Technische Universität Ilmenau

Fakultät für Informatik und Automatisierung

Fachgebiet Regelungstechnik

Dissertation

Observability Studies for Spacecraft Attitude Determination based on Temperature Data

zur Erlangung des akademischen Grades Doktoringenieur (Dr.-Ing)

von

Tobias Posielek

---

eingereicht am:	18.03.2022
Tag der wiss. Aussprache:	22.09.2022
1. Gutachter:	Prof. Dr.-Ing. Johann Reger
2. Gutachter:	Prof. Dr.-Ing. Jaime A. Moreno Pérez
3. Gutachter:	Prof. Dr.-Ing. habil. Dipl.-Math. Klaus Röbenack

DOI: 10.22032/dbt.53449
URN: urn:nbn:de:gbv:ilm1-2022000320

Zusammenfassung

Die Schätzung und Steuerung der Fluglage ist elementar für jede Raumfahrzeugmission. Die erforderliche Genauigkeit hängt von der jeweiligen Mission und ihren Nutzlasten ab. Ein funktionierendes Lageregelungssystem ist jedoch immer unverzichtbar, um die Zielgenauigkeit und Stabilität der Nutzlasten zu gewährleisten, die für den Erfolg der Mission entscheidend sind. Daher ist es sinnvoll, redundante Methoden zur Schätzung und Regelung der aktuellen Fluglage einzusetzen. Diese Arbeit fokussiert sich primär auf die Lageschätzung. Hierbei wird untersucht ob und wie Temperaturmessungen für die Lagebestimmung genutzt werden können. Diese Untersuchung wird durchgeführt, indem die zugrundeliegenden mathematischen Beschreibungen der Fluglage sowie der Temperaturdynamik betrachtet werden. Auf deren Grundlage wird dann ein Beobachter zur Lageschätzuung entwickelt, der sich hauptsächlich auf die Temperaturdaten von zwei verschiedenen Sensorkonfigurationen stützt.

In der ersten Konfiguration wird nur ein einziger Temperatursensor verwendet, dessen Informationen mit Gyroskopmessungen fusioniert werden, um die Lage zu bestimmen. Dies wird durch eine Transformation in Normalform und eine neuartige Lagebeschreibung erreicht. Auftretende Mehrdeutigkeiten bei der Lagebestimmung sowie alternative Beobachterdesigns werden vorgestellt. Die Analyse zeigt, dass mit dem vorgeschlagenen Beobachter lokale Aussagen zur Lageschätzung getroffen werden können - vorausgesetzt, die verwendeten Modelle und Messungen sind ausreichend genau und es steht genügend Rechenleistung zur Verfügung.

In der zweiten Konfiguration werden sechs Paare von Temperatursensoren betrachtet. Jedes Paar besteht aus zwei Sensoren mit unterschiedlichen physikalischen Eigenschaften und zeigt in Richtung einer anderen Raumfahrzeugachse. Diese Sensorsignale enthalten genügend Informationen, um die Fluglage zu rekonstruieren, ohne dass die Verwendung von Ableitungen höherer Ordnung erforderlich ist. Es wird ein Algorithmus vorgeschlagen, der die Position der Sonne und der Erde schätzt und diese zur Bestimmung der Lage verwendet.

Die Beobachter für beide Konfigurationen verwenden eine Transformation in eine kanonische Form, um ihre Schätzungen zu erhalten. Die resultierenden Beobachter sind daher sowohl in den transformierten als auch in den ursprünglichen Koordinaten formuliert. Während diese Beobachter unter Annahmen die häufig in der Literatur verwendeten werden äquivalent sind, kann es, sobald diese Annahmen fallengelassen werden, zu einer Reihe interessanter Phänomene wie Mehrdeutigkeit der Lösungen und sogar Instabilität kommen. Diese Phänomene werden an unserem vorgestelltem System veranschaulicht und es werden Methoden vorgeschlagen, um sie zu bewältigen.

Die für die zweite Konfiguration entworfenen Beobachter werden auf die von der Raumsondenmission GRACE erhaltenen Daten angewandt. Dabei hat sich gezeigt, dass die vorgeschlagenen Modelle für die Temperaturschätzung mit einem R^2-Wert zwischen $78,8\%$ und $99,9\%$ gut geeignet sind. Die vorgeschlagenen Algorithmen erlauben eine Genauigkeit mit einem mittleren Fehler über eine Umlaufbahn von weniger als fünf Grad und lassen sich nachweislich leicht durch zusätzliche Messungen ergänzen.

Abstract

Attitude estimation and control is fundamental for every spacecraft mission. Accuracy requirements are strongly dependant on mission level goals and the respective payloads and experiments. However, it is always essential for the mission success to have a functioning attitude control system to allow a high pointing accuracy and stability of the payloads. Therefore, it is useful to employ redundant means to estimate and control the current attitude. The estimation of the attitude is the main topic of this work in which the information contained in temperature measurements for attitude estimation is investigated. This investigation is carried out by providing the underlying mathematical descriptions of the attitude as well as temperature dynamics. Different observer designs are considered based on these models to estimate the attitude relying mostly on the temperature data obtained from two different sensor configurations.

In the first configuration, only a single temperature sensor is employed and the information is fused with gyroscope measurements to determine the attitude. This is achieved based on a transformation into normal form and a novel attitude description. Arising ambiguities in the attitude estimation, as well as alternative observer designs are presented. The analysis shows that with the proposed observer, it is possible to estimate the attitude provided that the employed models and measurements are sufficiently accurate and that enough computational power is available.

The second configuration considers six pairs of temperature sensors. Each pair consists of two sensors with different physical properties and every pair points into a different body axis. These sensor signals contain enough information to reconstruct the attitude without requiring the usage of higher-order derivatives. An algorithm is proposed that estimates the position of the Sun and Earth and uses these to estimate the attitude.

The observers for both configurations use a transformation of the system dynamics into canonical form to obtain a formulation of the problem that allows for estimation. The resulting observers are therefore formulated in transformed and original coordinates. While these observers are equivalent under assumptions widely used in literature, the moment these assumptions are dropped, a number of interesting phenomena such as ambiguity of the solutions and even instability can occur. These phenomena are illustrated by the system of interest and methods are proposed to deal with them.

The designed observers for the second configuration are applied to the data obtained from the spacecraft mission GRACE. The results indicate that the proposed models are well suited for the temperature estimation with a R^2 value between 78.8% and 99.9%. The proposed algorithms admit an accuracy with a mean error over an orbit of less than five degrees and are shown to be easily augmented with additional measurements.

Danksagung

Die vorliegende Arbeit entstand während meiner Zeit als wissenschaftlicher Mitarbeiter am Institut für Systemdynamik und Regelungstechnik (SR) des Deutschen Zentrums für Luft- und Raumfahrt (DLR).

Im Folgenden möchte ich mich recht herzlich bei allen Personen bedanken, welche mir die Erstellung dieser Arbeit erst ermöglicht haben.

Zunächst möchte ich meinem Doktorvater Prof. Dr-Ing. Johann Reger danken, der trotz externer Betreuung und vollem Terminkalender stets Zeit gefunden hat mich bei fachlichen Angelegenheiten zu unterstützen.

Des Weiteren möchte ich mich bei meinem Instituts- und Abteilungsleiter Prof. Dr.-Ing. Johann Bals bedanken, der mir erst durch die gewährten Freiheiten überhaupt das Schreiben dieser Arbeit ermöglicht hat.

Ein ganz besonderer Dank geht auch an die Mitarbeiter des Raumfahrtkontrollzentrums GSOC, deren Bereitstellung von Messdaten der Mission GRACE erheblich zur Bereicherung dieser Arbeit beigetragen haben. Besonders zu erwähnen sind hierbei Jacobus Herman, Michael Kirschner, Lukas Hoffmann, Kay Müller und Sebastian Löw für die Bereitstellung und Diskussion der Daten. Dank gebührt auch Tobias Lesch, der hier den Kontakt hergestellt hat.

Außerdem möchte ich mich besonders bei Dr. Kai Wulff bedanken, durch den ich von unserer gemeinsamen Zusammenarbeit viel über wissenschaftliches Arbeiten lernen konnte.

