algorithms can be used.

7.1. Aeromagnetic Surveying

Aeromagnetic surveying is the process of measuring the magnetic field from an aircraft. Often, multiple total-field and vector magnetometers are placed on the surface of the aircraft. Depending on the aircraft type, the sensors are placed on the wings, fuselage or tail. The positioning of the sensors is important, as the aircraft itself induces a magnetic field, called the platform field. The wings and tail are the most common places for the sensors, as they are extremities and therefore measure a smaller part of the platform field. Sometimes sensors are even placed on a tail stinger or attached to a rope as a so-called 'bird' installation (Figure 7.1), to reduce the influence of the platform field [54]. However, this is not feasible at high speeds because of the drag, and some aircraft types, like drones, do not allow for this kind of installation because of their limited payload.

The first step in finding the anomaly field is to remove the platform field from the measured magnetic field. A large portion of the platform field is induced by the Earth's magnetic field. A model of the platform field is therefore often based on the aircraft's position, speed and heading. The position and heading are usually given by a combination of a GPS receiver and an Inertial Measurement Unit (IMU). The IMU is used to calculate the roll, pitch and yaw of the aircraft. Roll, pitch and yaw are the angles between the aircraft's body axes and the Earth's axes. These values are often used to calculate the induced parts of the platform field. The roll, pitch and yaw can be calculated in an arbitrary coordinate system. The coordinate system commonly employed for calculating the roll, pitch, and yaw is often the aircraft's body axes, although the Earth's axes can also be utilized.

The interference of the platform is, among other things, dependent on the position and orientation of the aircraft in the Earth's magnetic field. This underlies many aeromagnetic compensation methods.

Figure 7.1: A bird installation on a helicopter. The magnetometer is attached to a rope, which is attached to the helicopter. The magnetometer is therefore less sensitive to the platform field of the helicopter. This figure is taken from [54].

7.1.1. Breakdown of the measured magnetic field

The field $\mathbf{B}_t$ that is measured by a sensor is a combination of the Earth's magnetic field $\mathbf{B}_e$, the anomaly field $\mathbf{B}_a$, the platform field $\mathbf{B}_p$ and noise of the sensors $\mathbf{B}_n$. When the fields are in the same coordinate system, the measured field can be written as

$$\mathbf{B}_t = \mathbf{B}_e + \mathbf{B}_a + \mathbf{B}_p + \mathbf{B}_n. \tag{7.1}$$

In case an aircraft is equipped with multiple sensors, one can also write the measured field for sensor i as

$$\mathbf{B}_t^i = \mathbf{B}_e + \mathbf{B}_a + \mathbf{B}_p^i + \mathbf{B}_n^i, \tag{7.2}$$

with the assumption that sufficiently close sensors measure the same Earth and anomaly field.

7.2. Calibration and Compensation

7.2.1. Calibration flights

A calibration flight is (a part of) a flight that is used to calibrate the magnetometers so that the magnetic field of the platform or aircraft is removed from the measurements. The platform field, $\mathbf{B}_p$, is the field that is induced by the movement of the aircraft and currents in the aircraft's electrical system. The goal of a calibration flight is to find coefficients for several (measured) features.

To calibrate a magnetometer for an aircraft, a noise-free environment is required. In the breakdown of the magnetic field (Equation 7.1), it is seen that the measured magnetic field can be described as a sum of 4 fields: the field $\mathbf{B}_e$ produced by the outer core of the Earth, the anomaly field $\mathbf{B}_a$, the platform $\mathbf{B}_p$ and sensor noise $\mathbf{B}_n$. In general, the field $\mathbf{B}_e$ is well-known on Earth. For this field, values can be taken from the World Magnetic Model (Section 6.1.1) or IGRF (Section 6.1.2). Since the anomaly field, $\mathbf{B}_a$, is our field of interest, and therefore unknown, it is difficult to subtract these from measurements to isolate the platform field $\mathbf{B}_p$. In Section 2.5, however, it was observed that the strength of a magnetic dipole (or multipole), degrades with a factor of r^{-3} (or faster), where r is the distance to the source. This gives rise to a simple trick that can be used to suppress the anomaly field $\mathbf{B}_a$ in the measurements: increasing the distance to the sources, by flying higher, reduces the contribution of the anomaly field $\mathbf{B}_a$ relatively more than the contribution of the

main field $\mathbf{B}_e$. This means that at sufficiently high altitudes, the anomaly field $\mathbf{B}_a$ can be neglected, and the platform field $\mathbf{B}_p$ can be isolated from the measurements when the noise is small enough:

$$\mathbf{B}_t^h = \mathbf{B}_e + \mathbf{B}_p + \mathbf{B}_n, \tag{7.3}$$

for sufficiently large distances to the sources of the anomaly field $\mathbf{B}_a$. This distance depends on the spatial frequencies of the anomaly field. In the literature, often a height $h > 3000$ m is used. A height of 3000 m ensures that sources of at most a depth of 1000 m are dampened by a factor of at least $4^{-3} = \frac{1}{64}$, which mainly affects the anomaly field. Deeper sources are dampened less. This dampening effect is further discussed in Chapter 10 and is illustrated in Figure 10.3.

For the calibration flight, the platform field $\mathbf{B}_p$ can now be separated from the measurements $\mathbf{B}_t^h$. The goal now is to find the coefficients for certain input variables, such that the platform field $\mathbf{B}_p$ can be approximated by a (linear) combination of these input variables. Finding these coefficients is called calibration.

A common way to calibrate the magnetometers is by flying the aircraft in a certain pattern, while the magnetometers record the magnetic field. This pattern is chosen such that it includes many maneuvers. The clover pattern is popular since it allows measurement of the effects of maneuvers in different heading directions [23][14]. On the straight sides of the clover pattern, the pilot performs a combination of roll and pitch maneuvers, while the corners are used to measure the effects of yaw maneuvers.

An example of these "clover" patterns can be seen in Flight 1002 of the MIT Challenge (Section 7.4). Flight 1002 contains 2 calibration flights, both around 3000m in height. The flights are visualized in Figure 7.2.

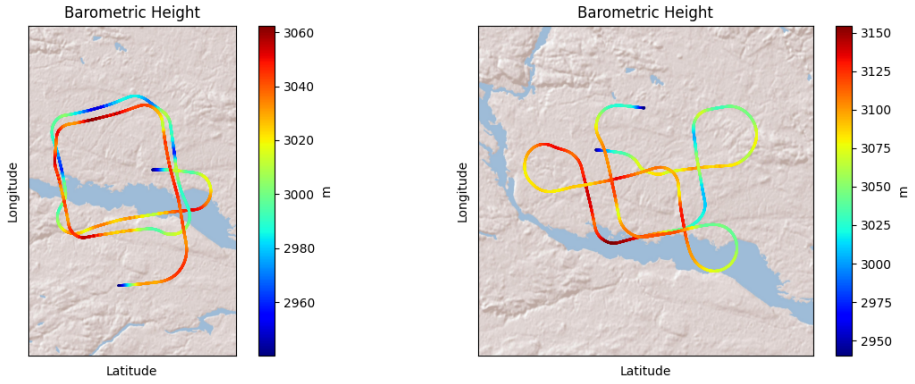


Figure 7.2: Calibration flights of flight 1002. Flt1002$_a$ is the first calibration flight, Flt1002$_b$ is the second calibration flight.

The unknown platform magnetic field $\mathbf{B}_p$ can be estimated by performing a least squares fit on the data. First, we need to define a model for the platform field $\mathbf{B}_p$. A linear model is often used. The model can be of the following form:

$$\mathbf{B}_p = \mathbf{A}\mathbf{c}, \tag{7.4}$$

where $\mathbf{A}$ is a matrix containing the input variables or features of the model as rows and $\mathbf{c}$ is a vector containing the coefficients corresponding to the features. The features are often prescribed by the model that is used for compensation. They can for example be dependent on the magnetic field strength, the maneuvers and the heading of the platform, but also other data such as altitude or electrical currents can be used. Common models are the Tolles-Lawson model (Section 7.2.2) and Leliak model (Section 7.2.3), but also more complex models such as Neural Networks (Section 7.2.4) are used in literature. Once the model is defined, the coefficients $\mathbf{c}$ can be estimated by performing a least squares fit on the data. The coefficients can then be used to compensate the measurements for the platform field $\mathbf{B}_p$.

7.2.2. Tolles-Lawson

Tolles and Lawson [67] developed a calibration method for the magnetic interference field of an aircraft. The Tolles-Lawson method is a linear model and uses a vector magnetometer to compensate a total field sensor for the platform field.

The vector magnetometer measures $\mathbf{B}_t$ (Equation 7.1) while a total-field sensor measures $|\mathbf{B}_t|$. The signal of interest is $|\mathbf{B}_e + \mathbf{B}_a|$: the magnitude of the Earth's magnetic field and the anomaly field. Restating 7.1 yields the following equation:

$$|\mathbf{B}_e + \mathbf{B}_a|^2 = |\mathbf{B}_t - \mathbf{B}_p - \mathbf{B}_n|^2 = |\mathbf{B}_t|^2 - 2\mathbf{B}_t \cdot (\mathbf{B}_p + \mathbf{B}_n) + |\mathbf{B}_p + \mathbf{B}_n|^2 \tag{7.5}$$

$$|\mathbf{B}_e + \mathbf{B}_a| = \sqrt{|\mathbf{B}_t|^2 - 2\mathbf{B}_t \cdot (\mathbf{B}_p + \mathbf{B}_n) + |\mathbf{B}_p + \mathbf{B}_n|^2} \tag{7.6}$$

$$= |\mathbf{B}_t| \sqrt{1 - \frac{2\mathbf{B}_t \cdot (\mathbf{B}_p + \mathbf{B}_n)}{|\mathbf{B}_t|^2 + |\mathbf{B}_p + \mathbf{B}_n|^2}} \approx |\mathbf{B}_t| \sqrt{1 - \frac{2\mathbf{B}_t \cdot \mathbf{B}_p}{|\mathbf{B}_t|^2}} \tag{7.7}$$

$$= |\mathbf{B}_t| - \frac{\mathbf{B}_t \cdot \mathbf{B}_p}{|\mathbf{B}_t|}, \tag{7.8}$$

where the first approximation is valid when the noise is neglected and $|\mathbf{B}_p + \mathbf{B}_n| \ll |\mathbf{B}_t|$. The Tolles-Lawson method is therefore valid when the platform field is much smaller than the measured field. The second approximation follows from a first-order linearization of the square root function.

The above derivations are partly taken from Chapter 3 in [23]. Now it remains to find a good approximation for $\mathbf{B}_p$. The Tolles-Lawson assumes that the interference field is a composition of three fields: the permanent field $\mathbf{B}_{\text{perm}}$, the induced field $\mathbf{B}_{\text{ind}}$ and the eddy current field $\mathbf{B}_{\text{eddy}}$. The interference field is then given by

$$\mathbf{B}_p = \mathbf{B}_{\text{perm}} + \mathbf{B}_{\text{ind}} + \mathbf{B}_{\text{eddy}}. \tag{7.9}$$

The fields represent the magnetic interference of the aircraft with the Earth's magnetic field. The permanent field is the magnetization of the aircraft's structure. Ferromagnetic parts in the aircraft magnetize when they have been in contact with magnets. This field is constant for the aircraft's body. A fixed magnetometer does not measure any change when the aircraft turns. However, a turn changes the orientation of the aircraft to the Earth's magnetic field. The induced field is the field that is aligned with the Earth's magnetic field and is therefore dependent on the orientation of the aircraft. It does not turn with the aircraft. A fixed magnetometer measures a change in the induced field when the aircraft turns. The eddy current field is the field that is induced by the current that flows through the aircraft's structure. This field is dependent on the orientation of the aircraft, as well as the rate of change of the orientation.

Tolles-Lawson is a simple linear method for aeromagnetic compensation. It does not require any knowledge about the roll, pitch and yaw of the aircraft, since these are computed in a relative coordinate system of the vector magnetometer. Besides that, the Tolles-Lawson method is only used for the compensation of total field sensors.

7.2.3. Leliak

The directional cosines used in the Tolles-Lawson method are not very intuitive. Paul Leliak [36] developed a method that uses the movements of the aircraft relative to its body. The directional cosines in his method can thus be replaced by a combination of the roll, pitch and yaw maneuvers of the aircraft. The three maneuvers are visualized in Figure 7.3 and can be described as follows.

- Roll (ψ): the aircraft turns around the transverse / x-axis of the aircraft.
- Pitch (λ): the aircraft turns around the longitudinal / y-axis of the aircraft.
- Yaw (ω): the aircraft turns around the vertical / z-axis of the aircraft.

Besides the three maneuvers, Leliak's method also considers the magnetic heading (θ) and the dip angle (ϕ). The magnetic heading is the angle between the magnetic north and the longitudinal axis of the aircraft. The dip angle is the angle between the magnetic field and the horizontal plane. The dip angle is also known as the inclination angle. If the roll, pitch and yaw are equal to zero, then the directional cosines are completely

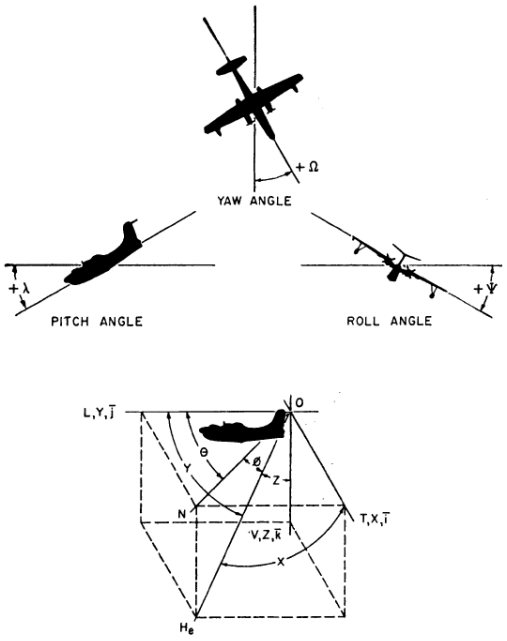


Figure 7.3: The three maneuvers of the aircraft. This illustration is taken from [36].

determined by the magnetic heading and the dip angle:

$$\cos X = \cos\phi \sin\theta \tag{7.10}$$
$$\cos Y = \cos\phi \cos\theta \tag{7.11}$$
$$\cos Z = \sin\phi. \tag{7.12}$$

If the aircraft is not in a level flight, then the directional cosines are also determined by the roll, pitch and yaw. These maneuvers can be described separately for each axis. A roll maneuver around the x-axis gives the following directional cosines:

$$\cos X = \cos\phi \sin\theta \cos\psi + \sin\phi \sin\psi \tag{7.13}$$
$$\cos Y = \cos\phi \cos\theta \tag{7.14}$$
$$\cos Z = \sin\phi \cos\psi - \cos\phi \sin\theta \sin\psi. \tag{7.15}$$

A pitch maneuver around the y-axis gives the following directional cosines:

$$\cos X = \cos\phi \sin\theta \tag{7.16}$$
$$\cos Y = \cos\phi \cos\theta \cos\lambda - \sin\phi \sin\lambda \tag{7.17}$$
$$\cos Z = \sin\phi \cos\lambda + \cos\phi \cos\theta \sin\lambda. \tag{7.18}$$

A yaw maneuver around the z-axis gives the following directional cosines:

$$\cos X = \cos\phi \sin\theta \cos\omega - \sin\phi \sin\omega \tag{7.19}$$
$$\cos Y = \cos\phi \cos\theta \cos\lambda + \sin\phi \sin\lambda \cos\omega \tag{7.20}$$
$$\cos Z = \sin\phi \cos\lambda \cos\omega - \cos\phi \cos\theta \sin\lambda. \tag{7.21}$$

Leliak's method finds similar coefficients as the Tolles-Lawson method. The permanent field is described by 3 coefficients (T, L and V). 5 linear independent coefficients are needed to describe the induced field ($TT-LL$), ($VV-LL$), ($TL+LT$), ($TV+VT$), ($LV+VL$). The eddy current field is described by 8 independent coefficients: ($tt-ll$), ($vv-ll$), lt, vt, tl, vl, tv and lv. Leliak suggests that the coefficients can be found by performing sinusoidal maneuvers.

7.2.4. Compensation using Neural Network

Aeromagnetic magnetic compensation using neural networks has been an active research topic since 1993 [69]. Using neural networks is a logical step in the development of aeromagnetic compensation methods. Neural networks are a powerful tool for pattern recognition and classification. It allows us to model complex relationships between the input and output variables. These relationships can be a lot more complex than the linear relationships that are used in the classical Tolles-Lawson and Leliak methods. This comes at a cost, however. Neural networks are computationally expensive and require a lot of data to train. But neural networks also lose physical interpretability. The weights of the neural network are not directly related to the physical properties of the Earth, the aircraft or magnetic fields. This makes it difficult to explain the results of the neural network and extent the results to other situations, something that is important in the field of aeromagnetic compensation.

Several attempts [69, 39, 24] have been done to use neural networks for aeromagnetic compensation. Gnadt briefly reviews the results of these attempts in [23]. He also noticed that the feature space is very large, resulting in slow training times and overfitting. As a solution, he performed a feature in combination with a Principal Component Analysis (PCA) to reduce the feature space. The results of different neural network architectures are compared for different training and validation sets. A model is found that outperforms Tolles-Lawson and the PLSR-based method. Noteworthy is that the good performance of his method is mainly due to a single sensor ($Mag5$). However, sufficient accuracy is achieved for a combination of different sensors. How the results of different sensors can be combined to improve the accuracy of aeromagnetic compensation is an interesting topic to research. At the moment of writing, there is still an open challenge to improve the accuracy of aeromagnetic compensation using neural networks [23] (Section 7.4).

7.3. Dealing with Multicollinearity

Multicollinearity is a problem in aeromagnetic compensation. The variables that are used to describe the platform field are highly correlated. Both the Tolles-Lawson and Leliak methods already contain steps to reduce the number of coefficients, such that most coefficients are linearly independent. This allowed us to describe the permanent, induced and eddy currents fields in 3, 5 and 8 coefficients respectively, instead of the 3, 9 and 9 coefficients that occur in the original formulas. However, the remaining 16 coefficients can still be highly correlated. This section discusses some methods to deal with multicollinearity.

For all models that assume that the platform field can be approximated by a linear combination of platform features, the compensation algorithm can be formulated as a matrix-vector product. For Tolles-Lawson and Leliak, the features of the platform, in the form of directional cosines, are placed in a matrix $\mathbf{A}$ and the coefficients $\mathbf{C}$ are placed in a vector. The reconstructed platform field is then given by the matrix-vector product:

$$\mathbf{B}_p = \mathbf{AC}. \tag{7.22}$$

A first approach to solve the problem is to invert the matrix $\mathbf{A}$:

$$\mathbf{C} = \mathbf{A}^{-1}\mathbf{B}_p = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T\mathbf{B}_p. \tag{7.23}$$

However, in a multicollinear situation, the matrix $\mathbf{A}$ is not invertible. A matrix is invertible if and only if it has full rank. For a multicollinear problem, some variables are not linearly independent, resulting in a rank smaller than the number of rows. Since this gives a non-unique solution, other methods are needed to solve find the coefficients $\mathbf{C}$.

7.3.1. Ridge Regression

Ridge regression, also known as Tikhonov regularization, regularizes the problem by adding a penalty term to the solution. The penalty term is added to the solution to prevent the coefficients from becoming too large and is a function of the coefficients $\mathbf{C}$. The loss function is defined as the sum of the squared error and the penalty term:

$$\mathscr{L} = \frac{1}{N}\|\mathbf{B}_p - \mathbf{AC}\|^2 + \lambda^2\|\mathbf{C}\|^2, \tag{7.24}$$

where λ is the regularization parameter and N is the number of measurements. Increasing the regularization parameter λ increases the penalty term, which results in a smaller solution $\mathbf{C}$. The accuracy of the solution is therefore dependent on the parameter λ. There are several methods to determine a good value for λ. The analytical way to find the optimal value for λ is to take the derivative of the loss function with respect to the

coefficients $\mathbf{C}$ and set this derivative equal to zero. Another more robust, method is to use cross-validation. We do not go into detail about the methods to find the optimal value for λ here but refer to [70] for more information.

If an optimal λ is available, a regularization term $\lambda^2\mathbf{I}$ can be added to the matrix $\mathbf{A}^T\mathbf{A}$, which acts as a "smoothener". This results in a new matrix $\mathbf{A}^T\mathbf{A}+\lambda^2\mathbf{I}$ which is invertible. The solution for the coefficients $\mathbf{C}$ after adding the regularization term is given by:

$$\mathbf{C} = (\mathbf{A}^T\mathbf{A}+\lambda^2\mathbf{I})^{-1}\mathbf{A}^T\mathbf{B}_p. \tag{7.25}$$

According to [70], the usage of ridge regression, as in Equation 7.25, results in a better solution than the solution found by using the least squares method (Equation 7.23). However, determining the optimal value for the regularization parameter λ is not trivial.

7.4. MIT Challenge

As part of a PhD thesis, [23] used neural networks to compensate for the magnetic field induced by the aircraft. The goal of the work was to remove noise from the sensors and thereby obtain more accurate results in magnetic navigation. The datasets that are used in his research, provided by the United States Air Force pursuant to Cooperative Agreement Number FA8750-19-2-1000, are publicly available. The datasets are also provided as part of an MIT Challenge [22].

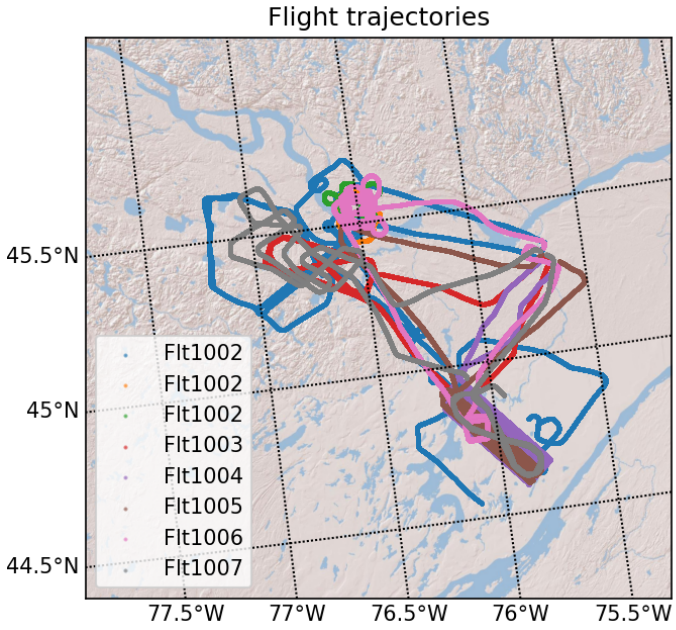


Figure 7.4: Trajectories of different flights in the MIT dataset performed in the region of Ottawa.

The data is provided as a .h5 file. The available data consists of 6 different flights that have been performed in the summer of 2020 in the region of Ottawa, Canada. The trajectories are visualized in Figure 7.4. The flight data is claimed to be unique since it contains a nearly perfect signal from a scalar magnetometer on the tail of a stinger, but also noisy signals from scalar magnetometers (Geometrics G-823A optically pumped, split-beam cesium vapor magnetometers) and vector magnetometers (Bartington Mag-03 and Billingsley TFM100) attached to and inside the aircraft. Figure 7.5 illustrates the position of the sensors on the aircraft.

Besides the measurements of the magnetometers (Figure 7.6), the dataset also contains information about the aircraft's position. The horizontal position is available in latitude and longitude, measured by a GPS receiver and by an INS. Several measurements for altitude are available, including barometric altitude, radar altitude and a digital elevation model. The roll, pitch and yaw angles of the aircraft are computed from the INS. Measurements of currents and voltages are also available.

The goal of the challenge [22] is to compensate the sensor data for the magnetic field that is induced by the platform. To remove the platform's magnetic field and yield a clean magnetic signal, one needs to identify

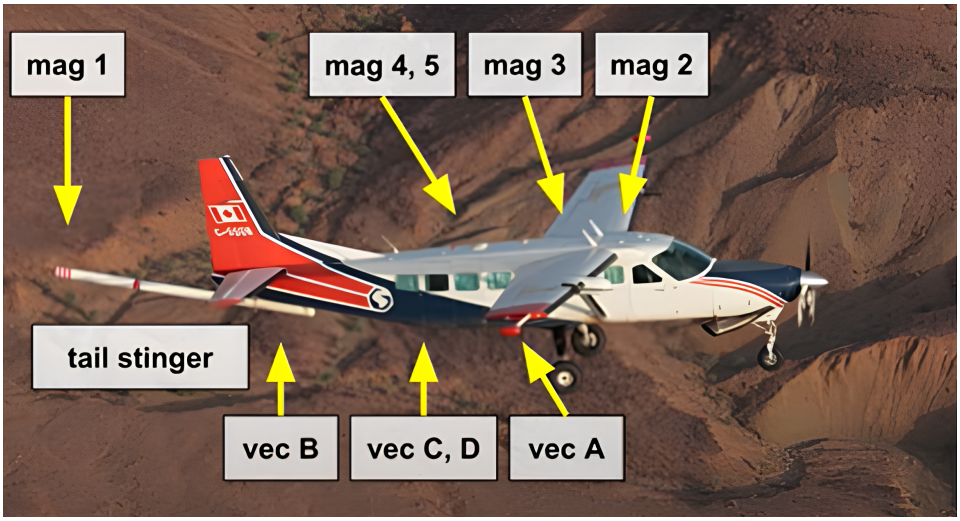


Figure 7.5: The positioning of the magnetometers on the aircraft. 'mag' refers to a Total-Field magnetometer, 'vec' refers to a tri-axial magnetometer.

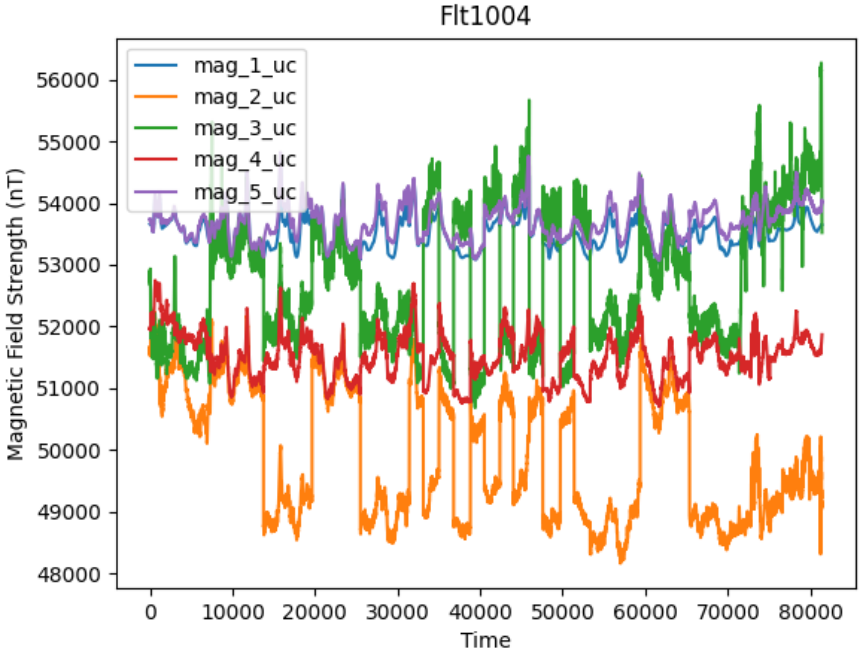


Figure 7.6: Measurements of the 5 scalar magnetometers during flight 1004. Magnetometer 1 (blue) is the uncompensated signal of the sensor on the tail stinger.

the desired truth signal. Two options are presented for the truth signal: the tail stinger or a magnetic anomaly map. The tail stinger has the advantage of being more accurate due to its location, but it is not available on all aircraft. A magnetic anomaly map has the advantage of being the actual desired signal, but it is undersampled, it can be difficult to interpolate and it is not available in unexplored regions. For the challenge, the sensor on the tail of the stinger is used as the true signal.

Compensation for the sensor data can be done with aeromagnetic compensation models. Most compensation models assume that the noise in the sensor data is caused by the magnetization of the platform (the aircraft). The models use measurements of different types of sensors to estimate the magnetic field that is induced by the platform. An example of a compensation model is Tolles-Lawson [67], which uses measurements of the aircraft's position, roll, pitch and yaw angles, and the magnetic field measured by a vector magnetometer. The model assumes that the platform's magnetic field consists of permanent and induced components. The Tolles-Lawson model is quite simple and is not able to compensate for the magnetic field of the aircraft in a very accurate way. More advanced models are available, such as the one by [23] that uses neural networks to compensate for the magnetic field. These models can for example find non-linear relationships between the current and voltage measurements and the magnetic field, which might be useful for electromagnetic interference. Some of these models are discussed in Section 7.2.4.

7.4.1. Example of Tolles-Lawson Compensation

The Tolles-Lawson compensation algorithm can be demonstrated on the dataset provided for this challenge. Flight 1002 in the dataset contains 2 calibration flights: one at the beginning of the survey and one at the end. Both flights are illustrated in Figure 7.2. The corresponding Tolles-Lawson coefficients for these flights can be estimated by performing a least-squares fit on the data obtained during the calibration flights:

$$\mathbf{c} = \arg\min_{\mathbf{c}} \sum_{i=1}^{n} ((\mathbf{B}_t - \mathbf{B}_e) - \mathbf{A}\mathbf{c})^2 \,. \tag{7.26}$$

The resulting coefficients for both calibration flights using data from sensor c are listed in the Tables 7.1 - 7.3. It is assumed that the coefficients of the Tolles-Lawson model are constant over the flight. This gives the possibility to estimate the platform field $\mathbf{B}_p$ at different altitudes and positions. The measured field $\mathbf{B}_t$ can then be corrected for the platform field $\mathbf{B}_p$ by subtracting the estimated platform field from the measured field:

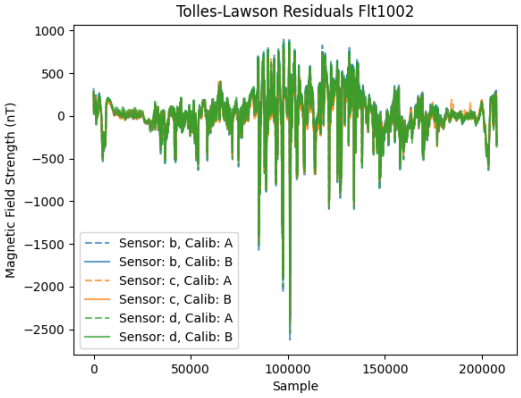
$$\mathbf{B}_t = \mathbf{B}_e + \mathbf{B}_a + \mathbf{B}_p + \mathbf{B}_n \tag{7.27}$$

$$\approx \mathbf{B}_e + \mathbf{B}_a + \mathbf{A}\mathbf{c}, \tag{7.28}$$

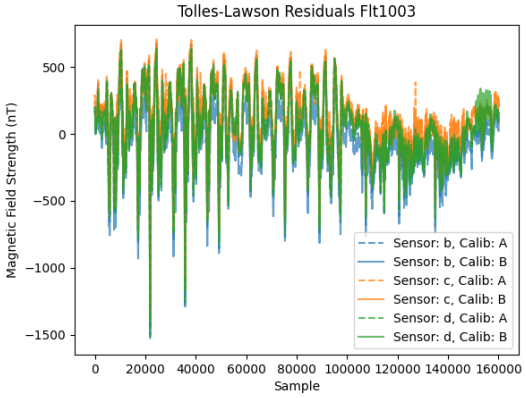
where $\mathbf{B}_a$ is the anomaly field and $\mathbf{A}$ are features depending on the measurements of a vector magnetometer. The anomaly field $\mathbf{B}_a$ can now be estimated by moving all the known terms to the left side of Equation 7.27. The correction is performed on Mag 1, the sensor on the stinger, which is assumed to measure $\mathbf{B}_t$. The coefficients are computed for both calibration flights and for every sensor. The data of sensor A is corrupted in Flt1002, so this sensor cannot be calibrated using Flt1002$_a$ and Flt1002$_b$. The Tolles-Lawson residuals $((\mathbf{B}_t - \mathbf{B}_e) - \mathbf{A}\mathbf{c})$ for different flights are plotted in Figure 7.7. These are the estimates of the anomaly field $\mathbf{B}_a$. Figure 7.7c indicates the compensated signal derived from the uncompensated signal of mag 1, illustrated in Figure 7.6 in blue.

Table 7.1: Permanent Tolles-Lawson coefficients of flight 1002 for sensor c.

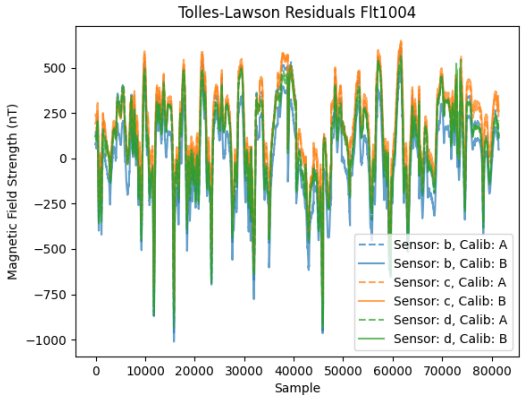
Flight	**Permanent**		
	X	Y	Z
1002 a	940.04	223.59	-1800.33
1002 b	2970.32	54.433	-1312.020



(a) Flt1002



(b) Flt1003



(c) Flt1004

Figure 7.7: Tolles-Lawson residuals for different flights which are obtained by subtracting the estimated platform field and the main field from the measured field. The residuals are an estimate of the anomaly field $\mathbf{B}_a$.

Table 7.2: Induced Tolles-Lawson coefficients of flight 1002 for sensor c.

Flight	**Induced**					
	XX	XY	XZ	YY	YZ	ZZ
1002 a	0.27076	0.17621	0.08189	0.28656	0.057066	0.34032
1002 b	0.19210	0.12166	0.05556	0.194579	0.038983	0.22416

Table 7.3: Eddy Tolles-Lawson coefficients of flight 1002 for sensor c.

Flight	**Eddies**								
	XX'	XY'	XZ'	YX'	YY'	YZ'	ZX'	ZY'	ZZ'
1002 a	0.34032	-0.09455	-0.09443	0.0	0.11034	-0.05957	0.0	-0.024828	0.12921
1002 b	0.22416	-0.070434	-0.068271	0.0	0.072912	-0.041341	0.0	-0.016575	0.084303

Table 7.4: Error in Tolles-Lawson coefficients of flight 1002 for sensor c.

Flight	**Error**		
	Bias	Variance	Total
1002 a	0.0	-0.02482	320.259
1002 b	0.0	-0.016575	592.865

8

Equivalent Layer

We have seen that the spherical harmonic model is a good model for representing Earth's magnetic field on large scales. On these scales, a core-field model or a simple crustal-field model is sufficient to represent the magnetic field. However, it is difficult to use the spherical harmonic model to make higher-resolution representations of the magnetic field in a certain area. This would require a very high degree of the spherical harmonic model, and therefore a lot of data over the surface of the Earth. We noticed that a higher-resolution model can not be made from satellite data alone. Aeromagnetic surveys play an important role in this, but they are time-consuming and have some complications. In the previous chapter, we discussed how to compensate the sensor data of a magnetometer for the platform of an aircraft. In this chapter, we assume that the data has been compensated and we discuss a technique that can be used to locally enhance the resolution of the magnetic field. This technique is called Equivalent Layer.

Equivalent Layer, or Equivalent Sources, is a technique that serves two purposes. It can be used as an interpolator to evaluate the magnetic field at points where no measurements are available and it can be used to extrapolate the measurements to find values of the magnetic field at higher altitudes. Equivalent Layer uses, as many other methods do, the assumption that a magnetic field can be approximated by the one generated by dipoles. In Section 2.5.1 it is discussed that this assumption is sufficient when the points of evaluation are far away from the sources. Given this assumption, Equivalent Layer is quite intuitive: a layer containing a large number of dipoles is placed in the ground below the measurements and the magnetic moments of the dipoles are adjusted to fit the measurements. If the layer fits the measurements well, we can speak of an Equivalent Layer.

Interpolating the magnetic field with Equivalent Layer can be done in three steps:

1. Measure the magnetic field at arbitrary positions in a plane.

2. Compute the magnetic moments of the dipoles in the layer that fit the measurements.

3. Evaluate the magnetic field at points where no measurements are available by computing the superposition of each dipole in the layer.

Using Equivalent Layer as an interpolator for magnetic fields has several advantages compared to using a linear or cubic spline interpolator. The main advantage is that a field generated with Equivalent Layer automatically satisfies the physical properties of a magnetic field since it is a superposition of magnetic dipoles. Another big advantage is that Equivalent Layer imposes no restrictions on how measurement points must be distributed. Quite the contrary, it is not necessary to take the measurements at the same height as the point of interest. Moreover, this allows Equivalent Layer to be used as an extrapolator: measurements at a certain height can be used the estimate the magnetic field at a higher altitude, as long as there exist no sources positioned above the points of measurement this is guaranteed by Green's third identity.

This chapter starts with the derivation of the Equivalent Layer problem (Section 8.1), followed by a section that describes how the problem can be solved when only Total-field measurements are available (Section 8.2). Finally, several methods and tricks are presented that can be used to solve the problem (Section 8.3).

8.1. Derivation of the Equivalent Layer problem

Finding an Equivalent Layer is a memory-intensive process. It often requires performing a minimization problem over all the measurement points and the dipoles in the layer. The problem can be derived from the equation of a dipole. In this section, we start with defining the problem for a single vectorial measurement and a layer consisting of a single dipole and build towards the Equivalent Layer problem that consists of multiple vectorial measurements of the magnetic field and has many dipoles in a layer.

8.1.1. One dipole and a single measurement

Let $\mathbf{B}$ be a magnetic field produced by a single dipole which is positioned at $\mathbf{x}_0$ and has a magnetic moment $\mathbf{m} = [m_x, m_y, m_z]^T$. In what follows, we show that the magnetic moment of an arbitrarily positioned dipole can be tweaked to fit the magnetic field at the point of measurement.

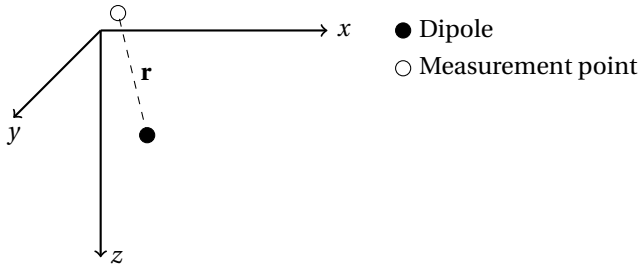


Figure 8.1: Illustration of a single dipole and a single measurement point. The dashed line shows the vector $\mathbf{r}$ that connects the dipole and the measurement point.

For now, fix the position of the dipole to the origin and assume that the measurement is taken at $\mathbf{r} = [x, y, z]^T$. The magnetic field at the point of measurement is given by the vector $\mathbf{B}(\mathbf{r})$ containing the magnetic field components in the x, y and z directions:

$$B_x(\mathbf{r}) = \frac{\mu_0}{4\pi}\left(\frac{3(m_x x + m_y y + m_z z)x}{|\mathbf{r}|^5} - \frac{m_x}{|\mathbf{r}|^3}\right) \tag{8.1}$$

$$B_y(\mathbf{r}) = \frac{\mu_0}{4\pi}\left(\frac{3(m_x x + m_y y + m_z z)y}{|\mathbf{r}|^5} - \frac{m_y}{|\mathbf{r}|^3}\right) \tag{8.2}$$

$$B_z(\mathbf{r}) = \frac{\mu_0}{4\pi}\left(\frac{3(m_x x + m_y y + m_z z)z}{|\mathbf{r}|^5} - \frac{m_z}{|\mathbf{r}|^3}\right). \tag{8.3}$$

These equations are derived from Equation 2.36 by fixing the position of the dipole $\mathbf{x}_0 = [0,0,0]^T$. This reduced the number of unknowns in the system from six (the dipole positions and their moments) to three (the dipole moments). If the dipole is not positioned at the origin, one can translate the coordinate system so that the dipole is positioned at the origin. This assumption does not affect the magnetic field at the point of measurement, since that is only dependent on the relative position of the dipole and the point of measurement. To simplify calculations, we further assume that a measurement is taken at $\mathbf{r} = [0,0,z_0]^T$ where $z_0 > 0$. One can justify this assumption by rotating the coordinate system so that the measurement is taken at the z-axis. This assumption leads to the following relation between the magnetic moments and the magnetic field at the point of measurement:

$$\begin{cases} B_x(\mathbf{r}) &= \frac{\mu_0}{4\pi}\left(-\frac{m_x}{|z_0|^3}\right) \\ B_y(\mathbf{r}) &= \frac{\mu_0}{4\pi}\left(-\frac{m_y}{|z_0|^3}\right) \\ B_z(\mathbf{r}) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_z z_0^2)}{|z_0|^5} - \frac{m_z}{|z_0|^3}\right), \end{cases} \tag{8.4}$$

which can be solved for the magnetic moment:

$$\begin{cases} m_x = -\frac{4\pi}{\mu_0}|z_0|^3 B_x(\mathbf{r}) \\ m_y = -\frac{4\pi}{\mu_0}|z_0|^3 B_y(\mathbf{r}) \\ m_z = \frac{1}{2}\frac{4\pi}{\mu_0}|z_0|^3 B_z(\mathbf{r}). \end{cases} \tag{8.5}$$

The magnetic moment can also be expressed as a matrix-vector product of the depth and the measurement vector:

$$\mathbf{m} = \frac{4\pi}{\mu_0}|z_0|^3 \begin{bmatrix} -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \mathbf{B}(\mathbf{r}), \tag{8.6}$$

for $\mathbf{r} = [0, 0, z_0]^T$. The magnetic moment $\mathbf{m} = (m_x, m_y, m_z)$ that is found is proportional to the magnetic field $\mathbf{B}(\mathbf{r})$ at the point of measurement and to the cube of the distance $|z_0|$ to the dipole. If a rotation was applied to the coordinate system, then the magnetic moment needs to be rotated back to the original coordinate system. This can be done by multiplying the magnetic moment with the inverse of the rotation matrix. Therefore, this example shows that a single dipole can be used to estimate the magnetic field at a single point in space, even if the relative position of the dipole and the point of measurement is fixed. If the relative position of the dipole is fixed, then this estimate has a unique solution. On the other hand, if $\mathbf{r}$ is a free variable, then the system is underdetermined and the solution is not unique.

8.1.2. One dipole and multiple measurements

Having more information about the problem we are solving can help us make the problem less underdetermined. The system for a single dipole and a single measurement has six unknowns and three equations. An additional measurement leads to a system of six equations (for the magnetic field) and nine unknowns: three for the magnetic moments of the dipole ($\mathbf{m} = [m_x, m_y, m_z]^T$) and six for the relative position of the measurement points to the dipole ($\mathbf{r}_1 = [x_1, y_1, z_1]^T$ and $\mathbf{r}_2 = [x_2, y_2, z_2]^T$):

$$\begin{cases} B_x(\mathbf{r}_1) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_1 + m_y y_1 + m_z z_1)x_1}{|\mathbf{r}_1|^5} - \frac{m_x}{|\mathbf{r}_1|^3}\right) \\ B_y(\mathbf{r}_1) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_1 + m_y y_1 + m_z z_1)y_1}{|\mathbf{r}_1|^5} - \frac{m_y}{|\mathbf{r}_1|^3}\right) \\ B_z(\mathbf{r}_1) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_1 + m_y y_1 + m_z z_1)z_1}{|\mathbf{r}_1|^5} - \frac{m_z}{|\mathbf{r}_1|^3}\right) \\ B_x(\mathbf{r}_2) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_2 + m_y y_2 + m_z z_2)x_2}{|\mathbf{r}_2|^5} - \frac{m_x}{|\mathbf{r}_2|^3}\right) \\ B_y(\mathbf{r}_2) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_2 + m_y y_2 + m_z z_2)y_2}{|\mathbf{r}_2|^5} - \frac{m_y}{|\mathbf{r}_2|^3}\right) \\ B_z(\mathbf{r}_2) &= \frac{\mu_0}{4\pi}\left(\frac{3(m_x x_2 + m_y y_2 + m_z z_2)z_2}{|\mathbf{r}_2|^5} - \frac{m_z}{|\mathbf{r}_2|^3}\right). \end{cases} \tag{8.7}$$

Increasing the number of measurements leads to an overdetermined system. In general, overdetermined systems are not solvable. For example, if multiple measurements are taken of a magnetic field that is produced by a single dipole, and the dipole position is not known, then the overdetermined system only has a solution if the dipole that estimates the field is positioned exactly at the place where the dipole that generated the field is located. If only a slight noise is present in the measurements, then the system does not have a solution.

The inversion of the magnetic field is therefore not possible if the dipole position is not known. What is possible is minimizing the error between the measured magnetic field and the estimated magnetic field. This can be done by for example minimizing the sum of the squared errors between the measured and estimated magnetic field:

$$\underset{\mathbf{m}\in\mathbb{R}^3}{\text{minimize}} \sum_{i=1}^{n} |\mathbf{B}(\mathbf{r}_i) - \mathbf{B}_{\text{est}}(\mathbf{r}_i, \mathbf{m})|^2, \tag{8.8}$$

where n is the number of measurements, $\mathbf{r}_i$ is the position of the ith measurement, $\mathbf{B}(\mathbf{r}_i)$ is the measured magnetic field at the ith measurement point and $\mathbf{B}_{\text{est}}(\mathbf{r}_i, \mathbf{m})$ is the estimated magnetic field at the ith measurement point using a dipole at the origin with magnetic moment $\mathbf{m}$.

8.1.3. Multiple dipoles and a single measurement

In the previous section, it was shown that an arbitrarily positioned dipole can be fitted to a single measurement of the magnetic field. Adding more dipoles to the system makes the problem even more underdetermined. If a single dipole can already fit a single measurement, then multiple dipoles can of course also fit a single measurement. One can simply set the magnetic moments of the other dipoles to the zero vector.

8.1.4. Multiple dipoles and multiple measurements

Using many measurements for a few dipoles leads to an overdetermined system while fitting multiple dipoles to a single measurement leads to an underdetermined system. Therefore, a proper trade-off between the number of measurements and the number of dipoles to estimate the field must be made. A layer of multiple dipoles becomes interesting when many measurements are taken.

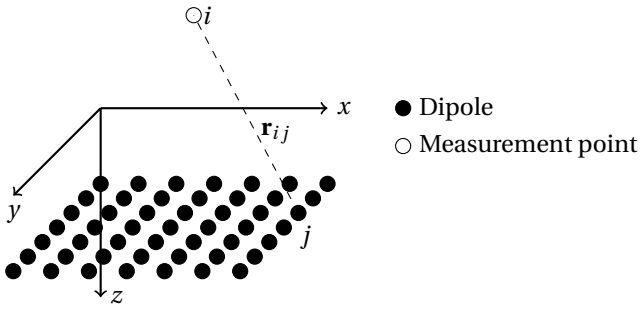


Figure 8.2: Illustration of a layer of dipoles and the j-th measurement point. The dashed line shows the vector $\mathbf{r}_{ij}$ that connects the j-th dipole to measurement point i.

A layer of dipoles can be constructed by for example placing dipoles on a regular grid in the xy-plane at a certain depth z, as illustrated in Figure 8.2. The estimated magnetic field at a measurement point $\mathbf{r}_i$ is then the superposition of the magnetic field of all dipoles in the layer:

$$\mathbf{B}_{\text{est}}(\mathbf{r}_i) = \sum_{j=1}^{m} \mathbf{B}_j(\mathbf{r}_{ij}, \mathbf{m}_j) = \sum_{j=1}^{m} A(\mathbf{r}_{ij})\mathbf{m}_j\,, \tag{8.9}$$

where m is the number of dipoles in the layer, $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{x}_j$ is the distance vector from the ith dipole to the measurement point $\mathbf{r}_i$ and $\mathbf{B}_j(\mathbf{r}_{ij}, \mathbf{m}_j)$ is the magnetic field produced by the jth dipole at the measurement point $\mathbf{r}_i$, this quantity can also be represented as a matrix-vector product $A(\mathbf{r}_{ij})\mathbf{m}_j$. We can write out the expression for $\mathbf{B}_j(\mathbf{r}_{ij}, \mathbf{m}_j)$ using Equation 2.36:

$$\mathbf{B}_j(\mathbf{r}_{ij}, \mathbf{m}_j) = \frac{\mu_0}{4\pi} \frac{(3(\mathbf{m}_j \cdot \mathbf{r}_{ij})\mathbf{r}_{ij} - \mathbf{m}_j \mathbf{r}_{ij}^2)}{|\mathbf{r}_{ij}|^5}\,. \tag{8.10}$$

We now introduce a 3×3 matrix $A(\mathbf{r}_{ij})$ to represent the influence of a dipole grid on the magnetic field at a measurement point at distance $\mathbf{r}_{ij}$:

$$A(\mathbf{r}_{ij}) = \frac{\mu_0}{4\pi} \frac{1}{|\mathbf{r}_{ij}|^3} \left[R(\mathbf{r}_{ij}) - I_3\right] \tag{8.11}$$

$$R(\mathbf{r}_{ij}) = 3\frac{\mathbf{r}_{ij}\mathbf{r}_{ij}^T}{|\mathbf{r}_{ij}|^2}\,, \tag{8.12}$$

which leads to the following linear form:

$$\mathbf{B}_j(\mathbf{r}_{ij}, \mathbf{m}_j) = A(\mathbf{r}_{ij})\mathbf{m}_j\,. \tag{8.13}$$

In many implementations of the Equivalent Layer algorithm, the positions of the dipoles are fixed. Therefore, the matrix $A(\mathbf{r}_{ij})$ can be precomputed. In general, for n measurement points and m dipoles, the estimated field can be written as a matrix-vector product:

$$\mathbf{B}_{\text{est}}(\mathbf{r}_i) = \mathbf{A}(\mathbf{r}_i)\mathbf{m}\,, \tag{8.14}$$

where $\mathbf{B}_{\text{est}}(\mathbf{r}_i)$ is a vector containing the estimated magnetic field at point i, $\mathbf{A}(\mathbf{r}_i)$ is a $3 \times 3m$ matrix and $\mathbf{m}$ is a vector of length $3m$ containing the magnetic moments of all dipoles in the layer. For a layer of dipoles with fixed positions $\mathbf{x}_j$, the magnetic moments $\mathbf{m}_j$ can be found using a least squares fit:

$$\underset{\mathbf{m}\in\mathbb{R}^{3\times m}}{\text{minimize}} \sum_{i=1}^{n} |\mathbf{B}(\mathbf{r}_i) - \mathbf{A}(\mathbf{r}_i)\mathbf{m}|^2\,. \tag{8.15}$$

Here, $\mathbf{m} = [\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_m]$ is a matrix with the magnetic moments of all dipoles in the layer.

8.2. The Equivalent Layer Problem for Total-Field Magnetic Anomalies

Before we discuss how the Equivalent Layer problem can be solved, we first need to discuss the type of magnetic field measurements that are used. The minimization problem, as defined in Equation 8.15, assumes that we have access to vectorial measurements $\mathbf{B}(\mathbf{r}_i)$ of the magnetic field. These are however not always available. Often a Total-Field Sensor is used that measures the intensity of the magnetic field. It can also be possible that a vectorial magnetometer is used, but that the measurements are too noisy or that the orientation of the magnetometer is not known. Since Total-Field sensors were in general more accurate than vectorial magnetometers, most of the work on Equivalent Layer has been done for Total-Field sensors. In this section, the Equivalent Layer problem for Total-Field sensors is described.

The minimization problem in Equation 8.15 can be adapted to include Total-Field measurements. The cost function in the minimization problem becomes the sum of the squared difference between the Total-Field measurements and the Total-Field produced by the dipoles:

$$\underset{\mathbf{m}\in\mathbb{R}^{3\times m}}{\text{minimize}} \sum_{i=1}^{n} \left[|\mathbf{B}(\mathbf{r}_i)| - |\sum_{j=1}^{m} \mathbf{B}_j(\mathbf{r}_i - \mathbf{x}_j, \mathbf{m}_j)| \right]^2, \tag{8.16}$$

where $|\mathbf{B}(\mathbf{r}_i)|$ is the Total-Field measurement at the ith measurement point $\mathbf{r}_i$ and $|\sum_{j=1}^{m} \mathbf{B}_j(\mathbf{r}_i - \mathbf{x}_j, \mathbf{m}_j)|$ is the Total-Field produced by the dipoles at the ith measurement point $\mathbf{r}_i$. One might argue that if we are only interested in the Total-Field produced by the dipoles, we do not need to know the orientation of the moment $\mathbf{m}_j$. However, the orientation of the moment is still required to calculate the Total-Field produced by the dipoles, knowing only the magnitude is not sufficient, since dipoles with for example opposite oriented moments can cancel each other out on the Total-Field and the intensity of a single dipole is not constant on the surface of a sphere, for example on the surface of the Earth, the total-field intensity at the poles is two to three times larger than at the equator (Section 3.2).

The computation of the norm in the second term of Equation 8.16 is a non-linear operation. Therefore, the minimization problem is non-linear and significantly more difficult to solve than the minimization problem in Equation 8.15. The problem can be reduced to a linear system by projecting the Total-Field measurements onto a reference field, as described in Equation 5.3 in Section 5.4. We assume that we know a reference field $\mathbf{B}_{\text{ref}}(\mathbf{r}_i)$, which satisfies the following conditions:

$$\mathbf{B} = \mathbf{B}_{\text{ref}} + \mathbf{B}_{\text{anomaly}} \tag{8.17}$$

$$|\mathbf{B}_{\text{ref}}| >> |\mathbf{B}_{\text{anomaly}}|. \tag{8.18}$$

To reduce the number of unknowns, the Total-Field measurements are often projected onto a reference field, as described in Section 5.4. The reference field can for example be chosen as the main field (Section 3.2) or IGRF (Section 6.1.2). As long as the reference field is sufficiently larger than the anomaly field, the projection is justified, and we can reduce the minimization problem to:

$$\underset{m\in\mathbb{R}^{1\times m}}{\text{minimize}} \sum_{i=1}^{n} \left[|\mathbf{B}(\mathbf{r}_i)| \cdot \hat{\mathbf{B}}_{\text{ref}}(\mathbf{r}_i) - \sum_{j=1}^{m} \mathbf{B}_j(\mathbf{r}_i - \mathbf{x}_j, \hat{\mathbf{m}}(\mathbf{r}_i) m_j) \right]^2, \tag{8.19}$$

where $\hat{\mathbf{B}}_{\text{ref}}(\mathbf{r}_i)$ is the unit vector of the reference field at the ith measurement point $\mathbf{r}_i$ and $\hat{\mathbf{m}}(\mathbf{r}_i)$ is the unit vector of the magnetic moment that corresponds with $\hat{\mathbf{B}}_{\text{ref}}(\mathbf{r}_i)$. The vector $m \in \mathbb{R}^{1\times m}$ contains the magnitudes of the magnetic moments of the dipoles. In many cases, the reference field is constant on the region of interest, resulting in the following simplification:

$$\underset{m\in\mathbb{R}^{1\times m}}{\text{minimize}} \sum_{i=1}^{n} \left[|\mathbf{B}(\mathbf{r}_i)| \cdot \hat{\mathbf{B}}_{\text{ref}} - \sum_{j=1}^{m} \mathbf{B}_j(\mathbf{r}_i - \mathbf{x}_j, \hat{\mathbf{m}} m_j) \right]^2. \tag{8.20}$$

This system is again linear but has a factor of three fewer unknowns than the system in Equation 8.15. The method of solving the system is now similar to solving the system in Equation 8.15 without the need of finding the magnetic moments.

8.3. Methods and Tricks for Solving the Equivalent Layer Problem

We derived the Equivalent Layer problem (Section 8.1) and found that the Equivalent Layer problem for Total-field measurements can also be reduced to a linear system (Equation 8.20). Now, we look at the formulated minimization problem (Equation 8.15), which can be solved by a variety of methods. In many cases, a QR-decomposition is sufficient to find a solution that minimizes the least squares error. The QR-decomposition might fail, for example when the number of dipoles is larger than the number of measurement points, or when the measurement points are not sufficiently distributed. In these cases, different methods and tricks are required to solve the Equivalent Layer problem.

The Equivalent Layer problem, as described in Section 8.1 is a compute-intensive problem. The dimensions of the matrix scale with the number of measurement points n and the number of dipoles m. Several studies have been conducted to improve the performance of the Equivalent Layer algorithm[48][37]. In this section, we summarize the most important suggestions and improvements for solving the Equivalent Layer problem.

8.3.1. Best Approximation

We first look at the best approximation of the linear system. Note that the influence of a single dipole positioned at $\mathbf{x}_0$ on a point $\mathbf{x} = \mathbf{x}_0 + \mathbf{r}$ can be written as:

$$B(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{3(\mathbf{r}\cdot\mathbf{m})\mathbf{r} - \mathbf{m}|\mathbf{r}|^2}{|\mathbf{r}|^5}. \tag{8.21}$$

This can be written as the inproduct of a matrix $A(\mathbf{r})$ that is independent of the moments $\mathbf{m}$ and the vector containing the moments $\mathbf{m}$:

$$\mathbf{B}(\mathbf{r}) = A(\mathbf{r})\mathbf{m}, \tag{8.22}$$

where $A(\mathbf{r})$ is a 3x3 matrix that can be written as:

$$A(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{1}{|\mathbf{r}|^3} \left(\frac{3\mathbf{r}\mathbf{r}^T}{|\mathbf{r}|^2} - I_3 \right), \tag{8.23}$$

and $\mathbf{m} = [m_x, m_y, m_z]^T$.

Now, the best approximation of a single measurement $B(\mathbf{r})$ by a single dipole is given by the least squares solution of the matrix-vector product. This can be written as:

$$\mathbf{m} = (A^T A)^{-1} A^T \mathbf{B}. \tag{8.24}$$

This can be extended to a layer with multiple dipoles and multiple measurements. Consider a layer of m dipoles with relative positions $\mathbf{r}_{1,k}, \mathbf{r}_{2,k}, \ldots, \mathbf{r}_{m,k}$ to measurement point k and magnetic moments $\mathbf{m}_1, \mathbf{m}_2, \ldots, \mathbf{m}_m$, as illustrated in Figure 8.2. The magnetic field at the k-th point is the superposition of the magnetic fields of all the dipoles:

$$\mathbf{B}_k = \sum_{i=1}^{m} A(\mathbf{r}_{i,k})\mathbf{m}_i. \tag{8.25}$$

8.3.2. Preconditioning

It is also possible to solve the Equivalent Layer problem with an iterative method. For example, the Jacobi or Gaus-Seidel method can be used to solve the system. We noticed that the convergence of these iterative methods is very slow. The cause of this slow convergence seems to be the high condition number of the matrix. The condition number of a matrix is a measure of the sensitivity of the solution of a system of linear equations to perturbations in the matrix. Lower condition numbers result in faster convergence of the iterative methods. Not only for iterative methods but also for direct methods, a lower condition number results in a more stable solution.

A helpful trick to improve the convergence of the solution is to pre-condition the system. The magnetic moments $\mathbf{m}$ that we are looking looking for can be very large numbers, especially when working with deeper positioned dipoles in the Equivalent Layer. The reason for this is the decay of the magnetic field strength that is proportional to $1/r^3$, where r is the distance between the dipole and the measurement point: dipoles positioned two times deeper have to be eight times stronger to have the same influence on the magnetic field at the measurement point. As a result, we know that we are not looking for deep-positioned dipoles with a small magnetic moment or shallow-positioned dipoles with a large magnetic moment. We can use this

knowledge to pre-condition the system. A pre-conditioner transforms a system $Ax = b$ into an equivalent system that is easier to solve. The equivalent system

$$M^{-1}Ax = M^{-1}b. \tag{8.26}$$

Here, M is a pre-conditioner. The pre-conditioner M is chosen such that the condition number of $M^{-1}A$ is smaller than the condition number of A.

We can pre-condition the Equivalent Layer system by scaling the magnetic moments $\mathbf{m}$ with a factor d_i^3, where d_i is the depth of the i-th dipole. This way, the magnetic moments are scaled to the same order of magnitude. Besides the depth-dependent scaling, we also scale the magnetic moments with a constant factor $c = 10^7/2$ to reduce the magnitude of the solution vector x. This factor comes from the $\mu_0/4\pi$ in Equation 8.23. Therefore, the pre-conditioner M is a diagonal matrix with diagonal elements $1/(d^3c)$:

$$M = \begin{bmatrix} \frac{1}{d_1^3 c} & & \\ & \ddots & \\ & & \frac{1}{d_n^3 c} \end{bmatrix}. \tag{8.27}$$

When measurements are taken at a constant altitude z_0 above the surface, and the dipoles are positioned at a constant depth z_d below the surface, the pre-conditioner can be simplified to:

$$M = \frac{2}{(z_0 + z_d)^3 10^7}. \tag{8.28}$$

8.3.3. Dense Equivalent Layer

To see why the Equivalent Layer method works, we go back to Green's third identity, discussed in Section 2.4.1. It followed that the magnetic scalar potential at a point inside a closed surface can be calculated as an integral of the magnetic scalar potential on the boundary of the surface. The assumption that all the sources are located below a plane allowed us to represent the magnetic scalar potential above the plane as an integral over the plane. Besides that, it was shown that a similar integral can be used to calculate the magnetic field as well.

A continuous magnetic induction field $\mathbf{B}$ on a plane Ω can therefore be approximated by an infinite sum of magnetic dipoles that are infinitesimally close to each other and placed directly under the plane. Let $\Omega = [0,1] \times [0,1]$ be a square plane positioned at depth z_0, then for a given point $(x, y, z_0) \in \Omega$, the magnetic field can be approximated as follows:

$$\mathbf{B}(x,y) = \lim_{n\to\infty} \sum_{i,j=1}^{n} A([x - x_i, y - y_j, z_0 + \frac{1}{n}])M_{i,j}, \tag{8.29}$$

where $x_i = \frac{i}{n}$ and $y_j = \frac{j}{n}$ are the coordinates of the i-th and j-th dipole, respectively and $M_{i,j}$ is the magnetic moment of the dipole at position $(x_i, y_j, z_0 + \frac{1}{n})$. The reason that the dipoles are not placed directly on the plane is to avoid the singularity at the plane.

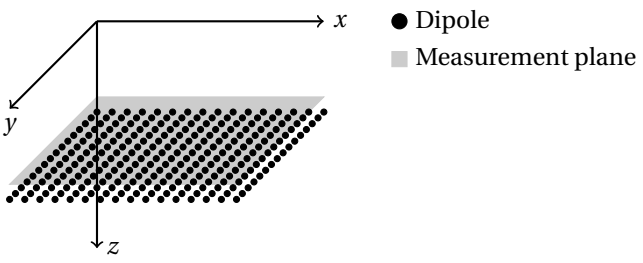


Figure 8.3: Illustration of a dense layer of dipoles and a measurement plane. The plane is represented by a transparent rectangle.

In the limit, the Equivalent Layer problem can be considered as a continuous layer of point masses. Relaxing the assumption that the dipoles are infinitesimally close to each other, we can approximate the continuous layer of point masses by a dense layer of dipoles. The magnetic field at a point (x, y, z_0) above the plane can then be approximated by the magnetic field of the dense layer of dipoles.

8.3.4. Polynomial Equivalent Layer

In [48], Vanderlei C. Oliveira found that the Equivalent Layer algorithm can be improved by using a piecewise polynomial layer instead of a layer consisting of dipoles, a so-called Polynomial Equivalent Layer (PEL). The piecewise polynomial layer is constructed by dividing the layer into regular grid cells and approximating the potential field in each cell with a bivariate polynomial (Figure 8.4). Here, the assumption is made that the distribution of the potential field in the Equivalent Layer can be approximated by a piecewise-polynomial function. The polynomial coefficients can be found by solving a linear regularized inversion problem for each cell separately. After the coefficients have been found, a piecewise-polynomial Equivalent Layer can be constructed by combining the polynomials of each cell. Finally, evaluation of the Total-Field at the point of interest is done by integrating over the piecewise-polynomial layer using Green's theorem (Section 2.4.1) and Upward Continuation (Section 10.1).

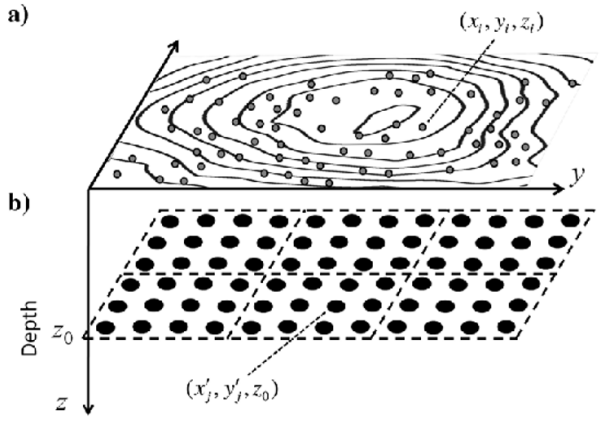


Figure 8.4: Illustration of the Polynomial Equivalent Layer. Instead of estimating the moments of the dipoles (black dots), the Polynomial Equivalent Layer fits a polynomial function to the magnetic potential in each cell (dashed rectangles). This figure is taken from [48]

A main benefit of the Polynomial Equivalent Layer is the minimum depth of the layer. Since a classical Equivalent Layer consists of dipoles, which are point sources, the minimum depth of the layer is dependent on the spatial resolution of the measurement data. In contrast, the Polynomial Equivalent Layer divides the magnetic potential more evenly into the cells using a piece-wise polynomial function. The accuracy of PEL remains dependent on the width of the cells and the degree of the polynomials that are used.

8.3.5. Dual Layer Equivalent Sources

Another significant improvement of the Equivalent Layer algorithm is the use of Dual-Layer Equivalent Sources [37]. The Dual-Layer Equivalent Sources (DLES) approach uses two layers of dipole sources instead of one. The spatial frequency that a layer of dipoles is able to reconstruct is dependent on the depth of the layer and the distance between the dipoles. A low frequent magnetic field can therefore be reconstructed by a layer close to the surface consisting of a large number of dipoles, or by a layer at a greater depth consisting of a smaller number of dipoles. The DLES approach uses two layers of dipoles, one close to the surface and one at a greater depth. The first layer is used to reconstruct the low-frequency components of the magnetic field and the second layer is used to reconstruct the high-frequency components.

8.3.6. Multi-layer Method

In Section 8.3.3 we showed that a continuous magnetic field on a plane can be approximated by an infinite sum of magnetic dipoles. In this section, we discretize the dense layer of dipoles into multiple layers. The idea is quite similar to Fourier series, where a continuous function can be approximated by an infinite sum of sinusoids. Now, the magnetic field on a plane can be approximated by a series of layers of dipoles.

This principle is visualized in Figure 8.5. The figure contains 3 layers. Deeper layers contain fewer dipoles, as they represent lower spatial frequencies. When solving the inverse problem, the deepest layer is solved first. The magnetic field of the deepest layer is then subtracted from the measured magnetic field. The next layer is then solved using the residual magnetic field. This process is repeated until all layers have been solved or until the residual magnetic field is sufficiently small. The advantage of this method is that the deepest layer can be solved using a small number of dipoles, which reduces the computational cost, we refer to this principle as layerwise solving.

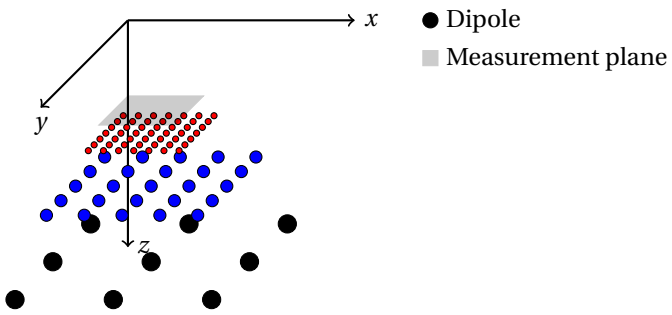


Figure 8.5: Illustration of a dense layer of dipoles and a measurement plane. The plane is represented by a transparent rectangle.

8.3.7. Regularized Equivalent Layer

In the above example about multi-layer methods, we only considered dipoles positioned in layers at fixed depths. However, the dipoles can also be positioned at arbitrary depths. In this case, it is still suggested to place more dipoles closer to the surface, as the spatial resolution of the magnetic field that has to be represented by these dipoles is higher. Placing dipoles at arbitrary depths has however a computational disadvantage: we can no longer solve the problem layer-for-layer. Instead, we need to find a way to drive the solver to a solution that prefers deeper dipoles over shallower dipoles. This can be done by adding a regularization term to the objective function. Regularization is a technique that is used to prevent overfitting of the model. In our case, overfitting means that we mainly use shallow dipoles that are close to the measurement plane to represent the magnetic field. This is not desirable, as it requires often a larger number of dipoles to represent the magnetic field, and it is also not physically correct.

Instead of solving the system $A\mathbf{m} = \mathbf{b}$, we solve the system $(A^T A + \lambda^2 I)\mathbf{m} = A^T \mathbf{b}$, where λ^2 is a regularization parameter. The regularization parameter λ^2 can be tuned by minimizing the L-curve, or by using cross-validation to find the optimal distribution of dipoles with minimal total magnetic moments. This works when we are working with a preconditioned system, which already scales the magnetic moments of the dipoles with their depth (Section 8.3.2) to give them the same order of magnitude.

9

A Wavelet Approach to Position Estimation and Magnetic Field Reconstruction

In the previous chapter, we have seen that Equivalent Layer can be used for modelling, interpolating and extrapolating magnetic fields. The accuracy of the methods depends on the number of sources in the layers and their positions. In this chapter, we try to use the measured signal to determine the positions of the sources. We introduce the so-called Anderson functions, extend these functions to two dimensions and derive a mother wavelet that can be used to construct a wavelet basis for the magnetic field. The goal is to find the positions and magnetic moments of the sources without a priori knowledge of the number of sources and their positions or depth.

In the coming sections, we first limit ourselves to signals that are one-dimensional in space. These signals are obtained by moving a sensor with a constant velocity in a straight line. The magnetometer outputs values as a function of time and the constant velocity gives the magnetic field as a function of position on the line. These so-called spatial signals can be analyzed using Anderson functions to obtain the positions of underlying magnetic sources. A wavelets family consisting of different shapes and shifts of the Anderson functions can be used to look for these sources.

Later on, we extend the theory to two-dimensional signals. These signals represent the magnetic field as a function of position in a plane. We derive the Anderson functions for this two-dimensional case and use them to construct a wavelet basis. This wavelet family is used to reconstruct the signal on the two-dimensional plane.

9.1. Wavelet Theory

One might be familiar with the concept of frequency in the context of time signals, for example, the frequency of a sound wave. Low frequencies, sounding like deep bass, are associated with long wavelengths. High frequencies, sounding like high-pitched whistles, are associated with short wavelengths. The same concept can be applied to spatial signals. Instead of waves that propagate in time, we have waves that propagate in space [8, pp. 4–6]. The spatial frequency indicates the number of waves that pass a certain point in space per unit of time.

9.1.1. Fourier Transform

Since the one-dimensional spatial signals are just time signals divided by the velocity, we can make use of the same signal processing techniques. A common technique is the Fourier transform, which decomposes a signal into a sum of sinusoids with different frequencies. The resulting sinusoids called the Fourier components, are the coefficients for periodic functions with different periods frequencies. In continuous time, the transformation into Fourier components gives all information about the frequency of the signal, at the cost of losing information about the time.

The Fourier Transform is not so useful for our purpose, because it is a global transformation that gives no information about the local properties of the signal. For example, if we are analyzing the signal of a single dipole source, the Fourier transform gives us the frequency of the signal, which can be related to the depth of the source. However, it does not provide any information about the spatial position of the source, or in terms

of time signals, the time at which the source is located. A short-time Fourier transform (STFT) can be used to obtain some local information, but this requires a fixed window size, which is not always desirable. The STFT is also primarily suited for analyzing stationary signals with a fixed frequency content. It does not inherently provide a multiscale analysis of the signal, which is helpful in the analysis of magnetic fields with sources at unknown depths.

9.1.2. Wavelet Transform

In contrast to the Fourier transforms the wavelet transform is a more local transformation. It uses functions with compact support instead of periodic functions to analyze the signal. These functions are often derived from a single function, called the mother wavelet. The mother wavelet is scaled and shifted to obtain a basis which we refer to as the daughter wavelet functions. To ensure that the daughter wavelet functions form an orthogonal basis, the mother wavelet must satisfy some requirements to be a valid wavelet. The requirements for a mother wavelet [6, p. 61] are the following:

- The wavelet must be absolutely integrable

$$\int_{-\infty}^{\infty} |\psi(x)| dx < \infty. \tag{9.1}$$

- The wavelet must be square integrable

$$\int_{-\infty}^{\infty} \psi(x)^2 dx < \infty. \tag{9.2}$$

- The wavelet must be mean-zero (Admissibility)

$$\int_{-\infty}^{\infty} \psi(x) dx = 0. \tag{9.3}$$

It is preferred that the mother wavelet is continuously differentiable and that the wavelet is compactly supported.

The compact support makes it easier to orthogonalize the wavelet upon shifting: if we compare a wavelet with the same wavelet shifted by its support, then the integral is zero because the wavelet is zero outside of its support. The support is dependent on the scale of the wavelet. A wavelet that represents a deeper source has larger support than a wavelet that represents a shallower source. In a signal, one can fit more wavelets with small support than with large support, therefore, wavelets with smaller support can give more information about the local properties of the signal.

9.1.3. Orthogonality of Wavelets

Let $\psi(x)$ be a mother wavelet with compact support $[-s,s]$ and let $\psi_{a,b}(x)$ represent the child wavelet with scaling parameter $a \in \mathbb{R}_+$ and shift parameter $b \in \mathbb{R}$. Then the wavelet transform of a function $f(x)$ is defined as

$$\mathcal{W} f(a,b) = \langle f, \psi_{a,b} \rangle = \int_{-\infty}^{\infty} f(x) \psi_{a,b}(x) dx. \tag{9.4}$$

If the wavelets $\psi_{a,b}$ form an orthonormal basis for the space where f is defined, then the wavelet transform gives a unique representation of f in terms of the wavelet coefficients $\mathcal{W} f(a,b)$. This is highly favorable for computations and reconstructions. When wavelets are derived from a mother wavelet, there is no guarantee that the wavelets are orthonormal. Some restrictions on the shift and scaling parameters of the wavelets are necessary to ensure that the wavelets are orthonormal.

We obviously have that a shifted wavelet $\psi(x-2s)$ is orthonormal to the mother wavelet $\psi(x)$, because the shifted wavelet has compact support at $[s,3s]$ and therefore

$$\int_{-\infty}^{\infty} \psi(x-2s)\psi(x) dx = \int_{-s}^{s} \psi(x-2s)\psi(x) dx + \int_{s}^{3s} \psi(x-2s)\psi(x) dx = 0. \tag{9.5}$$

With the scaling operator, it is less intuitive to generate orthonormal functions. For the Haar wavelet function

$$\psi(x) = \begin{cases} 1 & \text{if } 0 \le x < \frac{1}{2} \\ -1 & \text{if } \frac{1}{2} \le x < 1 \\ 0 & \text{otherwise}, \end{cases} \tag{9.6}$$

one can show that the scaling operator is orthogonal if we scale a with at least a factor of 2. This is because of the shape of the Haar wavelet, which is a piecewise constant function. In contrast, the Anderson wavelets we try to construct do not have this property. Therefore, Anderson wavelets require orthogonalization of the wavelets for each scaling parameter a. This is done by using the Gram-Schmidt process, which is discussed in Section 9.3.1.

When moving from continuous signals to discrete signals, the parameters a and b are replaced by integers $a \in \mathbb{Z}_+$ and $b \in \mathbb{Z}$. A general way to construct discrete wavelets from a mother wavelet is to use the affine group for the shift parameter and the scaling parameter. The resulting bases are:

$$\psi_{a,b}(x) = \frac{1}{\sqrt{2^a}} \psi\left(\frac{x - 2sb}{2^a}\right). \tag{9.7}$$

9.2. Anderson Functions

Anderson functions are a set of basis functions that can be used to model a one-dimensional signal that is obtained by moving in a straight line near a dipole source. The shape of the function is determined by the velocity between the source and the sensor, the closest distance between the source and the sensor (CPA) and the direction of projection.

9.2.1. Derivation of the 1D Anderson Functions

The Anderson functions can be derived directly from the magnetic dipole equation by evaluating them on a line.

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{m} \cdot \mathbf{r})\mathbf{r}}{|\mathbf{r}|^5} - \frac{\mathbf{m}}{|\mathbf{r}|^3} \right), \tag{9.8}$$

where $\mathbf{m}$ is the dipole moment and $\mathbf{r}$ is the distance vector between the dipole and the sensor.

Introduce the trajectory of the sensor as

$$\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t\,, \tag{9.9}$$

where $\mathbf{r}_0$ is the position of the sensor at time $t = 0$ and $\mathbf{v}$ is the velocity of the sensor. To simplify the derivation, we assume that $\mathbf{r}_0$ is the closest point of approach (CPA) between the sensor and the dipole so that the sensor is closest to the dipole at time $t = 0$. In this example, it is also assumed that the sensor is moving and the dipole is fixed, but this is analogous to the case where the sensor is fixed and the dipole is moving.

If we substitute this into the magnetic dipole equation, we get

$$\mathbf{B}(t) = \frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{m} \cdot (\mathbf{r}_0 + \mathbf{v}t))(\mathbf{r}_0 + \mathbf{v}t)}{|\mathbf{r}_0 + \mathbf{v}t|^5} - \frac{\mathbf{m}}{|\mathbf{r}_0 + \mathbf{v}t|^3} \right) \tag{9.10}$$

$$= \frac{\mu_0}{4\pi} \left(\frac{3(\mathbf{m} \cdot \mathbf{r}_0 + \mathbf{m} \cdot \mathbf{v}t)(\mathbf{r}_0 + \mathbf{v}t)}{|\mathbf{r}_0 + \mathbf{v}t|^5} - \frac{\mathbf{m}}{|\mathbf{r}_0 + \mathbf{v}t|^3} \right), \tag{9.11}$$

which can be simplified using a dimensionless quantity $\theta = \frac{|\mathbf{v}|t}{r_0}$ and other substitutions. We define $r_0 = |\mathbf{r}_0|$, $\hat{\mathbf{r}}_0 = \frac{\mathbf{r}_0}{r_0}$, $\hat{\mathbf{v}} = \frac{\mathbf{v}}{|\mathbf{v}|}$ and $\hat{\mathbf{m}} = \frac{\mathbf{m}}{|\mathbf{m}|}$ to get

$$\mathbf{B}(t) = \frac{\mu_0}{4\pi r_0^3} \left(\frac{3(\mathbf{m} \cdot (\hat{\mathbf{r}}_0 + \frac{\mathbf{v}t}{r_0}))(\hat{\mathbf{r}}_0 + \frac{\mathbf{v}t}{r_0})}{|\hat{\mathbf{r}}_0 + \frac{\mathbf{v}t}{r_0}|^5} - \frac{\mathbf{m}}{|\hat{\mathbf{r}}_0 + \frac{\mathbf{v}t}{r_0}|^3} \right) \tag{9.12}$$

$$\mathbf{B}(\theta) = \frac{\mu_0}{4\pi r_0^3} \left(\frac{3(\mathbf{m} \cdot (\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta))(\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta)}{|\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta|^5} - \frac{\mathbf{m}}{|\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta|^3} \right) \tag{9.13}$$

$$\mathbf{B}(\theta) = \frac{\mu_0 |\mathbf{m}|}{4\pi r_0^3} \left(\frac{3(\hat{\mathbf{m}} \cdot (\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta))(\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta)}{(1 + \theta^2)^{5/2}} - \frac{\hat{\mathbf{m}}}{(1 + \theta^2)^{3/2}} \right). \tag{9.14}$$

Expanding the terms $(\hat{\mathbf{m}} \cdot (\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta))(\hat{\mathbf{r}}_0 + \hat{\mathbf{v}}\theta)$ and combining terms with the same exponent in theta gives the representation of the magnetic field of a dipole on a line in terms of three Anderson functions

$$\mathbf{B}(\theta) = \frac{\mu_0 |\mathbf{m}|}{4\pi r_0^3} \left[f_0 \mathbf{C}_0 + f_1 \mathbf{C}_1 + f_2 \mathbf{C}_2 \right], \tag{9.15}$$

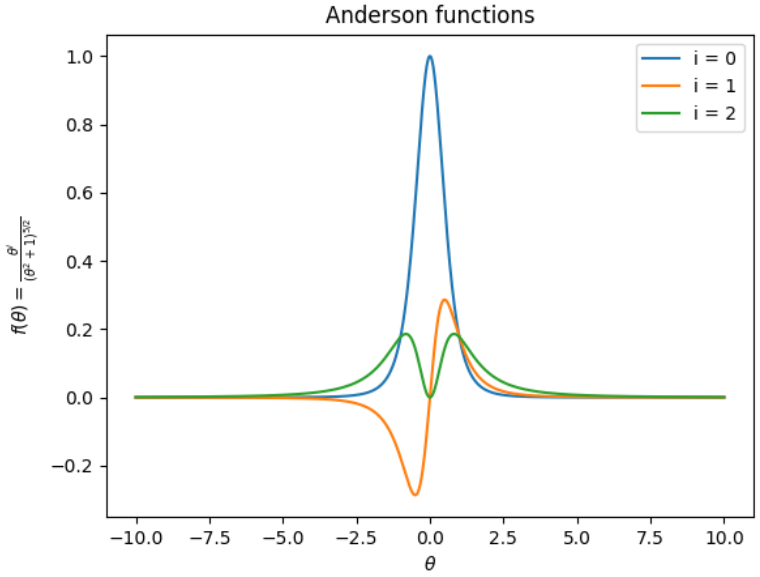


Figure 9.1: The three Anderson functions $f_i(\theta) = \frac{\theta^i}{(1+\theta^2)^{5/2}}$ for $i = 0,1,2$ as a function of θ in the range $[-10,10]$.

where

$$f_i(\theta) = \frac{\theta^i}{(1+\theta^2)^{5/2}} \quad i = 0,1,2, \tag{9.16}$$

are illustrated in Figure 9.1 and $\mathbf{C}_n$ are the coefficients of the Anderson functions

$$\mathbf{C}_0 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0 - \hat{\mathbf{m}} \tag{9.17}$$

$$\mathbf{C}_1 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{v}} + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{r}}_0 \tag{9.18}$$

$$\mathbf{C}_2 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}}. \tag{9.19}$$

Here, the vector $\mathbf{C}_i$ represents the contributions of the i-th Anderson function to the magnetic field vector $\mathbf{B}$, up to a scaling factor of $\frac{\mu_0|\mathbf{m}|}{4\pi r_0^3}$. The k-th components of the vectors $\mathbf{C}_i$ for $i = 0,1,2$ describe the k-th component of the magnetic field vector $\mathbf{B}$.

9.3. Construction of 1D Anderson Wavelets

The Anderson functions, as described in Equation 9.16, form a basis for the magnetic field of a single dipole measured on an arbitrary line in $\mathbb{R}^3$. In this section, we validate the properties of a Mother Wavelet for a Wavelet Family and see if we can use the Anderson functions as a Mother Wavelet. We also discuss what conditions are not met and how we resolve these issues.

9.3.1. Orthonormalized Anderson Functions

We want to use the Anderson functions as a Mother Wavelet, to construct a Wavelet Family that spans the whole range of magnetic fields created by dipoles on a line. The Anderson functions in Figure 9.1 are, however, three functions that are not orthonormal. If we want to use the Anderson functions as a Mother Wavelet, we need to orthonormalize the Anderson functions so that they all represent a different subspace. Otherwise, we would have an overlap between the Anderson Functions in parts of the space, which would result in a redundant representation of the magnetic field. Note that this step is different from the orthogonality property of the wavelets with their own shifted and scaled versions, which is a property of the wavelet family (Section 9.1.3). This step is necessary because we are working with Wavelets on a vector field, and we, therefore, have a Mother Wavelet that consists of three functions. The functions can be made orthonormal by using the Gram-Schmidt process. Equation 9.20 shows the orthonormalized Anderson functions.

$$\begin{aligned}\hat{f}_0(\theta) &= f_2(\theta)\sqrt{\frac{128}{3\pi}} \\ \hat{f}_1(\theta) &= f_1(\theta)\sqrt{\frac{128}{5\pi}} \\ \hat{f}_2(\theta) &= \left(f_0(\theta) - \frac{5}{3}f_2(\theta)\right).\end{aligned} \tag{9.20}$$

It must be noted that by orthonormalizing the Anderson functions, we lose the interpretation of the coefficients. The coefficients of the orthonormalized Anderson functions are not the same as the coefficients of the original Anderson functions. It is therefore important to keep track of the coefficients that are used during the orthonormalization process. A QR-decomposition of the matrix of the original Anderson functions can be used to represent the orthonormalized Anderson functions in terms of the original Anderson functions. Let $A = [f_0, f_1, f_2]$ be the matrix of the original Anderson functions, $Q = [\hat{f}_0, \hat{f}_1, \hat{f}_2]$ the matrix of the orthonormalized Anderson functions and R the QR-decomposition of A. Then R^{-1} for Equation 9.20 is given by

$$R^{-1} = \begin{pmatrix} 0 & 0 & \sqrt{\frac{128}{3\pi}} \\ 0 & \sqrt{\frac{128}{5\pi}} & 0 \\ \sqrt{\frac{24}{5\pi}} & 0 & -\sqrt{\frac{24}{5\pi}} \end{pmatrix}. \tag{9.21}$$

9.3.2. Support of Anderson Wavelets

The Anderson functions, as defined in Equation 9.16 do not have compact support. This is an issue because the wavelet transform only works well on functions with compact support. The Anderson functions, however, converge to zero for large θ values, because the denominator increases faster than the numerator. Therefore, the support of the wavelets can be made compact by truncating the support of the Anderson functions. We tried different truncation lengths, $[-4,4]$ and $[-2,2]$. The support levels are shown in Figure 9.2. The length of the support is important for the shift transformation of the mother wavelet. When truncating the support, we have to make sure that the functions are orthogonal to their shifted version. Further analysis of the support

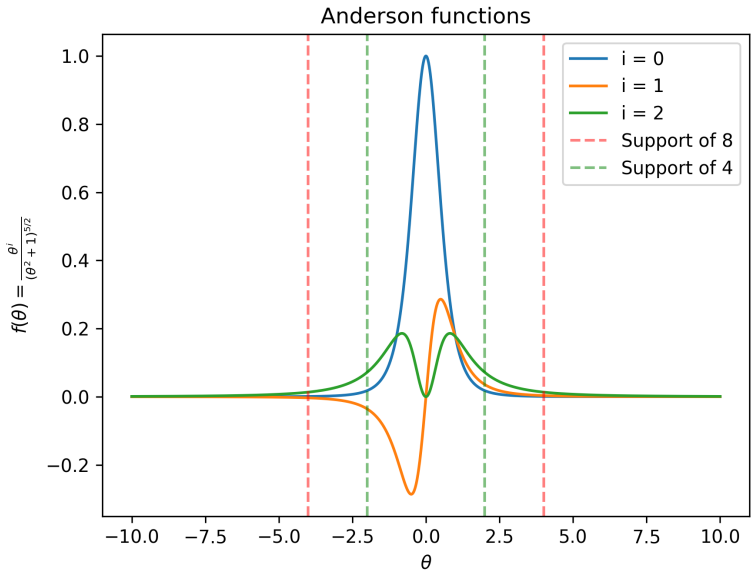


Figure 9.2: The Anderson functions for a source at 1 m depth. θ describes the position on the line where the magnetic field is measured. The support of the wavelets is not compact, but two truncated support lengths are indicated by the dashed lines.

of the wavelets showed indeed that a larger support results in more orthogonal functions after the shift transformation. Figure 9.3 shows the Anderson functions with different support lengths and their shifted version. The condition matrices, which contain the inner products of the functions, can be used to determine how orthogonal the basis functions are. It seems that the support of $[-4,4]$ is a reasonable choice for the support

of the wavelets. Depending on the application, a window of $[-2,2]$ might be sufficient. Larger support lengths result in more orthogonal functions after shifting, but that comes at the cost of resolution in the representation of the magnetic field. Therefore, we use a support of $[-4,4]$ for the wavelets in this thesis, if not stated otherwise. The support length we found seems to be in line with what the literature suggests. [54, pp. 4–3] states that a source at 1000 m depths shows most of its interesting features within a window of only 4000 m in width.

Besides restricting the support, we also tried adding a window function to the Anderson function, to make the support compact. The window functions did improve the orthogonality of the functions, but it resulted in wiggles in the reconstruction of the magnetic field, which is not desirable. Therefore, we decided to use the truncated support of the Anderson functions, instead of a window function.