Weiterhin gebührt ein großer Dank auch an die Mitarbeiter und Freunde, die Teile meiner Arbeit Korrektur gelesen haben. Insbesondere Johannes Robens, Christoph Schwarz, Björn Gäßler, Carsten Oldemeyer, Juliane Skibbe, Lâle Briese, Mihai Gînța, Nikolas Tekles, Paul Acquatella, Peter Kötting und Jonas Witschel sind dabei zu erwähnen. Letzterem sei besonderes gedankt, vor allem auch für seine Unterstützung, produktiven Diskussionen und immerwährende Geduld während unseres Studiums der Technischen Kybernetik und Systemtheorie.

Zu guter Letzt möchte ich mich noch bei all meinen Freunden und meiner Familie bedanken, die mich auf dem Weg zum Fertigstellen dieser Arbeit begleitet haben. In diesem Kontext gebührt natürlich besonderer Dank meinen Eltern, die mich zu jedem Zeitpunkt in meinem Leben unterstützt haben und mir stets vorbehaltlos Rückhalt gegeben haben.

Contents

1 Introduction

Figure 1.1: Russia's "Nauka" Multipurpose Laboratory Module is pictured shortly after docking to the Zvezda Service Module's Earth-facing port on the International Space Station. In the foreground is the Soyuz MS-18 crew ship docked to the Rassvet module on 29 July 2021 [1]. Courtesy: NASA.

Spacecraft attitude determination and control has undoubtedly been a topic of major importance since the beginnings of space exploration. Recent events on the International Space Station (ISS) have reinforced this statement. During the docking of the science laboratory Nauka, see Figure 1.1, the module's thrusters initiated an undesired rotation with a spin rate of 0.56 degrees per second [2]. These undesired rapid rotations may cause structural damage or let the antennas point away from the desired target and therefore interrupt potentially vital communication. In the addressed incident, the thrusters stopped after 15 minutes due to a shortage of fuel, requiring a reorientation manoeuvre of the ISS to bring it to its desired attitude without any damage.

Such an incident emphasises the importance of having an accurate attitude estimation at all times to identify failure scenarios and initiate counter-measures. This leads to the drive of finding all possible ways to estimate the attitude in order to introduce additional redundancy and complement already existing methods. Common designs use a combination of star trackers, sun sensors, Earth horizon sensors, magnetometers and accelerometer to

obtain an estimate of the attitude. The only purpose of all these sensors is the attitude determination. Therefore, the natural question arises if sensors that are already commonly available at the spacecraft might contain additional information that can be used for the attitude estimation. Temperature sensors are a promising candidate, as a high number of them is used throughout the whole spacecraft to monitor the temperature evolution of the individual electrical components. Some of these sensors are also located at the outside of the spacecraft and barely influenced by the inner heat flow. In this case, the temperature evolution is mostly governed by the Sun and Earth irradiation and, hence, by the attitude of the spacecraft. This leads to the topic of this work, namely the analysis of the attitude reconstruction using temperature data.

1.1 State of the Art

Attitude reconstruction of spacecraft using temperature data is closely related to two topics. The first is observer design which provides the means for state estimation. The second topic is attitude estimation itself. This can be considered as the application of observer design on the attitude estimation problem. The background and already existing methods of these two topics are discussed in the following.

1.1.1 Observer Design for Nonlinear Systems

Observer design is highly relevant for many applications. Whenever a state required for the control cannot be directly measured due to technical or economic reasons, the design of an observer is inevitable. In many applications, e.g. automotive, aeronautics or aerospace, observers are a commonly used tool to allow for a sophisticated controller design, see e.g. [3–9].

The literature about the theoretical background of observer design itself is not as vast as for controller design. One of the first works regarding observer design was published in the early 1960s by Kalman [10]. It deals with the topic of state observer synthesis for linear time, invariant, stochastic systems establishing the basis for the nowadays commonly used *Kalman filters*. Another pioneering work of Luenberger [11] deals with the observer design for linear, time invariant, deterministic systems and led to the well known *Luenberger observer*. The design of an observer for a nonlinear system is far more difficult than for its linear counterpart. The augmentation of the previously mentioned two observers is straightforward but stability and desired performance cannot be easily guaranteed any more. Thus, the first publications regarding this issue posed various assumptions on properties and structure of the nonlinear system [12–15]. Some of these assumptions are very specific and difficult to fulfil in practice. The authors of [16, 17] were the first to pose conditions for the observer design of a broad class of nonlinear systems. They have proposed conditions to transform the nonlinear system into a linear canonical form using output injection. This allows to design an observer for the transformed system and with the inverse transformation the nonlinear observer of the system in original coordinates. The idea of transforming the system into canonical form was also employed in the first work establishing high gain observers [18].

These kinds of observers use high gains to obtain fast convergence as well as robustness to model uncertainties. Details about high gain differentiators are discussed in Section 2.2.3. A recent overview of the existing canonical forms and their transformation can be found in [19].

A comprehensive overview about the topic of observability and observation can be found in [20] and [21]. In the sense of [21] we can distinguish between *Luenberger-like* and *Kalman-like* observer designs. Roughly speaking, both of them incorporate a model of the original system dynamics and introduce a correction term $K(y - \hat{y})$ determined by a gain matrix K and the error between the measured and estimated output y and $\hat{y}$. The main difference between these two observer concepts lies in the fact that Luenberger observers often use a constant gain matrix as they were first defined for linear time invariant systems [22]. Kalman-like observers however, use a time varying gain matrix $K(t)$ which is obtained by solving a Riccati equation at every point of time t. The latter observer concept assumes that the state dynamics are disturbed by independent random process and measurement noise with zero mean and covariance [23]. If these assumptions are valid, the observer gives for linear systems the optimal state estimate. Kalman-like observers are commonly employed in many applications as they provide optimal results and require only the tuning of two covariance matrices. They exist in various different forms with different advantages such as extended Kalman filters [24, 25], unscented Kalman filters [26, 27], moving horizon estimation [28, 29] and particle filters [30, 31]. Usually the methods of higher accuracy also exhibit higher computational costs. In [32], a detailed overview and derivation of the mentioned observers can be found.

Other observers with a different structure than the proposed Luenberger- and Kalman-like observers are sliding mode observers [33, 34], linear matrix inequality based designs [35, 36] or algebraic observers [37, 38]. Recent works such as [39] deal with the fusion of these observer designs. The idea is to combine a global and local observer where the former ensures convergence for any initial condition while the latter guarantees desired local behaviour. This shall yield optimal results with respect to domain of attraction, convergence speed, model robustness and sensitivity to noise.

In this work we do not focus on the choice of the observer or on the analysis of their specific advantages for the application. Instead, we put the focus on the system properties and the transformation of the system into a normal form which allows to apply the mentioned observers in a more advantageous way compared to the original formulation. This requires background on the topic of attitude dynamics and estimation.

1.1.2 Attitude Estimation

In order to estimate a spacecraft's attitude, there are a number of different sensors providing useful information for accurate estimates. In [40] and [41] a detailed overview about the most important sensors is given. The most common ones are Earth horizon sensors, sun sensors, magnetic field sensors, accelerometer, star trackers and gyroscopes. An overview of them can be found in Table 1.1. From an algorithmic point of view, the main differences lie in the fact that sun, Earth, magnetic field sensors and accelerometers provide vector

Table 1.1: Main sensors and their provided measurements [44]

Sensor	Measurement	Limitations	Attitude estimation $\hat{q}$
Sun sensor	Sun position s	Sensible to planet albedo	Fusion with other vector measurements
Earth sensor	Earth position r	Horizon uncertainty	Fusion with other vector measurements
Magnetometer	Magnetic field B	Earth field uncertainty	Fusion with other vector measurements
Accelerometer	Acceleration a	Sensible to external forces	Fusion with other vector measurements
Star tracker	Attitude q	Stray light, angular rotation, body frame misalignment bias	Filtering
Gyroscope	Angular velocity ω	Bias integration	Integration

measurements, while star sensors provide an explicit attitude estimate and gyroscopes only measurements for the angular velocity which require integration and knowledge of the initial value for the attitude estimation. All these sensors have some individual advantages and limitations. For the sensors providing vector measurements, at least two linear independent measurements are required to estimate the attitude of the spacecraft. Most star sensors have internal algorithms processing the acquired vector measurements to the stars and provide a direct attitude estimate. These are very accurate despite their sensitivity to stray light from Earth, Moon and Sun. However, the main disadvantages are their costs and small bandwidth. On the other side, gyroscopes provide measurements even for high frequencies but suffer under the problem of giving biased measurements. Usually, a combination of all these sensors is used to compensate the individual limitations. For example, the satellite TET-1 and GRACE use a combination of these to allow for valid estimates even in the case of a single point failure as described in [42] and [43].

This task of measurement fusion is described in detail in [41,45]. Basically, it is distinguished in these works between static and filtering attitude estimation. The first considers the vector measurements and formulates them into Wahba's problem. For this, a large number of possible algorithms exists that can solve this problem. Filtering approaches use the same vector measurements but use various variations of Kalman Filters to estimate the attitude. The main feature of these variations of the Kalman filter is that the error used for the correction step is not as in the conventional case the result of a subtraction but of a multiplication.

Generally, the topic of incorporating different measurements for the attitude estimation is called sensor fusing and is tackled using different approaches dependent on the available sensor data. As already mentioned, gyroscope measurements are usually accurate for high frequencies while star tracker measurements give the best results for low frequencies. Therefore, it is natural to combine these two measurements in an optimal sense using a complementary filter. In [46] the attitude estimate is obtained using only these two measurements. A multiplicative extended Kalman filter in four different estimation layouts is proposed to

obtain the estimate. In [47], additional GPS measurements are used for the attitude. In this work the observer has the form of a federated Kalman filter. As star trackers are in general rather on the expensive side of the sensors, it is desirable to replace their measurements by cheaper measurements. In the works [48] and [49], star trackers are substituted by accelerometer and magnetometer measurements. In combination with the gyroscope measurements, these works define different complementary filters to fuse the measurements in an optimal manner. Instead of accelerometer and magnetometer measurements, in [50] sun sensor and horizon sensor measurements are used. The sun sensor measurements are hereby similar to the temperature sensors measurements but with faster dynamics. The attitude estimate is obtained using a direct cosine matrix based steady state Kalman filter as well as a quaternion based extended Kalman filter. While the gyroscopes' measurements are required for the attitude control system they are not strictly needed for the attitude estimation itself. [51] proposes a method to combine only magnetometer and simple sun sensor measurements. The measurements are incorporated into an extended Kalman filter to obtain the attitude estimate. To sum up, all of the described sensor combinations are based on additional sensors that are solely used for attitude estimation.

We want to analyse the option of using instead or additional to these common sensors a single or a set of temperature sensors. In general, temperature measurements for reconstruction purposes is a vivid subject of research as they contain a lot of information. For example, it has been investigated in [52], if the position of the spacecraft can be estimated with an unscented Kalman filter using only temperature data. For the task of attitude determination using temperature data there exists some work that tackles this issue. First, it is to mention that estimating the attitude using twelve temperature sensors as discussed in Chapter 6 is in practice achieved using a set of sensors called coarse earth sun sensors (CESS). They can for example be used for an Earth-oriented safe mode as proposed in [53]. This set of sensors includes twelve temperature sensors which are arranged such that they have an unobstructed view to space as described in [54]. Half of the sensors are coated black while the other half is coated silver. Thus, all sensors have a similar infrared but different visual absorption. Every sensor measures temperatures driven by the Earth and solar irradiation. These measurements are weighted, averaged and extrapolated to steady-state equilibrium temperature. An algorithm calculates the resulting Earth and Sun vectors incorporating albedo irradiation. Unfortunately, the algorithm and its conditions are not described in detail which makes a comparison to the method proposed in this work difficult. However, the described working principle suggests that an explicit design of an observer relying on a temperature model and measurement derivatives has not been employed. As these suggest more accurate estimates we discuss these approaches in this work.

Model based attitude estimation with fewer temperature sensors is discussed in [55]. Here, an algorithm is proposed to determine the heat flux acting on a satellite, i.e. the superposition of the heat fluxes due to Sun and Earth. This is done by solving a system of differential equations to calculate the thermal sensitivity of temperature to the variation of heat flux as well as the surface temperature with no heat flux absorbed. In the follow-up work [56] these estimated heat fluxes are used to determine the attitude represented in Euler angles. This is done by solving a nonlinear equation system with the Levenberg-Marquardt algorithm. The system of equations results directly from the equations governing the solar, albedo and infra-red heat flow. The approach is carried out in a simulative fashion for six temperature

measurements. For even less sensors the problem is investigated in [57], [58] and [59]. Here, an extended Kalman filter and an unscented Kalman filter are proposed to estimate the attitude. Their performance is evaluated via Monte Carlo simulations. These works do not use the maximum function to avoid negative heat flows and use temperature rates as the output signal. The proposed works use nonlinear temperature dynamics for their estimations which can only be solved numerically.

Concluding, existing works regarding model based attitude estimation using temperature data have a straightforward observer design using mostly the standard Kalman filter framework. An in-depth analysis is missing that focuses on the specific dynamics and angles that influence the temperature evolution. Such analyses would give a better understanding of the actual steps that are behind the reconstruction and facilitate the verification and validation process. Additionally, no analysis has been made which proposes the minimal amount of temperature sensors required for the attitude estimation. This topic becomes extremely relevant in the case of sensor failure. Among others, we fill this gap as outlined in the following section.

1.2 Contributions

The main goal of this thesis is to investigate the attitude estimation of spacecraft using temperature data. The following notable contributions are made in order to achieve the aims of this work.

- A compact holistic physical model combining the thermal and attitude dynamics. The modelling has led to the design of a Modelica library as described in [60].
- A transformation between a conventional and a novel attitude description. The novel description consists of three variables where the first variable describes the angle between two desired vectors. This transformation is a useful tool for dynamics influenced by a time varying angle between two vectors. The transformation and some of its properties are published in [61].
- Observer synthesis for the attitude estimation based on a single temperature sensor. This is one of the first works using only a single temperature sensor for the attitude estimation. The notable contributions are in detail:
  - An observability analysis which discusses and analyses non-observable points which complicates the attitude estimation. This contribution is partially published in [62].
  - Observer design in transformed coordinates. This design uses a normal form and an optimisation to obtain the estimate based solely on a single temperature and angular velocity measurements. Design and simulation results are submitted for publication [63].
  - Observer design in natural coordinates. This design is the pendant to the observer in transformed coordinates augmented with an image extension. It is one

of the few application oriented designs that deals explicitly with the issue of non-observable points.

- Observer synthesis for the attitude estimation based on twelve temperature sensors. This is one of the first works proposing a model based attitude estimation approach for such a sensor configuration. The notable contributions are in detail:
  - Observer design in transformed coordinates. The design employs a singularity free transformation from the temperature measurements and their first derivatives. The design can be easily augmented to incorporate additional measurements for the attitude estimation.
  - Observer design in natural coordinates. For the corresponding design in natural coordinates, the best choice of states is proposed to avoid numerical difficulties and instability.
  - Verification and application of the models and the observers on the measurements obtained from the mission GRACE.

1.3 Outline

An overview of the Chapters can be found in Figure 1.2.

Chapters 2, 3 and 4 give the background and state of the art required to allow the mathematical analysis and solution of the proposed topic. In detail, Chapter 2 introduces the concept of observers and gives state of the art solutions for the observability problem of integrator as well as generic nonlinear systems. The latter is achieved based on a transformation into normal form which will be used throughout the complete work as the fundamental observer design method of choice. Chapter 3 discusses the basics specific to attitude estimation and control. Chapter 4 provides the background on spacecraft thermal modelling. This concludes into a mathematically sound nonlinear description of the considered holistic thermal attitude system. This description is then adapted for two different sensor configurations.

Chapter 5 addresses the first sensor configuration which uses only a single temperature sensor to estimate the attitude. In order to achieve this, we start with an observability analysis which leads to a description of the unobservable states. Then, the idea of the observer design using observability canonical form is pursued. To facilitate the resulting optimisation problem, a novel transformation is introduced describing the attitude as a set of three angles where one of them is the angle between two desired vectors. This leads to two different observers with almost identical dynamics with respect to a transformation into canonical form. These observers are applied for a sample attitude estimation scenario. For the second observer, the problem of instability due to singularities arises. This is avoided by designing an observer whose estimates are restricted on a subset of the state space without singularities.