9.3.3. Mean-zero Anderson Functions

Besides the support of the wavelets, other requirements have to be met by the mother wavelet. To see if we can use the three Anderson functions defined in Equation 9.16 as a mother wavelet, we need to check if they satisfy all the requirements. The Anderson functions have approximately compact support, because the denominator increases faster than the numerator. The first two requirements are trivially satisfied because the Anderson functions are absolutely integrable and square-integrable. The third requirement is not satisfied, because not all the Anderson functions are mean-zero. Only the second Anderson, which is an odd function of θ, is mean-zero. The first and third Anderson functions are both even functions (1 and θ^2 respectively) and therefore not mean-zero.

This means that we have to make some concessions to find mean-zero mother wavelets. There are several ways to do this:

- **Use only the second Anderson function** as a mother wavelet. This simplifies the construction of the wavelet family, at the cost of reducing the span of the wavelet family. Using only the second Anderson function makes it impossible to reconstruct all the magnetic fields of a dipole on a line.
- **Sacrifice the mean-zero property** of the mother wavelet.
- **Extending the basis** of the wavelet family with other functions that are mean-zero. This increases the span of the wavelet family.

We have decided to sacrifice the mean-zero property of the mother wavelet. This admissibility condition ensures that the wavelet transform is invertible. Relaxing this condition has as a consequence that the wavelet family is not orthonormal for shape transformations. The wavelet transform could start to mix different frequencies, resulting in aliasing. Therefore, the truncation length of the support (Figure 9.3) becomes an important parameter for wavelet families that are not mean-zero: larger support ensures that the wavelet family is more orthogonal. It is also possible to bring back the orthogonality of the wavelet family, by orthonormalizing the wavelet family. In Section 9.4.2 we discuss how orthonormalization can be done as a pre-processing step. Therefore, sacrificing the mean-zero property of the mother wavelet is not a problem.

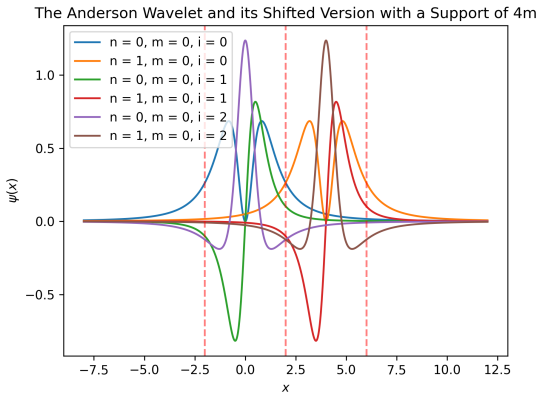
9.4. Reconstucting signals using Anderson wavelets

9.4.1. Signal-based Search-Space

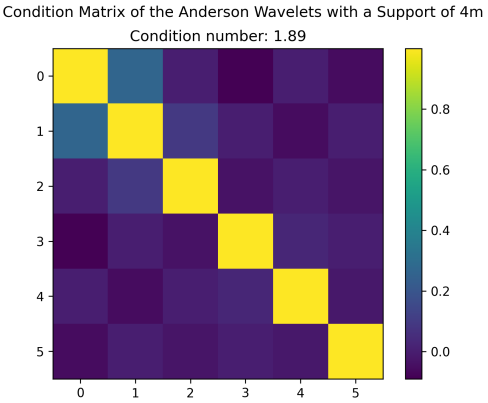
Minimum and Maximum Depth

Given a signal that is measured over a line of length L, In sampling theory, it is often not possible to reconstruct the original signal from measurements. This is not only caused by noise but also by the fact that the signal is sampled at a discrete set of points. The rate at which the signal is sampled (how far the samples lie apart) determines the maximum frequency that can be reconstructed. The minimum frequency that can be captured is determined by the length of the signal. The longer the signal, the lower the minimum frequency that can be captured.

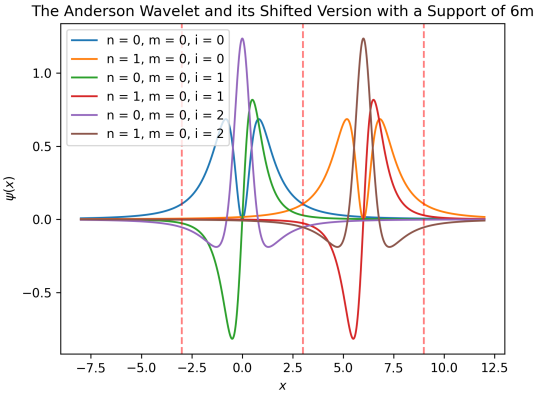
Similar reasoning can be applied to the Anderson Wavelets, which can help us effectively define our search-space. In Section 9.3.2, we have seen that the Anderson Wavelets do not have compact support, but that their support is small enough outside the domain of $[-2,2]$. This sets a lower bound on the depth of the sources that can be reconstructed. For example, if we have samples that are $dx = 32$ m apart, it is very difficult to distinguish between a source at 1 m and a source at 2 m or to determine the exact position of a source at 1 m depth. We therefore limit our search-space to a source that is at least $2^{m_0} = dx/s$ meters deep, where s is the total length of the support of the wavelet. This means that we only reconstruct sources that



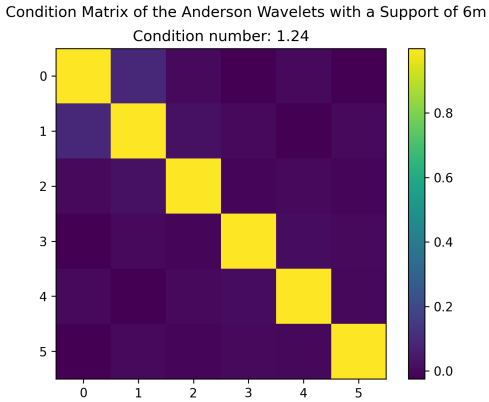
(a) The Anderson functions with a support of $[-2,2]$ and a 4 m shifted version.



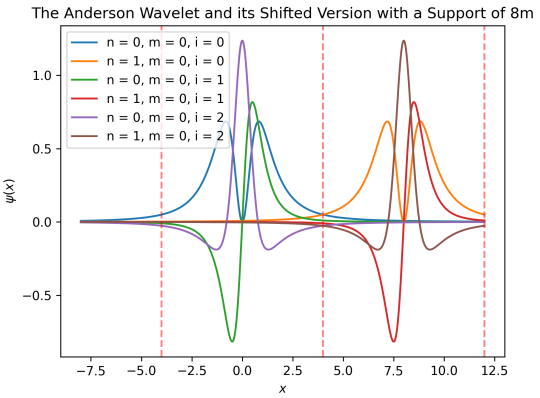
(b) The 'condition matrix' of the 6 Anderson functions illustrated in Figure 9.3a.



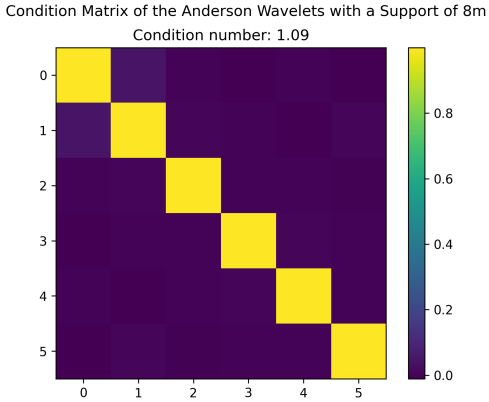
(c) The Anderson functions with a support of $[-3,3]$ and a 6 m shifted version.



(d) The 'condition matrix' of the 6 Anderson functions illustrated in Figure 9.3c.



(e) The Anderson functions with a support of $[-4,4]$ and a 8 m shifted version.



(f) The 'condition matrix' of the 6 Anderson functions illustrated in Figure 9.3e.

Figure 9.3: The Anderson functions with different support lengths and their shifted versions. n is the number of shifts that are made, $m = 0$ indicates that the signal originates from a source at depth $2^0 = 1$ m and i indicates the number of the anderson function. The 'condition matrix' is a matrix of the inner products of all the functions. The functions are orthogonal if the condition matrix is diagonal. Increasing the support length leads to less overlap between the functions when shifting with the same amount, and therefore a more orthogonal set of functions.

are at least $m_0 = 32/4 = 8$ m deep. Preferably, we would look for deeper sources, because they impact the signal on a larger part of the line and are therefore easier to be detected. A higher sampling rate would also help: if we take $2^{m_0} = 8dx/s$, we know for sure that we have 8 samples in the support of the wavelets we look for. In practice, for one-dimensional signals in space, the sampling rate is not a limiting factor, because the sensors have a high enough sample rate. In the two-dimensional case, the sampling rate is a more limiting factor, because there the spacing between the measurements is not only determined by the flight speed and the sampling rate but also by the distance between the survey lines.

The maximum depth we look for is also dependent on the support. We prefer that (almost) the full support is contained in the measured signal because this makes it easier to orthogonalize the wavelets and reconstruct the original signal. We use the following formula to determine the maximum depth:

$$m_{\max} = \left\lceil \log_a \left(\frac{L}{s} \right) \right\rceil , \tag{9.22}$$

where L is the length of the line, s is the total length of the support of the wavelet and $a = 2$ is the scaling parameter and $\lceil x \rceil$ is the ceiling function, which rounds x up to the nearest integer. This means that for the reconstruction of our 256 m long line, we look for sources that are at most $a^{m_{\max}} = 2^{\lceil \log_2 (\frac{256}{4}) \rceil} = 2^6 = 64$ m deep. This way, we are sure that the support of the wavelets is contained in the signal.

Depth-dependent shifts
Now that the minimum and maximum depths are known, we can determine the shifting factors of the wavelets. For this, we again make use of the support of the wavelets. In Section 9.1.3 it is shown that shifting a wavelet by its support ensures that a wavelet is orthogonal to its shifted version. We therefore also shift the wavelets with at least their support. For a depth of 2^m, the support is $4 \cdot 2^m$, resulting in half the number of shifts over the length of the signal when the depth is doubled. This distribution of the shifts is used in Figure 9.5 to visualize the spectrum.

9.4.2. Pre-computing the Wavelet Family

We have seen that the search-space for the depth of the sources is dependent on the length of the signal and the support of the wavelets. Now that we know the depth (shape) and the shifts (position) of the wavelets, we can now use the mother wavelets to construct the wavelet family. The bases can be found by applying Equation 9.7 to the three Anderson functions. The total number of wavelets in a family is dependent on the number of shifts and the number of depths. The number of wavelets is given by:

$$N = \sum_{m=m_0}^{m_{\max}} a^{m-m_0} \cdot n_0 , \tag{9.23}$$

where $m_{\max}$ and m_0 are the maximum and minimum depth factor and n_0 is the number of wavelets at the maximum depth. For a depth factor $a = 2$ and $n_0 = 1$ dipole in the deepest layer, a signal length of $L = 256$ m, a support of $s = 4$ m, and a total of 2048 samples, we get $N = 127$ shapes in the Wavelet Family for each of the 3 basis functions. This means that the total number of wavelets is $N_{\text{total}} = 3 \cdot 127 = 381$.

Since the Anderson functions are not mean-zero and do not have compact support, this wavelet family is not orthogonal. We orthogonalize the wavelet family by storing all the wavelets in a matrix A which has the wavelet functions as columns:

$$A = \begin{bmatrix} \psi_{m_{\max},0} & \psi_{m_{\max}-1,0} & \psi_{m_{\max}-1,1} & \cdots & \psi_{m_0,a^{m-m_0}-1} & \psi_{m_0,a^{m-m_0}} \end{bmatrix} . \tag{9.24}$$

The order of the wavelets in the matrix is important because we want to orthogonalize the wavelets in the order of the depth. We start with the wavelets that cause the lowest spatial frequencies and work our way up to the highest spatial frequencies. For each depth, we sort the wavelets by their shifts. As our domain is limited to positive shifts, we only need to sort the wavelets by their shifts. Therefore, the first wavelet is the deepest wavelet, the next two wavelets are for the second deepest depth, ordered from left-to-right by their shifts. This ordering can be found by iterating over the cells in the spectrum (Figure 9.5) from bottom to top and from left to right. The wavelets are then orthogonalized by applying the Gram-Schmidt process to the columns of the matrix A. The result is a matrix Q with orthogonal columns and a matrix R with the coefficients of the orthogonalization.

$$A = QR , \tag{9.25}$$

where

$$Q = \left[\psi^{\perp}_{m_{\max},0} \quad \psi^{\perp}_{m_{\max}-1,0} \quad \psi^{\perp}_{m_{\max}-1,1} \quad \cdots \quad \psi^{\perp}_{m_0,a^{m-m_0}-1} \quad \psi^{\perp}_{m_0,a^{m-m_0}} \right], \tag{9.26}$$

and R is an $N \times N$ matrix where N is the number of wavelets in the family (the number of columns in A).

This orthogonalization is computationally intensive, especially for higher-resolution wavelets, which contain more samples, more shifts and more depths. Let alone when the functions are extended to two dimensions. Therefore, we suggest to pre-compute the wavelet family. As the wavelets are dependent on the positions of the samples in the signal, which can be arbitrary, we decided to pre-compute the wavelets on predefined regular grids. The measurements can be interpolated to these grids using Cubic Splines. For the one-dimensional case, we pre-computed the Wavelet Families on a line with 64, 128, … and 2048 samples. One can then interpolate the signal to the nearest grid and use the pre-computed wavelets.

A major advantage is that the pre-computed wavelets can be used for multiple scales of the problem. For example, if we want to detect sources on a 1 m signal with 128 samples, we can use the pre-computed wavelets on the 128-sample grid. The same family can be used to detect sources on a 10 km signal of 128 samples but requires scaling of the wavelets by a factor of 10^4. Our wavelet family is scale-invariant, so we can use the same wavelets for both signals.

9.4.3. Detection and Localization by analyzing the Spectrum of the Anderson Wavelet Transform

Using our orthonormal basis functions in our orthonormalized wavelet family, we can find the coefficients of each of the wavelet functions by applying the Wavelet transform of Equation 9.4 to the basis functions and the signal. Hereby, we project the signal onto the orthonormal basis functions, resulting in a set of coefficients of how each basis function contributes to the signal.

In this section, we try to estimate the depth and position of dipole sources using only measurements on a line. A simple, simulated magnetic field is used to demonstrate the techniques. For simplicity, we assumed that the sources are located straight underneath the line of flight. In total, 1024 measurements are taken with a spacing of 0.25 m throughout 256 m. The measurements are illustrated in Figure 9.4.

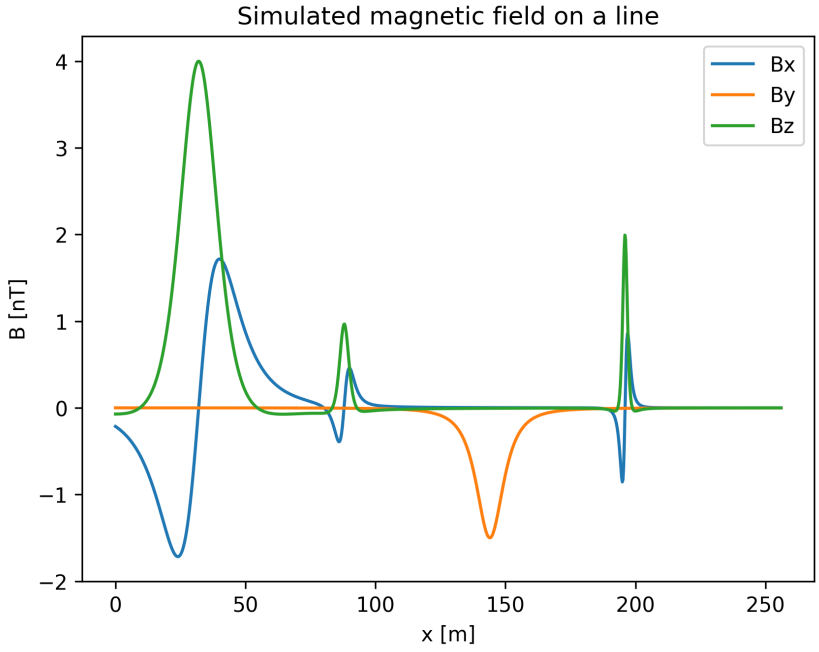


Figure 9.4: Vectorial measurements of a simulated magnetic field on a line of 256 m with a spacing of 0.25 m between the samples.

At the end of this section, we give the exact locations of the sources that are used to generate this magnetic field. But first, we try to estimate these locations just by analyzing the signal. The detection coefficients are used to determine if a source is present at a certain depth and position. For now, we are only interested in where the source is located, and not how strong the actual source is. The depth and position of the sources can be found by looking for local extrema in the spectrum.

We introduce a wavelet spectrum to analyze and visualize the position of the sources. In the one-dimensional case, the wavelet spectrum has two axes. The x-axis describes the discretized horizontal position of the estimated sources, in steps of the support. The y-axis indicates the depth of the sources. The color in the spectrum represents the weight of the Anderson coefficients of the corresponding wavelets.

In the one-dimensional case, the coefficients $\mathbf{C}_i$ are vectors of length 3 for each Anderson function (Equation 9.17). In total, there are 9 coefficients per cell in the spectrum. To assign a single value to these coefficients, we take the sum of the absolute values. Let $\mathbf{C}_i(a,b)$ represent the coefficients of the Anderson wavelet with shape parameter a and shifting parameter b derived from the i-th mother wavelet, then the value $\hat{C}(a,b)$ in the spectrum is given by

$$\hat{C}(a,b) = \sum_{i=0}^{2} \sum_{j=0}^{2} |\mathbf{C}_{i,j}(a,b)|, \tag{9.27}$$

where $\mathbf{C}_{i,j}(a,b)$ represents the j-th components of the Anderson coefficient vector $\mathbf{C}_i(a,b)$.

We need to consider two spectra. The first spectrum represents the intensity of the wavelet coefficients in the orthonormalized wavelet space ($\mathbf{Q}y = \mathbf{B}$). The coefficients are a result of projecting the signal onto the orthonormal basis functions: $y = \mathbf{BQ}$. Due to the orthonormalization, this spectrum shows "leakage" to cells to the left and below the actual source location. Figure 9.5a shows an example of such a spectrum. Local maxima in the spectrum correspond with possible locations of dipole sources. The second type of spectrum is obtained by converting the orthonormalized coefficients back to the original wavelet space. $\mathbf{AC} = \mathbf{B}$. Note that from our QR-decomposition, we know that $y = R\mathbf{C}$ and therefore

$$\mathbf{C} = R^{-1} y. \tag{9.28}$$

Figure 9.5b shows an example of such a spectrum. Both spectra in Figure 9.5 are related to the signal of Figure 9.4. Inspection of the spectrum gives us information about the location of the sources. From the clear local maxima that are present in the spectrum, we expect that in total four sources are located at the following positions:

- a 16 m deep source at $x = 32$,
- a 4 m deep source at $x = 88$,
- an 8 m deep source at $x = 144$,
- at 196 m a 2 m deep souce.

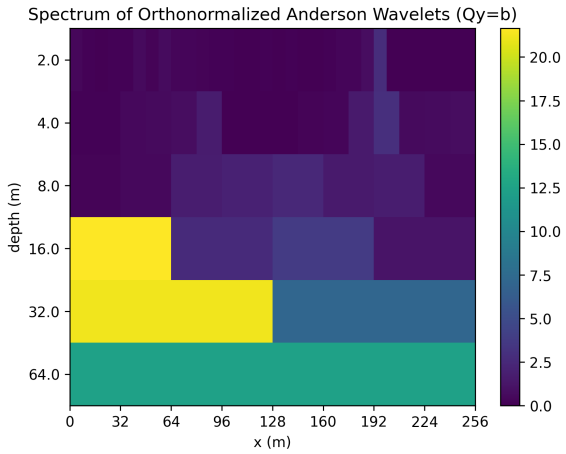
The method seems to work since the original simulated field consists of four dipoles which are positioned exactly at the places that we expect from inspecting the spectrum. The Orthonormalized Anderson wavelet family seems therefore a good candidate for detecting and localizing sources underneath a line of measurements. It must be remarked that, in this example, the sources were purposely positioned at depths and shifts that are contained in our search space.

9.4.4. Reconstructing the original signal from the Anderson Wavelet coefficients

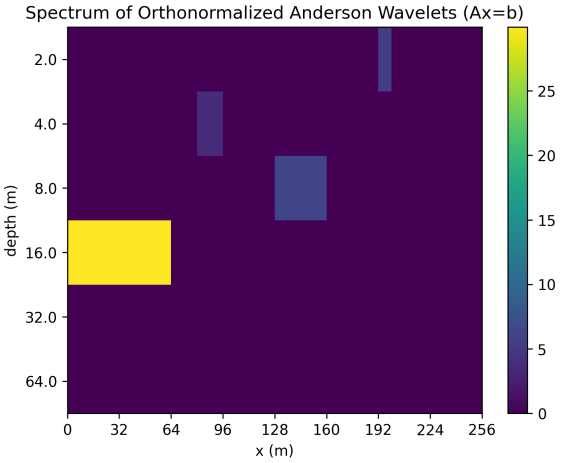
In the previous section, we showed that the Anderson Wavelets can be used to detect the location of the sources. They do, however, also form a basis for the space of magnetic fields originating from dipoles. A reconstruction of the signal can therefore be made by multiplying the wavelet functions with their coefficients. The wavelet transformation using the orthonormalized wavelet family outputs the orthonormalized Anderson coefficients y. These coefficients are transformed back to the original Anderson coefficients $\mathbf{C}'_i(a,b)$ by inverting the orthonormalization using the matrix R. Figure 9.6 shows that the wavelets are indeed able to reconstruct the original field.

9.4.5. Translating the Anderson coefficients to Dipole parameters

The Anderson wavelets are constructed in such a way that they form a basis for the magnetic field originating from dipoles. We have tested this by reconstructing a magnetic field originating from dipoles using the Anderson wavelets and their coefficients. As the Anderson functions form a basis for the magnetic field originating from dipoles, it must be possible to translate the coefficients to the parameters of the dipole. We show how this can be done in this section. We start with the original Anderson functions and extend them to the wavelet family.

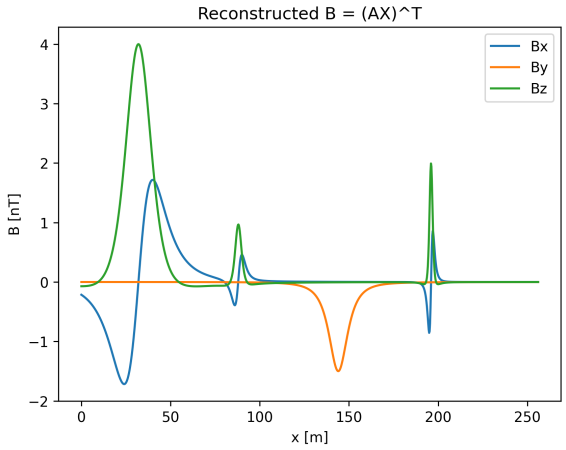


(a) Spectrum of the orthonormalized Anderson Wavelets. The values represent the sum of the absolute values of the Orthonormalized Anderson coefficients y in $\mathbf{Q}y = \mathbf{B}$.

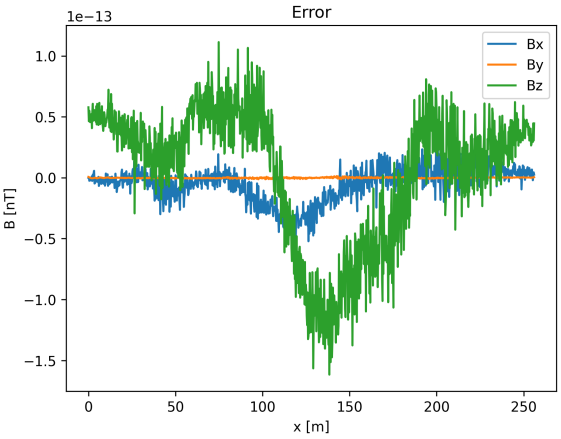


(b) Spectrum of the original Anderson Wavelets, obtained by inverting the orthonormalization using R. The values represent the sum of the absolute values of the original Anderson coefficients $\mathbf{C}$ in $\mathbf{AC} = \mathbf{B}$.

Figure 9.5: Example of the spectrum of (Orthonormalized) Anderson Wavelets obtained from a one-dimensional signal of 256 m. The wavelet support is set to 4 m, therefore, the maximum search depth is 64 m. The color in the spectrum indicates the total weight of the corresponding wavelet coefficients as defined in Equation 9.27.



(a) Reconstructed field using the original Anderson coefficients $\mathbf{C}'_i(a,b)$



(b) Error between the reconstructed field (Figure 9.6a) and the original field (Figure 9.4)

Figure 9.6: Reconstruction of the original field (Figure 9.4) using the Anderson wavelet coefficients. The error between the original field and the reconstructed field is shown in Figure 9.6b.

A dipole is characterized by its magnetic moment $\mathbf{m}$ and its (relative) position $\mathbf{r}$. These vectors can be retrieved from the Anderson coefficients $\mathbf{C}_i$. First, we solve for the magnetic moment $\mathbf{m}$ and then for the position $\mathbf{r}$. We assume that we know $\mathbf{C}_0$, $\mathbf{C}_1$ and $\mathbf{C}_2$, which are the Anderson coefficients of the three Anderson functions and can be found by projecting the magnetic field onto the Anderson functions. Every coefficient is a vector of length 3 and satisfies Equation 9.17:

$$\mathbf{C}_0 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0 - \hat{\mathbf{m}} \tag{9.29}$$

$$\mathbf{C}_1 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{v}} + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{r}}_0 \tag{9.30}$$

$$\mathbf{C}_2 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}}. \tag{9.31}$$

To isolate the magnetic moment $\hat{\mathbf{m}}$ from the coefficients $\mathbf{C}_i$, we only need to know the third coefficient $\mathbf{C}_2$ and the direction of movement of the sensor $\hat{\mathbf{v}}$. In the one-dimensional case, where we move along a line, we can easily find $\hat{\mathbf{v}}$ by taking the difference between two consecutive positions of the sensor. The vector $\hat{\mathbf{v}}$ is then the unit vector in the direction of the difference and looks probably something like $\hat{\mathbf{v}} = [1,0,0]$. First, we extract $\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}$ from $\mathbf{C}_2$, then we use this value to find $\hat{\mathbf{m}}$.

Given $\mathbf{C}_2 = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}}$, we want to find $\mathbf{C}_2 \cdot \hat{\mathbf{v}}$, which means:

$$\mathbf{C}_2 \cdot \hat{\mathbf{v}} = (3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}}) \cdot \hat{\mathbf{v}}.$$

Using the distributive property of the dot product, we can rewrite the expression as:

$$(3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}) \cdot \hat{\mathbf{v}} - (\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}).$$

Now, we can simplify the first term. Since $\hat{\mathbf{v}}$ is a unit vector, we have $\hat{\mathbf{v}} \cdot \hat{\mathbf{v}} = 1$:

$$\begin{aligned}(3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}}) \cdot \hat{\mathbf{v}} &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}) \\ &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})(1) \\ &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}).\end{aligned}$$

We substitute this expression in the original equation:

$$\mathbf{C}_2 \cdot \hat{\mathbf{v}} = 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}) - (\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}),$$

and factor out $(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})$:

$$\begin{aligned}\mathbf{C}_2 \cdot \hat{\mathbf{v}} &= (3-1)(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}) \\ &= 2(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}).\end{aligned}$$

Therefore, we can obtain $\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}$ as

$$\boxed{(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}) = \frac{(\mathbf{C}_2 \cdot \hat{\mathbf{v}})}{2}},$$

where $\mathbf{C}_2$ is the second Anderson coefficient and $\hat{\mathbf{v}}$ is the direction of movement of the sensor.

Now that we have $\hat{\mathbf{m}} \cdot \hat{\mathbf{v}}$, we can use it to find $\hat{\mathbf{m}}$:

$$\begin{aligned}\mathbf{C}_2 &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}} \\ \hat{\mathbf{m}} &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \mathbf{C}_2 \\ \boxed{\hat{\mathbf{m}} &= \frac{3}{2}(\mathbf{C}_2 \cdot \hat{\mathbf{v}})\hat{\mathbf{v}} - \mathbf{C}_2}.\end{aligned}$$

It is also possible to find the direction of the magnetic moment $\hat{\mathbf{m}}$ using the $\mathbf{C}_0$. This requires that we know the vector from the sensor to the source $\hat{\mathbf{r}}_0$. In general, this vector is unknown, but if we assume that the sources are positioned straight under the sensor, we can use the vector $\hat{\mathbf{r}}_0 = [0,0,-1]$. Resulting in the following expression for $\hat{\mathbf{m}}$:

$$\boxed{\hat{\mathbf{m}} = \frac{3}{2}(\mathbf{C}_0 \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0 - \mathbf{C}_0}.$$

Obtaining the direction vector $\hat{\mathbf{r}}_0$ is a bit more complicated. In theory, it should be possible to find $\hat{\mathbf{r}}_0$, even when $\hat{\mathbf{r}}_0 \cdot \hat{\mathbf{v}} \neq 0$. Allowing us to find the direction vector $\hat{\mathbf{r}}_0$ without the assumption that the sources are positioned straight under the sensor. However, this would require further investigation of the relationship between the Anderson coefficients and the dipole parameters. We also noted that the coefficients of the Anderson functions (Equation 9.17) have no $\hat{\mathbf{r}}_0 \cdot \hat{\mathbf{v}}$ terms. A lack of these terms could be the reason why it is difficult to find back the vector $\hat{\mathbf{r}}_0$ from the Anderson coefficients. Further study is required to determine if this is indeed the case. This is left as future work.

Under certain circumstances, we can however simplify the problem by assuming that we know the CPA (Closest Point of Approach) and that we centered the Anderson functions around that point. This means that the vector $\hat{\mathbf{r}}_0$ is of the shape $\hat{\mathbf{r}}_0 = [0, \cdot, \cdot]$ with $\hat{\mathbf{v}} \cdot \hat{\mathbf{r}}_0 = 0$. Under this assumption, we can find the direction vector $\hat{\mathbf{r}}_0$, using the already known $\hat{\mathbf{v}}$, $\mathbf{C}_i$ and $\hat{\mathbf{m}}$. We start by taking the dot product of $\mathbf{C}_1$ and $\hat{\mathbf{v}}$:

$$\begin{aligned}\mathbf{C}_1 \cdot \hat{\mathbf{v}} &= (3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{v}} + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{r}}_0) \cdot \hat{\mathbf{v}} \\ &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)(\hat{\mathbf{v}} \cdot \hat{\mathbf{v}}) + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})(\hat{\mathbf{r}}_0 \cdot \hat{\mathbf{v}}) \\ &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0),\end{aligned}$$

using the fact that $\hat{\mathbf{v}} \cdot \hat{\mathbf{v}} = 1$ and $\hat{\mathbf{v}} \cdot \hat{\mathbf{r}}_0 = 0$. Dividing both sides by 3, we obtain:

$$(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0) = \frac{\mathbf{C}_1 \cdot \hat{\mathbf{v}}}{3}.$$

We can now use this equation to find $\hat{\mathbf{r}}_0$:

$$\begin{aligned}\mathbf{C}_1 &= 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{r}}_0)\hat{\mathbf{v}} + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{r}}_0 \\ &= (\frac{\mathbf{C}_1 \cdot \hat{\mathbf{v}}}{3})\hat{\mathbf{v}} + 3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})\hat{\mathbf{r}}_0.\end{aligned}$$

And finally we can isolate $\hat{\mathbf{r}}_0$:

$$\boxed{\hat{\mathbf{r}}_0 = \frac{\mathbf{C}_1 - \left(\frac{\mathbf{C}_1 \cdot \hat{\mathbf{v}}}{3}\right)\hat{\mathbf{v}}}{3(\hat{\mathbf{m}} \cdot \hat{\mathbf{v}})}}.$$

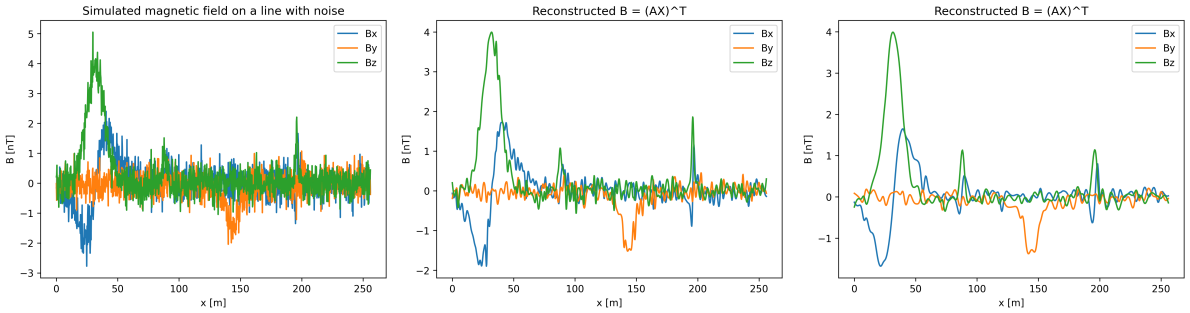
As we have not performed this derivation for a general case where sources are not always straight under the trajectory of movement, we restrict ourselves from now on to reconstructions in the wavelet space with Anderson coefficients and omit the translation step to the coefficients of a dipole model.

9.4.6. Wavelet Reconstructions of Different Signals

In this section, we look at the reconstructions of different one-dimensional signals. We first use the source distribution as in Figure 9.4, but add some noise to the measurements. Then, we increase the number of sources and place them randomly in the domain, to see how the reconstruction behaves in a more realistic scenario.

Reconstruction of a Noisy Signal

To see how noise affects the Wavelet reconstruction, we add some noise to our measurement data. We add noise with a uniform distribution between -1% and 1% of the maximum value of the original signal. The result is shown in Figure 9.7a. We then reconstruct the signal using all wavelets, and the result is shown in Figure 9.7b. We can see that the noise is mainly in the high frequencies, which are the shallower wavelets. We can limit the search space to deeper wavelets, which reduces the noise in the reconstruction. The result of this is shown in Figure 9.7c. We can see that the noise is reduced, but the resolution of the reconstruction is also reduced: the shallow positioned source at $x = 196$ m is barely reconstructed.

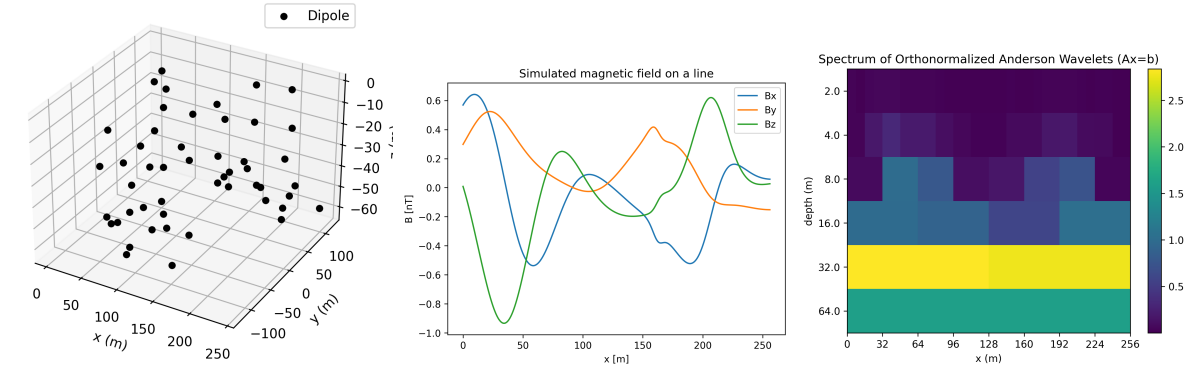


(a) Simulated magnetic field on a line with noise.

(b) Reconstructed magnetic field on a line with noise.

(c) Reconstructed magnetic field on a line with noise, limited to deeper wavelets.

Figure 9.7: Reconstruction of the original field (Figure 9.4) for a signal with 1% noise. The left figure shows the measured field, the middle figure shows the reconstruction using all wavelets, and the right figure shows the reconstruction using only wavelets with a minimum depth of 4 m.



(a) The positions of the dipoles that generate a more complex magnetic field.

(b) The measurements that are taken in the field by moving the sensor along the line $y = 0$ and $z = 0$.

(c) The spectrum of the measured field, obtained by projecting the signal on the orthonormal wavelet family and inverse transforming the result to the standard basis.

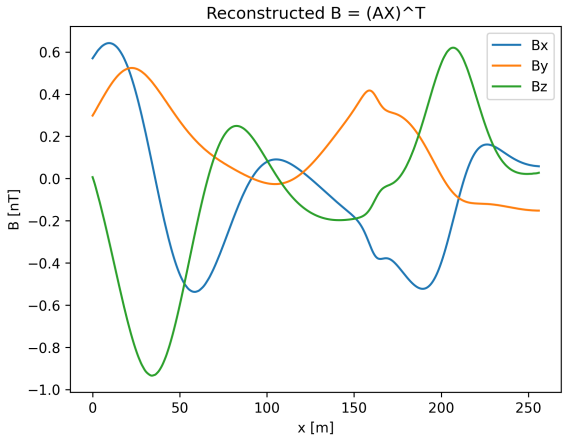
Figure 9.8: The analysis of the field generated by 50 dipoles with random positions and magnetic moments. The left figure shows the positions of the dipoles, the middle figure shows the measurements of the field, and the right figure shows the spectrum of the measurements.

9.4.7. Reconstruction of more complex signals

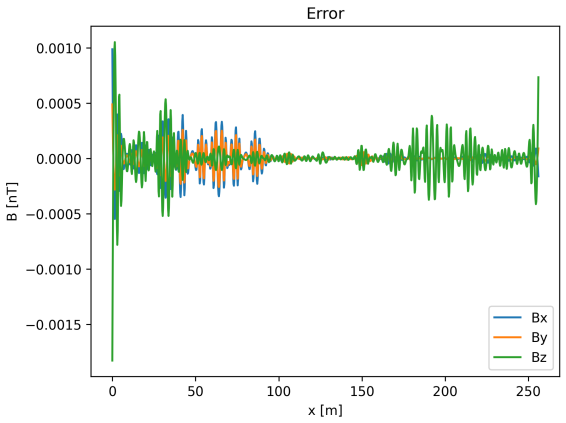
All the previous examples in this chapter concerned fields that originate from a few dipole sources, positioned at the position of the wavelets, and with almost no overlap in their support. In this section, we look at the reconstruction of a more complex signal. We place a total of 50 dipoles. Their positioned is sampled uniformly from the domain $[0,256] \times [-128,128] \times [-2^{m_{\max}}, -2^{m_0}]$ and their magnetic moments are chosen randomly.

The positions of the dipoles are shown in Figure 9.8a. The measurements of the field are shown in Figure 9.8b. The spectrum of the measurements is shown in Figure 9.8c. We can see that the spectrum is more complex than the previous examples. The spectrum is also noisier because the field is generated by more dipoles, the dipoles are positioned more closely together, and the dipoles are not positioned at the exact locations of the wavelets. The reconstruction of the field is shown in Figure 9.9a and the corresponding error is visualized in Figure 9.9b.

The error is significantly larger than in the previous example (Figure 9.6b). The main reason for this is that the positions of the wavelets that are used for reconstructing the signal do not completely align with the sources of the more complex field. The wavelets do still a good job in reconstructing the signal: the original signal is reconstructed with approximately a 0.1% error. The largest errors are made at the boundaries of the domain. This can be explained by the fact that our search space only contains wavelets that are completely contained in the domain. Extending the search space to wavelets that are partially contained in the domain would probably reduce the error at the boundaries.



(a) Reconstructed field using the Anderson Wavelets.



(b) Error between the reconstructed field (Figure 9.9a) and the measured field (Figure 9.8b)

Figure 9.9: Reconstruction of the field generated by 50 dipoles with random positions and magnetic moments. The left figure shows the reconstruction using the Anderson Wavelets, and the right figure shows the error between the reconstruction and the measurements.

9.5. Two-dimensional Anderson Functions

We have demonstrated the effectiveness of the Anderson Wavelets in the reconstruction of the magnetic field of a dipole on a line. If we want to use the same approach to make maps of the magnetic field, we need to extend the Anderson Wavelets to two dimensions. In this section, we derive a two-dimensional variant of the Anderson functions. This is a new result that has not been published before. Similar to the one-dimensional case, we start with the magnetic field of a dipole and substitute the position of the sensor in there, which is now a function of two variables.

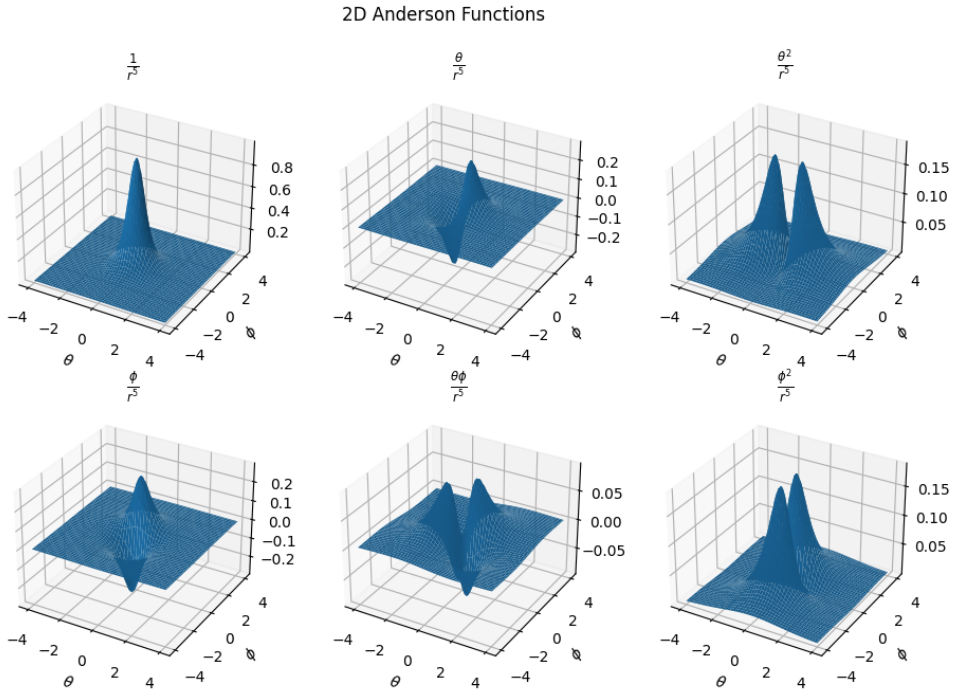


Figure 9.10: The 2D Anderson functions. $r = (1+\theta^2+\phi^2)$

Instead of a single vector that represents the movement of the sensor, we now have two vectors $\mathbf{v}$ and $\mathbf{w}$ that represent the movement of the sensor in the x and y direction respectively. The position of the sensor is now a function of two variables t and s that represent the movement in the x and y direction respectively. The variables t and s are on itself dependent on time t': the position of the sensor at time t' is given by $\mathbf{r}(t) = \mathbf{r}_0 + \mathbf{v}t(t') + \mathbf{w}s(t')$. In further calculations, we leave out the time t' from the notation for brevity. Using the magnetic field of a dipole (Equation 2.36), the following substitutions for the dimensionless quantities θ and ϕ

$$\mathbf{r} = \mathbf{r}_0 + \mathbf{v}t + \mathbf{w}s, \tag{9.32}$$

$$r_0 = |\mathbf{r}_0|, \tag{9.33}$$

$$\theta = \frac{\mathbf{v}t \cdot \mathbf{r}_0}{\mathbf{r}_0 \cdot \mathbf{r}_0} = \frac{(\mathbf{v} \cdot \mathbf{r}_0)\,t}{r_0^2} = \frac{vt}{r_0}, \tag{9.34}$$

$$\phi = \frac{\mathbf{w}s \cdot \mathbf{r}_0}{\mathbf{r}_0 \cdot \mathbf{r}_0} = \frac{(\mathbf{w} \cdot \mathbf{r}_0)\,s}{r_0^2} = \frac{ws}{r_0}, \tag{9.35}$$

$$\tag{9.36}$$

and the assumptions that $\mathbf{v}$, $\mathbf{w}$ and $\mathbf{r}_0$ are orthogonal to each other, so $\mathbf{v} \cdot \mathbf{r}_0 = \mathbf{w} \cdot \mathbf{r}_0 = 0$, and that $r_0 = |\mathbf{r}_0|$, we

can write the magnetic field at the position of the sensor as

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi}\left[\frac{3(\mathbf{m}\cdot(\mathbf{r}_0+\mathbf{v}t+\mathbf{w}s))(\mathbf{r}_0+\mathbf{v}t+\mathbf{w}s)}{|\mathbf{r}_0+\mathbf{v}t+\mathbf{w}s|^5} - \frac{\mathbf{m}}{|\mathbf{r}_0+\mathbf{v}t+\mathbf{w}s|^3}\right] \tag{9.37}$$

$$= \frac{\mu_0}{4\pi}\left[\frac{3r_0^2(\mathbf{m}\cdot(\hat{\mathbf{r}}_0+\frac{\mathbf{v}t}{r_0}+\frac{\mathbf{w}s}{r_0}))(\hat{\mathbf{r}}_0+\frac{\mathbf{v}t}{r_0}+\frac{\mathbf{w}s}{r_0})}{r_0^5|\hat{\mathbf{r}}_0+\frac{\mathbf{v}t}{r_0}+\frac{\mathbf{w}s}{r_0}|^5} - \frac{\mathbf{m}}{r_0^3|\hat{\mathbf{r}}_0+\frac{\mathbf{v}t}{r_0}+\frac{\mathbf{w}s}{r_0}|^3}\right] \tag{9.38}$$

$$= \frac{\mu_0 m}{4\pi r_0^3}\left[\frac{3(\hat{\mathbf{m}}\cdot(\hat{\mathbf{r}}_0+\hat{\mathbf{v}}\theta+\hat{\mathbf{w}}\phi))(\hat{\mathbf{r}}_0+\hat{\mathbf{v}}t+\hat{\mathbf{w}}s)}{(1+\theta^2+\phi^2)^{5/2}} - \frac{\hat{\mathbf{m}}}{(1+\theta^2+\phi^2)^{3/2}}\right] \tag{9.39}$$

$$= \frac{\mu_0 m}{4\pi r_0^3}\left[\frac{3(\hat{\mathbf{m}}\cdot(\hat{\mathbf{r}}_0+\hat{\mathbf{v}}\theta+\hat{\mathbf{w}}\phi))(\hat{\mathbf{r}}_0+\hat{\mathbf{v}}t+\hat{\mathbf{w}}s)}{(1+\theta^2+\phi^2)^{5/2}} - \frac{\hat{\mathbf{m}}(1+\theta^2+\phi^2)}{(1+\theta^2+\phi^2)^{5/2}}\right]. \tag{9.40}$$

Simplification of this expression gives the 2D Anderson function as a sum of six functions

$$\mathbf{B} = \frac{\mu_0 m}{4\pi r_0^3}\frac{1}{(1+\theta^2+\phi^2)^{5/2}}\left[\mathbf{C}_0+\theta\mathbf{C}_1+\theta^2\mathbf{C}_2+\phi\mathbf{C}_3+\theta\phi\mathbf{C}_4+\phi^2\mathbf{C}_5\right], \tag{9.41}$$

where

$$\mathbf{C}_0 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{r}}_0)\hat{\mathbf{r}}_0 - \hat{\mathbf{m}} \tag{9.42}$$

$$\mathbf{C}_1 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{r}}_0)\hat{\mathbf{v}} + 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{v}})\hat{\mathbf{r}}_0 \tag{9.43}$$

$$\mathbf{C}_2 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{v}})\hat{\mathbf{v}} - \hat{\mathbf{m}} \tag{9.44}$$

$$\mathbf{C}_3 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{r}}_0)\hat{\mathbf{w}} + 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{w}})\hat{\mathbf{r}}_0 \tag{9.45}$$

$$\mathbf{C}_4 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{v}})\hat{\mathbf{w}} + 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{w}})\hat{\mathbf{v}} \tag{9.46}$$

$$\mathbf{C}_5 = 3(\hat{\mathbf{m}}\cdot\hat{\mathbf{w}})\hat{\mathbf{w}} - \hat{\mathbf{m}}. \tag{9.47}$$

These functions are plotted in Figure 9.10.