The second sensor configuration is considered in Chapter 6 and uses real mission data to validate the algorithms. Here, twelve temperature sensors are available, two on each side of a cuboid spacecraft with different physical properties. The proposed mathematical system

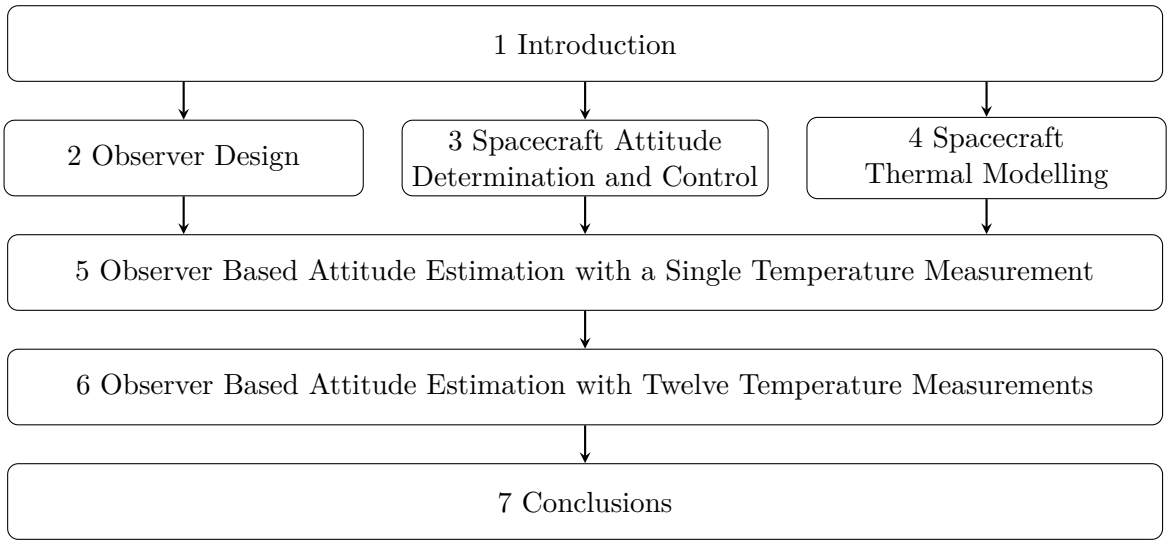


Figure 1.2: Structure of the thesis.

is fitted to the real data and the adequacy of the model is discussed. As in the previous chapter, two observers are designed using the transformation into observability form. The challenge lies thereby in the choice of the appropriate state to estimate and then identify the mapping required to obtain the desired attitude description. In contrast to the previous chapter, the second observer does not suffer from singularities but ambiguities occur if a state with discontinuous dynamics is chosen. This is discussed and solutions to this issue are proposed. Finally, an analysis of the fusion of the considered temperature measurements with other low cost measurements is presented to highlight the benefits and application possibilities of the proposed approach.

2 Observer Design for Nonlinear Systems

In this section, we introduce the main definitions and theorems that shall be used throughout this work to deal with the problem of attitude reconstruction using temperature data. Conventional differentiators and observers are introduced to enable state estimation of nonlinear systems.

2.1 Observability and Detectability of Nonlinear Systems

This section is divided into two parts. The first part introduces the fundamental definitions and theorems required for observer design. In the second part these notions are discussed in detail and their individual implications are highlighted.

2.1.1 Background and Definitions

Consider any nonlinear system

$$\dot{x} = f(x, u) \tag{2.1a}$$
$$y = h(x) \tag{2.1b}$$

with the state $x(t) \in \mathbb{R}^n$, the input $u(t) \in \mathbb{R}^m$ at time $t \in \mathbb{R}$, the usually nonlinear vector field $f : \mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$, the output $y(t) \in \mathbb{R}^p$ and the corresponding output function $h : \mathbb{R}^n \to \mathbb{R}^p$. The functions f and h are assumed to be sufficiently smooth and the input function $u \in \mathcal{U}$ to be bounded where $\mathcal{U}$ denotes the space of possible control functions. We denote the solution of (2.1a) for a given input function u and an initial state $x(0) = x_0$ as $x(\cdot, u, x_0)$.

We start by giving the definition of an observer in the sense of [21]. This definition will be the anchor throughout this section used to motivate the proposed definitions. Roughly speaking, an observer is a dynamical system that estimates or reconstructs the desired internal states using measurement variables.

Definition 2.1. *A **global observer** of* (2.1) *has the form*

$$\dot{X}(t) = F(X(t), u(t), y(t), t) \tag{2.2a}$$
$$\hat{x} = H(X(t), u(t), y(t), t) \tag{2.2b}$$

where for all initial values $x_0, x_0' \in \mathbb{R}^n$ and control functions $u \in \mathcal{U}$ the following holds

i) $\forall x_0 = x_0' \; \forall t \in \mathbb{R}^+ : \hat{x}(t, u, x_0) = x(t, u, x_0')$

ii) $\forall x_0 \neq x_0' : \lim_{t\to\infty} \|\hat{x}(t, u, x_0) - x(t, u, x_0')\| = 0\,.$

Thus, any system of the form (2.2) using only the input and output is an observer if the two proposed conditions i) and ii) are fulfilled. In informal language, any differential algebraic equation with the intention to fulfil these two equations is already called observer to easily transport the intention of the dynamical system considered. The first condition i) is usually simple to fulfil as it requires only to incorporate the dynamics of the system into the observer dynamics. The second condition ii) contains the main idea of the observer, namely that the estimate converges to the real state which is often difficult to prove. Before we go into details of the observer design, we need to introduce some basic concepts used when talking about observers. The first question that usually arises is whether it is even possible to design an observer to estimate every system state. To formalise this question we need to introduce the concept of *indistinguishable states*.

Two states are defined to be indistinguishable, if their associated output is identical for every input function u at every time t as formalised in the following definition.

Definition 2.2. *The pair* (x_0, x_0') *is* ***indistinguishable*** *for the system* (2.1) *if*

$$\forall u \in \mathcal{U}\, \forall t \geq 0 : h(x(t, u, x_0)) = h(x(t, u, x_0')). \tag{2.3}$$

A pair (x_0, x_0') *is* ***indistinguishable by u*** *if*

$$\forall t \geq 0 : h(x(t, u, x_0)) = h(x(t, u, x_0'))\,. \tag{2.4}$$

A pair (x_0, x_0') *that is not indistinguishable is called* ***distinguishable****.*

If no two states are indistinguishable, it is clear that for every pair of initial values there is an input function such that the outputs are not identical. In this case we call the system observable.

Definition 2.3. *The system* (2.1) *is* ***observable at*** $\boldsymbol{x_0}$ *if*

$$\forall x_0' \neq x_0\, \exists u \in \mathcal{U}\, \exists t \geq 0 : h(x(t, u, x_0)) \neq h(x(t, u, x_0')). \tag{2.5}$$

The system (2.1) *is* ***globally observable****, if it is observable at all* $x_0 \in \mathbb{R}^n$*.*

Note that for a global observer (2.1) it is often desirable to have the same input function u for all pairs (x_0, x_0'). Such an input is called *universal*.

Definition 2.4. *An input* $u \in \mathcal{U}$ *is* ***universal*** *for the system* (2.1) *if*

$$\forall x_0 \neq x_0'\, \exists t \geq 0 : h(x(t, u, x_0)) \neq h(x(t, u, x_0'))\,. \tag{2.6}$$

With such a universal input it is possible, at least in theory, to reconstruct the initial state from the output measurements. As the control law is often designed independent of the observer, the Condition (2.6) is often assumed to hold for all $u \in \mathcal{U}$.

Definition 2.5. *The system* (2.1) *is* ***uniformly observable*** *if every input is universal.*

However, often observability cannot be guaranteed on the complete state space. Therefore, a local definition is introduced that allows to rule out indistinguishable states at least in a neighbourhood of x_0.

Definition 2.6. *The system* (2.1) *is* ***weakly observable at*** $\boldsymbol{x_0}$ *if there exists a neighbourhood* $U(x_0)$ *of* x_0 *such that*

$$\forall x_0' \in U(x_0) \setminus \{x_0\} : \ x_0 \text{ and } x_0' \text{ are distinguishable}. \tag{2.7}$$

A system (2.1) *is* ***weakly observable*** *if it is weakly observable for all* $x_0 \in \mathbb{R}^n$.

A stronger condition that prevents trajectories from leaving the neighbourhood $U(x_0)$ before the states can be distinguished is introduced in the following as in [21].

Definition 2.7. *The system* (2.1) *is* ***locally weakly observable at*** $\boldsymbol{x_0}$ *if there exists a neighbourhood* $U(x_0)$ *of* x_0 *such that for all neighbourhoods of* x_0 *with* $V(x_0) \subset U(x_0)$ *holds for all* $x_0' \in V(x_0) \setminus \{x_0\}$ *that*

$$\exists u \in \mathcal{U} \ \exists t \geq 0 \ \forall \tau \leq t : x(\tau, u, x_0), x(\tau, u, x_0') \in V(x_0) \ \wedge \ h(x(t, u, x_0)) \neq h(x(t, u, x_0')) \,. \tag{2.8}$$

A system (2.1) *is* ***locally weakly observable*** *if it is weakly observable for all* $x_0 \in \mathbb{R}^n$.

In order to characterise the set of indistinguishable states, it is useful to define the *observability mapping*, i.e. the mapping that maps every state to its outputs and all its derivatives. In order to do so, we define the vector field for a given constant input $u \in \mathbb{R}$ as $f_u := f(\cdot, u)$. This allows to write the output and its derivatives as a function of the states, i.e. $y^{(i)} = \mathcal{L}_{f_u}^i h(x)$ for $i \in \mathbb{N}$, where $\mathcal{L}$ denotes the Lie derivative as defined in Appendix A.1.

Definition 2.8. *The* ***observability mapping*** $\mathcal{O}(x)$ *of system* (2.1) *for* $u \in \mathbb{R}^m$ *is defined as*

$$\mathcal{O}(x) = \begin{bmatrix} h(x) & \mathcal{L}_{f_u} h(x) & \mathcal{L}_{f_u}^2 h(x) & \dots \end{bmatrix} . \tag{2.9}$$

Note that the dimension of the matrix defined by the observability mapping is not finite. In practice, it is sufficient for most systems to consider the matrix with finite dimension. Hereby, the dimension must be at least the system order n, but may be arbitrarily high if more derivatives are required to reconstruct the states. A special case form systems defined by polynomials [64] where a decidable criterion for observability can always be given. For these

kind of systems, the Lie derivative is augmented on polynomial ideals. Each Lie derivative of order m for $m \in \mathbb{N} \cup \infty$ of a polynomial ideal is defined to be a set. The Lie derivative form an ascending chain with a maximal set which can be obtained for a finite number $M \in \mathbb{N}$. This can be used to formulate an algebraic condition to determine if the polynomial system is observable.

The observability mapping allows to characterise the points at which the system is locally weakly observable using the Jacobian matrix of the observability mapping

Theorem 2.9. *If the system* (2.1) *satisfies the observability rank condition*

$$\operatorname{rk}\left(\frac{\partial \mathcal{O}(x)}{\partial x}\right) = n \tag{2.10}$$

*at $x \in \mathbb{R}^n$ for $u \in \mathbb{R}^m$, the system is **locally weakly observable** at x.*

Proof. Can be found in [65]. □

This condition is the nonlinear pendant to the linear observability matrix rank condition. Note that this condition also guarantees by the implicit function theorem the existence of an inverse function which reconstructs the state x from the output and its derivatives $(y, \dot{y}, \dots)$.

All these definitions were introduced in view of the observer design and the possibility to reconstruct the states based on the input and output trajectories. However, observability is not necessary for the existence of an observer in the sense of Definition 2.1. Indeed, for indistinguishable states the convergence property ii) may be obtained due to the system dynamics itself. This requires that indistinguishable states converge to the same trajectory. In this case the system is called *detectable*.

Definition 2.10. *The system* (2.1) *is **detectable** if for all controls $u \in \mathcal{U}$ and all pairs (x_0, x_0') with $x_0, x_0' \in \mathbb{R}^n$ which are indistinguishable by u the following holds*

$$\lim_{t \to \infty} x(t, u, x_0) - x(t, u, x_0') = 0. \tag{2.11}$$

An overview of the introduced notions can be found in Figure 2.1. The variables V and U denote neighbourhoods of x_0 in the displayed definitions.

2.1.2 Interpretation and Discussion

It can be seen that *observability* ensures for every pair of initial states the existence of an input which leads to a different output trajectory. As a consequence every initial state can be reconstructed from output and input trajectories. Note that the definition does not ensure the existence of a *universal input*, i.e. that a single combination of input and output trajectory is enough for the reconstruction. This depends on the dynamics f, h, the control space $\mathcal{U}$ and the controllability of the system. Consider for example the control space $\mathcal{U}$

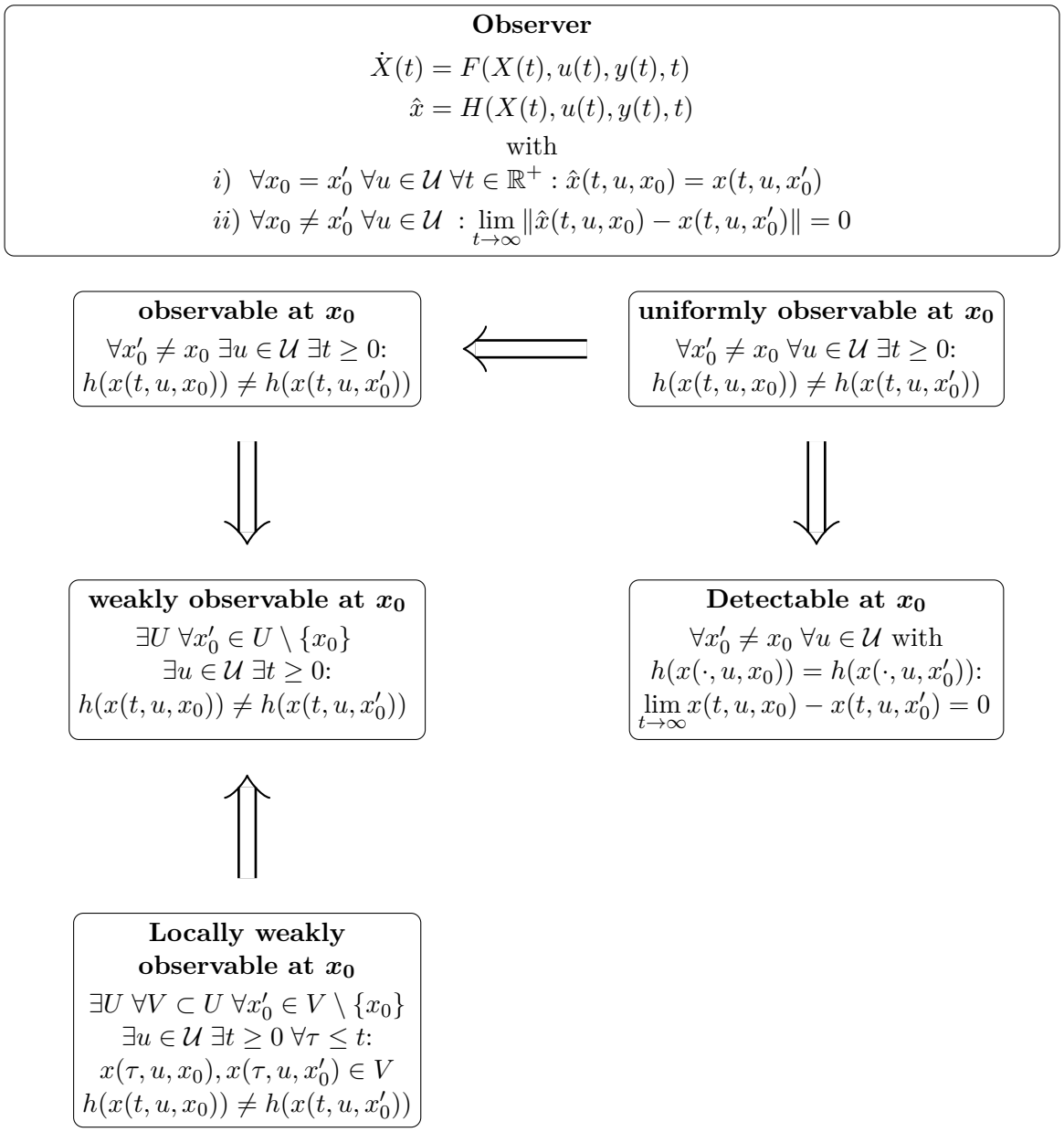


Figure 2.1: Overview of definitions related to observer design.

which is identical to the space of all possible functions and the system (2.1) which is globally observable and controllable. Due to the controllability, there is for all $x_0, x_0' \in \mathbb{R}^n$ a control function $u(\cdot, x_0, x_0') \in \mathcal{U}$ such that $x(t', x_0, u(,\cdot, x_0, x_0')) = x_0'$ for $t' \in \mathbb{R}_{\geq 0}$. Consequently, for any three states $x_0, x_0', x_0'' \in \mathbb{R}$ the common input u with $u = u(t, x_0, x_0')$ for $t \leq t'$ and $u = u(t, x_0', x_0'')$ for $t' < t \leq t''$ can be used to show that these three states are distinguishable based on the same input function. This idea can be augmented to show the existence of a uniform input. If every input is universal the system is *uniformly observable*. Example systems with this property due to their structure can be found in [21].