9.6. Reconstruction of magnetic field using a two-dimensional Anderson Wavelet Family

In the previous section, we derived the 2D Anderson function. The same procedure as in Section 9.4 can be used to obtain a Wavelet Family, the only difference is that the shifting parameters are now two-dimensional. Once we have computed and orthogonalized the Wavelet Family, we can use it for the reconstruction of the magnetic field.

Similar to Section 9.4.7, we try to reconstruct a magnetic field that is generated by a set of sources. The sources are uniformly distributed over the domain $[0,256]\times[0,128]$ at different depths. The magnetic field is measured at the surface. The magnetic field is shown in Figure 9.11 and Figure 9.12.

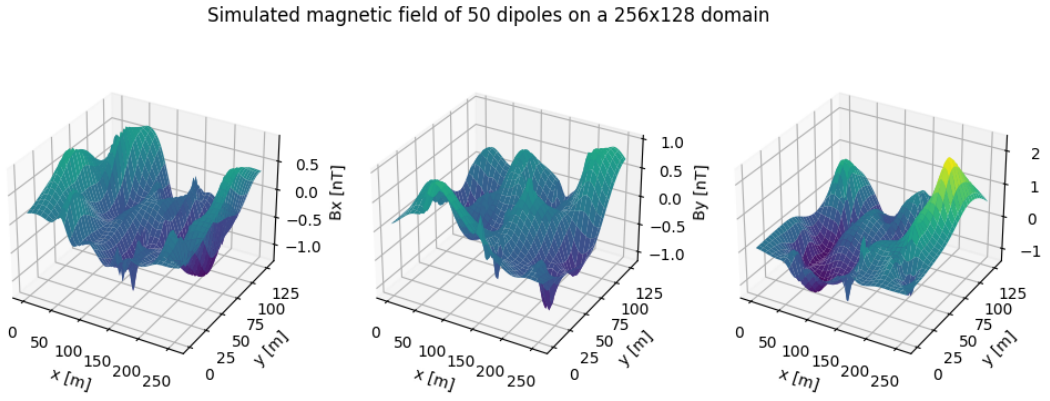


Figure 9.11: Measured magnetic field in the x-y plane. The sources are uniformly distributed over the domain $[0,256]\times[0,128]$ at different depths. The magnetic field is measured at the surface.

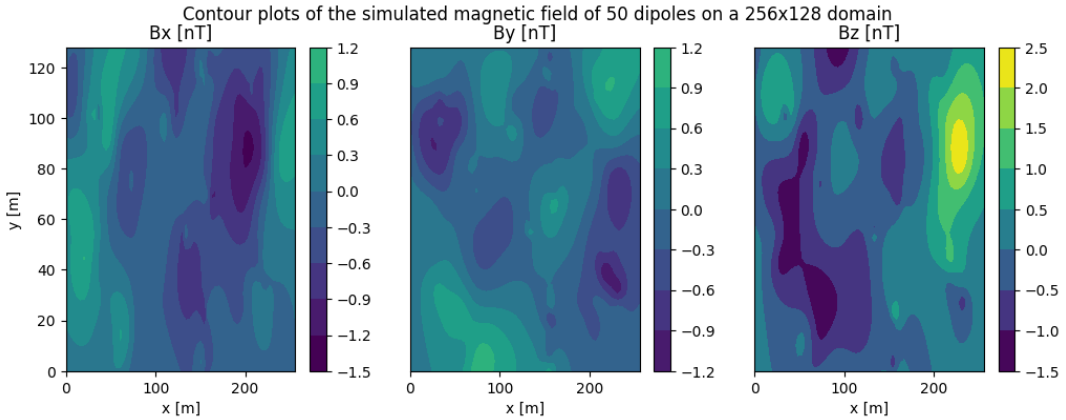


Figure 9.12: Contour plot of the simulated magnetic field in the x-y plane. The sources are uniformly distributed over the domain $[0,256] \times [0,128]$ at different depths. The magnetic field is measured at the surface.

The wavelet transformation of the two-dimensional Orthogonal Anderson Wavelet family is applied to the measured magnetic field. The resulting spectrum, shown in Figure 9.13, shows the intensities for the different wavelets. In the two-dimensional case, the spectrum has shifted in two dimensions and is, therefore, a three-dimensional plot. To visualize the spectrum, we have plotted the spectrum for each depth separately.

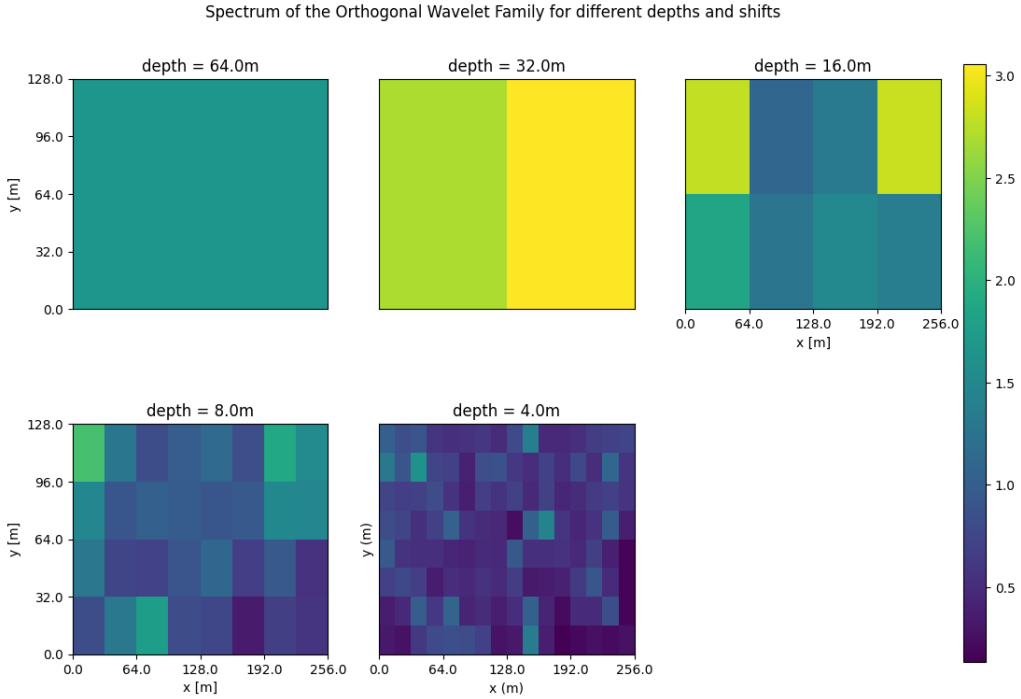


Figure 9.13: Spectrum of the measured magnetic field. The spectrum is plotted for each depth separately.

The coefficients, whose norm is visualized in the spectrum (Figure 9.13), can be used to reconstruct the signal. This is done by multiplying the coefficients with the corresponding wavelet and summing over all wavelets. The reconstructed magnetic field is shown in Figure 9.14 and Figure 9.15. The error between the reconstructed and the simulated magnetic field is shown in Figure 9.16.

An error analysis for the reconstructed magnetic field provides valuable insight into the accuracy of our reconstruction method. As can be seen in Figure 9.16, the general structure of the simulated magnetic field (Figure 9.12) is visible in the reconstruction. This is indicative of a successful retention of the signal's most important characteristics, demonstrating that our reconstruction method effectively captures the primary aspects of the magnetic field.

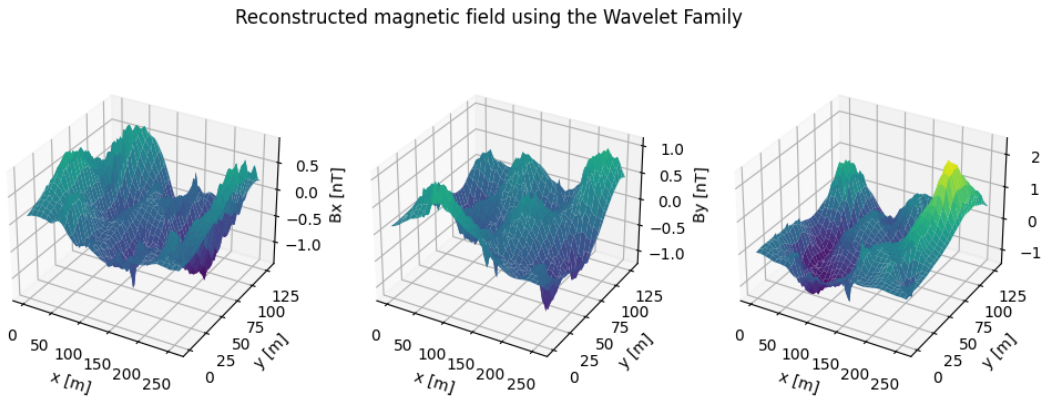


Figure 9.14: Reconstructed magnetic field. The reconstruction is done by multiplying the coefficients with the corresponding wavelet and summing over all wavelets.

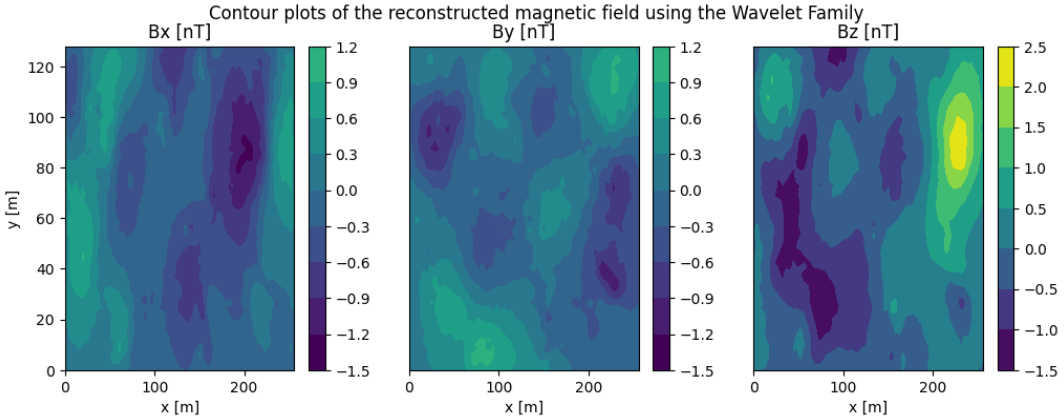


Figure 9.15: Contour plot of the reconstructed magnetic field.

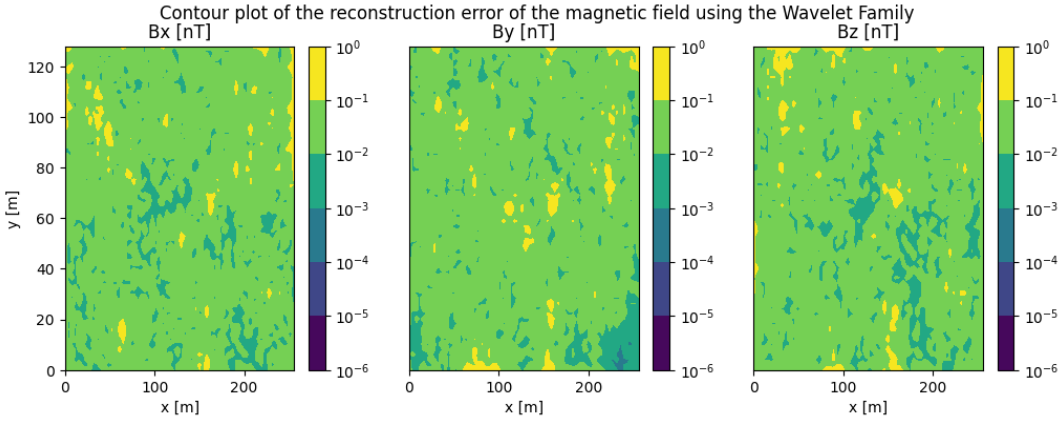


Figure 9.16: Contour plot of the error between the reconstructed and the simulated magnetic field.

Moreover, the location of the maxima and minima in the reconstructed field closely align with those in the simulated field, implying that the method can accurately pinpoint the regions of highest and lowest magnetic field strength. This precision in locating the magnetic field's extrema is a crucial attribute, as these are the regions of greatest interest in several applications, such as in the detection of magnetic anomalies in geophysical surveys.

However, it should be noted that the reconstructed solution appears slightly less smooth than the original. This could potentially be due to the inclusion of wavelets that represent sources close to the surface, which may introduce high-frequency components and thus result in a less smooth reconstruction. A smoother solution might be obtained by limiting the wavelet search space and neglecting these particular wavelets, which could help to reduce the presence of unwanted, high-frequency artifacts in the reconstructed field.

On a quantitative note, the average error between the reconstructed and simulated magnetic fields is approximately 2%. While this is generally a low error level, some local peaks in the error can be observed. In conclusion, the results depicted in Figures 9.14 and 9.16 show promising fidelity in the magnetic field reconstruction using our wavelet-based method, albeit with room for further improvement in the smoothness of the solution.

9.7. Discussion and Conclusion

In this chapter, we introduced a new method for the detection and localization of sources in the magnetic field. The method is based on the Anderson functions, which are a set of three functions that form a basis for the magnetic field of a single dipole on a line. We constructed a wavelet family of these functions, allowing us to detect and localize multiple sources at the same time. For a good reconstruction of signals, it is required that wavelets do not overlap. We therefore orthogonalized the wavelet family using the Gram-Schmidt procedure and stored the orthogonalized wavelet functions and the orthogonalization coefficients in the matrix Q and R of our QR-decomposition respectively. The orthogonal wavelet family is then given by Q and the orthogonalization coefficients found by projecting the original signal onto the orthogonal wavelets are given by Q.

The combination of wavelet coefficient and wavelet function can be used to reconstruct the signal. We have seen that, depending on the number of wavelets used, the reconstructed signal can be a good approximation of the original signal. This reconstruction takes place in the orthogonal wavelet domain. The coefficients can be transformed to the original domain by multiplying them with the inverse of the orthogonalization matrix R. The wavelets in this domain are derived by shifting and scaling the original Anderson functions, and thus form a basis for the magnetic field of multiple dipoles on a line. In theory, it should be possible to find a translation from the wavelet coefficients to the position and magnetic moment of the dipoles that correspond with the wavelet coefficients. However, we have not been able to find such a translation in every case yet. We were able to find the magnetic moments and positions of the dipoles for the case where the CPA (Closest Point of Approach) is known. This is an important step to use the method as an alternative to other inversion methods such as Equivalent Layer. Further research is required to find a translation for the general case.

After the reconstruction of several one-dimensional signals, we extended the method to two dimensions. We first derived the two-dimensional Anderson function, which is a function that forms a basis for the magnetic field of a single dipole on a plane. This function is then shifted and scaled to obtain a two-dimensional wavelet family. The wavelet family is again orthonormalized, which can be a time-intensive procedure. We, therefore, suggest to store the orthonormalized wavelet families. This can be done by pre-computing the wavelet families on regular grids and interpolating the measurement data to the grid that best matches the measurement data. The measurement data can then be projected on the orthonormal wavelet family, to find the wavelet coefficients that represent the signal. Simulations show that these two-dimensional Anderson wavelets can be used to reconstruct the magnetic field of multiple dipoles on a plane reasonably well. Future work could focus on optimizing the wavelet selection process to further minimize the reconstruction error.

10

Upward Continuation

Upward continuation is a method to estimate the magnetic field at a certain height above the ground. The method uses properties of a magnetic field and measurements at lower altitudes to do this. The benefit is that the magnetic field can be estimated at a height where no measurements are available. This is useful, for example, when the magnetic field is measured just above the surface, but one wants to evaluate the magnetic field at a certain height. A nice benefit is that Upward Continuation makes it unnecessary to take measurements at different altitudes, reducing the area to be covered by measurements from 3D to 2D. This is beneficial since taking measurements in large areas is expensive and time-consuming.

An important condition for Upward Continuation is that the sources of the field must be below the measurement plane. This is a good assumption for the magnetic field on Earth in the context of aeromagnetic surveys since the main sources of the magnetic field are the Earth's core and the magnetized rocks in the crust. These are all located below a measurement plane in the air.

10.1. Upward continuing the Potential Field

An essential principle derived in [5, p. 313] is that a potential field U at a point within a region R can be calculated as an integral over the surface enclosing this region ∂R. This follows from Green's third identity (Section 2.4.1) and the assumption that U is harmonic, continuous, and continuously differentiable on R. A textbook example of the region R surrounding a point P is a hemisphere of radius r. The field defined at the points inside the hemisphere can be represented as an integral over the surface (Figure 10.1).

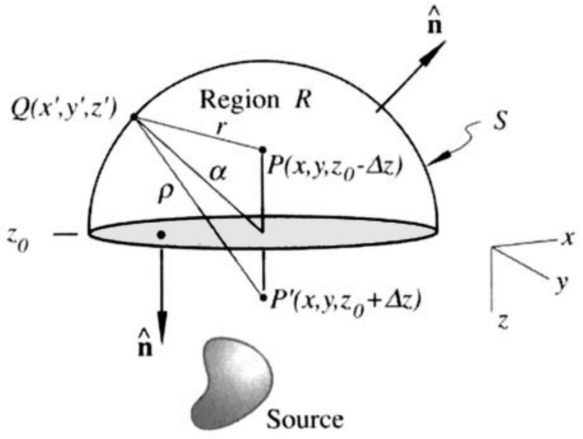


Figure 10.1: Hemisphere of radius r enclosing a point P. The region R is the volume enclosed by the hemisphere. If a function is harmonic throughout R, then it can be evaluated at any point in R by integrating over the surface $\partial R = S$. No knowledge is required about the sources of the field, except that they are located outside the hemisphere. This figure is adapted from [5, p. 314].

In practice, this principle is used to calculate the potential field in an unknown point $P \in \Omega \subset \mathbb{R}^3$ by using the potential field on the surface $\partial\Omega$ enclosing the domain of interest Ω. Unfortunately, measuring the potential field in a hemisphere around the domain of interest is often not feasible. Luckily, it is possible to increase the radius of the hemisphere. As $r \to \infty$, the influence of the spherical parts of the hemisphere on the points of interest becomes negligible, and the flat part of the hemisphere turns into an infinite plane that can be integrated over to obtain the potential field in the points of interest. This step is only allowed if the sources lie outside the hemisphere. The final expression derived in [5, pp. 315–316] is:

$$U(P) = U(x, y, z_0 + \Delta z) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{U(x', y', z_0)\Delta z}{2\pi[(x-x')^2 + (y-y')^2 + \Delta z^2]^{3/2}}\, dx'\, dy', \tag{10.1}$$

where z_0 is the height of the plane, and Δz is the distance between the plane and the point P. $\Delta z > 0$ if the sources lie below the plane. An illustration to visualize this principle is shown in Figure 10.2.

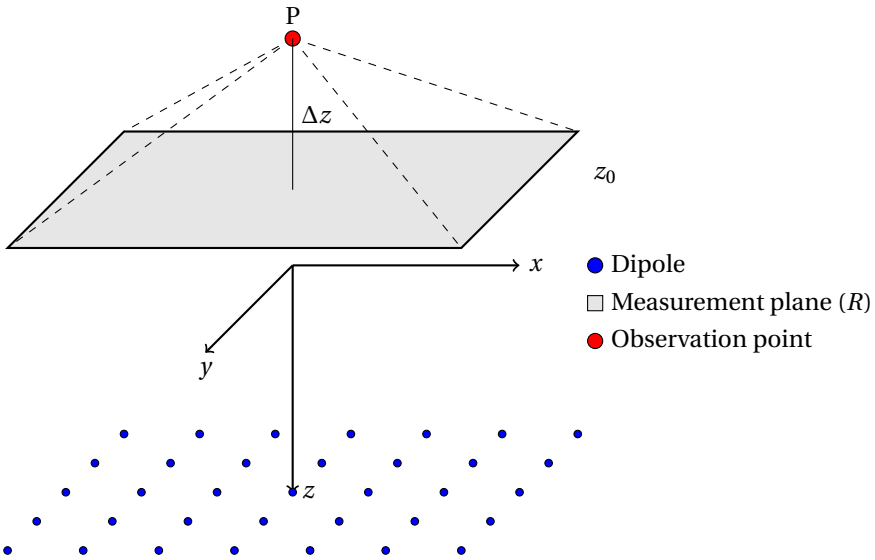


Figure 10.2: The principle of Upward Continuation. The potential or magnetic induction field at a point P above a plane can be calculated by integrating over the plane. The sources of the field are located below the plane.

It is easy to see that Equation 10.1 is a special case of the general formula for the potential field in a point P in a region R:

$$U(P) = \int_{\partial R} \frac{U(x', y', z')}{2\pi[(x-x')^2 + (y-y')^2 + (z-z')^2]^{3/2}}\, dx'\, dy'\, dz'. \tag{10.2}$$

10.2. Upward continuing the Magnetic Induction Field

We validate if we can apply the same principle of Upward Continuation to the magnetic induction field as we did to the potential field. In Section 2.4 it is shown that the magnetic induction field $\mathbf{B}$ is harmonic under certain circumstances. The magnetic induction field can be calculated from the potential field using the following formula:

$$\mathbf{B} = -\nabla U. \tag{10.3}$$

It remains to be shown if the magnetic induction field can be upward continued using the same formula as the potential field. The magnetic induction field is a vector field, so the integral in Equation 10.1 is not valid. However, the magnetic induction field can be written as a sum of three scalar fields:

$$\mathbf{B} = B_x\mathbf{i} + B_y\mathbf{j} + B_z\mathbf{k}, \tag{10.4}$$

where $\mathbf{i}$, $\mathbf{j}$, and $\mathbf{k}$ are the unit vectors in the x, y, and z direction respectively. The magnetic field can be upward continued using the same formula as the scalar potential field U, but applying the same formula separately on each of the three components of the magnetic field:

$$\mathbf{B}_i(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{B}_i(x', y', z_0)\Delta z}{2\pi[(x-x')^2 + (y-y')^2 + \Delta z^2]^{3/2}}\, dx'\, dy'. \tag{10.5}$$

To prove this, we can use the following formula for the gradient of a scalar field:

$$\nabla U = \frac{\partial U}{\partial x}\mathbf{i} + \frac{\partial U}{\partial y}\mathbf{j} + \frac{\partial U}{\partial z}\mathbf{k}, \tag{10.6}$$

and thus

$$\mathbf{B}_i = -\frac{\partial U}{\partial i}, \tag{10.7}$$

where i is the x, y, or z direction.

It remains to be shown that $\mathbf{B}_i(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \frac{\mathbf{B}_i(x',y',z_0)\Delta z}{2\pi[(x-x')^2+(y-y')^2+\Delta z^2]^{3/2}}\,dx'\,dy'$. This can be shown by differentiating Equation 10.2 with respect to i. For $i = x$, the x-component of the **B**-field, this can be done as follows:

$$B_x = -\frac{\partial U}{\partial x} = \frac{\partial}{\partial x}\int_{\partial R}\frac{-U(x',y',z')}{2\pi\mathbf{r}^3}\,dx'\,dy'\,dz' \tag{10.8}$$

$$= \int_{\partial R} -U(x',y',z')\frac{\partial}{\partial x}\frac{1}{2\pi\mathbf{r}^3}\,dx'\,dy'\,dz' \tag{10.9}$$

$$= -\int_{\partial R}\frac{\partial U(x',y',z')}{\partial x}\frac{1}{2\pi\mathbf{r}^3}\,dx'\,dy'\,dz' + \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\left[\frac{-U(x,y',z')}{2\pi\mathbf{r}^3}\right]_{x=-\infty}^{x=\infty} dy'\,dz' \tag{10.10}$$

$$= -\int_{\partial R}\frac{\partial U(x',y',z')}{\partial x}\frac{1}{2\pi\mathbf{r}^3}\,dx'\,dy'\,dz' \tag{10.11}$$

$$= \int_{\partial R}\frac{B_x(x',y',z')}{2\pi\mathbf{r}^3}\,dx'\,dy'\,dz', \tag{10.12}$$

where

$$\mathbf{r} = \mathbf{r}(x',y',z') = \sqrt{(x-x')^2+(y-y')^2+(z-z')^2}. \tag{10.13}$$

Equation 10.10 follows from applying the integration by parts formula to the inner integral over x'. The term $\left[\frac{-U(x,y',z')}{2\pi\mathbf{r}^3}\right]$ is equal to zero for $x = \pm\infty$ due to the factor $\mathbf{r}^3$ in the denominator. The same result can be obtained for the y and z components of the magnetic field. The same result can be obtained by proving that the components of the magnetic field are harmonic. This is done in Section 2.4. We can therefore conclude that the magnetic field can be upward continued using the same formula as the potential field by applying the formula separately on each of the three components of the magnetic field.

10.2.1. Upward continuation of the magnetic field in a 2D plane

In Section 10.2 it is shown that the magnetic field can be upward continued using the same formula as the potential field. However, for every point of interest P, the integral must be calculated over the entire surface ∂R. This is not feasible for large surfaces. In this section, it is shown that the magnetic field can be upward continued in a 2D plane using Fourier Transforms.

10.2.2. Upward continuing a constant field

Before we introduce the Fourier Transform for Upward Continuation, we investigate the case where the magnetic field is constant over the entire surface ∂R. This is a special case that does not occur often in nature, but an important case to consider when working with the Fourier Transform on truncated surfaces.

Assume that the magnetic induction field **B** is constant over the entire surface ∂R with a value of $\mathbf{B}_c$. The magnetic field can be upward continued using the following formula:

$$\mathbf{B}(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{\mathbf{B}_c(x',y',z_0)\Delta z}{2\pi[(x-x')^2+(y-y')^2+\Delta z^2]^{3/2}}\,dx'\,dy' \tag{10.14}$$

$$= \frac{\mathbf{B}_c\Delta z}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}\frac{1}{[(x-x')^2+(y-y')^2+\Delta z^2]^{3/2}}\,dx'\,dy' \tag{10.15}$$

$$= \frac{\mathbf{B}_c\Delta z}{2\pi}\int_{-\infty}^{\infty}\frac{2}{(y-y')^2+\Delta z^2}\,dy' \tag{10.16}$$

$$= \frac{\mathbf{B}_c\Delta z}{2\pi}\frac{2\pi}{\Delta z} = \mathbf{B}_c. \tag{10.17}$$

Therefore, we can conclude that a constant magnetic field is not affected by Upward Continuation.

That the values of a constant function are not dampened in Upward Continuation is not surprising. A constant magnetic field requires the field lines to be parallel, in the same direction, and equally spaced on the whole infinite plane. This is impossible for dipole sources since the field lines of dipoles are closed loops. It would require an infinite number of magnetic monopoles to create a constant magnetic field on an infinite plane. An infinitely strong dipole source placed at an infinite distance would also do the trick, but this is not physically possible.

Although constant magnetic fields on an infinite plane do not exist in nature, it seems that they can occur when taking measurements in the field. This is mainly because the magnetic field is measured with samples on a finite domain. The truncated domain allows for field lines that are non-closed loops.

For the same reason as for constant fields, fields with a non-zero mean value are also less affected by Upward Continuation. A field with a non-zero mean can be separated into a constant field and a field with a zero mean value. The constant field is not affected by Upward Continuation, while the field with a zero mean value is. Especially when taking measurements on a finite domain, fields with a non-zero mean occur more often than completely constant fields. For example, when waves of lower spatial frequencies do not completely fall within the domain, the magnetic field has a non-zero mean value. These lower spatial frequency signals are produced by sources that are much deeper than the size of the domain.

The difficulty here is that it is not possible to determine the distance to this deep source by using the observation points. Without the distance to the source, we cannot use the property that the magnetic field decays with a third power of the distance, and therefore the dampening of the field cannot be determined. For this reason, it is preferable to do computations on fields that are zero-mean. This can be achieved by subtracting the mean value of the magnetic field from the magnetic field, leading to an anomaly field where the mean is used as a reference value. By doing this, the exact field at evaluation points can no longer be determined, but the dampening of the anomaly field can be determined for different heights of Upward Continuation.

10.2.3. Upward continuing a vector magnetic field in a 2D-plane

The theory for Upward Continuation of the potential in a 2D-plane has been described in [5, pp. 316–319]. Here they use the Fourier Transform as a dampening function that dampens higher spatial frequencies more than lower spatial frequencies when upward continuing the potential. The same theory can be applied to the magnetic field, as we have shown in Section 10.2 that the magnetic field itself can also be upward continued.

First, we note that Equation 10.5 can be rewritten as a convolution:

$$\mathbf{B}_i(P) = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} \mathbf{B}_i(x', y', z_0)\Phi_u(x - x', y - y', \Delta z)\, dx'\, dy', \tag{10.18}$$

where Φ_u is given by

$$\Phi_u(x, y, \Delta z) = \frac{\Delta z}{2\pi[x^2 + y^2 + \Delta z^2]^{3/2}} = -\frac{1}{2\pi}\frac{\partial}{\partial \Delta z}\frac{1}{\sqrt{x^2 + y^2 + \Delta z^2}}. \tag{10.19}$$

We can also write this in a compact form:

$$\mathbf{B}_u = \mathbf{B} \circ \Phi_u, \tag{10.20}$$

where $\mathbf{B} \circ \Phi_u$ is the convolution of $\mathbf{B}$ and Φ_u and $\mathbf{B}_u$ is the upward continued magnetic field.

The upward-continued value of a 2D horizontal plane can now be calculated using convolution 10.20. A convolution in the time domain is equivalent to a multiplication in the frequency domain. Let $\mathscr{F}[\mathbf{B}]$ be the vector containing the Fourier Transforms of the magnetic field components $\mathbf{B}$ and $\mathscr{F}[\Phi_u]$ be the Fourier Transform of the upward-continuation function Φ_u. Then the Fourier Transform of the upward-continued magnetic field is given by

$$\mathscr{F}[\mathbf{B}_u] = \mathscr{F}[\mathbf{B}]\mathscr{F}[\Phi_u], \tag{10.21}$$

with the following relation between the Time and Frequency domain:

$$\mathbf{B}_u = \mathbf{B} \circ \Phi_u \quad \Leftrightarrow \quad \mathscr{F}[\mathbf{B}_u] = \mathscr{F}[\mathbf{B}] \cdot \mathscr{F}[\Phi_u]. \tag{10.22}$$

This convolution can be calculated using the Fourier Transform. The Fourier Transform of $\mathbf{B}$ is dependent on the observations and is therefore not given here. The Fourier Transform of Φ_u is given by

$$\mathscr{F}[\Phi_u] = -\frac{1}{2\pi}\frac{\partial}{\partial \Delta z}\mathscr{F}\left[\frac{1}{|r|}\right] \tag{10.23}$$

$$= -\frac{\partial}{\partial \Delta z}\frac{e^{-|k|\Delta z}}{|k|} \tag{10.24}$$

$$= e^{-\Delta z|k|}. \tag{10.25}$$

This makes the computation of the upward-continued magnetic field very fast. The Fourier Transform $\mathscr{F}[\mathbf{B}]$ of the measured magnetic field can be calculated using for example the Fast Fourier Transform. The Fourier Transform $\mathscr{F}[\Phi_u]$ of the upward-continuation function Φ_u can be pre-computed for a range of heights. Upward continuing the field to different heights is then a matter of multiplying the Fourier Transform of the magnetic field with the Fourier Transform of the upward-continuation function. The upward continued magnetic field can be calculated using for example the Inverse Fast Fourier Transform.

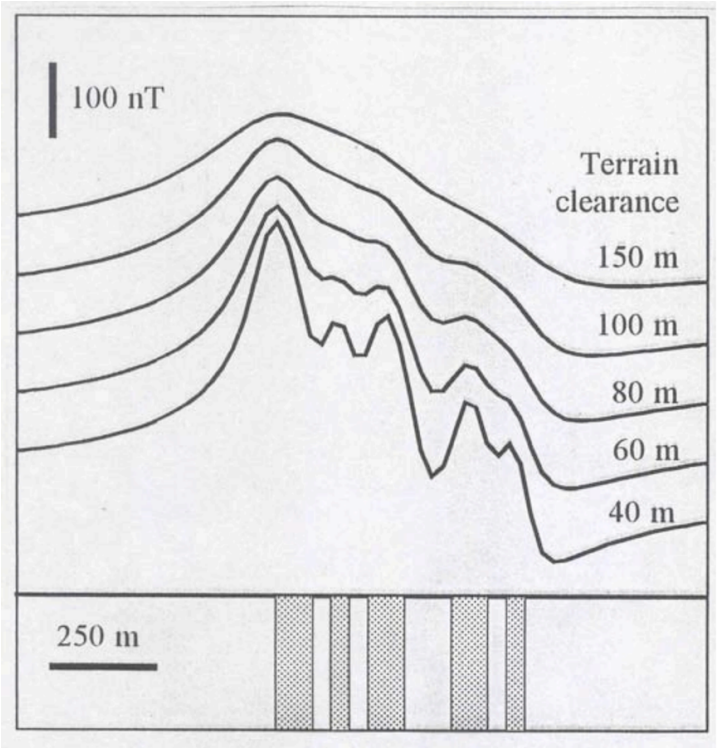


Figure 10.3: The effect of altitude on the spatial variation in the magnetic field. This figure is taken from [54, pp. 4–8].

The fact that higher frequencies are more dampened than lower frequencies when upward continuing the magnetic field has some consequences for the process of magnetic map-making. If one wants to make a high-resolution map of the magnetic field that captures detailed variations in the magnetic field, then the measurements should be taken at a low altitude close to the sources. Figure 10.3 shows the effect of altitude on the spatial variation in magnetic field measurements of a simulated geological magnetic field. It emphasizes the importance of maintaining low terrain clearance, the distance between the sensor and the ground, when mapping areas with near-surface magnetic features. Reeves analyzed the theoretical magnetic profiles calculated for different flying heights over a simplified geological model and concluded that distinct peaks for individual sources are only identifiable when the terrain clearance is smaller than the dimensions of the feature under investigation[54, pp. 4–8]. This suggests that most anomalies appearing as single features at survey altitudes likely result from multiple, closely spaced magnetic units within the subsurface.

10.2.4. Upward continuing an anomaly field in a 2D plane

In Section 2.4 it is shown that the field $\Delta\mathbf{T}$ is approximately harmonic. This means that Equation 10.21 can also be used to upward continue an Anomaly Field $\Delta\mathbf{T}$ (Section 5.4). To prove this, the Fourier Transform of $\Delta\mathbf{T}$ is calculated. First, $\Delta\mathbf{T}^0$ is defined as

$$\Delta\mathbf{T}^0 = \mathbf{B}_a^0 \cdot \hat{\mathbf{B}}_e^0 = -\hat{f}^0 \cdot \nabla U_a^0, \tag{10.26}$$

where the superscript ^{0} indicates that the field is measured at height z_0, $\hat{f}^0$ is the unit vector in the direction of the magnetic field and ∇U_a^0 is the gradient of the anomaly potential U_a^0.

The Fourier Transform of $\Delta\mathbf{T}^0$ is given by

$$\begin{aligned}
\mathscr{F}[\Delta\mathbf{T}^0] &= \mathscr{F}[-\hat{f}^0 \cdot \nabla U_a^0] && (10.27)\\
&= -\hat{f}^0 \cdot \mathscr{F}[\nabla U_a^0] && (10.28)\\
&= (-\hat{f}^0 \cdot \kappa)\mathscr{F}[U_a^0], && (10.29)
\end{aligned}$$

where κ is the wavenumber and the first step is valid when $\hat{f}^0$ is constant over the area of interest.

Upward continuation of a scalar potential U_a^0 is given by

$$\mathscr{F}[U_a^z] = \mathscr{F}[U_a^0]e^{-|\kappa|\Delta z}, \tag{10.30}$$

where Δz is the height difference between the measured and the upward-continued field and superscript z indicates that the field is at height $z_0 + \Delta z$. Substituting Equation 10.30 into Equation 10.27 gives

$$\mathscr{F}[\Delta\mathbf{T}^0] = (-\hat{f}^0 \cdot \kappa)\frac{\mathscr{F}[U_a^z]}{e^{-|\kappa|\Delta z}}. \tag{10.31}$$

Analogous to Equation 10.27, one can find the Fourier Transform of $\Delta\mathbf{T}^z$:

$$\mathscr{F}[\Delta\mathbf{T}^z] = (-\hat{f}^z \cdot \kappa)\mathscr{F}[U_a^z], \tag{10.32}$$

where $\hat{f}^z$ is unit vector indicating the direction of the magnetic field at height $z_0 + \Delta z$. $\mathscr{F}[U_a^z]$ is unknown in this case, but can be expressed in known terms using Equation 10.31:

$$\begin{aligned}
\mathscr{F}[\Delta\mathbf{T}^z] &= (-\hat{f}^z \cdot \kappa)\mathscr{F}[U_a^z] && (10.33)\\
&= (-\hat{f}^z \cdot \kappa)\frac{1}{-\hat{f}^0 \cdot \kappa}\mathscr{F}[\Delta\mathbf{T}^0]e^{-|\kappa|\Delta z} && (10.34)\\
&= \frac{-\hat{f}^z \cdot \kappa}{-\hat{f}^0 \cdot \kappa}\mathscr{F}[\Delta\mathbf{T}^0]e^{-|\kappa|\Delta z}. && (10.35)
\end{aligned}$$

This shows that the Fourier Transform of the upward continued field $\Delta\mathbf{T}^z$ is equal to the Fourier Transform of the measured field $\Delta\mathbf{T}^0$ multiplied by a factor $\frac{-\hat{f}^z \cdot \kappa}{-\hat{f}^0 \cdot \kappa}$. The dampening factor $e^{-|\kappa|\Delta z}$ is dependent on the wavenumber κ and the height difference Δz. Higher spatial frequencies have a higher wavenumber and are therefore more dampened.

10.3. Discrete Upward Continuation

A drawback of Upward Continuation in integral form is that its computations are based on a continuous function. The integral over the infinite or truncated plane requires continuous measurements of the magnetic field. This is not possible in practice. It is only possible to measure the magnetic field at a discrete set of points. There are several ways to discretize the integral. In this section, we discuss two meshing methods. The first method assumes a regular grid. The second method uses a polar-coordinate-based mesh. The polar-coordinate-based mesh is more accurate than the regular grid when upward continuing a single point, but it is less practical for upward continuing a large area since it requires a specific distribution of the measurement points.

10.3.1. Upward Continuation on a regular grid

The most simple way to discretize the integral is to use a regular grid. Let $[x_1, \ldots, x_n] \times [y_1, \ldots, y_m]$ be a regular grid of points with spacing Δx and Δy in the x and y direction respectively. Then, the upward continued value at a point $(x_i, y_j, z_0 + \Delta z)$ can be approximated by a double sum over the values at points (x_k, y_l, z_0) for k, l in K, L:

$$U_h(i, j, z_0 + \Delta z) = \frac{\Delta z}{2\pi} \sum_{k \in K} \sum_{j \in J} \frac{U(x_k, x_l, z_0)}{[(x_i - x_k)^2 + (y_j - y_l)^2 + \Delta z^2]^{\frac{3}{2}}} \Delta x \Delta y, \tag{10.36}$$

where K and L are indices centered around i and j. This discretization is easy to implement, but not very accurate. The influence of a point (x_k, y_l, z_0) on the point $(x_i, y_j, z_0 + \Delta z)$ decreases with the third power of the distance. If Δz is chosen to be small, then almost the whole integral is dominated by the point (x_i, y_j, z_0) straight under the point of interest. At the same time, if $\Delta z \gg \Delta x, \Delta y$, then the domain of dependence of the integral is very large. This requires a large number of measurements to get a good approximation.

Despite these drawbacks, this method is still useful. Surveys are often done in lines, so that the points are close to each other in the x direction and have a constant spacing in the y direction: $\Delta y \gg \Delta x$. In this case, the main discretization error is the one in the y direction. This can be reduced by using a higher order interpolation scheme in the direction perpendicular to the lines. Canciani suggests [14] using linear interpolation in the x direction and cubic interpolation in the y direction. Interpolating the measurement points allows us to use a finer grid.

10.3.2. Upward Continuation on a logarithmic polar mesh

As stated before, the function that is integrated decays with the third power of the distance. This motivates the use of a polar coordinate system. The domain can be divided into several concentric circles with increasing radii. The smallest circle is centered underneath the point of interest and has a single measurement point. The next circle forms a ring around the first circle. On this ring, 3 points are spaced evenly. The number of points on the n-th ring is $2n + 1$. The radius of the rings can be chosen to be constant or to increase with the number of the ring. The latter is called a logarithmic mesh and is visualized in Figure 10.4.

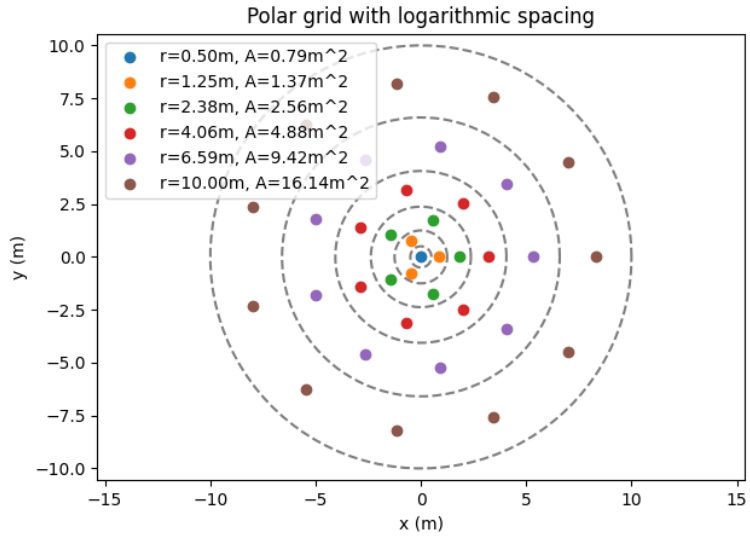


Figure 10.4: A logarithmic mesh with 5 rings, initial radius $r_0 = 0.5$ and radius $r_{n+1} = r_n + r_0 * dr^{n+1}$ with radius increment $dr = 1.5$.