The proposed *observability* definition 2.3 does not ensure that the initial state reconstruction can be made in finite time. In [66], indistinguishability is introduced on a finite time interval $[0, T]$. If this definition is used for the proposed observability definition, all states can be reconstructed in finite time T.

The proposed properties are considered to be *weak* if the conditions hold only in a neighbourhood of the initial state of interest. Additionally, the state is considered to be *locally weakly observable* if the reconstruction of the initial state can be made before the state trajectory leaves a neighbourhood of the initial state. This definition is in particular of interest because it follows directly from the observability mapping condition (2.10). From the definitions it can be seen that *local weak observability* implies *weak observability*. While this implication appears counter intuitive when compared to the common notions of e.g. attractivity, it shall be emphasised that *local* in this context relates to the property that the trajectories evolve only in a neighbourhood of the considered initial state x_0 while still admitting the distinguishability property. For other properties such as stability or attractivity, the term *local* refers to the fact that only points in a neighbourhood of the considered initial equilibrium point admit the desired property. This property is described in this context by the word *weak*.

By the observer definition it becomes clear that observability is not necessary for the existence of an observer as *detectability* already provides the property required for the estimation. In this case, merely a set of indistinguishable initial states can be reconstructed from the input and output trajectory. However, detectability guarantees that all these initial states converge to the same state trajectory. Note that detectability requires this convergence for *all* inputs that lead for a pair of initial states to the same output trajectory. This makes observability not a sufficient condition for detectability, as it provides no information about inputs that lead to indistinguishable states.

For a system that is detectable but not observable, it is not possible to identify the initial state of the system solely by the output measurements. However, this is also not required for the observer design as only the estimation of the current state is of interest. In this case the convergence dynamics may not be influenced by the observer design but are given by the system dynamics. This property plays a major role for systems in Byrnes-Isidori normal form [67]. In this form, the states are separated into external and internal states. The dynamics of the external states are described by an integrator system which allows a simple observer design as discussed in the next section. If the dynamics of the internal states fulfil the property (2.11) their dynamics can simply be copied to obtain a complete observer as discussed in Section 2.3.

The proposed observability and detectability definitions suggest an estimation of the state x by reconstruction of the initial state x_0 and subsequent forward integration. This has the disadvantage that the time required for the reconstruction is introduced in the estimation as a delay. Thus, the convergence property of an observer stated in Definition 2.1 ii) cannot be ensured. Consequently, it is difficult to obtain an estimator of the form (2.2) solely based on the observability and detectability property. However, the observability mapping and its Jacobian can be used for an explicit observer design as investigated in the upcoming section.

2.2 Observer for Integrator Systems: Differentiator Design

In this section, we investigate the observer design for systems described by integrator chains. For these systems, the state estimation corresponds to an estimation of the derivatives of the measurement signal with respect to time. Therefore, this kind of observer is commonly referred to as a differentiator and plays a major role for the observer design introduced in the next section. The main issue occurring for this kind of estimation is noise impairing the ideal measurements. This requires a careful parameter choice when designing the observer to deal with this imperfection. A lesser attenuation of the noise usually leads to estimates with undesired high dynamics while a larger attenuation leads to larger time delays introduced in the estimate. Methods for signal differentiation can be divided into three categories. The first category includes linear differentiators out of which we discuss the two most common representatives. The first representative is obtained by a concatenation of an ideal differentiator with a low pass filter introduced in Section 2.2.2. The second representative is the high gain differentiators considered in Section 2.2.3. The second category of differentiators includes sliding mode differentiators which we discuss in Section 2.2.4. The third category contains algebraic differentiators [38, 68–72] which calculate a time-derivative estimate by using a finite weighted combination of time integrations of the measurement signal. These are not discussed in detail in this thesis but are worth to be investigated in future work due to their good robustness properties with respect to interfering noises. We start this section by stating the differentiation problem.

2.2.1 The Differentiator Problem

Differentiation deals with the problem of estimating the derivatives $u^{(i)} \in \mathcal{C}^{n-i}(\mathbb{R}, \mathbb{R})$ of an at least n-times differentiable signal $u \in \mathcal{C}^n(\mathbb{R}, \mathbb{R})$ under the presence of measurement noise v. This can be formulated as an observability problem. Suppose the signal u is n-times

differentiable and consider the system

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} x + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ u^{(n)} \end{pmatrix}, \qquad x(0) = \begin{pmatrix} u(0) \\ u^{(1)}(0) \\ \vdots \\ u^{(n-2)}(0) \\ u^{(n-1)}(0) \end{pmatrix} \tag{2.12a}$$

$$y = x_1 + v\,. \tag{2.12b}$$

It can be readily verified that $x_i = u^{(i-1)}$ for $i \in \{1, \dots, n\}$. Thus, the design of an observer (2.1) for system (2.12) is the approximation of the first $n-1$ derivatives of u.

2.2.2 Linear Differentiator

The most intuitive and most commonly used approach uses the Laplace space, in which an ideal differentiation $y = \dot{u}$ is represented by the transfer function $G(s) = s$. As this transfer function is not proper, it is commonly approximated by cascading the differentiator with a low pass filter, i.e. by the transfer function $G(s) = \frac{s}{as+1}$ with sufficiently small $a > 0$ resulting in a high pass filter. A concatenation of these high pass filters to approximate higher derivatives leads to the transfer function

$$G(s) = \frac{s^n}{a_n s^n + \dots + a_1 s + 1}\,,$$

with $a_n \neq 0$ but close to zero and the remaining parameters $a_{n-1}, \dots, a_1$ close or identical to zero. Transfer functions are usually designed such that the corresponding state space realisation has initial states identical to zero. Since it is convenient to prescribe non zero initial values to reduce transient behaviour, we calculate the state space presentation as in Section A.6 with the coefficients

$$N_0 = 0,\ \dots\ ,\ N_{n-1} = 0,\ N_n = \frac{1}{a_n}$$
$$D_0 = \frac{1}{a_n},\ D_1 = \frac{a_1}{a_n}\ \dots\ ,\ D_{n-1} = \frac{a_{n-1}}{a_n}\,.$$

By introducing the first n states as output this leads to the linear differentiator

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -\frac{1}{a_n} & -\frac{a_1}{a_n} & \cdots & \cdots & -\frac{a_{n-1}}{a_n} \end{pmatrix} x + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix} u \tag{2.13a}$$

$$y = \begin{pmatrix} & \frac{1}{a_n} I_n & \\ -\frac{1}{a_n^2} & \cdots & -\frac{a_{n-1}}{a_n^2} \end{pmatrix} x + \begin{pmatrix} 0 \\ \frac{1}{a_n} \end{pmatrix} u\,. \tag{2.13b}$$

We show that this differentiator has the expected behaviour, i.e. the outputs converge to the desired derivatives of the input u as time goes to infinity and the parameters converge to zero.

Theorem 2.11. *Let $a = [1, a_1, \dots a_n]$ be the coefficients of the Hurwitz polynomial given by $p(s) = 1 + a_1 s + \cdots + a_n s^n$. Further, let u be an at least $n+1$ times differentiable function with bounded derivatives $u^{(i)}$ for $i \in \{0, \dots, n+1\}$. Then the states of the linear differentiator* (2.13) *converge to a multiple of the derivatives of the input, i.e.*

$$\lim_{a\to 0} \lim_{t\to\infty} \frac{1}{a_n} x_i(t) - u^{(i-1)}(t) = 0$$

for $i \in \{1, \dots n\}$ where $\lim_{a\to 0}$ denotes all series of coefficients of Hurwitz polynomials $a_1, \dots, a_n$ converging to zero. Furthermore, the outputs converge to the derivatives of the input

$$\lim_{a\to 0} \lim_{t\to\infty} y_i(t) - u^{(i-1)}(t) = 0$$

for $i \in \{1, \dots n+1\}$.

Proof. In order to prove the claim, we calculate the solution of the system explicitly. To do so we need the inverse of the dynamics matrix A which can be readily verified to have the form

$$A^{-1} = \begin{pmatrix} -a_1 & -a_2 & \cdots & \cdots & -a_n \\ 1 & 0 & \dots & 0 & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & 1 & 0 \end{pmatrix}.$$

We write every column of this matrix as the sum of the unit vectors

$$\begin{aligned} A^{-1} e_i &= e_{i+1} - a_i e_1 \\ A^{-1} e_n &= -a_n e_1 \,. \end{aligned}$$

Consequently, the first column of the k-th power of the matrix has the form

$$\begin{aligned} (A^{-1})^k e_1 &= e_{k+1} - \sum_{i=1}^{k} (A^{-1})^{k-i} a_i e_i \\ (A^{-1})^n e_1 &= -a_n e_1 - \sum_{i=1}^{n-1} (A^{-1})^{n-i} a_i e_i \,. \end{aligned}$$

This allows to determine the second summand of the solution of x using n-times partial

integration

$$
\begin{aligned}
\frac{1}{a_n}\int_0^t \mathrm{e}^{A(t-\tau)}Bu\mathrm{d}\tau &= \mathrm{e}^{A(t-\tau)}e_1u|_0^t - \int_0^t \mathrm{e}^{A(t-\tau)}e_1\dot{u}\mathrm{d}\tau \\
&= \mathrm{e}^{A(t-\tau)}e_1u|_0^t + \sum_{i=1}^{n-1}\mathrm{e}^{A(t-\tau)}(A^{-1})^ie_1u^{(i)}|_0^t \\
&\quad - \int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau \\
&= \mathrm{e}^{A(t-\tau)}e_1u|_0^t + \sum_{i=1}^{n-1}\mathrm{e}^{A(t-\tau)}(e_{i+1} + \sum_{j=1}^{i}(A^{-1})^{i-j}a_je_j)u^{(i)}|_0^t \\
&\quad - \int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau \\
&= \mathrm{e}^{A(t-\tau)}e_1u|_0^t + \sum_{i=1}^{n-1}\mathrm{e}^{A(t-\tau)}e_{i+1}u^{(i)}|_0^t \\
&\quad - \sum_{i=1}^{n-1}\mathrm{e}^{A(t-\tau)}\sum_{j=1}^{i}(A^{-1})^{i-j}a_je_ju^{(i)}|_0^t - \int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau \\
&= \sum_{i=0}^{n-1}\mathrm{e}^{A(t-\tau)}e_{i+1}u^{(i)}|_0^t - \sum_{i=1}^{n-1}\mathrm{e}^{A(t-\tau)}\sum_{j=1}^{i}(A^{-1})^{i-j}a_je_ju^{(i)}|_0^t \\
&\quad - \int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau \\
&= \sum_{i=0}^{n-1}(u^{(i)}(t)I - u^{(i)}(0)\mathrm{e}^{At})e_{i+1} \\
&\quad - \sum_{i=1}^{n-1}(u^{(i)}(t)I - u^{(i)}(0)\mathrm{e}^{At})\sum_{j=1}^{i}(A^{-1})^{i-j}a_je_j \\
&\quad - \int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau\,.
\end{aligned}
$$

An upper bound for the norm of the integral containing the n-th derivative of u can be found using the boundedness of $u^{(n)}$

$$
\begin{aligned}
\Big\|\int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau\Big\| &\le \Big\|\int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1\mathrm{d}\tau\Big\|\|u^{(n)}\|_\infty \\
&= \|-\mathrm{e}^{A(t-\tau)}(A^{-1})^ne_1|_0^t\|\|u^{(n)}\|_\infty \\
&= \|(I-\mathrm{e}^{At})(A^{-1})^ne_1\|\|u^{(n)}\|_\infty \\
&= \Big\|(I-\mathrm{e}^{At})(-a_ne_1 - \sum_{i=1}^{n-1}(A^{-1})^{n-i}a_ie_i)\Big\|\|u^{(n)}\|_\infty\,.
\end{aligned}
$$

We use that A is Hurwitz and so the exponential of At converges to zero for rising t, i.e. $\lim_{t\to\infty}\mathrm{e}^{At} = 0$, in order to calculate the limits of the integral containing the n-th derivative

of u

$$\lim_{a\to 0}\lim_{t\to\infty}\left\|\int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau\right\| \leq \lim_{a\to 0}\left\|-a_ne_1-\sum_{i=1}^{n-1}(A^{-1})^{n-i}a_ie_i\right\|\|u^{(n)}\|_\infty$$
$$= 0\,.$$

As the norm of the integral converges to zero, the integral itself converges to zero as well. This allows to calculate the limits

$$\begin{aligned}
&\lim_{a\to 0}\lim_{t\to\infty}\frac{1}{a_n}\int_0^t \mathrm{e}^{A(t-\tau)}Bu\mathrm{d}\tau-\sum_{i=0}^{n-1}u^{(i)}(t)e_{i+1}\\
&=\lim_{a\to 0}\lim_{t\to\infty}-\sum_{i=0}^{n-1}u^{(i)}(0)\mathrm{e}^{At}e_{i+1}-\sum_{i=1}^{n-1}(u^{(i)}(t)I-u^{(i)}(0)\mathrm{e}^{At})\sum_{j=1}^{i}(A^{-1})^{i-j}a_je_j\\
&\quad-\lim_{a\to 0}\lim_{t\to\infty}\int_0^t \mathrm{e}^{A(t-\tau)}(A^{-1})^{n-1}e_1u^{(n)}\mathrm{d}\tau\\
&=-\lim_{a\to 0}\lim_{t\to\infty}\sum_{i=1}^{n-1}u^{(i)}(t)\sum_{j=1}^{i}(A^{-1})^{i-j}a_je_j\\
&=0\,.
\end{aligned}$$

Note that the two limits do generally not commute. The previous equation holds because the input u and its derivatives are bounded for all t. Finally, we can calculate the limits of the solutions of the system (2.13)

$$\begin{aligned}
&\lim_{a\to 0}\lim_{t\to\infty}\frac{1}{a_n}x(t)-\sum_{i=0}^{n-1}u^{(i)}(t)e_{i+1}\\
&=\lim_{a\to 0}\lim_{t\to\infty}\frac{1}{a_n}\mathrm{e}^{At}x(0)+\frac{1}{a_n}\int_0^t \mathrm{e}^{A(t-\tau)}Bu(\tau)\mathrm{d}\tau-\sum_{i=0}^{n-1}u^{(i)}(t)e_{i+1}\\
&=0
\end{aligned}$$

and the first part of the claim is shown. It follows directly that the second claim is true for the first $n-1$ derivatives, i.e.

$$\lim_{a\to 0}\lim_{t\to\infty}y_i(t)-u^{(i-1)}(t)=0$$

for $i\in\{1,\dots n\}$. Furthermore, the function $\ddot{x}_n-u^{(n+1)}$ is bounded because $x_1,\dots,x_n$ and $u^{(0)},\dots,u^{(n+1)}$ are bounded. Thus, we can infer from

$$\lim_{a\to 0}\lim_{t\to\infty}\frac{1}{a_n}x_n(t)-u^{(n-1)}(t)=0$$

with Barbalat's Lemma that

$$\lim_{a\to 0}\lim_{t\to\infty}\frac{1}{a_n}\dot{x}_n(t)-u^{(n)}(t)=0\,.$$

Finally, $\frac{1}{a_n}\dot{x}_n = y_{n+1}$ finishes the proof. □

Remark 1. As in conventional observer design it is possible to consider the error $e = x - \hat{x}$ with x from (2.12) and $\hat{x}$ from (2.13). The resulting error dynamics $\dot{e} = \dot{x} - \frac{1}{a_n}\dot{\hat{x}}$ are a non-trivial function of the error e and the estimated state $\hat{x}$. Thus, it is not possible to evaluate the stability properties of the observer by solely considering these error dynamics. Additionally, the coupled dynamics $(e, \hat{x})$ have poles on the imaginary axis. Thus, any argument for stability needs to incorporate the possible initial values, which is why the proof using the explicit solution is presented instead.

Remark 2. The transfer function of the outputs has the form $\frac{Y_i(s)}{U(s)} = \frac{s^i}{a_n s^n + \cdots + a_1 s + 1}$ with $i \in \{1, \ldots, n\}$.

The theorem shows that (2.13) functions as an observer for the system (2.12). The trade-off between accuracy and robustness against the measurement noise is made by the choice of the parameters a. Generally, smaller coefficients $a_1, \ldots a_n$ guarantee a better approximation at the cost of aggressive behaviour and less robustness. The parameters can be chosen in the same manner as for conventional low-pass filters such as e.g. Bessel filters, Butterworth filters or Chebyshev filters.