The logarithmic mesh gives a very good approximation of the integral. The finer mesh on the inside allows for good accuracy for small Δz. The coarser mesh on the outside allows for fast computation and also for a good approximation for large Δz. Unfortunately, the mesh is not practical, because the mesh is centered around the point of interest. This means that the mesh has to be moved for every point of interest. For a plane survey, this means that the coarser points have to be moved around the plane, which ends up being similar to the regular grid, but with a much higher computational cost.

10.3.3. Linear and Cubic interpolation of irregularly spaced data

Irregularly spaced data can be interpolated to a regular grid using an interpolation scheme. Linear interpolation assumes that the function behaves linearly between two points. In the context of magnetic fields, linear interpolation is only useful for really small distances. Cubic interpolation assumes that the function behaves like a polynomial of degree 3, which is more in line with the behavior of a magnetic dipole field (see Equation 10.51).

The magnetic dipole field behaves like $1/r^3$ at great distances. As long as the plane of interest is not too close to the origin, the singularity at the origin can be ignored. The factor theorem can be applied to the denominator of Equation 10.51 to get a polynomial of positive degree and a factor $1/r^3$. The polynomial is of degree 3 so that cubic interpolation is a good approximation.

The underlying assumption that our field is a superposition of magnetic dipoles makes it smarter to limit ourselves to the maximum degree of 3 that the dipole field in Equation 10.51 has. A cubic function can be represented exactly with a cubic spline. Using higher-order interpolation schemes is possible, but it is not necessary and it does not improve the accuracy.

10.4. Domain of Dependence for Upward Continuation

Equation 10.5 assumes that the exact value of the magnetic field is known on an infinite plane below the point of interest P. It is impossible to measure the magnetic field on an infinite plane. Therefore, the magnetic field is measured on a finite plane. For now, it is assumed that the magnetic field can be measured continuously on a finite plane, this domain of dependence gives an upper bound on the Upward Continuation error depending on the size of the plane that is integrated over.

The analysis of the Domain of Dependence is done using a finite plane with a circular shape. The circular plane is chosen to have a center with the same x- and y-coordinates as the point of interest, but a different z-coordinate. The radius of the circle describes the minimum set of points that is used in the integral for Upward Continuation. The motivation for using a circular shape is that the influence of a point on the point of interest decreases with the third power of the distance. This means that, when for example using a square plane, the influence of the points on the corners is negligible compared to the influence of the points on the sides. If one wants to use the estimates from this section on a domain that is not circular, one can use the largest circle that fits in the domain to get a bound on the error.

10.4.1. An upper bound on the error

Equation 10.5 can also be written in polar coordinates using substitution $(x-x') = r\cos(\phi)$ and $(y-y') = r\sin(\phi)$:

$$\mathbf{B}(z_0+\Delta z) = \int_0^{\infty}\int_0^{2\pi} \frac{\mathbf{B}(r,\phi,z_0)\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}}\, r\, d\phi\, dr\,. \tag{10.37}$$

We define the error in the Upward Continuation as the difference between the exact value of the magnetic field and the computed value using a finite plane with a radius r_0:

$$\epsilon(P) = |\mathbf{B}(P) - \mathbf{B}_{r_0}(P)|\,, \tag{10.38}$$

where

$$\mathbf{B}_{r_0}(P) = \int_0^{r_0}\int_0^{2\pi} \frac{\mathbf{B}(r,\phi,z_0)\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}}\, r\, d\phi\, dr\,, \tag{10.39}$$

is the magnetic field computed using a finite plane with radius r_0. The maximum error is given by:

$$\epsilon(P) = \left|\mathbf{B}(P) - \mathbf{B}_{r_0}(P)\right| \tag{10.40}$$

$$= \left|\int_0^{\infty}\int_0^{2\pi} \frac{\mathbf{B}(r,\phi,z_0)\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}}\, r\, d\phi\, dr - \int_0^{r_0}\int_0^{2\pi} \frac{\mathbf{B}_{r_0}(r,\phi,z_0)\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}}\, r\, d\phi\, dr\right| \tag{10.41}$$

$$= \left|\int_{r_0}^{\infty}\int_0^{2\pi} \frac{\mathbf{B}(r,\phi,z_0)\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}}\, r\, d\phi\, dr\right|. \tag{10.42}$$

The value $\mathbf{B}(r,\phi,z_0)$ can be approximated by the maximum absolute value of the magnetic field on the open set $S = \{(r,\phi)|r \geq r_0 \wedge 0 \leq \phi < 2\pi\}$. Let $M = \max_{(r,\phi)\in S} |\mathbf{B}(r,\phi,z_0)|$ be the maximum absolute value of the

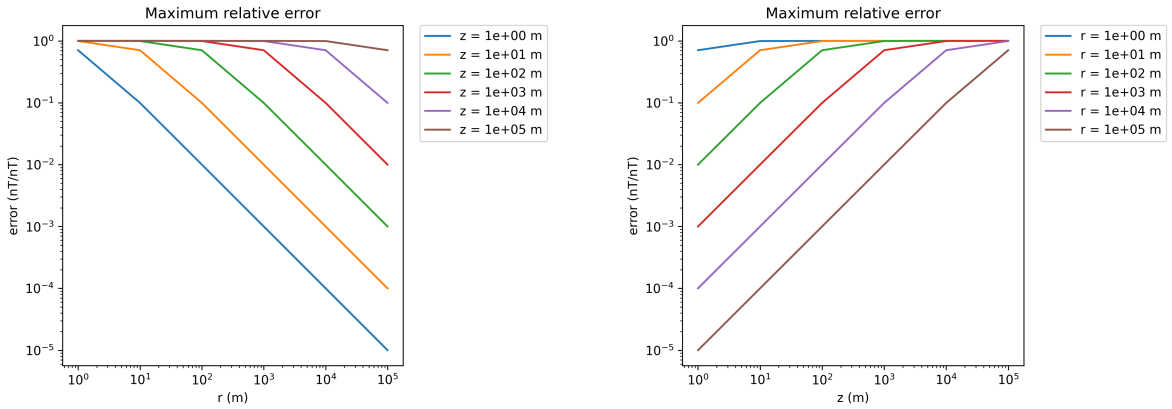
magnetic field outside the measured domain, but on the plane. Then, an upper bound for the error is given by:

$$\epsilon(P) \leq \int_{r_0}^{\infty} \int_0^{2\pi} \frac{M\Delta z}{2\pi[r^2+\Delta z^2]^{3/2}} r \, d\phi \, dr \tag{10.43}$$

$$= M\Delta z \int_{r_0}^{\infty} \frac{1}{[r^2+\Delta z^2]^{3/2}} r \, dr \tag{10.44}$$

$$= \frac{M\Delta z}{\sqrt{r_0^2+\Delta z^2}}, \tag{10.45}$$

for $\Delta z > 0$. An upper bound for the error is therefore proportional to the radius r_0 of the plane ∂R.



(a) Relative error of Upward Continuation for a single point as a function of the radius of the plane that is integrated over.

(b) Relative error of Upward Continuation for a single point as a function of the distance between the plane and the point of interest

Figure 10.5: Maximum relative error as a function of the radius of the plane and the distance between the plane and the point of interest.

The maximum error is given by Equation 10.43. We want to express the error as a relative error, independent of the strength of the magnetic field. To do so, we use the assumption that the maximum absolute of the magnetic field outside the circle of radius r_0 is less than the average absolute value $\overline{M}$ of the magnetic field on the circle of radius r_0:

$$M = \max_{(r,\phi)\in S} |\mathbf{B}(r,\phi,z_0)| \leq \frac{1}{2\pi r_0} \int_0^{2\pi} \int_0^{r_0} |\mathbf{B}(r,\phi,z_0)| \, dr \, d\phi = \overline{M}. \tag{10.46}$$

The maximum relative error $\overline{\epsilon(P)}$ is then given by:

$$\overline{\epsilon(P)} = \frac{1}{\overline{M}}\epsilon(P) = \frac{M}{\overline{M}} \frac{\Delta z}{\sqrt{r_0^2+\Delta z^2}} \leq \frac{\Delta z}{\sqrt{r_0^2+\Delta z^2}}. \tag{10.47}$$

Figure 10.5 shows the maximum relative error as a function of the radius of the plane and the distance between the plane and the point of interest. One can read from Figure 10.5a that the maximum relative error is proportional to the radius of the plane: one needs a radius r_0 of approximately 10^4m to have a relative error of 0.01 in a point that is 100m above the plane. To obtain an extra order of magnitude, the radius of the plane should be increased by a factor of 10. Relative in this case means that the error is divided by the average value of the magnetic field in the plane ∂R. If the point of interest has a magnetic field strength of 100nT, then the maximum error is 1nT for a radius of 10^4m and a distance of 100m between the plane and the point of interest. This is also immediately a motivation to do these computations on the magnetic anomaly field rather than the complete magnetic field: subtracting a constant from the magnetic field reduces the error by a constant since the value of M is reduced by the same constant. This is in line with what we have seen in Section 10.2.2. There, we noticed that computing the upward continued value of a constant magnetic field using a double integral over a plane is not a good idea, since we know that constant fields are not dampened. The integral can easily be computed analytically and the solution only depends on the distance between the plane and the point of interest. Therefore, subtracting constants from a magnetic field and performing

Upward Continuation separately on the magnetic anomaly field and the constant field reduces the error by a factor proportional to the constant:

$$\begin{aligned} \mathrm{UC}(\mathbf{B}) &= \mathrm{UC}(\overline{\mathbf{B}} + \mathbf{B}_A) && (10.48)\\ &= \mathrm{UC}(\overline{\mathbf{B}}) + \mathrm{UC}(\mathbf{B}_A) && (10.49)\\ &= \overline{\mathbf{B}} + \mathrm{UC}(\mathbf{B}_A)\,, && (10.50) \end{aligned}$$

where $\overline{\mathbf{B}}$ is the constant magnetic field and $\mathbf{B}_A$ is the magnetic anomaly field and UC is a shorthand notation for Upward Continuation. The error made by upward continuing the constant magnetic field $\overline{\mathbf{B}}$ is zero, and the maximum error made by upward continuing the magnetic anomaly field $\mathbf{B}_A$ is proportional to the maximum value of the magnetic anomaly field. This is again a good motivation to perform computations on an anomaly field with zero mean value.

10.4.2. A lower bound on the error

The estimation of the maximum value M is an obstacle in retrieving a stricter bound on the error. In the above example, M is assumed to be the maximum magnitude in the plane ∂R. This describes a worst-case scenario: the magnetic field can attain this maximum value in the whole plane. This is not the case in general. In this section, we consider another case: the magnetic field is induced by a single dipole, which leads to a lower bound on the error.

The bound $M = \max_{(r,\phi)\in S} |\mathbf{B}(r,\phi,z_0)|$ helps to find an upper bound on the error. However, the bound M is not a good estimation for the value of the magnetic field in a plane. M is only a good estimation if the magnetic field is constant in the plane. A constant magnetic field only occurs if the magnetic field is induced by a dipole position at infinity, or if dipoles are positioned infinitesimally close to each and the plane. This is not the case in general, but it helps us to physically understand the bound: the magnetic field of a dipole degrades with a third power of the distance. The number of points in the plane ∂R is proportional to r_0^2. Therefore, the error is proportional to r_0.

The magnetic field induced by a single dipole is given by:

$$\mathbf{B}(\mathbf{r}) = \frac{\mu_0}{4\pi} \frac{3(\mathbf{m}\cdot\mathbf{r})\mathbf{r} - \mathbf{m}|\mathbf{r}|^2}{|\mathbf{r}|^5} \tag{10.51}$$

where $\mathbf{r}$ is the position vector of the point of interest and $\mathbf{m}$ is the dipole moment. At a great distance, this behaves as $1/r^3$ where $r = |\mathbf{r}|$. If the plane ∂R is located at a distance z_d from the dipole, and the point of interest is located at a distance z_0 from the plane, then the influence of the dipole on an infinite plane is similar to the contribution of the points on the plane on the point of interest: the strength decays with a third power of the distance, and the number of points is again proportional to the radius of the plane. The approximated field strength in the plane is therefore given by:

$$\begin{aligned} \mathbf{B}(r,\phi,z_0) &= \frac{\mu_0}{4\pi} \frac{3((\mathbf{m}\cdot\mathbf{r}')\mathbf{r}') - \mathbf{m}|\mathbf{r}'|^2}{|\mathbf{r}'|^5} && (10.52)\\ &\approx \frac{M_d}{[r^2 + z_d^2]^{3/2}}\,, && (10.53) \end{aligned}$$

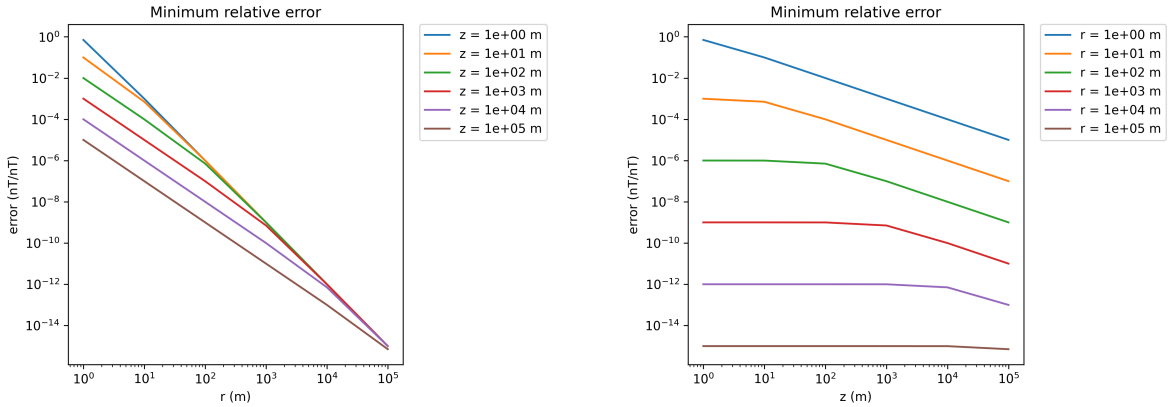
where $r' = \sqrt{r^2 + z_d^2}$.

Substituting Equation 10.52 in Equation 10.40 gives:

$$\begin{aligned} \epsilon(P) &= \int_{r_0}^{\infty} \int_0^{2\pi} \frac{\mathbf{B}(r,\phi,z_0)\Delta z}{2\pi[r^2 + \Delta z^2]^{3/2}}\, r\, d\phi\, dr && (10.54)\\ &= M_d \Delta z \int_{r_0}^{\infty} \frac{1}{[r^2 + \Delta z^2]^{3/2}[r^2 + z_d^2]^{3/2}}\, r\, dr && (10.55)\\ &\propto \frac{1}{r^4}\,, && (10.56) \end{aligned}$$

where the proportionality follows if $|\Delta z| \approx |z_d|$.

In Figure 10.6 the relative error obtained by upward continuing the magnetic field of a single dipole, is plotted as a function of the radius of the plane and the distance between the plane and the point of interest.



(a) Relative error of Upward Continuation for a single point as a function of the radius of the plane that is integrated over.

(b) Relative error of Upward Continuation for a single point as a function of the distance between the plane and the point of interest

Figure 10.6: Minimum relative error as a function of the radius of the plane and the distance between the plane and the point of interest. The minimum error is achieved when the field is generated by a single dipole underneath the point of interest. The error is proportional to the fourth power of the radius of the plane.

This is a much better bound than the one given by Equation 10.43. Note that this estimate is only valid in the ideal situation of a single dipole at depth z_0 and a point of interest that is located straight above the dipole at a distance z_0 from the plane. This implies that the theoretical maximum decrease $1/r^3$ of the dipole is reached in all directions in the plane. This is the theoretical minimum error since a magnetic dipole field cannot decay faster.

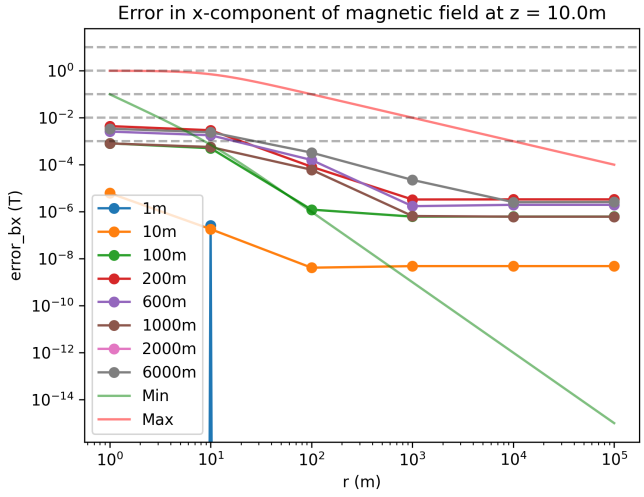


Figure 10.7: Numerical computation of the absolute error in the x-component of the magnetic field as a function of the radius of the plane for different depths of a single dipole.

Figure 10.7 shows the absolute error in the x-component of the magnetic field as a function of the radius of the plane for different depths of a single dipole. The error is computed by upward continuing the magnetic field of a single dipole to a single point of interest that lies 10 meters above the plane and comparing it to the analytical solution. The solution is approximated with the discrete Upward Continuation method using the logarithmic-polar mesh, as described in Section 10.3.2. For small radii, the minimum error estimate seems to be in good agreement with the numerical error. However, for larger radii, the numerical error does not attain the minimum error estimate of Section 10.4.2. This can be explained by the fact that the numerical error is computed using a discretized mesh, with a finite number of points and therefore a limited precision in the computation of the integral.

III

Benchmarking and Validation

11

Benchmarking

In this chapter, we discuss the benchmarking of the various map-making techniques. The goal of the benchmarking is to compare the performance of the various map-making techniques. We have chosen to benchmark the techniques by applying them to a simulated magnetic field. A `MagMap` Python package has been developed to simulate magnetic fields and apply map-making techniques to these fields. This chapter starts with a description of the modelling tool and the simulation parameters (Section 11.1). This is followed by a description of the setup of the benchmarking experiment (Section 11.2). Thereafter, we introduce several metrics that we use to compare the performance of the map-making techniques (Section 11.3). Finally, we discuss the results of the benchmarking experiment on the simulated magnetic field (Section 11.4).

11.1. Simulating Earth's Magnetic Field

11.1.1. Modelling Tool

This report contains various map-making techniques that can be applied to magnetic fields. As part of this research, we have implemented these techniques and combined them in a single Python package called `MagMap`. An interface is defined for each method, which allows us to call the methods with similar input parameters. This modular approach makes it easy to switch between the methods and compare the results.

The Python package includes classes and tools that simplify working with magnetic fields. One such class is the 'Field' class, which can accommodate various sources, such as single dipoles, layers of dipoles, noise terms, and real-world data from IGRF. A unique feature of the Field class is that it allows fixing the seed of the random number generator, which ensures the creation of reproducible random fields.

Fields can be easily combined and exported to a compact JSON file. Moreover, we can assess fields at specific points by providing a path of x-, y-, and z-coordinates to the Field object. The three magnetic field components at each point are returned, and we can specify an additional parameter to obtain only a Total-field measurement.

All the magnetic techniques implemented in this report can be directly applied to an object of class Fields. These methods apply to Total-field measurements or vector measurements.

11.1.2. Simulation Parameters

Many setups can be used to demonstrate the performance of the various map-making techniques. The performance of the map-making techniques is dependent on the type of field that is being mapped. To make the results comparable, we have chosen to limit ourselves to fields that are close to natural fields on a regional scale. This field is generated using a deep dipole to represent the core of the Earth, and a grid of dipoles to represent the crust.

Core-field models and crustal-field models are often represented in a geodetic coordinate system. The methods that we apply are defined in cartesian coordinates, preferably centered around the origin. Therefore, we have to convert the geodetic coordinates to cartesian coordinates. We chose to represent the core field by a single dipole which is placed at a depth of 6371 km straight underneath the origin of the cartesian coordinate system. The dipole is given a magnitude such that it matches the Earth's magnetic field at the Equator (3.12×10^4 nT). The dipole is rotated by adjusting its moment, such that the value at $(0,0,0)$ is the magnetic field at the center of the Netherlands (52.0° N, 4.3° E).

The crustal field consists of a grid of 61 × 61 dipoles with a spacing of 2 km. In total, they span a region of 120 km by 120 km centered around the origin. The dipoles in the grid are assigned a random depth ranging from 100 m to 9900 m. The random depth is taken from a triangular distribution, which ensures that the dipoles are more likely to be located at shallow depths, resulting in enough spatial variation in the map to demonstrate the performance of the map-making techniques. The intensity of the dipoles is chosen such their intensity at the surface is uniformly distributed between −300 nT and 300 nT.

Simulation Area and Resolution

The area of the regional magnetic field we simulate has a spatial extent of tens of kilometers. All the sources are positioned in a region of 120 km by 120 km centered around the origin. The minimum depth of the sources in the simulated crustal field is 100 m, and the maximum depth is 9900 m, making the field detailed enough to represent a regional magnetic field.

The resolution of the simulated field is determined by the size of the grid of dipoles. The grid of dipoles has a spacing of 2 km at a depth of at least 100 m, therefore, the wavelengths originating from the grid of dipoles are at least 4 km. It is important to note that the resolution of the simulated field is not the same as the resolution of the map. The resolution of the map is determined by the spacing of the measurement locations and how well they represent the field.

Simulation Noise

To make the simulations a bit more realistic, we add noise to the simulated field, as if it were measured in the field. We use a simple, white noise model, which is added to the simulated field. Each measurement of the field is perturbed by a random number drawn from a normal distribution with a mean of 0 nT and a standard deviation of 20 nT.

The noise model we used is very basic. We have seen in Chapter 7, that the compensation of noise is a complex problem. A white noise model is a simplification of the real noise that is present in the field, which is often colored, platform-dependent, and direction dependent.

Simulated Field

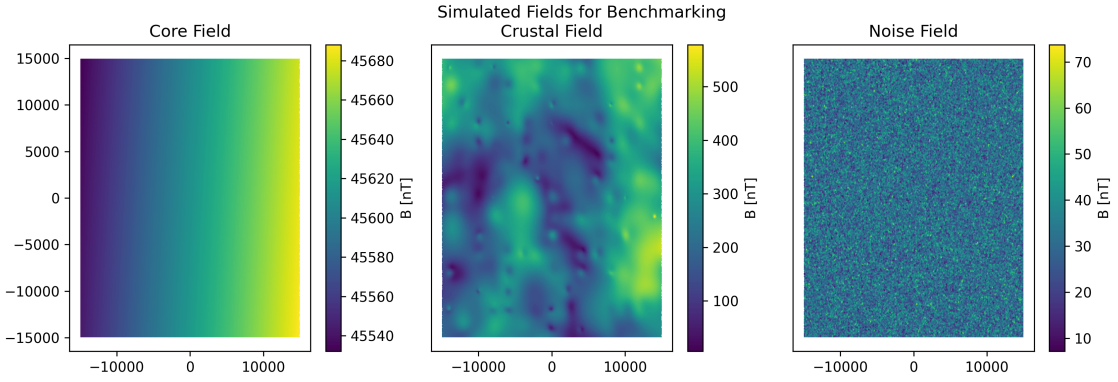


Figure 11.1: Total-field intensities of the magnetic fields that are simulated for the benchmarking experiment. The left panel shows the simulated core field, the middle panel the simulated crustal field, and the right panel shows an example of random noise that is added to the simulated field.

Using the parameters described above, we have simulated a regional magnetic field. The simulated field consists of a core field, a crustal field, and noise. These fields are illustrated in Figure 11.1. The core field changes slowly over the region of interest, while the crustal field has higher spatial variation. The noise is random and has no spatial correlation.

11.2. Setup of the Benchmarking Experiment

In the previous section, we have described how we simulate a regional magnetic field. In this section, the simulated magnetic fields are used as input for a benchmarking experiment. We have made a benchmarker as a tool to evaluate and compare the performance of different interpolation and extrapolation techniques of magnetic fields for a given set of measurements. The benchmarker takes in a set of input data, such as training

data and test data, and a list of map-making techniques (interpolators or extrapolators) to be evaluated. It then runs each map-making technique on the input data and interpolates (or extrapolates) the magnetic field from the training measurements to the positions of the test measurements. The benchmarker then evaluates the performance of the map-making techniques by comparing the interpolated (or extrapolated) magnetic field to the test magnetic field using a set of predefined metrics.

The benchmarker also allows for the visualization of the data: with a single command, the benchmarker can generate plots of the training measurements, the test measurements, the interpolated (or extrapolated) magnetic field, and the error between the interpolated (or extrapolated) magnetic field and the test magnetic field. It also supports exporting the benchmarking results to a `csv` file.

11.2.1. Measurements (Training Data)

To apply the benchmarker to the simulated field of Figure 11.1 we need to define a set of measurement points. The measurements are taken on a regular grid in the x-y plane at the surface ($z = 0$). To make sure that the measurements are taken in the region where the simulated field is defined (120 × 120 km, centered around the origin), we choose the measurement points to be in a region spans 30 × 30 km, centered around the origin of our coordinate system. All the experiments are performed with two different spacings of the measurement grid: a coarse grid with a spacing of 600 m consisting of 50 × 50 measurement points, and a fine grid with a spacing of 300 m consisting of 100 × 100 measurement points.

In our simulated environment, we can control everything: e.g. the measurement locations, the noise and the depth of the sources. Which makes it more suitable for benchmarking than real data. In practice, measurements are often taken using survey lines, which are generally not completely straight. Some extrapolation methods do however require a regular grid of measurements. We can obtain a regular grid of measurements by using one of the interpolation techniques. It is also better to measure at a constant height above the ground, e.g. 20 m, to avoid noise from objects on the ground.

11.2.2. Map-making Techniques

The benchmarker supports a wide range of map-making techniques. The map-making techniques are divided into two categories: interpolators and extrapolators. Table 11.1 shows an overview of the map-making techniques that are implemented in the benchmarker. The table also shows whether the map-making technique is parametric, an interpolator, and/or an extrapolator.

The methods that are incorporated in the benchmark are Cubic Splines for interpolation, Equivalent Layer for interpolation and extrapolation (Chapter 8), Multi-Layer Equivalent Layer for interpolation and extrapolation (Section 8.3.6), Upward Continuation as a discrete integral (Section 10.3) and Upward Continuation using the two-dimensional Fourier Transform (Section 10.2.3). The Anderson Wavelets (Chapter 9) are not implemented in the benchmarker, because their implementation so far only supports sources that are located straight underneath the wavelet center (Section 9.4.5). The Spherical Harmonic Expansion (Section 6.5) and the Kriging method are not used in the benchmark.

Map-Making Technique	Non-parametric	Interpolator	Extrapolator
Cubic Splines	✓	✓	✗
Equivalent Layer	✗	✓	✓
Equivalent Layer (Multi-Layer)	✗	✓	✓
Upward Continuation (Integral)	✓	✗	✓
Upward Continuation (FT)	✓	✗	✓
Anderson Wavelets	✓	✓	✓
Spherical Harmonic Expansion	✓	✓	✓
Kriging	✗	✓	✗

Table 11.1: Map-making techniques that are implemented in the benchmarker. The table shows whether the map-making technique is non-parametric, whether it can be used as an interpolator, and whether it can be used as an extrapolator. The first six techniques are used in the benchmarking experiment. The last three techniques are not used in the benchmarking experiment but are included in the table for completeness.

Parametric Techniques
Parametric techniques require a set of parameters to be defined before the map-making technique can be applied. The parameters are often dependent on the type of field that the method is applied to. For example, the Equivalent Layer method requires the depth of the layers to be defined. Parametric methods can be very fast and accurate, but their required a priori assumptions about the field can make them less suitable for general use.

In our benchmarking experiment, we use the Equivalent Layer method as a parametric technique. We did make a rough estimate of the depth of the sources in the simulated field (1600 m), but we did not tune the parameters to find the best possible result for Equivalent Layer. We also benchmarked a second Equivalent Layer method, which uses a multi-layered Equivalent Layer structure (Section 8.3.6). The results of these methods are discussed in the next section and are visualized in Figure 11.3a and Figure 11.3b.

Interpolators
Interpolators are map-making techniques that can be used to interpolate a magnetic field from a set of measurements to a set of evaluation points that are located in between the measurement points. In our work, we speak of interpolating when we have measurements at a certain altitude and we want to find the magnetic field at different points at the same altitude. The evaluation points must be surrounded by measurement points in the measurement plane, otherwise, the interpolator cannot be applied. It is not a mandatory requirement that the measurements are taken at the same altitude. For example, we can interpolate a magnetic field from a set of measurements at $z = 0$ m and $z = 2$ m to a set of evaluation points at $z = 1$ m. The interpolation techniques that we use in the benchmarking experiment are Cubic Splines and Equivalent Layer.

Extrapolators
By extrapolators we mean map-making techniques that take measurements of one height as input and extrapolate the magnetic field to evaluation points at a higher altitude. Equivalent Layer and Upward Continuation are techniques that can be used for extrapolating. Some extrapolation techniques require measurements on a regular grid, for example, for the Fourier Transform method for Upward Continuation. When the measurements are not available on a regular grid, we can use one of the interpolation techniques to obtain a regular grid of measurements and then extrapolate the magnetic field to the desired altitude. In our benchmarking experiment, the interpolation step before extrapolating is performed with Cubic Splines, as they are fast and accurate enough for this purpose.

Boundary Points
In our benchmarking experiments, we decided to remove the boundary of the domain from the evaluation points when extrapolating. The reason for this is that some methods, such as Upward Continuation using and Integral, are dependent on a larger domain of measurements (Section 10.4). We, therefore, restrict the evaluation points to a smaller part of the domain, which allows us to compare the extrapolation methods more fairly and express their performance in terms of the same metrics (Section 11.3).

11.3. Metrics

It is not trivial to define what a good map is, since the use-case of the map plays a big role. For example, a map that is used for navigation purposes may require different properties than a map that is used for anomaly detection. In this chapter, some global metrics are defined that can be used to compare the quality of maps in a general way, we also discuss how and where these metrics are applicable.

The quality of a map can be measured in many different ways. In this research, we limit ourselves to benchmarking two-dimensional maps. If we further restrict ourselves to maps that are defined on a regular grid, we can use the same metrics that are used to evaluate the quality of images. Under the assumption that we have access to the exact field that is being mapped, we can compare the map to the exact field, as if we are comparing the reconstruction of an image to the original image. The following metrics are commonly used to evaluate the quality of reconstructed images.

11.3.1. RMSE

The RMSE or root mean squared error is a way to quantify the difference between two vectors. It involves taking the square root of the average of the squared differences between the elements of the two vectors. The

average of the squared differences, also known as the mean squared error (MSE), is calculated as follows:

$$\mathrm{MSE}(x,y) = \frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2\,, \tag{11.1}$$

where x is a flattened vector of the original data, y is a flattened vector of the reconstructed data, and N is the number of elements in the vector.

The RMSE is then calculated by taking the square root of the MSE:

$$\mathrm{RMSE}(x,y) = \sqrt{\mathrm{MSE}(x,y)} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x_i - y_i)^2}\,. \tag{11.2}$$

In simpler terms, RMSE is a statistical tool that measures how far apart two sets of data are from each other. A low RMSE indicates that the reconstructed data is close to the original data.

11.3.2. PSNR

Peak Signal-to-Noise Ratio (PSNR [60]) is a quantitative measure that is often used in the field of image and video processing to assess the efficacy of compression algorithms. By comparing the original, uncompressed signal with the compressed version, PSNR yields a numerical value that quantifies the extent of distortion or noise introduced during the compression process. If we interpret magnetic maps as compressed versions of the exact magnetic field, we can use PSNR to compare the quality of different maps.

PSNR is derived from the mean squared error (MSE), which calculates the average squared difference between corresponding pixel values in the original and reconstructed images. The formula for MSE is given by Equation 11.1. The PSNR is then calculated by taking the ratio of the maximum possible pixel value L and the MSE and converting it to a logarithmic decibel (dB) scale:

$$\mathrm{PSNR}(x,y) = 10\log_{10}\left(\frac{(L-1)^2}{\mathrm{MSE}(x,y)}\right), \tag{11.3}$$

where

$$L = \max(x)\,. \tag{11.4}$$

As PSNR is inversely proportional to the MSE, a higher PSNR indicates a lower MSE and thus a lower error. The L in the denominator of Equation 11.3 is for images the maximum pixel value, for a total-field map of the magnetic field this is the maximum magnetic field strength. Scaling by L ensures that the PSNR is independent of the maximum field strength, which makes it a more intuitive metric. The PSNR can also be applied to vectorial maps, in which case the MSE is calculated for each component of the vector and the PSNR is calculated for each component separately. And if one prefers, the PSNR can be expressed in a single value by taking the average of the PSNRs of the individual components.

11.3.3. SSIM

Structural Similarity (SSIM [60]) is a method for measuring the similarity between two images. It is a popular image quality assessment method that has gained much attention in the field of image processing and computer vision [68]. It is a weighted combination of three properties of an image: luminance, contrast, and structure. These are defined as follows:

- Luminance:

$$L(x,y) = \frac{2\mu_x\mu_y + C_1}{\mu_x^2 + \mu_y^2 + C_1}\,, \tag{11.5}$$

  where μ_x and μ_y are the average pixel values of the two images.

- Contrast:

$$C(x,y) = \frac{2\sigma_{xy} + C_2}{\sigma_x^2 + \sigma_y^2 + C_2}\,, \tag{11.6}$$

  wjere σ_x and σ_y are the standard deviations of the two images and σ_{xy} is the covariance of the two images.

- Structure:

$$S(x,y) = \frac{\sigma_{xy} + C_3}{\sigma_x \sigma_y + C_3}, \tag{11.7}$$

where C_1 and C_2 are constants that are used to stabilize the division when the standard deviation is zero. Suggested default values for C_1 and C_2 are $C_1 = (0.01L)^2$ and $C_2 = (0.03L)^2$, where L is the maximum pixel value of the image. C_3 is often set equal to $C_2/2$.

The SSIM is then calculated as a weighted combination of the three properties:

$$\mathrm{SSIM}(x,y) = L(x,y)^{\alpha} C(x,y)^{\beta} S(x,y)^{\gamma}, \tag{11.8}$$

where α, β, and γ are often set equal to one, leading to the following formula:

$$\mathrm{SSIM}(x,y) = \frac{(2\mu_x\mu_y + C_1)(2\sigma_{xy} + C_2)}{(\mu_x^2 + \mu_y^2 + C_1)(\sigma_x^2 + \sigma_y^2 + C_2)}. \tag{11.9}$$

When an image has a mean value but low variations, it indicates the image exhibits a rather homogeneous distribution with minor discrepancies. This could lead to a high SSIM score for the brightness evaluation component. However, when it comes to contrast and structure assessments, the SSIM takes into account the dispersion and joint variability. Low values of these statistics suggest minimal variation or structural intricacy within the image, potentially leading to a lower SSIM score. Therefore, merely having a high average with tiny variations doesn't ensure a high SSIM score. The SSIM score ranges from -1 to 1, where 1 indicates a perfect match between the two images.

11.4. Results

In this section, we present the results of our benchmarking experiments. We benchmarked several types of fields, with different resolutions, two types of measurements (total-field and vectorial), and with and without noise. This resulted in a large number of benchmarking results. We discuss the most important results in this section.

11.4.1. Regularization of Equivalent Layer

Figure 11.2 shows the results of extrapolating total-field measurements of the magnetic field at height $z = 0$ m to height $z = 400$ m. Two methods are compared: the Equivalent Layer method, with a total of 21×21 sources at a depth of $z = -1600$ m and with 1500 m spacing. Equivalent Layer 1 shows slightly better results compared to Upward Continuation with the Fourier Transform, but these differences are not significant. Table 11.2 shows a more detailed overview of the results. The Equivalent Layer 2 and the Discrete Integral for Upward Continuation are not performing well and are excluded from the figure.

A noticeable result is the difference in performance between the single-layer Equivalent Layer 1 and multi-layered Equivalent Layer 2. Equivalent Layer 1 consists of a single layer of dipoles at 1600 m depth (Figure 11.3a), while Equivalent Layer 2 has a more varied structure with dipoles positioned at multiple depths (Figure 11.3b). An analysis of the magnetic moments that were assigned to the dipoles by the Equivalent Layer 2 method (Figure 11.4a) showed that mainly the upper dipoles, close to the surface, had a significant magnetic moment. This results in a fast decay of the magnetic field as we increase the altitude above the surface, and thus a poor extrapolation: the **SSIM** of 0.0795 and the **RMSE** of 1713 are both clear indicators that the reconstruction is fairly dissimilar to the desired magnetic field.

The current approach using the multi-layered Equivalent Layer 2 method leaves room for enhancement. Ideally, we aim to reconstruct the training data with fewer surface dipoles close to the surface. This goal is achievable because, provided the spatial frequencies permit, a single dipole situated deeper can represent multiple surface-proximate dipoles. To facilitate this, we can incorporate a regularization term into the Equivalent Layer method's cost function, as described in Section 8.3.7. In our experiment, we use $\lambda^2 = 10^{-3}$ as the regularization parameter.

When the regularization term is introduced ("Equivalent Layer 2 (Reg.)" in Table 11.2), it noticeably improves the extrapolation for Equivalent Layer 2. The contrast in the magnetic moments' distribution post-regularization is displayed in Figure 11.4a and 11.4b. This regularization term constrains the magnetic moments to prevent them from becoming excessively large, as seen in the unregularized scenario. Moreover, it diminishes the magnetic moments of the dipoles near the surface and enhances the magnetic moments of the dipoles located deeper within the Earth, resulting in a significantly better extrapolation. The **SSIM** of 0.965 and the **RMSE** of 20.667 are both clear indicators that the regularization term improves the extrapolation.

11.4.2. Interpolation of Fields without Noise

When interpolating a noise-free field, Cubic Splines outperform Equivalent Layer 1 (Table 11.3) in our benchmarking experiment in every metric. A more complex source distribution, such as the multi-layered Equivalent Layer 2, shows more promise, but elaborate tuning of the Equivalent Layer parameters is required to obtain a better result. All the methods give, after rounding, an **SSIM** of one, which means that the reconstructions are nearly identical to the test measurements in terms of their structural information. Regarding the **PSNR**, values above 30 dB are often considered acceptable, while values above 40 dB are considered very good, and values above 50 dB are considered excellent. According to this, all the interpolation methods give excellent results. However, the Equivalent Layer method is slower than Cubic Splines. As both methods give similar results, but Equivalent Layer is slower and requires more tuning, we conclude that Cubic Splines are

Table 11.2: Scores for extrapolating Core + Crustal + Noise Field from 100x100 - 300m to grid 64x64 400m + 400m (Vectorial)

Extrapolator	**RMSE** (nT)	**L2Norm** (nT)	**PSNR** (dB)	**SSIM** (1)	**Time** (s)
Equivalent Layer 1	26.843	671.066	64.659	0.959	2.006
Equivalent Layer 2	1713.912	42847.788	28.555	0.0795	4.545
Equivalent Layer 2 (Reg.)	20.667	516.666	66.930	0.965	4.486
UC Integral	1146.391	28659.772	32.048	0.847	0.099
UC FT	21.614	540.356	66.540	0.965	0.098

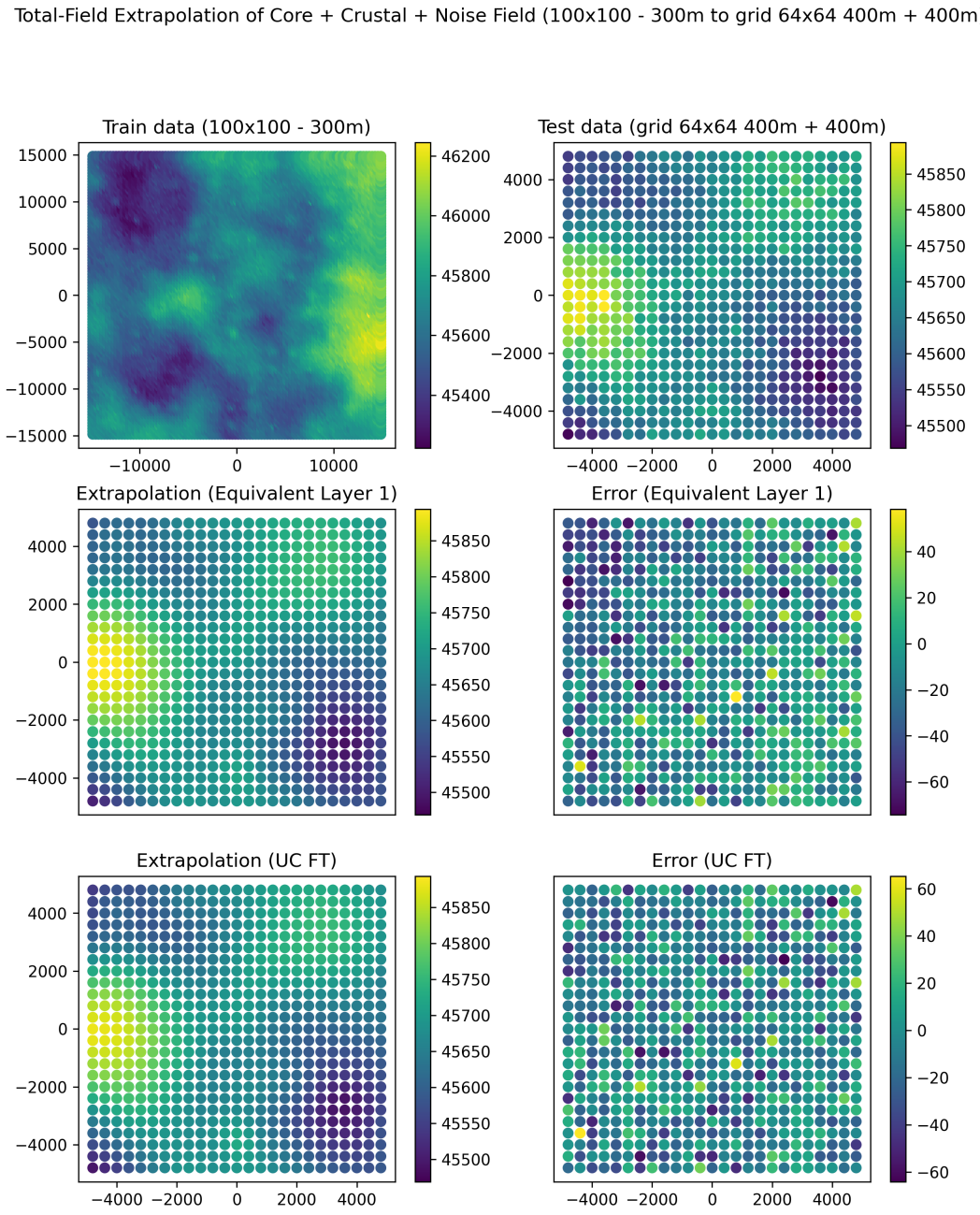
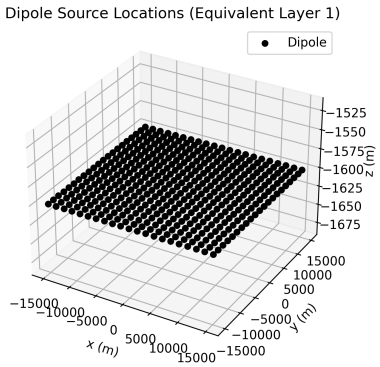
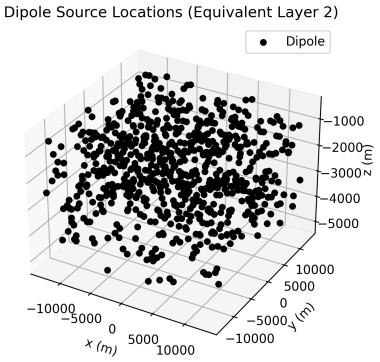


Figure 11.2: Benchmarking results for extrapolating a noisy simulated field from a 100×100 grid of total measurements with a spacing of 300 m to a 64×64 grid of evaluation points with a spacing of 400 m. The figure shows the training data (top left), the test data (top right) and the results of several map-making techniques.

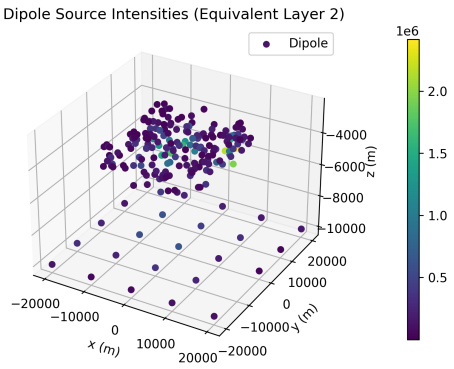


(a) Dipole locations for Equivalent Layer 1. The dipoles are located on a 21×21 grid at a depth of 1600 m with a spacing of 1500 m.

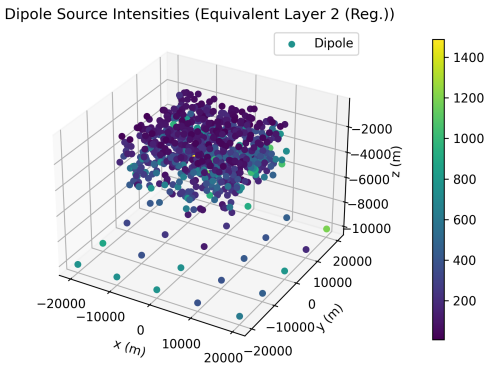


(b) Dipole locations for multi-layered Equivalent Layer 2. The dipoles are located on a 27×27 grid at depths ranging from 500 m to 5500 m. The dipoles are spaced 1000 m apart.

Figure 11.3: Dipole locations for the two Equivalent Layer models in the benchmarking experiments. The dipoles are shown as black dots. The first configuration (a) is used for the Equivalent Layer 1 model and has dipoles positioned in a plane. The second configuration (b) is used for the Equivalent Layer 2 model and has dipoles positioned in a volume, with more dipoles at shallower depths.



(a) Magnitudes of the magnetic moments of the dipoles in Equivalent Layer 2 without regularization.



(b) Magnitudes of the magnetic moments of the dipoles in Equivalent Layer 2 with regularization.

Figure 11.4: Magnitudes of the magnetic moments of the dipoles in Equivalent Layer 2 without regularization (a) and with regularization (b). The magnetic moments are scaled with the pre-conditioning factor (Section 8.3.2) to make them comparable. The dipoles with magnetic moments smaller than 1% of the maximum magnetic moment are hidden in the figure.

Table 11.3: Scores for interpolating Core + Crustal Field from 100x100 - 300m to 24x24 - 1000m (Vectorial)

Interpolator	**RMSE** (nT)	**L2Norm** (nT)	**PSNR** (dB)	**SSIM** (1)	**Time** (s)
Cubic Spline	0.304	12.633	102.125	1.000	0.186
Equivalent Layer 1	5.356	222.665	77.202	1.000	2.825
Equivalent Layer 2	2.713	112.778	83.111	1.000	6.926
Equivalent Layer 2 (Reg.)	3.223	133.995	81.613	1.000	4.077

Table 11.4: Scores for interpolating Core + Crustal + Noise Field from 100x100 - 300m to 24x24 - 1000m (Vectorial)

Interpolator	**RMSE** (nT)	**L2Norm** (nT)	**PSNR** (dB)	**SSIM** (1)	**Time** (s)
Cubic Spline	25.973	1079.661	63.489	1.000	0.183
Equivalent Layer 1	21.131	878.395	65.280	1.000	3.070
Equivalent Layer 2	21.093	876.832	65.296	1.000	7.003
Equivalent Layer 2 (Reg.)	20.806	864.892	65.415	1.000	5.978

a better choice than Equivalent Layer when it comes solely to interpolating magnetic fields.

11.4.3. Interpolation of Fields with Noise

Table 11.4 and 11.3 show the results of interpolating vectorial measurements of the magnetic field at height $z = 0$ m to a regular grid of 24×24 evaluation points. The first table shows the results for a noisy field, the second table shows the results for a noise-free field. The results show that the Equivalent Layer method performs better than Cubic Splines when interpolating a noisy field. The difference in performance is not significant, while the Equivalent Layer method is significantly slower. All methods seem to have an **RMSE** that is close to the noise level of the field (20 nT). Variations of the same experiment with different resolutions of the training data and different noise levels showed that increasing the sample rate of the training data does not have a significant effect on the accuracy while decreasing the noise level was directly reflected in a decrease of the **RMSE**. Even though the **RMSE** has increased by a factor of 10 when interpolating a noisy field, the values of **PSNR** and **SSIM** are still very high. The **PSNR** is still above 50 dB but has decreased since the noise is larger. The **SSIM** is, after rounding, still 1, which means that the structural information of the field is still preserved. Therefore, we can conclude that interpolating the magnetic field with Cubic Splines or Equivalent Layer is a viable option even when the field has white noise added to it.

11.5. Discussion and Conclusions

A disadvantage of the simulated environment is that it is a very simplified version of the real world. Every real-world dependency that is not included in the simulation, can influence the results of the benchmarking experiment. Better simulations of the measurement positions, the measurement noise, and the underlying magnetic field lead to more realistic results. For further research, it would be interesting how to get more realistic simulations of the magnetic field and the measurements.

The benchmark showed that the parametric method Equivalent Layer requires more tuning to obtain a good result. This holds for both interpolation and extrapolation. The choice of the depth, position and number of dipoles in the Equivalent Layer affect the accuracy of the interpolation significantly. The distribution of the magnetic moments of the dipoles has a considerable effect on the accuracy of the extrapolation. To obtain a good result, the magnetic moments of the dipoles should be distributed in such a way that the magnetic field decays slowly with increasing altitude. This can be achieved by introducing a regularization term in the cost function of the Equivalent Layer method. This regularization term can limit the total magnetic moments given to the dipoles, such that deeper dipoles are preferred over shallower dipoles.

11.5.1. Limitations

Accurate Modelling of Noise

The noise model that we have used in our simulations is very basic. We have seen in Chapter 7, that the compensation of noise is a complex problem. A model that is closer to reality would include colored noise, platform noise, and noise in the position of the measurements. Better modelling of the noise leads to more realistic benchmarking results.

Measurement Locations
The simulated environment allowed us to choose the measurement locations freely. We have chosen to measure at a fixed altitude and to distribute the measurements evenly over the region. In practice, measurements are often taken using survey lines, which are generally not completely straight. Constant altitude is not always possible, for example, when measuring in a mountainous area, but it is also very difficult to maintain a constant altitude when measuring from a moving aircraft. Mountains, buildings, and other obstacles can also cause holes in the measurement grid. It would be interesting to see how the map-making techniques perform when the measurements are not taken at a constant altitude and when there are holes in the measurement grid.

12

Validation

In the previous chapter, we benchmarked the performance of the map-making techniques on simulated data. In this chapter, we validate the map-making techniques on real-world data. We limit ourselves to the extrapolation techniques: finding the magnetic field at an altitude that is different from where the measurements are taken. The main goal of this chapter is to see what pre-processing steps are necessary to make real-world sensor data suitable for map-making and how well the map-making techniques perform on real-world data.

We use the data from a measurement campaign we performed on the campus of TU Delft. In this campaign, measurements at two altitudes were taken along parallel survey lines. The campaign is described in detail in Section 12.1. Section 12.2 describes how the raw sensor data is processed into a magnetic field map at a regular grid. Section 12.3 describes how the magnetic map at one altitude is extrapolated to another altitude. Section 12.4 presents the results of the validation. Section 12.5 discusses the results and concludes this chapter.

12.1. Measurement Campaign

12.1.1. Location

The best way of validating map-making techniques for aeromagnetic navigation is by flying and taking measurements at different altitudes in the air on a regional scale. Unfortunately, the duration of this thesis precluded the possibility of conducting a measurement campaign in an area where flying drones or aircraft is permitted. Therefore, we had to search for a smaller area that mimics the situation of a drone flying over a regional-scale magnetic field. We have established some conditions that this field should meet.

We are looking for a location which has the following properties:

1. The magnetic field is interesting enough to validate the map-making techniques.
2. All the sources of the magnetic field are below the altitude at which the measurements are taken. This is a requirement for Green's third identity to hold (Section 2.4.1).
3. The surface is flat and walkable.
4. GPS for ground-truth position measurements is available.
5. The location allows for measurements at different altitudes.
6. The location does not have too much temporal variation in the magnetic field.

Sometimes, the solution is right in front of you. The campus of TU Delft, where this thesis is written, has a parking garage that meets all of these conditions. The parking garage concerned, P2 Sports (Figure 12.1), is located at the following coordinates: 51.9963°N, 4.3789°E. The parking garage is made of steel and concrete, which makes it magnetically interesting. The upper parking deck is not covered, is accessible and often free of cars. Making it a walkable surface where measurements can be taken at different altitudes. The parking garage (Figure 12.2) is located in an area where GPS is available and has a clear view of the sky.

The only thing that cannot be controlled is the time-dependent variations in the magnetic field. The parking garage itself does not run off, but cars may be driving off and on and disturbing the field. We solved

Figure 12.1: Parking garage P2 Sports at TU Delft.

(a) The second quartile of the parking garage.

(b) The first quartile of the parking garage.

(c) The third quartile of the parking garage.

(d) The fourth quartile of the parking garage.

Figure 12.2: Tiled plot of the quartiles of the upper parking deck of the parking garage.

this by having the measurement campaign take place early, on a Saturday morning. We made good notes of each floor with where the cars were parked, and after the measurements were completed we checked to see if any cars had moved. One car, on the second highest floor, had left.

12.1.2. Sensors and Measurement Setup

The measurements are taken with the Sensys MagDrone R3 [59]. The MagDrone consists of a carbon fiber tube with on both ends a flux-gate magnetometer. These sensors are 1,000 mm apart. Both sensors have 3 axes and a measurement range of $\pm 75,000$ nT. A small box is attached to the middle of the tube, which handles GPS-tracking, accelerometer data, data storage and power. The datalogger samples the magnetometer data at 200 Hz and the GNSS data at 5 Hz. The noise level of the sensors due to the electronics is at most 50 pT/$\sqrt{\text{Hz}}$.

Figure 12.3: The Sensys MagDrone R3 attached to a telescopic stick.

The MagDrone is usually attached to a drone, but in this case, it is held in hand. The MagDrone is held perpendicular to the direction of movement, resulting in two measurements: one for the left sensor and one for the right sensor, which are both 50 cm apart from our walking path. We attached the MagDrone to a custom-made stick to take measurements at different altitudes (Figure 12.3). The stick is a telescopic stick, which can be adjusted in length from 1 to 4 m. The stick is made of aluminum, which has a relatively low magnetic permeability and is therefore not expected to disturb the magnetic field too much. The stick is, however, not very stable at 4 m and therefore we did not use the full length of the stick. We used the stick to extend the height of the MagDrone from 1 m to 3.5 m above the surface of the upper parking deck.

12.1.3. Data Collection

We collected two datasets in the measurement campaign. The first dataset is used as the training set, and the second dataset is used as the validation set. The training set is collected at 1 m above the surface of the upper parking deck, and the validation set is collected at 3.5 m above the surface of the upper parking deck. The measurements are taken in a grid-like pattern, with survey lines parallel to each other and the traverses perpendicular to the survey lines. The survey lines are walked over the edges of the tiles of the parking garage, which are 2.45 m apart. Eight traverses are walked at the following relative positions: 1 m, 5 m, 8.4 m, 10.2 m, 22.1 m, 25 m, 28.4 m, 32 m. The two ramps in the center of the parking garage are not flat, and therefore no measurements are taken there. This results in two holes in the dataset.

12.2. Data Processing

Working with real measurement data presents some difficulties. In the simulated environment (Chapter 11), we could choose the full path of the sensor ourselves and we have perfect information about the position. In

the real world, we have to deal with the fact that the sensor is not always in the same position: the body frame rotates in the world frame, the sensor is not always at the same height, and there is uncertainty in the position of the sensor, as this is determined by GNSS measurements. In this section, we describe the steps we take to preprocess the data.

12.2.1. Preprocessing

Trimming and Filtering

The first step is to remove the parts of the data that are not of interest. The first and last parts of the data are not of interest, as the sensor is not yet in the correct position. We inspected the data and determined the start and end of the data by hand. The data is then trimmed to the area of interest. The data is sampled at 200 Hz for the magnetic field and 5 Hz for the GPS. We downsample the data to 5 Hz to match the sampling rate of the magnetic field and GPS.

Geodetic Coordinates to Local Cartesian Coordinates

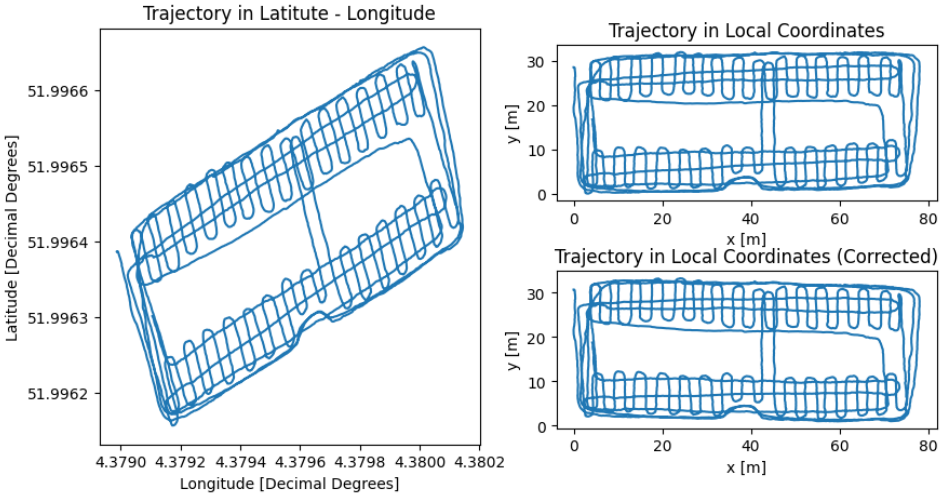


Figure 12.4: Three trajectories of the first traverse of the training set. The trajectory is shown in the geodetic coordinate system (Left), in the local cartesian coordinate system (Top Right) and in the corrected local cartesian coordinate system (Bottom Right).

The datalogger contains a GPS sensor which gives us our position. The position is given in Latitude, Longitude and Altitude. We convert the position to the local cartesian coordinate system to perform computations. First, we transform the position from the WGS84 coordinate system to the EPSG:28992 coordinate system. The EPSG:28992 coordinate system is the coordinate system used in the Netherlands. The transformation is done using the pyproj library. Then, we subtract the minimum x- and y-position from the position, so that the minimum position is at $(0,0)$ and rotate the data with an angle θ so that the survey lines align with the y-axis. The result is shown in the top-right plot of Figure 12.4. Despite that our survey had traversals that were parallel to the survey lines, the survey lines are not perfectly parallel to the y-axis. The path has a small slope when moving in the x-direction. We think this is caused by the projection of the GNSS measurements to the local cartesian coordinate system. We decided to correct this with a linear transformation. The final result is shown in the bottom-right plot of Figure 12.4. The slope is now removed, the survey lines are almost parallel to the y-axis, and the traverses are almost parallel to the x-axis. The drift in the y-coordiate of the traversals can be explained by the fact that the GNSS measurements are not very precise: approximately 1 m.

Straightening the Trajectories

Our survey was performed in a rectangular area. The survey lines are parallel to each other, and the traverses are perpendicular to the survey lines. We can use this information to straighten the trajectories of the measurements. We start by removing the turning points of our survey lines from the measurements. We detect the direction of movement by computing the average change in $x-$ and $y-$ direction over the previous and next n samples. Due to the precision of the GNSS measurements, it is not sufficient to use only two samples to compute the direction. We used $n = 3$ in our experiment, which corresponds to data from 1.4 s. We classify a point as moving in x-direction when the average change in x-direction is twice as large as in y-direction.

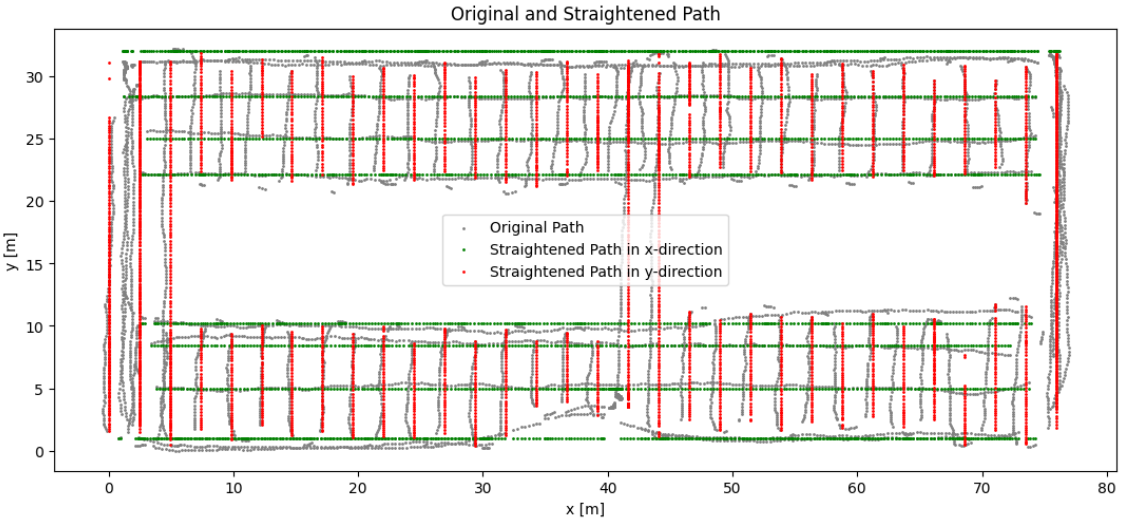


Figure 12.5: The original path, provided by the GNSS measurements (Gray) after the transformation to the local cartesian coordinate system and the straightened path in x-direction (Green) and y-direction (Red).

We then use the fact that our survey lines are 2.45 m apart to round the x-coordinate of all the points that move in y-direction to the nearest multiple of 2.45 m. We do the same for the y-coordinates, but there we round the y-coordinate to the closest position traversal in the following list: $y = 1$ m, $y = 5$ m, $y = 8.4$ m, $y = 10.2$ m, $y = 22.1$ m, $y = 25$ m, $y = 28.4$ m, $y = 32$ m. The results are illustrated in Figure 12.5. The gray line shows the original path, the green line shows the path after straightening in x-direction, and the red line shows the path after straightening in y-direction.

Rotating the Sensor Data

The next step is to rotate the sensor data to the world frame. The sensors are fixed in the body frame: the sensors stay aligned with the direction of movement of the person taking the measurements. That also means that the sensors rotate in the world frame when the person taking the measurements change direction. The goal is to align B_x with the x-axis, B_y with the y-axis, and B_z downwards. The z-axis is already aligned with the direction of gravity, so we do not need to rotate the z-axis. We can rotate the sensor data by multiplying the sensor data with a rotation matrix. We only consider the samples from the previous section which have a clear direction of movement, and filter out other samples. The direction of movement, including its sign, is used to rotate the sensor data from the body frame to the world frame. We then have the following two rotation matrices:

$$\mathbf{R}_x = \begin{bmatrix} -sgn(\hat{x}) & 0 & 0 \\ 0 & -sgn(\hat{x}) & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{R}_y = \begin{bmatrix} 0 & -sgn(\hat{y}) & 0 \\ sgn(\hat{y}) & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \tag{12.1}$$

where $\hat{x}$ and $\hat{y}$ are the x- and y-components of the direction of movement, and sgn is the sign function. The rotation matrix $\mathbf{R}_x$ is applied to the samples where the direction of movement is in x-direction, and the rotation matrix $\mathbf{R}_y$ is applied to the samples where the direction of movement is in y-direction, this is the traverses and survey lines respectively. Finally, we need to rotate the sensor data from the geodetic coordinate system to the local cartesian coordinate system. For this, we use the same angle θ as in Section 12.2.1:

$$\mathbf{R}_{\text{world}}(\theta) = \begin{bmatrix} \cos(\theta) & -\sin(\theta) & 0 \\ \sin(\theta) & \cos(\theta) & 0 \\ 0 & 0 & 1 \end{bmatrix}. \tag{12.2}$$

In the measurement setup at a higher altitude, the sensor is upward down. This requires an additional rotation of 180 ° around the y-axis:

$$\mathbf{R}_{\text{UC}} = \begin{bmatrix} -1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & -1 \end{bmatrix}. \tag{12.3}$$

Left and Right Sensor

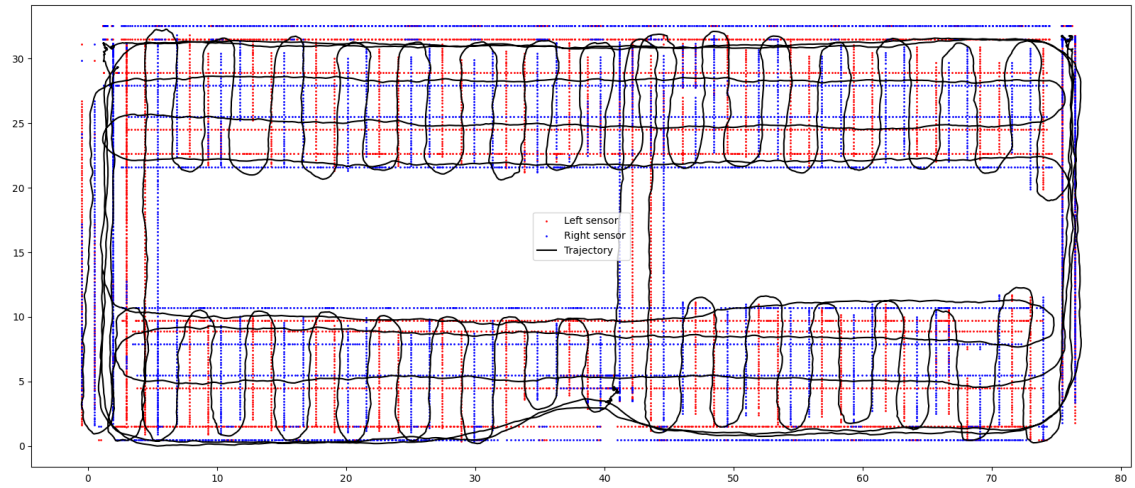


Figure 12.6: The position of the left sensor (Red) and the right sensor (Blue) in the world frame after correcting the path to be straight. The original path is shown in black.

The MagDrone setup consists of two sensors, one on the left and one on the right. The left sensor becomes the right sensor in the world frame when the direction of movement is reversed. We again use the direction of movement to determine which sensor is on the left and which sensor is on the right. We then shift the measurement data either 50 cm to the left or 50 cm to the right. The shift is done in the local cartesian coordinate system. The result is shown in Figure 12.6.

Platform Compensation
When the sensor is mounted on a platform, the platform can cause magnetic disturbances. The platform can be a car, a drone, or an aircraft. The platform can cause magnetic disturbances because of the magnetic materials used in the platform, or because of the electric currents flowing through the platform. The MagDrone is a platform itself, but its tube is made of non-magnetic material and the effects of the box with electronics are negligible (less than 50 pT/$\sqrt{\text{Hz}}$). The stick is also made of non-magnetic material. The survey is taken by walking around, and therefore no engines, other magnetic sources, or electric currents are involved. Platform compensation is therefore not necessary. When the sensor is mounted on an aircraft or drone, platform compensation is necessary.

Reference Field
We often consider anomaly maps, which represent the local deviation from a more global reference field. The IGRF model (Section 6.1.2) can be used to get a value of the average magnetic field at the location of the parking garage. On the scale of our measurements, IGRF has a constant value of $\mathbf{B}_0 = \begin{bmatrix}19183 & 720 & 45452\end{bmatrix}^\top$ nT. We can subtract this value from the measurements to get the anomaly field.

Preprocessing Results
Some results of the pre-processed data are visualized in figures 12.7 and 12.8. The figures show the total-field anomaly at 10.5 m and 13.0 m respectively. These values are found by taking the intensity of pre-processed components of the magnetic field, that are measured by the flux-gate sensors on the MagDrone.

12.3. Map-Making

The data is now ready to be used for map-making. We use the Equivalent Layer and Upward Continuation methods to make maps of the anomaly at a height of 13.0 m above sea level using measurement data at 10.5 m. For Upward Continuation a regular grid of measurement points is required. We obtained a regular grid of 288×128 points by interpolating the measurement data using cubic splines. The Equivalent Layer method requires some parameters. We ran the method with 2 sets of parameters. The first configuration used a layer positioned 5 m under the training surface, spanning the whole area, and with a 2 m spacing between the dipoles. The regularization parameter was $\lambda = 1$. The second configuration used the same spacing and

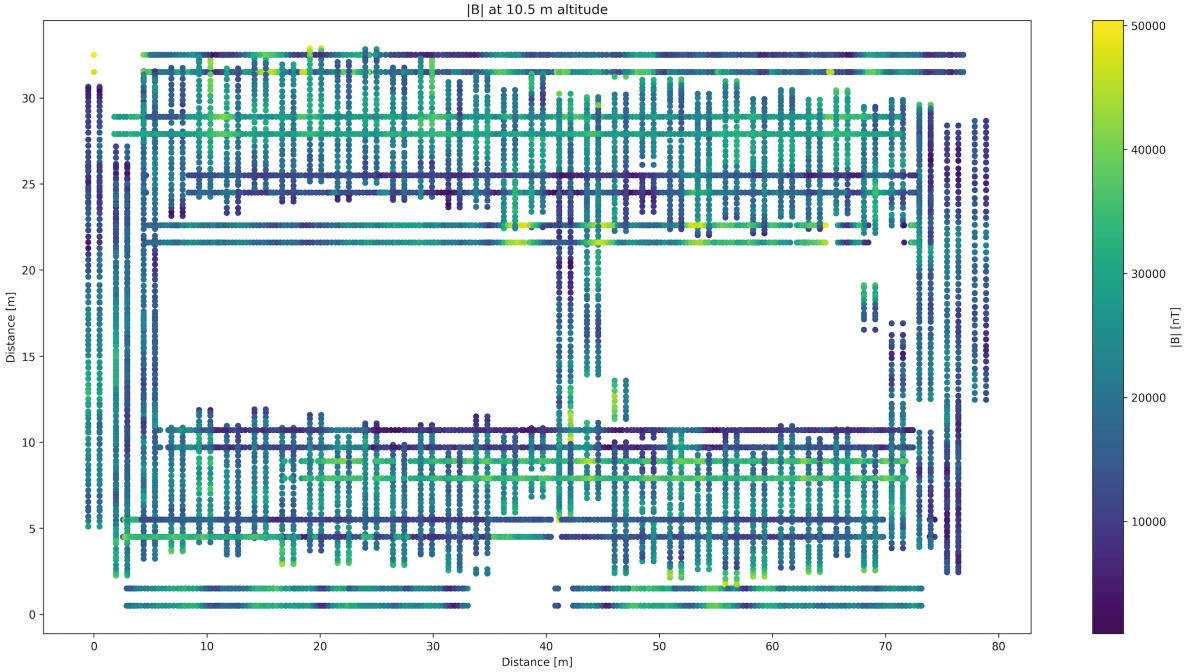


Figure 12.7: Preprocessed total-field anomaly of the upper parking deck of the parking garage at an altitude of 10.5 m above sea level. The measurements were taken at approximately 1 m above the surface.

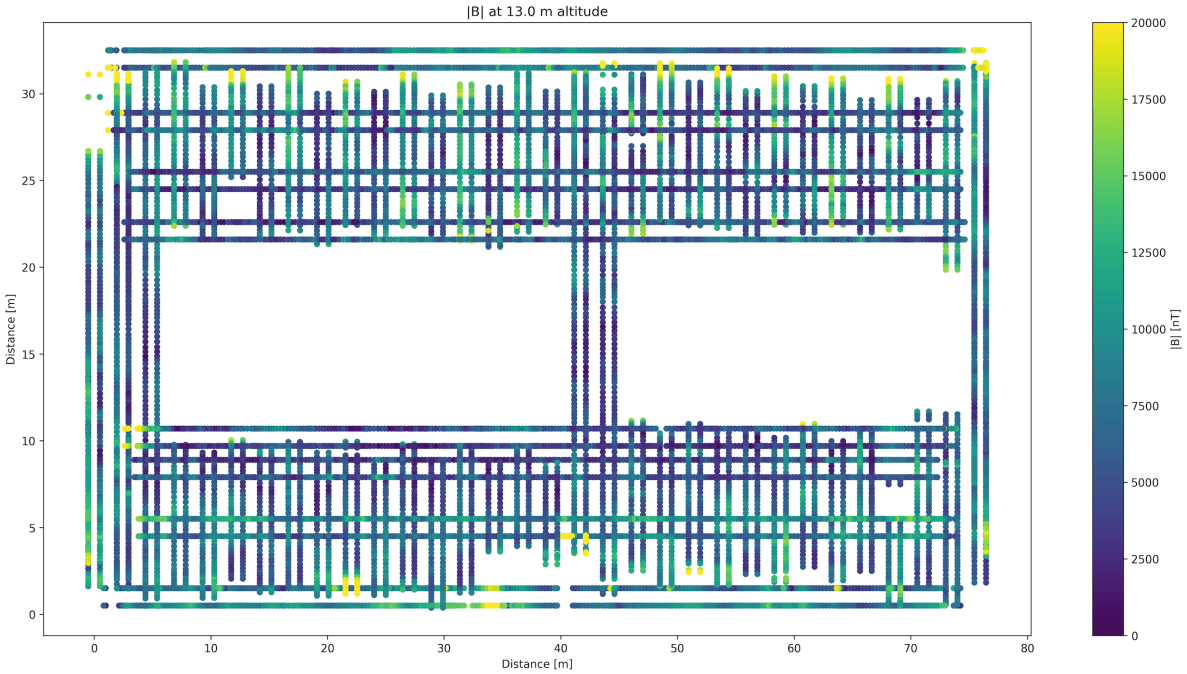


Figure 12.8: Preprocessed total-field anomaly of the upper parking deck of the parking garage at an altitude of 13.0 m above sea level. The measurements were taken at approximately 3.5 m above the surface.

regularization parameter, but the layer was positioned 2 m under the training surface. The 2 m depth was chosen because the spatial frequencies of the variations in the field are high, and therefore originate from a shallow layer. We found the depth of 5 m by having a small look at the data at a higher altitude and found that the values were dampened with approximately a factor of 3.3 when moving 2.5 m up. Solving $(d+2.5)^3 = 3.3a^3$ gives $d \approx 5$ m. The regularization parameter is chosen by trial and error. We found that a value of $\lambda = 1$ gives a good result.

12.4. Results

Figure 12.9 and 12.10 show the extrapolation results for Equivalent Layer and Upward Continuation using the Fourier Transform, respectively. The results are not what we expected. The methods both seem to find slow-varying or almost constant fields, which is not what you expect from a parking garage. The first Equivalent Layer configuration, with a layer positioned 5 m under the measurement plane, has too low spatial frequencies to accurately reconstruct the magnetic field. The dampening of the field is also not fast enough, which indicates that the layer is placed too deep. Conversely, the layer at 2 m under the measurement plane seems to be placed too shallow, as almost the whole intensity of the field is dampened out when extrapolating to 13.0 m. For completeness, we ran a third Equivalent Layer experiment with a layer positioned at 3.3 m. This layer showed slightly better results but is still not good enough to be used for map-making.

The Upward Continuation method also does not seem to work well. The interpolated anomaly field was not mean-zero and therefore the extrapolated field takes values around the mean of the lower field. There seems to be a bit more detail in the Upward Continuation results than in the Equivalent Layer results: some peaks that occur at the lower altitude are also visible in the extrapolated data. However, the Upward Continuation method still has an RMSE of around 10,000 nT. The scores for the extrapolation results are shown in Table 12.1.

Table 12.1: Scores for Extrapolating P2 Sports from z=10.5 to z=13.0 (Total-Field)

Extrapolator	**RMSE**	**L2Norm**	**PSNR**	**SSIM**	**Time (s)**
UC Integral	8495.629	189968.043	13.634	-0.070	0.201
UC FT	10964.552	245174.846	11.418	-0.019	0.289
Equivalent Layer 1(5m)	12435.856	278074.188	10.325	0.016	1.883
Equivalent Layer 2(2m)	8414.712	188158.687	13.717	0.045	2.636
Equivalent Layer 3(3.3m)	7095.058	158650.319	15.199	0.141	2.973

12.5. Discussion and Conclusion

In the pre-processing step of our data, we removed the IGRF from the data. This is a good step to take, as it removes the influence of the Earth's magnetic field from the data. However, this requires that the data is perfectly aligned with the geodetic coordinate system. We did our best in rotating the sensor data from the body frame to the world frame and removed all the measurement points that had an unclear direction, but there is still a chance that the data is not perfectly aligned. In a next measurement campaign, we should either keep better track of the rotation of the sensor, by using a gyroscope and accelerometer, or we should try to keep our sensor orientation constant to the world frame, by rotating the sensor counter to the rotation of the body.

Furthermore, we walked straight survey lines, but these lines were not straight in the GNSS data of the measurements. This is because the GNSS data is not perfectly accurate. It has approximately a 1 m circular positioning error. To obtain a better position estimate, we suggest using a higher precision positioning system, such as RTK-GPS, which combines the GNSS data with a correction signal from a base station. This can reduce the error to 1 cm. Another option is scaling up the size of the survey area so that the error in the GPS data is less significant.

Since we carried the sensor in our hands or on a stick, there was almost no noise in the data originating from the platform. This led the more accurate measurements of the magnetic field. However, this is not a realistic scenario. In a real-life situation, the sensor is mounted to a drone or another aircraft and is subject to permanent, induced and eddy current noise in the data. It is therefore important to test the algorithms in a more realistic scenario, where the sensor is mounted to a drone or another aircraft.

The performance of the Equivalent Layer method is a bit disappointing. We noticed that the layer at 5 m

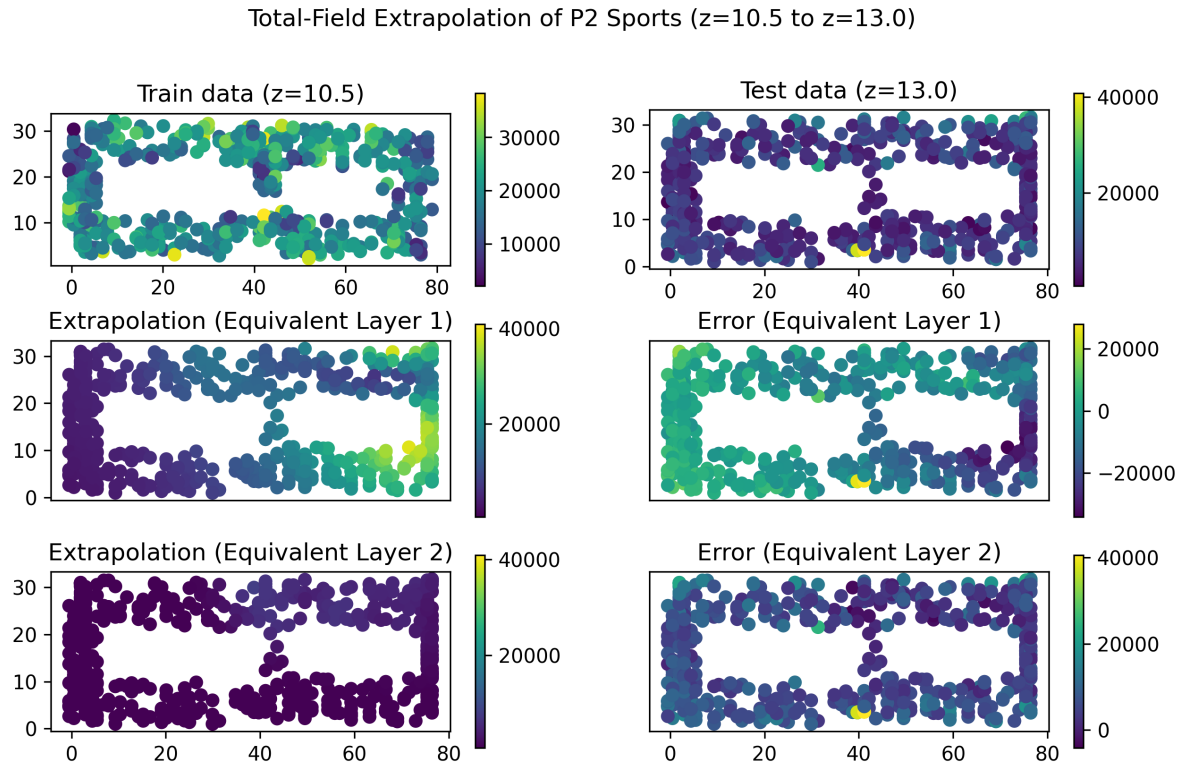


Figure 12.9: Extrapolation results for the total-field anomaly from 10.5 m to 13.0 m above sea level using Equivalent Layer.

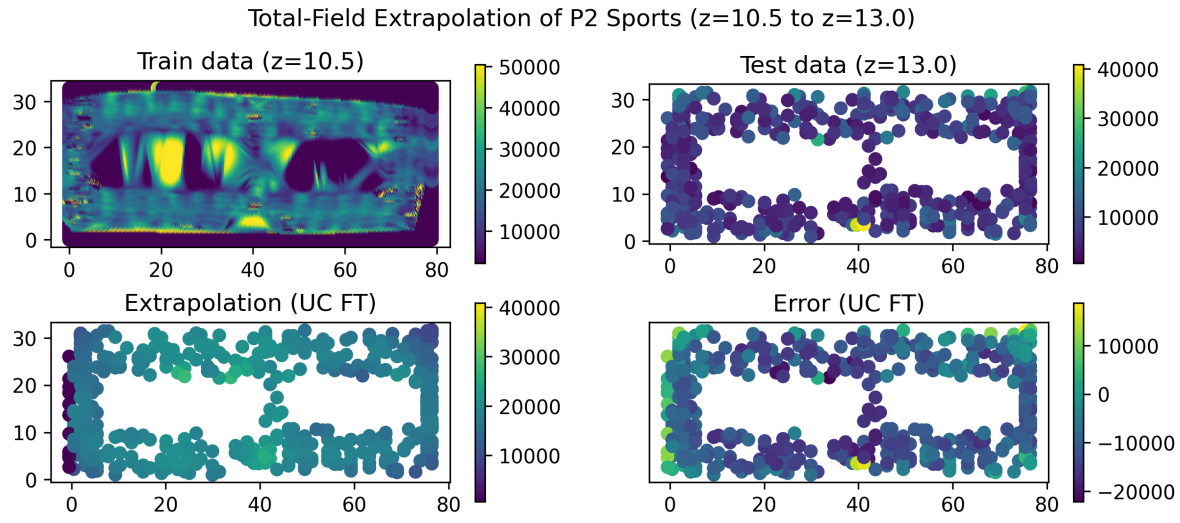


Figure 12.10: Extrapolation results for the total-field anomaly from 10.5 m to 13.0 m above sea level using Upward Continuation.

was positioned too deep, while the magnetic field from shallower layers decayed too fast when the field was evaluated at a higher altitude. It could be possible that the magnetic field cannot be accurately represented by magnetic dipoles. In Section 2.5.1, we remarked that dipoles dominate the magnetic field when the distance to the sources is sufficiently large. However, in our validation experiment, the distance to the sources is possibly not large enough to justify this approximation. We, therefore, suggest using the Equivalent Layer method with multipole sources, which can represent the magnetic field more accurately at shorter distances, or increasing the distance to the sources by measuring at a higher altitude, for example by using a drone. The same goes for the Upward Continuation method. The method is based on the Fourier Transform, which dampens the signal under the assumption that the field is dominated by dipoles. This assumption might not hold for our data, which could explain the poor performance of the method.

In conclusion, the preprocessing steps taken to prepare the measurement data for map-making were successful. The data is trimmed and filtered, converted to local Cartesian coordinates, straightened, and rotated to align with the world frame. However, the map-making methods applied did not yield the expected results. Both the Equivalent Layer and Upward Continuation methods produced maps that were too smooth and lacked the necessary detail to accurately represent the parking garage.

There is room for improvement. The pre-processing step involved removing the influence of the Earth's magnetic field from the data, but the data may not be perfectly aligned with the geodetic coordinate system. To obtain a better position estimate, a higher precision positioning system such as RTK-GPS can be used. Furthermore, the absence of noise in the data originating from the platform led to more accurate measurements of the magnetic field, but it is important to test the algorithms in a more realistic scenario, where the sensor is mounted to a drone or another aircraft. Finally, the Equivalent Layer and Upward Continuation methods had disappointing performances, possibly because the assumption that the field is dominated by dipoles might not hold for the data, and the suggestion is to use the Equivalent Layer method with multipole sources or increase the distance to the sources by measuring at a higher altitude.

IV

Applications

13

General Applications

One of the goals of this thesis is to get a better understanding of what magnetic maps are and how they can be used in navigation and localization. In previous chapters, we discussed the theory behind making magnetic maps. For this chapter, we assume that magnetic maps are globally available. We briefly discuss the possible applications of these maps when we have access to them. First, we refer to some research on aeromagnetic navigation. Then we discuss magnetic anomaly detection. Finally, we discuss gold, mineral and other types of exploration.

13.1. Aeromagnetic Navigation

Many navigation systems are dependent on GPS. However, GPS is not reliable in all situations. Therefore, several studies have been conducted to investigate the possibility of GPS-independent navigation. Robust, GPS-independent navigation systems are important for applications such as autonomous vehicles, drones, and aircraft. Aeromagnetic Navigation is one of the methods that has been investigated [13]. It refers to the use of magnetic maps to navigate an aircraft.

Aaron Canciani has many contributions to the field of Aeromagnetic Navigation [13, 14, 16, 15]. He is an old Air Force pilot who is now actively researching this topic at the Air Force Institute of Technology. In [14] he describes how magnetic anomaly maps can be used to passively navigate an aircraft by combining data from an inertial navigation system and barometric data with measurements from a total-field magnetometer. It was concluded that an accuracy of 13 m can be attained under perfect conditions. According to the paper, the three main challenges for navigation on magnetic anomaly maps are: (1) making high-quality magnetic anomaly maps, (2) performing platform compensation to account for the magnetic field of the aircraft in the measurements, and (3) the resolution and availability of maps at different heights. The first challenge is discussed in Part II of this thesis. The second challenge is briefly discussed in Chapter 7, but more research is required on platform compensation to make it more robust. The third challenge is partly resolved by using Upward Continuation (Chapter 10). The resolution aspect remains a problem: flying at a higher altitude reduces the spatial frequencies of the magnetic anomaly field, which makes it more difficult to navigate.

The maximum navigation accuracy of magnetic maps on a survey aircraft is around 10 m drms (Distance Root Mean Square Error). The limiting factors in the accuracy are the compensation of the platform and the spatial variations in the magnetic field. This accuracy seems sufficient to navigate an aircraft that travels large distances at a high speed. At the moment of writing this thesis, we did not have access to a working F-16 aircraft to test the navigation system. Fortunately, Canciani did test the navigation of an F-16 on a magnetic anomaly map [15]. It required a new compensation model to account for the platform field, but it was concluded that magnetic anomaly maps can give modest (59 m drms) navigation accuracy in the F-16.

13.2. Magnetic Anomaly Detection

The availability of high-resolution magnetic (anomaly) maps can be very valuable for magnetic anomaly detection (MAD). Magnetic Anomaly Detection is the process of using a signal to detect magnetic field variations caused by ferromagnetic objects. When a metallic object is present, it disturbs the Earth's magnetic field, creating an anomaly that can be detected by sensitive magnetometers (Figure 13.1). The Anderson Functions,

described in Chapter 9, play an important role in this process since these functions are a basis for the magnetic field measurements on a line originating from a single dipole. And because the magnetic field of many magnetic objects can be approximated by a single dipole (Section 2.5.1), it is possible to detect and locate magnetic objects using measurements. The magnetic field intensity that originates from a ferromagnetic object on Earth's surface is often of a much smaller order than the magnetic field intensity that originates from the Earth's core and crustal field.

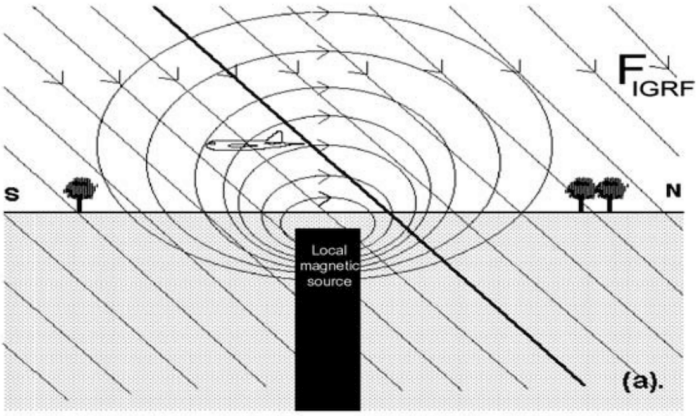


Figure 13.1: In a particular position, the airborne magnetometer captures the combined vector of the prevailing geomagnetic field and any anomalous inputs from one or multiple nearby sources. This figure is obtained from Reeves' work [54, pp. 1–10].

A key challenge in magnetic anomaly detection is differentiating the magnetic anomalies caused by the target object from background noise caused by the ambient geomagnetic field and other environmental factors [72]. Here is where high-resolution magnetic models (Chapter 6) and anomaly maps (Section 5.4) come in handy: they can be used to subtract the crustal field from the measurements and on the remaining signal, magnetic anomaly detection can be performed to look for magnetic objects. Therefore, a better map of the magnetic 'background' field allows for better detection of changes in the magnetic field caused by magnetic objects.

One of the leading companies in the MAD field is CAE, which has developed a highly advanced MAD system for military defense applications. Their latest system, the ASQ-252(V) MAD-XR, is more compact [12], allowing it to be used on smaller platforms like unmanned aerial systems, helicopters, and small fixed-wing aircraft. The MAD-XR detects changes in the Earth's magnetic field due to metallic objects nearby, providing valuable information for anti-submarine warfare and other applications.

13.3. Gold, Mineral and Oil and Gas Exploration

Magnetic anomaly maps also have significant value in discovering gold and other precious minerals [44, 57]. The high value of gold and other minerals is primarily due to their rarity. Gold is a diamagnetic metal, repelling magnetic fields, which makes its detection with magnetometers challenging. Nonetheless, examining host rocks and their magnetic characteristics can help locate gold and diamond deposits.

A notable instance of this approach is the research conducted by Shahri et al. [61], where they successfully utilized magnetic data to identify gold deposits in Iran: specifically, their focus was on Hired, a large gold prospecting area in the South Khorasan province of eastern Iran, which features gold mineralization in four target areas covering about 24 km^2. The host rocks mainly consist of Tertiary volcanic and some Jurassic and Cretaceous sediments. There is a strong correlation between gold grade, amount of pyrrhotite, and magnetic susceptibility in the stockwork mineralization east of target 1. Ground magnetic surveys were selected as a suitable geophysical method for drill target identification, with the total magnetic field intensity measured in 780 points along 25 lines, revealing large anomalies representing the magnetic responses of gold ore. Similar approaches have been used in the search for diamonds. For example, the diamond-rich kimberlite pipes in South Africa are often associated with magnetic anomalies [40].

Magnetic maps can be used for oil and gas exploration by providing valuable information about subsurface structures and geological formations that are associated with hydrocarbon reservoirs [65].

14

Localization using a Particle Filter

Magnetic Maps can play an important role in GPS-independent navigation. They can be used as a passive form of localization in conflict zones, an alternative to GPS in indoor or underground environments, or as a backup for GPS in case of jamming or spoofing. In this chapter, we show some examples of how magnetic maps can be used for localization using a particle filter and how map-matching can be used to find the trajectory of an object in a magnetic map. We introduce a particle filter and show how it can be used to estimate the position of a robot using a magnetic map on a floorplan. We also show how multiple maps can be combined to improve the localization. Finally, we demonstrate a map-matching algorithm that uses a basic evolutionary algorithm to find the trajectory of the robot by matching the magnetic field measurements to the magnetic map.

14.1. Particle Filter

A particle filter is a Monte Carlo method for estimating the state of a system. It is a recursive algorithm that uses a set of particles to represent the state of the system. The particles are initialized with a prior distribution and then updated with the current measurement. The particles are then resampled to represent the posterior distribution of the system. The algorithm is summarized in Algorithm 1.

Algorithm 1 Particle Filter

1: Initialize particles $x_1, x_2, \ldots, x_N$ with prior distribution $p(x)$
2: **for** $t = 1$ to T **do**
3: Update particles $x_1, x_2, \ldots, x_N$ with measurement z_t using $p(z_t|x_t)$
4: Resample particles $x_1, x_2, \ldots, x_N$ from posterior distribution $p(x_t|z_{1:t})$
5: **end for**

14.2. Particle Filter for Position Estimation using a Magnetic Field Map

A particle filter can be used to estimate the position of an object using a magnetic field map. The initial state of the system is the position of the particles. The particles can be initialized randomly within the map m if the prior distribution is unknown. The measurement z is the magnetic field at the current position of the object. Given a magnetic field map m and a set of particles $x_1, x_2, \ldots, x_N$, the likelihood of a measurement z at a position x is given by

$$p(z|x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{(z - m(x))^2}{2\sigma^2}\right), \tag{14.1}$$

where σ is the standard deviation of the measurement noise.

The particles are initialized with a prior distribution $p(x)$ and then updated with the current measurement z_t using the likelihood function $p(z_t|x_t)$. The particles are then resampled to represent the posterior distribution $p(x_t|z_{1:t})$.

14.3. Particle Filter for Position Estimation using Multiple Maps

The particle filter methodology can be extended to use multiple maps. For example, a particle filter that uses a magnetic field map and a topographic map can be used to estimate the position of an object, or a filter that uses maps of separate magnetic field components. Measurements from multiple sensors must be combined so that the likelihood function $p(z_t|x_t)$ can be calculated Let z_k be the measurement for map m_k and $p_k(z_k|x_t)$ be the likelihood function for map m_k. The likelihood function for the combined measurement is given by

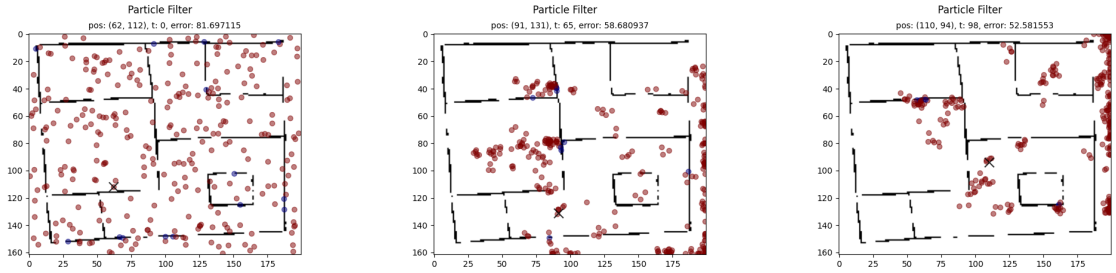
$$p(z_t|x_t) = \prod_{k=1}^{K} p_k(z_k|x_t)\,, \tag{14.2}$$

where K is the number of maps. This likelihood function is only valid if the measurements are independent. If the measurements are not independent, then the likelihood function must be modified to account for the correlation between the measurements.

14.4. Particle Filter for Localization using a Floorplan and a Magnetic Map

In the following examples, we demonstrate how localization can be performed using a particle filter. To make it a bit more interesting, we performed this demonstration in the context of indoor navigation for robots. A similar approach can be used for aeromagnetic navigation, by removing the walls. For now, we assume that we have only access to a floorplan of the building and that our robot has an odometry error of 0.1, meaning that for every meter the robot travels, it has a 10 cm standard deviation error. Later, we also add a Total-field magnetometer to the robot to improve the localization. Finally, measurements from a vector magnetometer are used to improve the localization even further.

The floorplan in our example is drawn by hand and is stored as an image. We wrote a Python script to binarize the image using a threshold value to detect 'walls' in the image. It is therefore not difficult to use a different floorplan, as long as it is stored as an image. After processing, the floorplan is stored in an object and can easily be visualized or accessed by the particle filter.



(a) Initial particle distribution in the building. The cross indicates the true, but unknown, position of the robot. The walls are shown in black.

(b) The state of the particle filter after moving the robot around for a while. Some particles have crashed into the walls and are removed from the set of particles. The remaining particles are resampled to represent the posterior distribution of the robot's position.

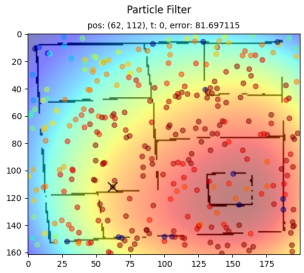
(c) The state of the particle filter after more movement. The particles are resampled again to represent the posterior distribution of the robot's position.

Figure 14.1: Particle filter for localization using only a floorplan.

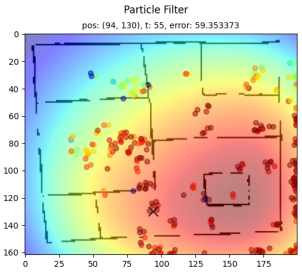
First, we demonstrate how the particle filter can be used to localize a robot in a building using only the floorplan. We initialize the particles randomly within the building (Figure 14.1a) and update their position with the relative motion of the robot obtained from the odometry sensors. Particles that crash into a wall are removed from the set of particles, resulting in a 'binary' likelihood function. The remaining particles are resampled to represent the posterior distribution of the robot's position (Figure 14.1b). Again, the particles are updated with the relative motion of the robot to get its final position. The same procedure is repeated, and the resulting posterior distribution is shown in Figure 14.1c. There is still a lot of uncertainty in the position of the robot, but there are some particles that are close to the true position of the robot. Moving the robot more around reduces the uncertainty even further.

Now, we provide a Total-field magnetic map of the environment to our robot. In our example, we have simulated a field using a single dipole. In general, these fields are more complex, especially when we consider in-building navigation, there are many factors originating from human activities that cause the magnetic field to change. Therefore, it might be interesting to perform in-building navigation on artificially generated

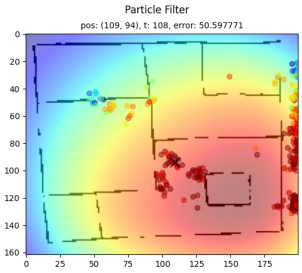
magnetic fields, by using beacons that overpower the local field (Section 14.5.6). This way, the field is better known and can therefore be used for navigation. For aeromagnetic navigation, it is not possible to overpower the field, but the noise and variation in the field are much lower, which allows for making accurate maps of the field (Part II).



(a) Initial particle distribution in the building. The cross indicates the true, but unknown, position of the robot. The walls are shown in black.



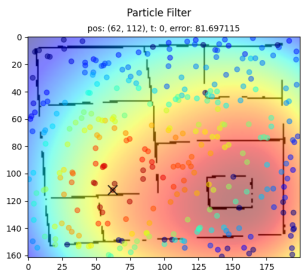
(b) The state of the particle filter after moving the robot around for a while. Some particles have crashed into the walls and are removed from the set of particles. The likelihood of the particles is calculated using the magnetic map. The remaining particles are resampled to represent the posterior distribution of the robot's position.



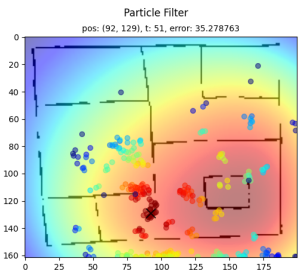
(c) The state of the particle filter after more movement. The particles are resampled again to represent the posterior distribution of the robot's position.

Figure 14.2: Particle filter for localization using a floorplan and a map of the total-field intensities of the magnetic induction field **B**. The underlying magnetic map that is used for localization is shown in the background. The color of the particles indicates the likelihood of the particles.

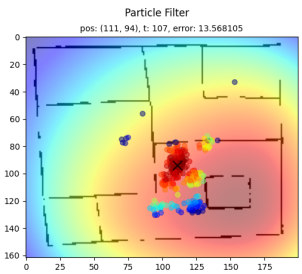
The same particle filter as before is used, but now the likelihood function is calculated using the Total-field magnetic map $|\mathbf{B}|$, as described in Section 14.2. Figure 14.2a shows the initial state of the particle filter. After movement, the resampling is done using a likelihood function that is based on the difference between the magnetic field value $m(x)$ at the position of a particle and a total field measurement (z) at the position of the robot. The likelihood ranges from zero to one, where zero means that the particle is in a location where the magnetic field is very different from the measured field, and one means that the particle is in a location where the magnetic field is the same as the measured field. This resampling strategy leads to the posterior distribution shown in Figure 14.2b. After another step, we get the posterior distribution shown in Figure 14.2c. The uncertainty in the position of the robot is reduced significantly. The particles are now concentrated around the true position of the robot.



(a) Initial particle distribution in the building. The cross indicates the true, but unknown, position of the robot. The walls are shown in black.



(b) The state of the particle filter after moving the robot around for a while. Some particles have crashed into the walls and are removed from the set of particles. The likelihood of the particles is calculated using multiple magnetic maps. The remaining particles are resampled to represent the posterior distribution of the robot's position.



(c) The state of the particle filter after more movement. The particles are resampled again to represent the posterior distribution of the robot's position.

Figure 14.3: Particle filter for localization using only a floorplan, a map of the total-field intensities **B** and three separate maps of the components of the magnetic field B_x, B_y and B_z. The map of total-field intensities is shown in the background. The color of the particles indicates the likelihood of the particles.

The convergence of using a single map of the total-field intensities of the magnetic induction field **B** is shown in Figure 14.2 is significantly better than the convergence of using only the floorplan, as shown in Figure 14.1. The total-field map, however, does not contain any information about the direction of the magnetic field. It happens often that the magnetic field value at several locations in the building is the same.

One can see in the background of Figure 14.2, that when using beacons, the total-field value is the same on circles around the beacons. In an ideal world, we navigate on a map that is continuous and has unique values. The total-field map is not unique. We can improve the localization by using more information about the magnetic field. We perform the same experiment, but we add a vector magnetometer to the robot. This sensor measures the three separate components of the magnetic field, which all lead to a map that can be used as input for our particle filter. The likelihood function is given by the product of the likelihood functions on the separate maps $|\mathbf{B}|$, B_x, B_y and B_z, as described in Section 14.3. The same movements and resamplings are performed as before. The initial state and posterior distributions of the particles are visualized in Figure 14.3a, 14.3b and 14.3c. The posterior distribution of the particles after using multiple maps in the particle filter lies closer to the actual position of the robot.

14.5. Wake-up Robot Problem: Map-matching for noisy measurements on a heatmap

This section focuses on solving the wake-up robot problem in a magnetic field. A wake-up robot [20] is unaware of its current location. To find out its location, the robot is allowed to drive around for small bits and make measurements. To tackle this problem, a map-matching algorithm is applied to the values of the magnetic field measured at different locations. Map-matching algorithms are commonly used in navigation to map inaccurate GPS data to a road network [10]. They expressed the trajectory of a vehicle in samples of GPS data. The samples are evaluated one by one. The roads around the current location are found by evaluating the GPS measurements on a map. The distances to these roads are computed. Minimizing a weighted sum of the distances and a measure for the orientation (derived using the previous location) that was traveled, gives an estimate of the new position. This section focuses on a similar algorithm for map-matching, using magnetic fields. Such a method would require robots to have magnetic sensors and gather their odometry data.

14.5.1. Particle Swarm Initialization

A measurement of a magnetic field does not necessarily give information about the location: the map ($B: \mathbb{R}^2 \to \mathbb{R}^3$) can be non-unique, thus there might be different locations, far away from each other, that have similar magnetic field strengths. In contrast, GPS data provided by satellites gives an approximation of where the vehicle is, due to the unique mapping of the positioning data on the Earth. The initial position of the wake-up robot is thus unclear. Two approaches can help find the initial particle positions. The first approach makes use of an initial measurement (Figure 14.4a). The second approach samples uniformly on a grid (Figure 14.4b).

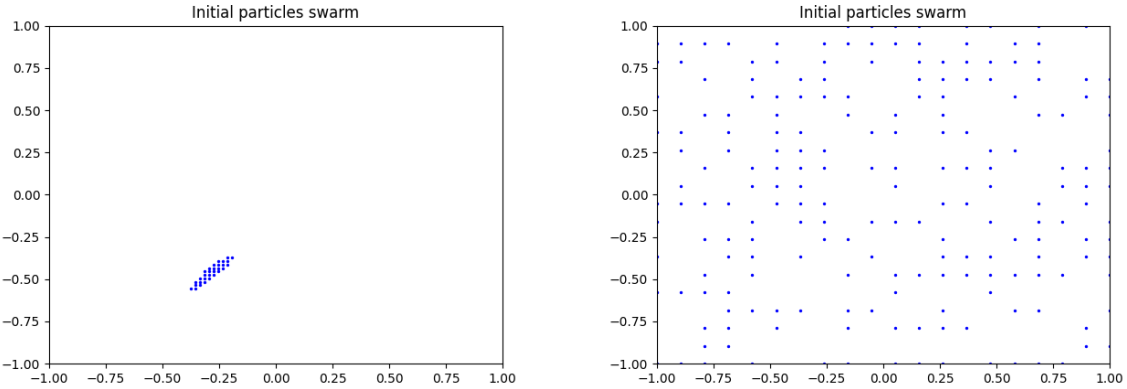
A single measurement of the magnetic field could exclude many parts from the domain. Due to the continuity of the domain, it is not possible to find the points where the values of the magnetic field are close to the measurement. However, assuming that our magnetic field is continuous and sufficiently smooth, evaluations of the magnetic field on a coarse grid could give us some initial position estimates for the algorithm.

To initialize the particle swarm, the following data is required:

- field: a map with the magnetic field components in x-, y- and z-direction.
- points: a selection of x- and y-coordinates in the space, e.g. a coarse grid.
- m_0: an initial measurement of the magnetic field consisting of three components: the magnetic field strength in x-, y- and z-direction.
- n: the number of particles to find.

The list of initial particles is obtained by selecting the n particles from $points$ that minimize the distance, in the L^2-norm, between the three magnetic field components of the point and the measurement m_0. Pseudocode for the implementation is given in Pseudocode 0 and an example of initial particle swarms are visualized in Figure 14.4.

The second approach uses only a selection of x- and y-coordinates in the space. Preferably, these coordinates are uniformly distributed. This can for example be done by sampling points from a grid, without replacement. The preferred initialization method depends on the restrictions that are applied in the map-matching process. This is elaborated in Section 14.5.5.



(a) Example of an initial particle swarm fitted to a measurement with $n = 30$ on a grid of 100×100 points.

(b) Example of a uniform initialized particle swarm with $n = 200$ on a grid with 20×20 points.

Figure 14.4: Examples of two different initializations of the particle swarm. The blue dots represent the particles in the swarm.

Algorithm 2 Particle swarm initialization

function INITIALIZEPARTICLES($field, points, m_0, n$)
  initialize an empty list of particles.
  initialize an empty list of distances.
  for p in $points$ **do**
    $m_p \leftarrow$ magnetic field components of p in $field$.
    distance $\leftarrow \|m_p - m_0\|_2$
    if distance < distances[-1] **then**
      distance[-1] = distance
      particles[-1] = p
      sort particles in ascending order of distances
      sort distances in ascending order of distances
    end if
  end for
end function

14.5.2. Trajectory of the particle swarm

The odometry data of the robot can be used to determine its trajectory. In this section, we assume that the robot has perfect odometry, meaning that the relative change in position between two measurements is known exactly. In Section 14.5.4, we compute the trajectory of a robot that has inaccurate odometry.

With perfect odometry, it is trivial to find the trajectory of the robot. The same trajectory is used for the particles. Let p_i^k be the position of particle k at time i, then

$$\mathbf{p}_{i+1}^k = \mathbf{p}_i^k + \mathbf{o}_i\,, \tag{14.3}$$

where o_i is the relative change in position between timestamp i and $i+1$. In Figure 14.5 one can see that each particle has the same (relative) trajectory.

14.5.3. Fitness of a trajectory

In practice, the odometry data of a robot is far from perfect. This means that we cannot rely on only the odometry data to determine the trajectory of the robot. Instead, we extend our particle swarm algorithm with a fitness function that compares the trajectory of the particles to the measurements of the magnetic field. Fitness functions for GPS algorithms for vehicles are easy to define: the distance to the closest road in combination with the orientation of the vehicle is a good measure. Finding a good fitness function for magnetic field measurements is more difficult: the measurements live in a continuous space consisting of field strengths in 3 dimensions and without any landmarks. Therefore, there is no 'closest' point to a measurement, we can however compare the measurement to the values of the magnetic field at certain points in the map.

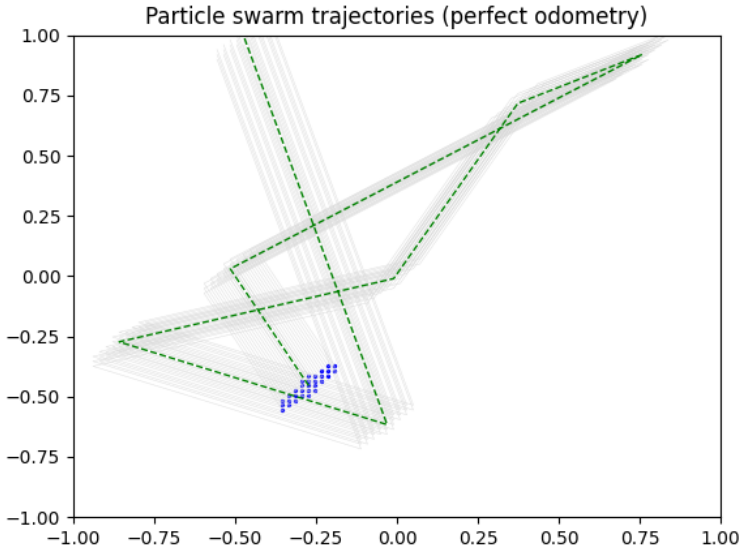


Figure 14.5: Trajectories of particles in a swarm of size $n = 30$ assuming perfect odometry. The trajectory with the best fitness is indicated in green.

Let $\mathbf{b}_i^k := B(\mathbf{p}_i^k)$ be the three values of the magnetic field map at location $\mathbf{p}_i^k$. The local error $\mathbf{e}_i^k$ is given by:

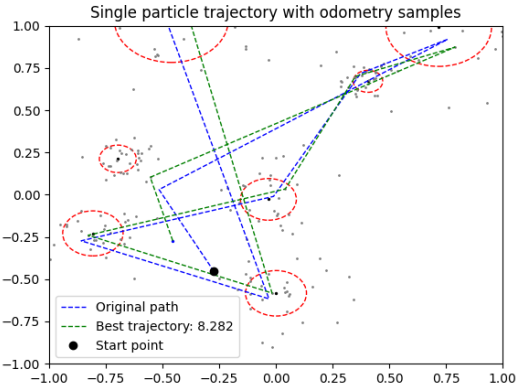
$$\mathbf{e}_i^k := \|\mathbf{b}_i^k - \mathbf{m}_i\|_2\,, \tag{14.4}$$

where m_i is the i-th measurement. Note that the exact values of the magnetic field at position $\mathbf{p}_i^k$ are unknown and that B is thus a generated map of the magnetic field.

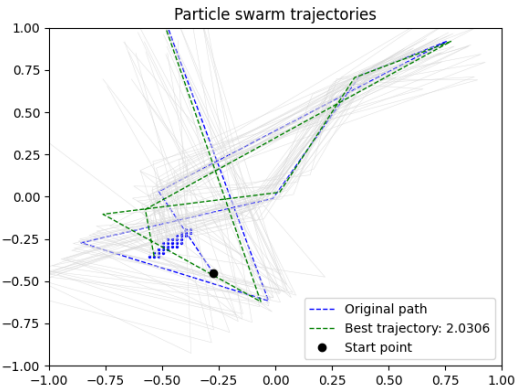
The fitness of trajectory i is now given by:

$$f_i := \|\mathbf{e}_i^k\|_2\,. \tag{14.5}$$

The optimal trajectory (Figure 14.5) can now be found by minimizing the fitness function.



(a) Trajectory of a single particle with an odometry error of 15%, indicating 1 standard deviation of the bivariate normal distribution with a red circle and the samples $\mathbf{e}_{odom}^{i,j}$ with gray dots.



(b) Trajectories of 30 particles with an odometry error of 15%.

14.5.4. Odometry error correction

Until now, it was assumed that a robot navigating through an environment can maintain an accurate relative position using odometry data. However, in real-world situations, odometry data may not be flawless. To account for this, an odometry error can be introduced when modeling the trajectories. The magnitude of this error depends on the traveled distance, represented by $|o_i|$, and the accuracy a of the odometry sensors. The odometry error is assumed to follow a bivariate normal distribution:

$$\mathbf{e}_{odom}^{i} \sim \mathcal{N}_2(0, 2a|o_i|)\,. \tag{14.6}$$

By drawing samples from this distribution, we can explore variations in particle positions as defined by Equation 14.3. This helps us investigate the vicinity of the new particle positions. The sample that minimizes the local error, $\mathbf{e}i^k$, is chosen for further trajectory calculations. Let's take S samples and denote the j-th sample as $\mathbf{e}_{odom}^{i,j} \sim \mathcal{N}_2(0, 2a|o_i|)$. Then, the selected sample is the $\hat{j}$-th sample, given by:

$$\hat{j} = \underset{j=1\ldots S}{\arg\min} \left| \mathbf{B}(\mathbf{p}i^k + \mathbf{o}i + \mathbf{e}_{odom}^{i,j}) - \mathbf{m}_i \right|_2 . \tag{14.7}$$

Here, B represents the magnetic field map. Since the sample $\hat{j}$ minimizes the local error, it can be incorporated into Equation 14.3 to obtain a better approximation of the magnetic field. Figure 14.7a shows the random samples, with the indicated samples corresponding to the points where the trajectory changes direction. These samples represent the new particle positions, now given by:

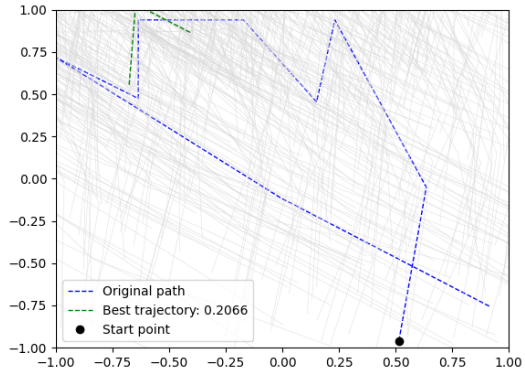
$$\mathbf{p}_{i+1}^{k} := \mathbf{p}_i^k + \mathbf{o}_i + \mathbf{e}_{odom}^{i,\hat{j}}\,. \tag{14.8}$$

We can use the same approach as in Section 14.5.3 to find the best trajectory. The best trajectory is the one that minimizes errors in the magnetic field values at the measurement points. Figure 14.6b displays the trajectories of 30 particles with a 15% odometry error. Comparing this figure to Figure 14.5, we can observe that the trajectories are no longer parallel. The odometry error correction explores the vicinity of particle positions, leading to minor changes in the direction and length of the relative paths.

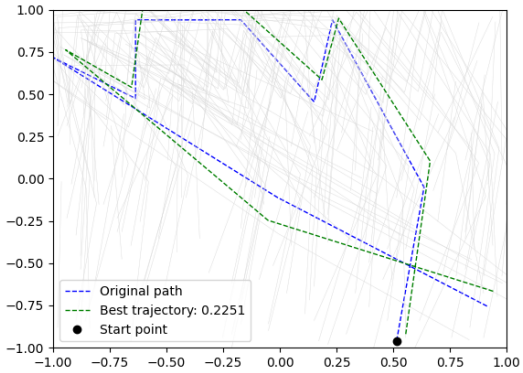
14.5.5. Domain restriction

When a robot moves within a specific environment, its possible positions are often limited by the constraints of that space. For instance, the domain may be enclosed by walls, or contain obstacles that create restricted areas. These domain constraints can play a significant role in determining the optimal trajectory for the robot on the map. To account for these restrictions, certain positioning samples or entire trajectories might be disregarded if they fall outside of the allowable domain.

A restriction factor is introduced to manage small movements outside the domain, which is denoted as $\Omega = [-1, 1]^2$. With a given restriction factor of $r > 0$, the restricted domain becomes $\hat{\Omega} = [-r, r]^2$, allowing a $(1-r)$% overshoot at the domain boundaries. Trajectories that have points outside the restricted domain are considered infeasible.



(a) Trajectories of particles in a swarm of size $n = 200$ assuming an odometry error of 5%. The trajectory with the best fitness is indicated in green.

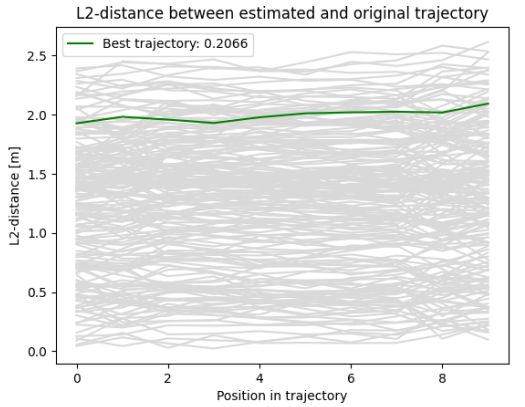


(b) Domain restricted trajectories of particles in a swarm of size $n = 200$ assuming an odometry error of 5%. The used domain restriction factor is 1.05. The trajectory with the best fitness is indicated in green.

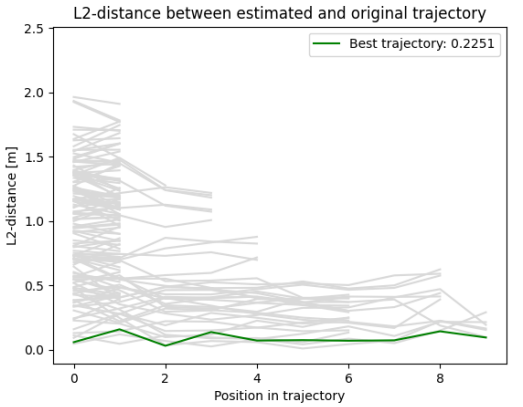
Figure 14.7: Unrestricted and restricted domains.

Incorporating domain restrictions can significantly improve the performance of the map-matching algorithm. This is particularly noticeable in cases where many trajectories extend beyond the domain, such as uniformly initialized particle swarms with long trajectories. Figure 14.7 provides an example of this, comparing 200 trajectories without domain restrictions to those with domain restrictions.

It is worth noting that the fitness of the best trajectory is higher in the unrestricted domain, but the distance between the best trajectory and the original (unknown) path we are trying to estimate is significantly smaller for the trajectory in the restricted domain. Figure 14.8 (a) illustrates the spatial distances between



(a) Unrestricted domain.



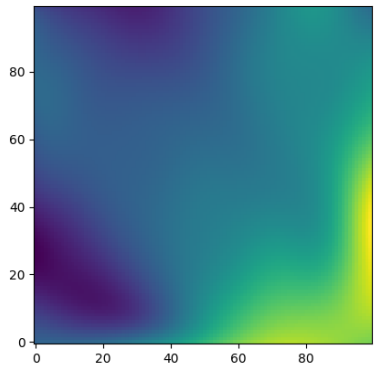
(b) Restricted domain with restriction factor is 1.05.

Figure 14.8: Distance between the original path and the domain restricted trajectories of particles in a swarm of size $n = 200$ assuming an odometry error of 5%. The trajectory with the best fitness is indicated in green.

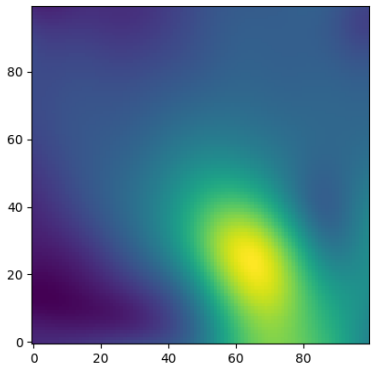
the points of the trajectories shown in Figure 14.7 and the original path. Figure 14.8 (b) presents the results for trajectories with domain restrictions, leading to cut-off trajectories. Evidently, the algorithm with domain restrictions performs significantly better in minimizing the distance to the original path.

14.5.6. Map-matching with beacons

In this section, we examine how the placement of beacons influences the accuracy of the map-matching algorithm. Beacons are used to (partially) overpower the noisy and challenging-to-model local magnetic field. For the experiment, a single beacon is placed in the room, and magnetic field maps are generated both with and without the beacon turned on. Figure 14.9a presents the map of the field without beacons. Meanwhile, Figure 14.9b displays the map of the magnitudes of the local field when the beacon is present



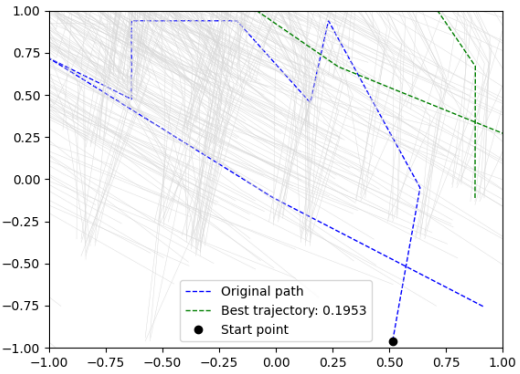
(a) Magnitude of the local magnetic field.



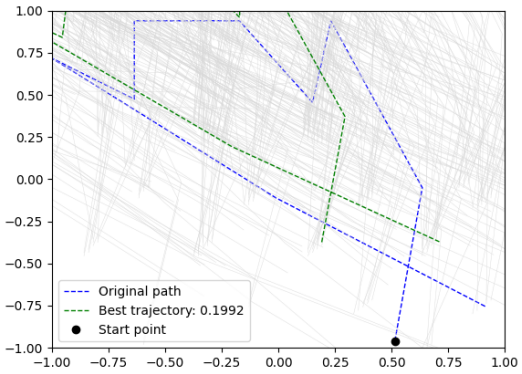
(b) Magnitude of the local magnetic field overpowered by a beacon.

Figure 14.9: Magnetic field maps with and without a beacon.

and overpowering the local magnetic field.



(a) Trajectories on a local magnetic field of particles in a swarm of size $n = 200$ initialized using an initial measurement. Assuming an odometry error of 5% and an unrestricted domain. The trajectory with the best fitness is indicated in green.



(b) Trajectories on an overpowered magnetic field of particles in a swarm of size $n = 200$ initialized using an initial measurement. Assuming an odometry error of 5% and an unrestricted domain. The trajectory with the best fitness is indicated in green.

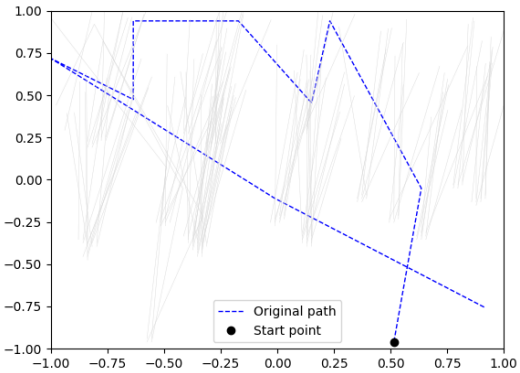
Figure 14.10: Trajectories on an unrestricted domain with measurement-based particle swarm initialization.

In this part, the map-matching algorithm is applied to find the trajectories in scenarios with and without a beacon. Figures 14.10 and 14.11 illustrate the differences in the map-matching algorithm for a magnetic field with and without a beacon. The trajectories are constructed from an initialized particle swarm based on measurements. Figures 14.12 and 14.13 present the results of the same configuration but with a uniform initialization.

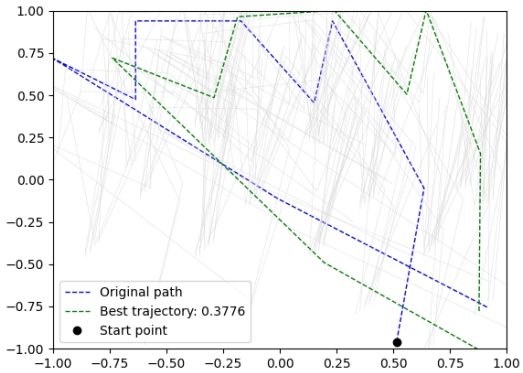
When observing Figure 14.11 (a) and (b), it becomes apparent that the restricted domain without a beacon and with initialization based on an initial measurement fails to find a feasible solution. However, the same configuration with a beacon successfully finds a solution. Comparing subfigure (a) of the local field with subfigure (b) of the overpowered field in Figure 14.10 - 14.13 reveals that placing a beacon improves the accuracy of the estimated trajectory position. Beacons are most effective when the particle swarm is initialized with an initial measurement.

14.5.7. Discussion and Conclusion

The magnetic map-matching approach using particle swarm optimization demonstrates potential, but there are some challenges to address. Firstly, as observed in Section 14.5.5, without domain restrictions, the model struggles to consistently reproduce the robot's trajectory. This implies that the method is most suitable for

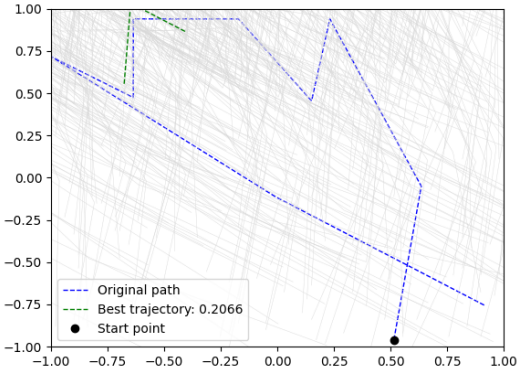


(a) Trajectories on a local magnetic field of particles in a swarm of size $n = 200$ initialized using an initial measurement. Assuming an odometry error of 5% and a restricted domain. The trajectory with the best fitness is indicated in green.

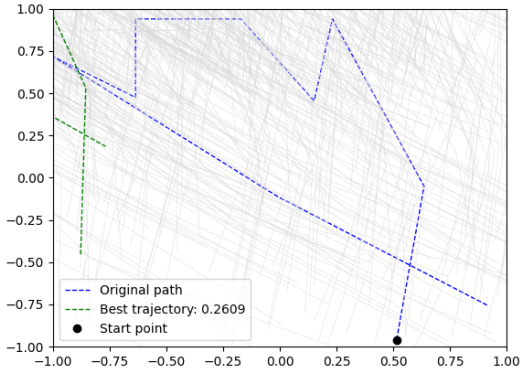


(b) Trajectories on an overpowered magnetic field of particles in a swarm of size $n = 200$ initialized using an initial measurement. Assuming an odometry error of 5% and a restricted domain. The trajectory with the best fitness is indicated in green.

Figure 14.11: Trajectories on a restricted domain with restriction factor 1.05 and a measurement-based particle swarm initialization.

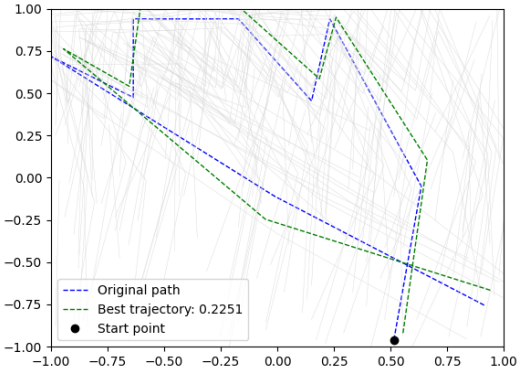


(a) Trajectories on a local magnetic field of particles in a swarm of size $n = 200$ uniformly initialized. Assuming an odometry error of 5% and an unrestricted domain. The trajectory with the best fitness is indicated in green.

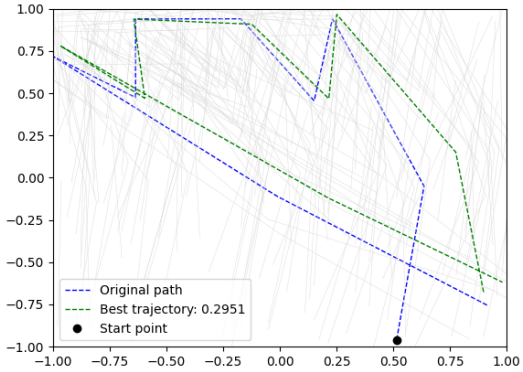


(b) Trajectories on an overpowered magnetic field of particles in a swarm of size $n = 200$ uniformly initialized. Assuming an odometry error of 5% and an unrestricted domain. The trajectory with the best fitness is indicated in green.

Figure 14.12: Trajectories on an unrestricted domain with uniform particle swarm initialization.



(a) Trajectories on a local magnetic field of particles in a swarm of size $n = 200$ uniformly initialized. Assuming an odometry error of 5% and a restricted domain. The trajectory with the best fitness is indicated in green.



(b) Trajectories on an overpowered magnetic field of particles in a swarm of size $n = 200$ uniformly initialized. Assuming an odometry error of 5% and a restricted domain. The trajectory with the best fitness is indicated in green.

Figure 14.13: Trajectories on a restricted domain with restriction factor 1.05 and a uniform particle swarm initialization.

known environments. The findings in Section 14.5.6 suggest that incorporating magnetic beacons within the domain can enhance the accuracy of the technique. Additionally, increasing the swarm size also improves the method's precision. However, it is important to exercise caution to avoid overfitting.

In summary, the magnetic map-matching technique with particle swarm optimization presents a promising approach to estimating robot trajectories, particularly when domain restrictions are applied and magnetic beacons are used in known environments. Further optimization, such as refining swarm size and initialization, can help to balance the trade-off between accuracy and overfitting, ensuring a more reliable and robust estimation of robot trajectories.

The principles discussed in the context of magnetic map-matching using particle swarm optimization can also be applied to other types of maps, such as altitude maps, gravitational maps, and more. These maps represent different physical properties of the environment but can similarly benefit from optimization techniques to improve the accuracy of trajectory estimation or navigation tasks. For example, when working with altitude maps, the optimization algorithm could help in estimating the most efficient path over varying terrain, taking into account factors like elevation changes and obstacles. Similarly, gravitational maps could be utilized to understand variations in gravitational forces in a region, which could be valuable for geological or geophysical studies.

The algorithms and techniques discussed can also be employed for aeromagnetic navigation, which involves using magnetic data collected from airborne surveys to identify and track geological structures and mineral deposits. By applying particle swarm optimization or other optimization techniques, the accuracy of the estimated flight paths and the interpretation of the collected magnetic data can be improved. This, in turn, can lead to more efficient and precise identification of geological features, mineral deposits or ferromagnetic objects.

In conclusion, the concepts and techniques explored in the context of magnetic map-matching using particle swarm optimization are versatile and can be extended to various types of maps and applications, including aeromagnetic navigation and beyond.

V

Conclusions and Recommendations

15

Conclusions and Recommendations

15.1. Conclusions

The goal of this thesis was to develop a better understanding of the theory and practice of magnetic map-making techniques. Throughout this thesis, we have tried to answer the question: What are the fundamental principles and concepts underlying magnetic map-making? To answer this question, we have formulated the following research questions:

1. What is the definition and purpose of a magnetic (anomaly) map?
2. What are the established methodologies and techniques utilized in the creation of magnetic maps?
3. How can the depth of dipole sources be accurately determined from the measured signal?
4. What is the empirical performance of current interpolation and extrapolation techniques for magnetic maps?
5. What is the impact of environmental and platform noise on measurement data in the process of map-making, and what strategies can be employed to minimize this influence?
6. In what ways can magnetic (anomaly) maps be utilized for navigation, localization, and exploration?

We started with a study of magnetic fields (Chapter 2) and the structure of Earth's magnetic field (Chapter 3) and look into existing techniques for representing the Earth's magnetic field. The first of these uses spherical harmonics to model the field with Legendre polynomials (Chapter 6). This is very effective on larger, continental, scales with wavelengths starting at 2500 km. Higher-resolution methods, using a 720-degree spherical harmonic expression, allow for a 56 km spatial resolution. This resolution is not enough for navigation purposes, therefore, we looked for methods that allow for a higher resolution. This is where magnetic anomaly maps play an important role. The definition of a magnetic (anomaly map), and thus Research Question 1, is handled in Chapter 5. The maps can be used to enhance coarser, global, reference fields with local information by mapping the deviation from the lower-resolution reference field. These maps can be made using a variety of techniques, which mainly consist of interpolating the measurement data.

Interpolation and Extrapolation Techniques

In practice, especially in the context of aeromagnetic maps, it is not possible to take measurements at every altitude above Earth's surface. Therefore, methods for extrapolating the measurements to higher altitudes are needed. The answer to Research Question 2 is twofold: we need interpolation techniques for mapping measurements to a different set of points, and extrapolation techniques, for finding the magnetic field at altitudes where no measurements are available. The methods concerned with interpolation are not extensively discussed in this thesis but are mentioned in Chapter 8.

The extrapolation methods are studied in more detail and make use of Green's third identity (Chapter 2.4.1). Green's third identity tells us that the magnetic field above an infinite plane is completely defined by the magnetic field on the plane, as long as all the sources lie below the plane. This theorem is the underlying

assumption for the extrapolation methods studied in our work. The simplest of these methods is the Upward Continuation method (Chapter 10), which exists in an integral form and as a Fourier transformation over the plane. Another method, Equivalent Layer (Chapter 8), tries to solve the inverse problem of finding the magnetic sources that could have generated the measured field. This is done by positioning one, or multiple, layer(s) of dipoles underneath the surface. The position and strength of these dipoles are then optimized to minimize the error between the measured field and the field generated by the dipoles.

A signal-based approach

In the literature, we found that Anderson functions can be used to find the position and magnetic moments of single magnetic sources from a signal. This prompted us to delve deeper into these functions, as they may offer a potential solution for estimating the position of multiple sources using the signal alone. This eliminates the requirement for a priori knowledge in the Equivalent Layer method (Chapter 9).

To do so, we constructed a Wavelet Family based on the Anderson functions, by shifting and scaling the functions. A QR-decomposition was used to orthonormalize the functions and to find the coefficients of the functions in the basis. The wavelet coefficients can be obtained by projecting the measured signal on the orthonormal wavelet family. We tested this projection on several types of fields with different complexities, with and without noise. The results show that the Anderson wavelet family is capable of accurately reconstructing the signal of the magnetic field generated by multiple sources.

We extended the three Anderson functions, which form a basis for the magnetic field of a dipole on a line, to two dimensions, to form a basis on a plane. To our knowledge, the Anderson Wavelet Family, and the representation of two-dimensional Anderson functions is a novel technique. Again, it was shown, that the resulting wavelet family, consisting of two-dimensional Anderson functions, can be used to represent the magnetic field of multiple sources, which extends the applicability of the Anderson functions from one-dimensional measurements on a line to two-dimensional measurements on a plane. Under certain circumstances, we managed to find a translation from the Anderson coefficients to the position and magnetic moments, but a general solution is still lacking. Research Question 3 is therefore only partially answered by the results of Chapter 9. Further research is required to find back the exact position of the sources that are used for reconstructing the signal using no a priori knowledge.

Benchmarking Map-making Techniques

We have implemented several of the interpolation and extrapolation methods in a Python package `MagMap` and benchmarked them on simulated magnetic fields (Chapter 11). We concluded that Cubic Splines are the best candidate for accurate and fast interpolation. Extrapolation is a more difficult problem. It seems that Upward Continuation is superior to other extrapolation methods for situations where we have no information about the distribution of the magnetic sources that generate the field. The Equivalent Layer requires some a priori knowledge about the sources to make an accurate model since the positioning of the layers significantly affects the accuracy of the method. The benchmarking results answer Research Question 4.

Real-world data and applications

The map-making techniques discussed in this thesis are mainly tested on simulated data of magnetic fields. Working with real-world data is more complicated. One has to measure the magnetic field with a magnetometer (Chapter 4), compensate for temporal variations in the magnetic field, and one has to deal with different types of noise. Especially platform noise, when attached to an aircraft, is a problem that needs to be addressed. The engine, currents, and ferromagnetic materials in an aircraft have a significant impact on the magnetic field measured with a magnetometer. The reduction of this noise, called aeromagnetic compensation, is studied in Chapter 7. To validate the map-making techniques on real-world data, we performed a measurement campaign and processed the data to create a magnetic map of a parking garage on the campus of TU Delft at two different altitudes (Chapter 12). The Upward Continuation and Equivalent Layer techniques were used to extrapolate the measurements from one altitude to the other. The extrapolated data was quite different from the measured field at the higher altitude. A clear explanation for this difference is still lacking, but we suspect that the type of sources and the preprocessing steps, such as rotating the sensor data from body to world frame, have a significant impact on the results. The steps taken to create the magnetic map of the parking garage, combined with the magnetic sensing and aeromagnetic compensation techniques, give a good indication of the challenges that one faces when working with real-world data. This answers Research Question 5.

To conclude this work, we looked at the applicability of magnetic maps. Chapter 13 briefly discusses the broad applications of magnetic maps. Among others, mineral exploration and navigation of an F-16 aircraft

are discussed. Finally, a small demonstration of localization and path-finding algorithms using particle filters is given in Chapter 14. These two chapters give an answer to Research Question 6.

15.2. Recommendations

Validity of High-resolution Magnetic Anomaly Maps
This research did not look at how to combine magnetic maps when several are available at different resolutions, or how to tile them together to create a larger map. Some work has been done on this topic, but mainly in the context of core and crustal fields [2]. It has to be verified if these methods are also applicable to higher-resolution anomaly maps. Moreover, the effect of temporal variations on magnetic anomaly maps is not studied in this thesis. The type of variations that high-resolution magnetic anomaly maps undergo takes place on a much shorter time scale than the variations of the core and crustal fields. Therefore, the methods used for core and crustal field models may not be directly applicable to high-resolution magnetic anomaly maps. Further research on the temporal variations in higher-resolution magnetic anomaly maps is needed to determine the validity of these maps and to find ways to update invalidated parts with new measurements.

Krylov methods for Equivalent Layer
In addition, a lot can still be gained on the computational side of the extrapolation methods. For example, the performance of Equivalent Layer seems to be dependent on two factors: the number of sources and the position of the sources. For a small number of sources, the linear system can be solved quite fast. But larger problem scales demand some optimizations. The EL problem is often solved using direct methods. Finding a suitable iterative (Krylov) method that solves the EL problem in a fast way seems to be a valuable step. This would require a sparse matrix representation of the problem, which requires some relaxation of the problem, but we think that this is possible since the strength of a dipole decays with a factor $1/r^3$.

Compact storage of magnetic maps
Another interesting topic is finding an efficient way to store high-resolution magnetic maps. Maps of fields at a fixed height can be interpolated to a regular grid and can then be stored in a (compressed) image. For maps of fields at different heights, it might be more efficient to only store the sources that can be used to create the map. These sources can for example be found with the Equivalent Layer or Anderson Wavelet methods. Extra research on compact storage of magnetic maps is required to see if this way of storing maps is indeed more efficient than storing the maps as images.

Relationship between the Anderson wavelet coefficients and Dipole parameters
In Chapter 9, we constructed a wavelet family based on the Anderson functions. The Anderson functions themselves form a basis for the magnetic field of a dipole evaluated on a line (Section 9.2). The wavelets in the basis are shifted and scaled versions of the Anderson function and have therefore also a relationship with the magnetic field of a dipole. We have found that, in specific cases, the position and magnetic moment of a dipole can be found from the Anderson wavelet coefficients. However, a general solution is still lacking. Further research is required to find back the exact position of the sources that are used for reconstructing the signal using no a priori knowledge.

Optimal Path-finding Algorithms for Magnetic Maps
In Chapter 14, we showed that a particle filter can be used to localize a magnetometer in a magnetic map. We noticed that spatial variations in the magnetic field play an important role in the localization. For example, the particle filter converges fast when the three components of the magnetic field form a unique pattern. This raises the question of whether there exist optimal paths in a magnetic map if one wants to move from one point to another. The optimal path would be the path that minimizes the uncertainty of the localization. Imagine a path that goes through a region where the magnetic field is constant, then the magnetic field measurements do not give any information about the position and one would have to rely on other inertial sensors. On the other hand, if there is another path that goes through a region where the magnetic field is spatially varying, then one can use the magnetic field measurements to localize the object. This can be very useful for the navigation of a drone. Especially in urban environments, it can be dangerous to lose information about the position of the drone. We think that the optimal path-finding problem can be solved using an evolutionary algorithm. The fitness function would be the uncertainty of the localization. The algorithm would then try to find the path that minimizes the uncertainty and reaches the destination in the shortest time possible.

Appendices

A

An introduction to Spherical Harmonics

Spherical harmonics are a set of functions that are used to model functions on a sphere. Spherical harmonic functions are orthogonal and form a complete basis set. This means that any function on a sphere can be represented as a linear combination of spherical harmonics. There are different sets of spherical harmonics, but the most common ones are the solid harmonics, functions on a sphere that satisfy the Laplace equation $\Delta f = 0$:

$$\Delta f = \frac{1}{r^2}\frac{\partial}{\partial r}\left(r^2\frac{\partial f}{\partial r}\right) + \frac{1}{r^2\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial f}{\partial\theta}\right) + \frac{1}{r^2\sin^2\theta}\frac{\partial^2 f}{\partial\phi^2} = 0\,, \tag{A.1}$$

where $f : \mathbb{R}^3 \to \mathbb{R}$ is a function of r, θ and ϕ. Often, the function f is rewritten as the product of a radial function R and a spherical harmonic function on the unit-sphere Y:

$$f(r,\theta,\phi) = R(r)\,Y(\theta,\phi)\,. \tag{A.2}$$

Separating the function into a radial function and a spherical harmonic function allows for finding an analytical solution to the Laplace equation. The radial function R is found by solving the radial part of the Laplace equation:

$$\frac{1}{R}\frac{\partial}{\partial r}\left(r^2\frac{\partial R}{\partial r}\right) = \lambda\,. \tag{A.3}$$

The spherical harmonic function Y is found by solving the spherical part of the Laplace equation:

$$\frac{1}{Y}\frac{1}{\sin\theta}\frac{\partial}{\partial\theta}\left(\sin\theta\frac{\partial Y}{\partial\theta}\right) + \frac{1}{Y}\frac{1}{\sin^2\theta}\frac{\partial^2 Y}{\partial\phi^2} = \lambda Y\,. \tag{A.4}$$

This equation can be separated into two equations, one for the θ-derivative and one for the ϕ-derivative:

$$\frac{1}{\Phi}\frac{\partial^2\Phi}{\partial\phi^2} = -m^2\,, \tag{A.5}$$

$$\lambda\sin^2\theta + \frac{\sin\theta}{\Theta}\frac{\partial\Theta}{\partial\theta}\left(\sin\theta\frac{\partial\Theta}{\partial\theta}\right) = m^2\,. \tag{A.6}$$

The solution to Equation A.5 for integer m is a linear combination of exponentials $\exp\left(im\phi\right)$. The solution to Equation A.6 is given by the Associated Legendre Polynomials $P_l^m(\cos\theta)$.

The general solution to the spherical harmonic equation is given by:

$$Y_l^m(\theta,\phi) = N\exp\left(im\phi\right)P_l^m(\cos\theta)\,, \tag{A.7}$$

where N is a normalization constant.

The general solution to the radial equation is given by:

$$R_l(r) = Ar^l + Br^{-l-1}\,, \tag{A.8}$$

where A and B are constants.

The general solution to the Laplace equation on a sphere is given by:

$$f(r,\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l(r) Y_l^m(\theta,\phi) \tag{A.9}$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} (Ar^l + Br^{-l-1}) Y_l^m(\theta,\phi) \tag{A.10}$$

$$= \sum_{l=0}^{\infty} \sum_{m=-l}^{l} R_l^m(r) + I_l^m(r), \tag{A.11}$$

where $R_l^m(r) = Ar^l Y_l^m(\theta,\phi)$ and $I_l^m(r) = Br^{-l-1} Y_l^m(\theta,\phi)$ are functions from $\mathbb{R}^3 \to \mathbb{C}$. Here, R_l^m is the regular solution and I_l^m is the irregular solution.

If the solution is required to be regular in $\mathbb{R}^3$, then $B = \mathbf{0}$ and the solution is given by:

$$f(r,\theta,\phi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m r^l Y_l^m(\theta,\phi), \tag{A.12}$$

where f_l^m are complex coefficients.

A.1. Legendre Polynomials

Legendre polynomials are a set of orthogonal polynomials. They are useful for representing functions as a linear combination of polynomials, in a similar way to how trigonometric functions are used in the Fourier series. The Legendre polynomials are defined as:

$$P_n(x) = \frac{1}{2^n n!} \frac{\partial^n}{\partial x^n} \left(x^2 - 1\right)^n, \tag{A.13}$$

where $n \in \mathbb{N}$ is the degree and $x \in [-1,1]$. The Legendre polynomials are orthonormal on the interval $[-1,1]$. This means that the inner product of two Legendre polynomials is zero, except when the polynomials are the same:

$$\int_{-1}^{1} P_n(x) P_m(x) dx = \begin{cases} 1 & n = m \\ 0 & n \neq m \end{cases}. \tag{A.14}$$

All the Legendre polynomials satisfy the following differential equation:

$$(1-x^2) \frac{\partial^2 P_n(x)}{\partial x^2} + 2x \frac{\partial P_n(x)}{\partial x} - n(n+1) P_n(x) = 0. \tag{A.15}$$

A.1.1. Associated Legendre Polynomials

There also exists a generalization of the Legendre polynomials, the so-called associated Legendre polynomials. Like the Legendre polynomials, the associated Legendre polynomials are orthogonal on the interval $[-1,1]$. Associated Legendre polynomials have a degree l and an order m where $l \geq m \geq 0$. The extra parameter m allows for the polynomials to represent more complex functions. The associated Legendre polynomials are given by:

$$P_l^m(x) = (-1)^m (1-x^2)^{m/2} \frac{\partial^m}{\partial x^m} \left(P_l(x)\right), \tag{A.16}$$

where $P_l(x)$ is the Legendre polynomial of degree l. The order m is the number of times the function is differentiated with respect to x. For negative m, the associated Legendre polynomials are given by the natural normalized (Equation A.18) polynomials:

$$P_l^{-m}(x) = (-1)^m \frac{(l-m)!}{(l+m)!} P_l^m(x). \tag{A.17}$$

The Legendre polynomials are thus a special case of the associated Legendre polynomials where $m = 0$.

A.1.2. Normalization

The associated Legendre polynomials are not normalized. Normalization is crucial for the associated Legendre polynomials to be used as a basis for representing functions. Depending on the application, different normalization constants are used. The following normalization constants are common for the associated Legendre polynomials:

$$\text{Natural} \qquad (-1)^m \frac{(l-m)!}{(l+m)!} \tag{A.18}$$

$$\text{Acoustic} \qquad \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \tag{A.19}$$

$$\text{Quantum} \qquad (-1)^m \sqrt{\frac{(2l+1)(l-m)!}{4\pi(l+m)!}} \tag{A.20}$$

$$\text{Geodesy} \qquad \sqrt{(2l+1)\frac{(l-m)!}{(l+m)!}} \tag{A.21}$$

$$\text{Schmidt} \qquad \sqrt{\frac{(l-m)!}{(l+m)!}}. \tag{A.22}$$

A.1.3. Orthogonality of First Derivatives of Harmonic Basis Functions on the Unit Sphere

Functions that satisfy the Laplace equation on a sphere can be used in many physical applications. The fact that they form an orthonormal basis is also very useful and allows efficient extraction of coefficients (Appendix B.4). One application of spherical harmonics is modelling the Earth's magnetic core field (Section 6.5). The magnetic field is a potential field, meaning that it is a function of the gradient of a scalar potential. For magnetic fields, this scalar potential is harmonic and satisfies the Laplace equation. These properties are further discussed in Chapter 2. For now, it is sufficient to note that knowing (the properties) of the gradient of spherical harmonic functions is useful from a computational perspective. The gradient of a spherical harmonic function is given by:

$$\nabla f = \left(\frac{\partial f}{\partial r}, \frac{\partial f}{\partial \theta}, \frac{\partial f}{\partial \phi} \right). \tag{A.23}$$

The solution to the Laplace equation on the unit sphere can be derived from Equation A.12 by fixing $r = 1$:

$$f(\theta,\phi) = f(r,\theta,\phi)\big|_{r=1} = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} f_l^m Y_l^m(\theta,\phi)',. \tag{A.24}$$

The coefficients f_l^m can be extracted from the solution by using the orthonormality of the spherical harmonics (Appendix B.4):

$$f_l^m = \langle f, Y_l^m \rangle = \int_{S^2} f(\theta,\phi) Y_l^m(\theta,\phi) d\Omega. \tag{A.25}$$

B

Gaussian Processes

Gaussian processes are stochastic processes, popular for supervised learning tasks like regression and classification. They are based on the idea that a set of data points can be approximated by a set of Gaussian distributions: every linear combination of the random variables has a multivariate normal distribution. Gaussian Processes (GPs) are powerful because they can capture the underlying structure of the data, even when the data is noisy or non-linear. GPs model non-linearities by using a kernel function to capture the underlying relationships between data points. A kernel function is a mathematical function that takes two inputs and returns a measure of similarity between them.

B.1. Kernel Functions

Kernel functions are the underlying functions that define the covariance matrix K in a Gaussian process. In [4] it is shown that if a GP has zero mean, then the behavior of the GP is completely determined by its covariance matrix. The kernel function thus plays a crucial role in defining the behavior of the GP. Kernel functions are measures of similarity between two inputs x and x'. A kernel function maps two inputs to a scalar value. If the inputs are close, the kernel function returns a high value. A low value indicates that the inputs are not similar. The kernel function is defined as

$$k(x, x') = \text{"similarity between } x \text{ and } x'\text{"}, \tag{B.1}$$

and can include any function that is positive semi-definite. Kernels are also symmetric: $k(x, x') = k(x', x)$, this is intuitive because the similarity between two inputs should be the same regardless of the order of the inputs.

Some common kernel functions are listed below.

$$\text{Linear Kernel: } k(x, x') = x^T x' \tag{B.2}$$

$$\text{Polynomial Kernel: } k(x, x') = (x^T x' + c)^d \tag{B.3}$$

$$\text{Gaussian Kernel: } k(x, x') = \exp\left(-\frac{||x - x'||^2}{2l^2}\right) \tag{B.4}$$

$$\text{Laplacian Kernel: } k(x, x') = \exp\left(-\frac{||x - x'||}{\sigma}\right) \tag{B.5}$$

$$\text{Periodic Kernel: } k(x, x') = \exp\left(-\frac{2\sin^2(\pi||x - x'||)}{l^2}\right). \tag{B.6}$$

The linear kernel is the simplest kernel function. It is defined as the dot product of the two inputs. Orthogonal input vectors have a similarity of zero. The polynomial kernel is a generalization of the linear kernel and is defined as the dot product of the two inputs raised to a power. The Gaussian kernel is the most commonly used kernel function in GPs. They are also known as the Radial Basis Function (RBF) kernel. Gaussian kernels are smooth and infinitely differentiable, allowing for a good fit of continuous functions.

A kernel is called isotropic if it is translation invariant: $k(x,x') = k(x-x', x-x')$. This means that the kernel function is the same regardless of the location of the inputs and occurs when the kernel function is a function of the distance between the inputs: $k(x,x') = k(\|x-x'\|)$. The benefit of an isotropic kernel is that they often have a single parameter that defines the kernel function. This parameter, the length scale l, controls the smoothness of the kernel function. If the length scale is large, then points that are far apart are still considered somewhat similar, resulting in a smooth kernel function, while a small length scale results in a more jagged kernel function. The performance of isotropic GPs heavily depends on the choice of the length scale, this is further discussed in Section B.1.2.

B.1.1. Covariance Matrix

A kernel function expresses the statistical relationship between two random variables. This is similar to the covariance between two random variables:

$$\text{Cov}(X,Y) = E[(X-E[X])(Y-E[Y])]\,. \tag{B.7}$$

Given a set of random variables $\mathbf{X} = \{X_1, X_2, \ldots, X_n\}$, the covariance matrix is a square matrix that contains the covariance between every pair of random variables. The covariance matrix generated by a kernel function $k(x,x')$ is given by the matrix K where $K_{ij} = k(x_i, x_j)$.

B.1.2. Hyperparameters

B.1.3. Mercer's Theorem

Mercer's theorem states that under certain conditions, a kernel function $k(x,x')$ can be expressed as a linear combination of a set of basis eigenfunctions $\phi_i(x)$ and eigenvalues λ_i [41]. For any kernel function $k : \mathscr{X} \times \mathscr{X} \to \mathbb{R}$ that is square integrable over the domain $\mathscr{D}$ and is symmetric and positive semi-definite

$$\int_{\mathscr{X}\times\mathscr{X}} f(x) f(x') k(x,x') dx dx' \geq 0\,, \tag{B.8}$$

for all $f(x) \in L^1(\mathscr{D})$, there exists a transformation $\phi_i : \mathscr{X} \to \mathscr{V}$ and $\lambda_i \geq 0$ such that

$$k(x,x') = \sum_{i=1}^{\infty} \lambda_i \phi_i(x) \phi_i(x')\,, \tag{B.9}$$

for all $x, x' \in \mathscr{D}$.

B.1.4. The Kernel Trick

The Kernel Trick is a method of transforming a non-linear problem into a linear problem. It is a direct consequence of Mercer's theorem: if a kernel function is SPD, then the kernel function can be expressed as an inner product in $\mathscr{V}$, where $\mathscr{V}$ is the Hilbert space of the eigenfunctions $\phi_i(x)$ [41].

A clear example of the kernel trick occurs in Support Vector Machines (SVMs). In SVMs, the objective function is to find the hyperplane that maximizes the margin between the two classes. Suppose that the data is not linearly separable, e.g. the boundary between the two classes is a circle. In this case, the SVM objective function is non-linear. However, the kernel trick can be used to map the data into a higher dimensional space, where the data is linearly separable. In the case of the circle, a suitable mapping would be

$$\phi(x) = (x_1, x_2, x_1^2 + x_2^2)\,, \tag{B.10}$$

which maps the data into a 3D space. In this space, the data is linearly separable, and the SVM objective function can be solved.

B.1.5. Orthonormal Basis Functions as Kernel Functions

It is also possible to construct a kernel function directly from a set of orthonormal basis functions. An orthonormal basis in a Hilbert space $\mathscr{V}$ is defined[9] as follows:

Definition B.1.1 (Orthonormal basis) *A basis $\boldsymbol{e}_k{}_{k=1}^{\infty}$ in a Hilbert space $\mathscr{V}$ is an orthonormal basis for $\mathscr{V}$ if $\boldsymbol{e}_k{}^{\infty}$ is an orthonormal system.*

Theorem 4.7.2 in [9] states that for an orthonormal system $\{\mathbf{e}_k\}_{k=1}^{\infty}$, the following are equivalent:

1. $\{\mathbf{e}_k\}^\infty$ is an orthonormal basis for $\mathcal{V}$.
2. $\mathbf{v} = \sum_{k=1}^\infty \langle \mathbf{v}, \mathbf{e}_k \rangle \mathbf{e}_k$ for all $\mathbf{v} \in \mathcal{V}$.
3. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$ for all $\mathbf{v}, \mathbf{w} \in \mathcal{V}$.
4. $\sum_{k=1}^\infty |\langle \mathbf{v}, \mathbf{e}_k \rangle|^2 = \|\mathbf{v}\|^2$ for all $\mathbf{v} \in \mathcal{V}$.
5. $\overline{span}\{\mathbf{e}_k\}_{k=1}^\infty = \mathcal{V}$.
6. If $\mathbf{v} \in \mathcal{V}$ and $\langle \mathbf{v}, \mathbf{e}_k \rangle = 0$ for all $k \in \mathbb{N}$, then $\mathbf{v} = 0$.

Item (2) shows that any vector $\mathbf{v}$ can be expressed as a linear combination of the basis vectors $\mathbf{e}_k$. The coefficients of the linear combination are the inner products $\langle \mathbf{v}, \mathbf{e}_k \rangle$, which is the projection of $\mathbf{v}$ onto the orthogonal complement of $\mathbf{e}_k$. Item (3) shows that the inner product is symmetric, i.e. $\langle \mathbf{v}, \mathbf{w} \rangle = \langle \mathbf{w}, \mathbf{v} \rangle$, which is what we want for a kernel function. Item (4) shows that the inner product is positive semi-definite, i.e. $\langle \mathbf{v}, \mathbf{v} \rangle \geq 0$ for all $\mathbf{v} \in \mathcal{V}$.

Now let $\phi_i(x)$ be a set of orthonormal basis functions that map $\mathcal{X} \in \mathbb{R}^n$ into a Hilbert space $\mathcal{V}$. Then, an inner product is defined as

$$\langle \phi_i(x), \phi_j(x') \rangle = \phi_i(x)\phi_j(x') = \delta_{ij}\phi_i(x)\phi_j(x), \tag{B.11}$$

where δ_{ij} is the Kronecker delta and $x, x' \in \mathcal{X}$. The kernel function can then be defined as

$$k(x, x') = \sum_{i=1}^{n} \phi_i(x)\phi_i(x'). \tag{B.12}$$

Currying

For the application of kernel functions in high-dimensional Hilbert spaces F, it is useful to define a curried version of the kernel function. Currying a function is simply performing a partial evaluation, it transforms a function that takes multiple arguments into a function that takes one argument at a time [31]. For example, the function $f(x, y) = x + y$ can be curried as $f(x)(y) = x + y$. In the case of our kernel function $k : \mathcal{X} \times \mathcal{X} \to \mathbb{R}$ and corresponding map $\Phi : \mathcal{X} \to V$, we can define a curried version as $k(x, \cdot) : \mathcal{X} \to \mathbb{R}$. Now we can introduce partial evaluations of our functions, which are mappings to higher-dimensional function space:

$$\Phi_k(y) = k(y, \cdot) : \mathcal{X} \to \mathbb{R}^X, \tag{B.13}$$

where $\mathbb{R}^X$ is space of functions that map $\mathcal{X} \to \mathbb{R}$.

The map Φ was a map from a low dimensional space $\mathcal{X}$ e.g. $\mathbb{R}^n$ to a much higher dimensional space $\mathbb{R}^d$ with $d \gg n$. The new expansion $\Phi_k(y)$ is also a map from $\mathcal{X}$, but now the expansion is taken to a function space $\mathbb{R}^X$. The result of $\Phi_k(y)$ is a function.

B.1.6. Reproducing Kernel Hilbert Space (RKHS)

Definition B.1.2 (Reproducing Kernel Hilbert Space (RKHS)) *A Hilbert space F of functions $f : \mathcal{X} \to \mathbb{R}$ is a Reproducing Kernel Hilbert Space if there exists a kernel function $k : \mathcal{X} \to \mathbb{R}$ such that*

1. $\forall x \in \mathcal{X}, \quad k(x, \cdot) \in F,$
2. $\forall f \in F, \quad \langle f, k(x, \cdot) \rangle = f(x).$

A RKHS projects the features into a function space and linear classification can be performed in this function space.

B.1.7. Zonal Kernel

A zonal kernel can be defined as a kernel function that is defined on a set of points $\mathcal{X}$ and is invariant to translations and rotations. In the context of a globe, zonal means that the kernel function is invariant to rotations about the north pole. Zonal kernels are useful for defining similarities between spherical harmonic functions (Appendix A). In equation Equation A.7, an orthonormal basis Y_l^m is defined on the unit sphere $\mathcal{S}^2$. This basis can be used to define a zonal kernel function using Equation B.12. The zonal kernel function is defined as follows:

$$k(x, x') = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} Y_l^m(x) Y_l^m(x'). \tag{B.14}$$

B.2. Bayesian Inference

Bayesian inference is a method of inferring the posterior distribution of a parameter given some data. A probability distribution is defined over the parameter θ and the data x. Then, using Bayes' theorem, the posterior distribution is defined as

$$p(\theta|x) = \frac{p(x|\theta)p(\theta)}{p(x)}, \tag{B.15}$$

where $p(x|\theta)$ is the likelihood function, $p(\theta)$ is the prior distribution, and $p(x)$ is the marginal likelihood. The posterior distribution is then used to make predictions about the parameter θ given some data x. This can be done by weighting the predictions by the posterior distribution.

B.3. Gaussian Process Regression

Gaussian process regression (GPR) is a method of approximating a function $f(x)$ using a Gaussian process. GPR is a non-parametric method, meaning that there are no assumptions about the form of the function $f(x)$. GPR is based on Bayesian inference: in GPR, the parameter θ is the function of interest, i.e. the function $f(x)$, and the data x is the training data.

For a linear function $y = wx + \epsilon$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$, the posterior distribution is defined as

$$p(w|y, X) = \frac{p(y|w, X)p(w)}{p(y|X)}. \tag{B.16}$$

Given unseen data x', the prediction $f'(x')$ can be calculated as

$$p(f'|x', y, X) = \int p(f'|x', w)p(w|y, X)dw. \tag{B.17}$$

. The prior distribution for a Gaussian process is defined as

$$\begin{bmatrix} y \\ f' \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu \\ \mu' \end{bmatrix}, \begin{bmatrix} k(x,x') + \sigma_n^2 & k(x,x') \\ k(x',x) & k(x',x') \end{bmatrix}\right). \tag{B.18}$$

The predictive distribution $p(f'|x', y, X)$ can be calculated by factoring the data and test observations:

$$p(f'|x', y, X) = \mathcal{N}(\bar{f}', \Sigma^2) \tag{B.19}$$

$$\bar{f}' = \mu' + k(x', x)[k(x,x) + \sigma_n^2]^{-1}(y - \mu) \tag{B.20}$$

$$\Sigma^2 = k(x', x') - k(x', x)[k(x,x) + \sigma_n^2]^{-1}k(x, x'), \tag{B.21}$$

where $\bar{f}'$ are the predictions with the corresponding variance on the main diagonal of Σ^2.

B.4. Extracting Coefficients using Orthogonal Basis Functions

Section B.3 describes how to use this kernel function to make predictions using Gaussian processes. The result obtained is a linear combination of the basis functions $\phi_i(x)$ evaluated at the points a set of points. Extracting the weights of the linear combination is useful for computations. For a continuous function f that is a linear combination of the basis functions ϕ_i:

$$f = \sum_{i=1}^{n} \alpha_i \phi_i, \tag{B.22}$$

one can easily find the coefficients α_i by taking the inner product of f with each basis function ϕ_i:

$$\langle f, \phi_i \rangle = \sum_{j=1}^{n} \alpha_j \langle \phi_j, \phi_i \rangle, \tag{B.23}$$

and solving for α_i:

$$\alpha_i = \frac{1}{\langle \phi_i, \phi_i \rangle} \langle f, \phi_i \rangle. \tag{B.24}$$

In the above equations, the inner product is defined as the definite integral of the product of the two functions:

$$\langle f, g \rangle = \int_{\mathcal{X}} f(x)g(x)dx, \tag{B.25}$$

this requires that f and g can be measured at all points in $\mathscr{X}$. Often, in practice, it is not possible to measure a function on all points in e.g. $\mathbb{R}^n$. In this case, the integral can be discretized using a grid of points $\mathscr{X} = \{x_1, x_2, \ldots, x_n\}$ with spacing Δx, and the inner product can be approximated as

$$\langle f, g \rangle \approx \sum_{i=1}^{n} f(x_i) g(x_i) \Delta x. \tag{B.26}$$

B.4.1. Monte Carlo integration

Another way to approximate the integral is using Monte Carlo integration. This method does not require evenly spaced points, but it requires a large number of samples to get a good approximation. The Monte Carlo integration uses a probability density function $p(x)$ to sample points from $\mathscr{X}$. The expectation of an integral $\int_{\mathscr{X}} f(x) dx$ can then be approximated as

$$E_p[f(X)] = \int_{\mathscr{X}} f(x) p(x) dx, \tag{B.27}$$

where E_p is the expectation with respect to the distribution p. The Monte Carlo approximation is then

$$\hat{f}_N = \frac{1}{N} \sum_{i=1}^{N} f(x_i), \tag{B.28}$$

where x_i are i.i.d. samples from $p(x)$ and N is the number of samples. The Monte Carlo approximation is unbiased, but it has a high variance. From the strong law of large numbers, the Monte Carlo estimate $\hat{f}_N$ converges to the true expectation $E_p[f(X)]$ as $N \to \infty$ and the central limit theorem states that the distribution of $\hat{f}_N$ converges to a normal distribution $N(E_p[f(X)], \mathrm{Var}[\hat{f}_N])$ as $N \to \infty$. The variance of the Monte Carlo approximation is given by

$$\mathrm{Var}[\hat{f}_N] = \frac{1}{N^2} \sum_{i=1}^{N} [f(x_i) - \hat{f}_N]^2. \tag{B.29}$$

In the case of using a uniform distribution $p(x) = 1/\mathscr{X}$, the probability function $p(x)$ is constant and equal to $1/\mathscr{X}$. This is similar to the case of using a grid of points with spacing Δx, but has one major advantage: sampling functions or signals with evenly spaced points limits the frequency content of the signal. This is because the Nyquist frequency is given by $f_{\text{Nyquist}} = 1/(2\Delta x)$. Sampling a signal with a frequency higher than f_{Nyquist} leads to alias in the signal. This is not a problem when using Monte Carlo integration, since the probability density function $p(x)$ can be chosen to be non-uniform, allowing for capturing higher frequency components, which are missed when using a grid of points.

Bibliography

[1] P. Alken et al. “International Geomagnetic Reference Field: the thirteenth generation”. In: *Earth, Planets and Space* 73.1 (Dec. 2021), pp. 1–25.

[2] Patrick Alken, Nils Olsen, and Christopher C. Finlay. “Co-estimation of geomagnetic field and in-orbit fluxgate magnetometer calibration parameters”. In: *Earth, Planets and Space* 72.1 (Dec. 2020), pp. 1–32.

[3] Franco Aubert. “Airborne geomagnetic mapping with an unmanned aerial vehicle: development of a sensor calibration scheme”. PhD thesis. 2022.

[4] Christopher M Bishop. “Pattern Recognition and Machine Learning”. In: (2006).

[5] Richard J. Blakely. “Potential theory in gravity and magnetic applications”. In: (1995), p. 441.

[6] Christian Blatter. *Wavelets : a primer*. A.K. Peters, 1998, p. 202.

[7] Arnold L. Bloom. “Principles of Operation of the Rubidium Vapor Magnetometer”. In: *Applied Optics, Vol. 1, Issue 1, pp. 61-68* 1.1 (Jan. 1962), pp. 61–68.

[8] G. D. (Glenn D.) Boreman. “Modulation transfer function in optical and electro-optical systems”. In: (2001).

[9] Birkhäuser Boston, • Basel, and • Berlin. “Ole Christensen Functions, Spaces, and Expansions Mathematical Tools in Physics and Engineering”. In: (2010).

[10] Sotiris Brakatsoulas et al. “On Map-Matching Vehicle Tracking Data.” In: *VLDB 2005 - Proceedings of 31st International Conference on Very Large Data Bases* 2 (2005), pp. 853–864.

[11] J. B. Bronzan. “The Magnetic Scalar Potential”. In: *American Journal of Physics* 39.11 (Nov. 1971), pp. 1357–1359.

[12] CAE. “CAEMAD-XR”. In: (2023).

[13] Aaron Canciani and John Raquet. “Absolute Positioning Using the Earth’s Magnetic Anomaly Field”. In: *Navigation* 63.2 (June 2016), pp. 111–126.

[14] Aaron Canciani and John Raquet. “Airborne Magnetic Anomaly Navigation”. In: *IEEE Transactions on Aerospace and Electronic Systems* 53.1 (2017), pp. 67–80.

[15] Aaron J. Canciani. “Magnetic Navigation on an F-16 Aircraft Using Online Calibration”. In: *IEEE Transactions on Aerospace and Electronic Systems* 58.1 (Feb. 2022), pp. 420–434.

[16] Aaron Joseph Canciani and Christopher J. Brennan. “An Analysis of the Benefits and Difficulties of Aerial Magnetic Vector Navigation”. In: *IEEE Transactions on Aerospace and Electronic Systems* 56.6 (Dec. 2020), pp. 4161–4176.

[17] L. R. Cander and S. J. Mihajlovic. “Forecasting ionospheric structure during the great geomagnetic storms”. In: *Journal of Geophysical Research: Space Physics* 103.A1 (Jan. 1998), pp. 391–398.

[18] A. Chulliat et al. “The US/UK World Magnetic Model for 2020-2025 : Technical Report”. In: (2020).

[19] ILLUSTRATION BY MARY CROOKS. *Earth’s Interior*. URL: `https://education.nationalgeographic.org/resource/core` (visited on 02/08/2023).

[20] S P Engelson and D V McDermott. “Error correction in mobile robot map learning”. In: *Proceedings 1992 IEEE International Conference on Robotics and Automation* 3 (1992), pp. 2555–2560.

[21] Yves Gallet et al. “On the use of archeology in geomagnetism, and vice-versa: Recent developments in archeomagnetism”. In: *Comptes Rendus Physique* 10.7 (Sept. 2009), pp. 630–648.

[22] Albert R. Gnadt et al. “Signal Enhancement for Magnetic Navigation Challenge Problem”. In: (July 2020).

[23] Albert Reuben Gnadt. “Advanced Aeromagnetic Compensation Models for Airborne Magnetic Anomaly Navigation”. In: (2022).

[24] Mitchell Hezel. “Improving Aeromagnetic Calibration Using Artificial Neural Networks”. In: *Theses and Dissertations* (Mar. 2020).

[25] Hrvoic. I. “Requirements for Obtaining High Accuracy With Proton Magnetometers”. In: (2010).

[26] G. (Gauthier) Hulot. *Terrestrial magnetism*. Springer, 2011, p. 410.

[27] L. Hurwitz and J. H. Nelson. “Proton vector magnetometer”. In: *Journal of Geophysical Research* 65.6 (June 1960), pp. 1759–1765.

[28] John David Jackson and John Wiley. “Classical Electrodynamics Third Edition”. In: (1998).

[29] David Jiles. “Introduction to Magnetism and Magnetic Materials”. In: (Sept. 2015).

[30] Wilfred Kaplan. “Advanced calculus”. In: (2003), p. 741.

[31] KIT. “Optimization Methods for Machine Learning and Engineering (Webmaster)”. In: (2022).

[32] Manon Kok, Jeroen D. Hol, and Thomas B. Schon. “Indoor positioning using ultrawideband and inertial measurements”. In: *IEEE Transactions on Vehicular Technology* 64.4 (Apr. 2015), pp. 1293–1303.

[33] R. A. Langel and R. H. Estes. “A geomagnetic field spectrum”. In: *Geophysical Research Letters* 9.4 (Apr. 1982), pp. 250–253.

[34] R. A. Langel and R. H. Estes. “The near-Earth magnetic field at 1980 determined from Magsat data”. In: *Journal of Geophysical Research: Solid Earth* 90.B3 (Feb. 1985), pp. 2495–2509.

[35] R. A. Langel and W. J. Hinze. “The Magnetic Field of the Earth’s Lithosphere: The Satellite Perspective”. In: *Magnetic Field of the Earths Lithosphere The* (1998).

[36] Paul Leliak. “Identification and Evaluation of Magnetic-Field Sources of Magnetic Airborne Detector Equipped Aircraft”. In: *IRE Transactions on Aerospace and Navigational Electronics* ANE-8.3 (1961), pp. 95–105.

[37] Duan Li et al. “Transforming Total-Field Magnetic Anomalies Into Three Components Using Dual-Layer Equivalent Sources”. In: *Geophysical Research Letters* 47.3 (Feb. 2020).

[38] Paul. Lorrain, Dale R. Corson, and Paul. Lorrain. “Electromagnetism : principles and applications”. In: (1990), p. 472.

[39] Ming Ma et al. “Uncertainty Estimation in the Neural Model for Aeromagnetic Compensation”. In: *IEEE Geoscience and Remote Sensing Letters* 15.12 (Dec. 2018), pp. 1942–1946.

[40] James C. Macnae. “Kimberlites and exploration geophysics”. In: *Geophysics* 44.8 (1979), pp. 1395–1416.

[41] Andreas Maier. *Can my Kernel really be Implemented as a Transform?* URL: `https://medium.com/codex/can-my-kernel-really-be-implemented-as-a-transform-c38a3a85e9e9` (visited on 05/21/2023).

[42] Stefan Maus et al. “High definition geomagnetic models: A new perspective for improved wellbore positioning”. In: *SPE/IADC Drilling Conference, Proceedings* 2 (2012), pp. 997–1011.

[43] Dmitry Murzin et al. “Ultrasensitive magnetic field sensors for biomedical applications”. In: *Sensors (Switzerland)* 20.6 (Mar. 2020).

[44] M. N. Nabighian et al. “The historical development of the magnetic method in exploration”. In: *Geophysics* 70.6 (Nov. 2005).

[45] Norman F. Ness. “Space exploration of planetary magnetism”. In: *Space Science Reviews* 152.1-4 (May 2010), pp. 5–22.

[46] NLOG. *Zwaartekracht en magnetisch veld*. URL: `https://www.nlog.nl/index.php/zwaartekracht-en-magnetisch-veld` (visited on 05/10/2023).

[47] NOAA and NCEI. *World Magnetic Model (WMM) | National Centers for Environmental Information (NCEI)*. URL: `https://www.ncei.noaa.gov/products/world-magnetic-model` (visited on 02/10/2023).

[48] Vanderlei C. Oliveira, Valéria C.F. Barbosa, and Leonardo Uieda. “Polynomial equivalent layer”. In: *Geophysics* 78.1 (Jan. 2013), G1–G13.

[49] F. Javier Pavón-Carrasco et al. “Bootstrapping Swarm and observatory data to generate candidates for the DGRF and IGRF-13”. In: *Earth, Planets and Space* 72.1 (Dec. 2020), pp. 1–18.

[50] R. Prigl et al. “A high precision magnetometer based on pulsed NMR”. In: *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment* 374.1 (May 1996), pp. 118–126.

[51] F. Primdahl. “The fluxgate magnetometer”. In: *Journal of Physics E: Scientific Instruments* 12.4 (Apr. 1979), p. 241.

[52] John M Quinn et al. “Naval Oceanographic Office THE JOINT US/UK 1995 EPOCH WORLD MAGNETIC MODEL”. In: (1995).

[53] QuSpin. *QTFM – QuSpin*. URL: https://quspin.com/qtfm/ (visited on 03/22/2023).

[54] C Reeves. *Aeromagnetic Surveys-Principles, Practice & Interpretation. Geosoft E-publication, Netherlands*.

[55] Ch Reigber, H. Lühr, and P. Schwintzer. “CHAMP mission status”. In: *Advances in Space Research* 30.2 (July 2002), pp. 129–134.

[56] Dennis W. Ricker. “Echo Signal Processing”. In: *Echo Signal Processing* (2003).

[57] A T Roux. “The application of geophysics to gold exploration in South Africa”. In: (1967), s.

[58] D. Schmitz. “Spherical harmonic analysis”. In: *Geophysics* (Feb. 1989), pp. 1217–1221.

[59] Sensys. *MagDrone R3 - Reliable magnetometer for drones (UAV)*. URL: https://sensysmagnetometer.com/products/magdrone-r3-magnetometer-for-drone/ (visited on 05/05/2023).

[60] De Rosal Igantius Moses Setiadi. “PSNR vs SSIM: imperceptibility quality assessment for image steganography”. In: *Multimedia Tools and Applications* 80.6 (Mar. 2021), pp. 8423–8444.

[61] H. Shahri. “The exploration of gold by magnetic method in Hired Area, South Khorasan, a case study”. In: (2010).

[62] SharkD. *File:Geocentric vs geodetic latitude.svg - Wikipedia*. URL: https://en.wikipedia.org/wiki/File:Geocentric%7B%5C_%7Dvs%7B%5C_%7Dgeodetic%7B%5C_%7Dlatitude.svg (visited on 02/10/2023).

[63] Arie Sheinker et al. “Magnetic anomaly detection using a three-axis magnetometer”. In: *IEEE Transactions on Magnetics* 45.1 (2009), pp. 160–167.

[64] Stan Stankov RMI SStankov et al. *Local Operational Geomagnetic Index K Calculation (K-LOGIC) from digital ground-based magnetic measurements*. Tech. rep. 2010.

[65] Nelson C. Steenland. “Oil fields and aeromagnetic anomalies”. In: *Geophysics* 30.5 (1965), pp. 706–739.

[66] Erwan Thébault et al. “The magnetic field of the earth’s lithosphere”. In: *Space Science Reviews* 155.1-4 (Aug. 2010), pp. 95–127.

[67] Walter E Tolles and J D Lawson. “Magnetic compensation of MAD equipped aircraft”. In: *Airborne Instruments Lab. Inc., Mineola, NY, Rept* (1950), p. 201.

[68] Zhou Wang et al. “Image quality assessment: from error visibility to structural similarity”. In: *IEEE transactions on image processing : a publication of the IEEE Signal Processing Society* 13.4 (Apr. 2004), pp. 600–612.

[69] Peter M. Williams. “Aeromagnetic compensation using neural networks”. In: *Neural Computing & Applications 1993 1:3* 1.3 (Sept. 1993), pp. 207–214.

[70] Peilin Wu et al. “Aeromagnetic compensation algorithm based on principal component analysis”. In: *Journal of Sensors* 2018 (2018).

[71] Shoujun Xu et al. “Magnetic resonance imaging with an optical atomic magnetometer”. In: *Proceedings of the National Academy of Sciences* 103.34 (Aug. 2006), pp. 12668–12671.

[72] Xin Xu et al. “DeepMAD: Deep Learning for Magnetic Anomaly Detection and Denoising”. In: *IEEE Access* 8 (2020), pp. 121257–121266.