2.2.3 High Gain Differentiator

High gain observers as described in [67] provide an exact approximation in zero time if their gains tend to infinity. They are widely used in e.g. [73–77] to name but a few. One main motivation for their usage is the principle of separation of estimation and control, stating that under certain assumptions the design of the observer and of the control law can be broken down into two separate parts and that the resulting combined observer-controller output feedback stabilization preserves the main features of the full state controller. This separation principle was proven for asymptotic stabilization of feedback-linearisable systems with high-gain observers in [78, 79]. However, the high gain of the observer leads to a high sensitivity to measurement noise and the so called peaking phenomenon, i.e. very high estimation errors during the transient phase. While the latter is usually tackled by the usage of a saturated control law, the measurement noise issue limits the maximal choice of the gains possible. High gain differentiators are a special case of high gain observers. Their error bounds in the differentiation of noisy signals are studied in [80] and [81]. A high gain differentiator as proposed in [67] has the form

$$\dot{\hat{x}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \hat{x} + \begin{pmatrix} \frac{a_{n-1}}{\varepsilon} \\ \frac{a_{n-2}}{\varepsilon^2} \\ \vdots \\ \frac{a_1}{\varepsilon^{n-1}} \\ \frac{a_0}{\varepsilon^n} \end{pmatrix} (y - \hat{y}) \tag{2.14a}$$

$$\hat{y} = \hat{x}_1 \,. \tag{2.14b}$$

where $a_0, a_1, \dots a_{n-2}, a_{n-1}$ are the coefficients that define the Hurwitz polynomial of the form $D(s) = a_0 + a_1 s + \dots + a_{n-1}s^{n-1} + s^n$ and $\varepsilon > 0$ is a design parameter close to zero. The scaled error between the estimated and the nominal states is defined by $e_i^{\mathrm{HG}} = x_i - \frac{1}{\varepsilon^i}\hat{x}_i$. This leads to the error dynamics

$$\dot{e}^{\mathrm{HG}} = \frac{1}{\varepsilon}\begin{pmatrix} -a_{n-1} & 1 & 0 & \dots & 0 \\ -a_{n-2} & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ -a_1 & 0 & \cdots & 0 & 1 \\ -a_0 & 0 & \cdots & \cdots & 0 \end{pmatrix} e^{\mathrm{HG}} + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ u^{(n)} \end{pmatrix} + \begin{pmatrix} \frac{a_{n-1}}{\varepsilon} \\ \frac{a_{n-2}}{\varepsilon^2} \\ \vdots \\ \frac{a_1}{\varepsilon^{n-1}} \\ \frac{a_0}{\varepsilon^n} \end{pmatrix} \nu\,.$$

It can be easily seen that these dynamics are asymptotically stable in the absence of measurement noise. A higher gain, i.e. smaller ε leads to a faster convergence rate. However, a high gain leads to the peaking phenomenon, i.e. a very aggressive transient behaviour with high estimation errors. Further, it can be seen that a higher gain also amplifies the influence of the measurement noise.

Remark 3. The transfer function of the final state estimating the derivative of $u^{(n-1)}$ has the simple form $\frac{X_n}{U} = \frac{s^{n-1}}{\frac{1}{a_0}\left(\frac{a_0}{\varepsilon^n} + \frac{a_1}{\varepsilon^{n-1}}s + \dots + \frac{a_{n-1}}{\varepsilon}s^{n-1} + s^n\right)}$. This can be considered as the concatenation of an ideal differentiator and a filter with exactly one order higher than the order of the differentiator. A filter of the same order as the differentiator would be sufficient as discussed in the previous section with linear differentiators. However, the higher order usually admits higher robustness at the price of worse performance.

Remark 4. The high gain differentiator and the linear differentiator are equivalent in the following sense. Consider for a given set of parameters a_i^{HG} and ε the high gain observer (2.14). Then the transformation into canonical form leads to the equivalent linear differentiator with the parameters $a_i = \frac{a_i^{\mathrm{HG}}}{\varepsilon^{n-i}}$ for $i \in \{1, \dots, n\}$. Indeed, the transfer function for $\frac{X_n}{U}$ leads to the same transfer function with these parameters. However, the other transfer functions $\frac{X_i}{U}$ generally differ and make these differentiators different if the estimations of smaller derivatives u^i for $i \in \{1, \dots, n-2\}$ shall also be used.

2.2.4 Sliding Mode Differentiator

Sliding mode differentiators are widely used because of their robustness properties. In sliding mode control a switching function is defined which is made to converge to zero in finite time. This is achieved by the usage of a discontinuous control law. Discontinuous control laws often introduce a loss of existence of the classical Carathéodory solution at the points where the control is discontinuous. These points define a set which is commonly referred to as the sliding manifold. Therefore, the differential equation is replaced by a differential inclusion which solution is defined such that it is tangential to this sliding manifold [82]. By using these solutions in the sense of Filippov, the resulting trajectory and control signal is well defined and continuous [83]. Furthermore, it is robust to a wide class of so called matched

uncertainties. In practical applications however, the discontinuous control law leads to high-frequency control signal called 'chattering' which is undesirable for the actuator. In order to retain the finite convergence while attenuating the chattering, higher order sliding modes were introduced, see e.g. [82, 84, 85]. This has lead to sliding mode differentiators, a class of higher order sliding mode differentiators that can converge in the absence of noise within finite time to the exact derivative provided an upper bound exists for the highest derivative, see e.g. [86–90].

A sliding mode differentiation as in [87, 91] can be written in the form

$$\dot{\hat{x}} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} \hat{x} + \begin{pmatrix} \frac{\lambda_1}{\left(\frac{1}{L}|y-\hat{y}|\right)^{\frac{1}{n}}} \\ \frac{\lambda_2}{\left(\frac{1}{L}|y-\hat{y}|\right)^{\frac{2}{n}}} \\ \vdots \\ \frac{\lambda_{n-1}}{\left(\frac{1}{L}|y-\hat{y}|\right)^{\frac{n-1}{n}}} \\ \frac{\lambda_n}{\left(\frac{1}{L}|y-\hat{y}|\right)^{\frac{n}{n}}} \end{pmatrix} (y-\hat{y})$$

$$\hat{y} = x_1 \,.$$

If we compare this form to a high gain observer as in Equation (2.14) we can see that the parameters $\lambda_1 \dots \lambda_n$ account for $a_{n-1} \dots a_0$ which are designed to make the system asymptotically stable. The influence of the gain ε is represented by $\left(\frac{1}{L}|y-\hat{y}|\right)^{\frac{1}{n}}$. While ε and L are parameters, $|y-\hat{y}|$ is a function of time which will eventually converge to zero if the observer goal is reached. Roughly speaking, a sliding mode observer is a high gain observer with adaptive gains that tend to infinity for diminishing estimation error.

The resulting error dynamics for the error $e = x - \hat{x}$ have the form

$$\dot{e} = \begin{pmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & \ddots & \ddots & \vdots \\ \vdots & \vdots & \ddots & \ddots & 0 \\ 0 & 0 & \cdots & 0 & 1 \\ 0 & 0 & \cdots & \cdots & 0 \end{pmatrix} e + \begin{pmatrix} \lambda_1 L^{\frac{1}{n}} \lfloor e_1 + \nu \rceil^{\frac{n-1}{n}} \\ \lambda_2 L^{\frac{2}{n}} \lfloor e_1 + \nu \rceil^{\frac{n-2}{n}} \\ \vdots \\ \lambda_{n-1} L^{\frac{n-1}{n}} \lfloor e_1 + \nu \rceil^{\frac{1}{n}} \\ \lambda_n L \operatorname{sgn}(e_1 + \nu) \end{pmatrix} + \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ u^{(n)} \end{pmatrix} . \tag{2.16a}$$

$$\hat{y} = x_1 \,. \tag{2.16b}$$

with $\lfloor \omega \rceil^\gamma = |\omega|^\gamma \operatorname{sgn} \omega$. Stability of this system can be shown using the homogeneity properties as introduced in Appendix A.8. The system (2.16) is a homogeneous system with homogeneity degree $d = -1$ and weights $m_i = n + 1 - i$ for $i = 1, \dots n$. The stability properties of this system depend heavily on the chosen parameters λ. Without measurement noise it has been shown that for an appropriate set of parameters λ and L the observer error converges to zero in finite time. The proof presented for this property in [85] uses homogeneity arguments and the construction of multiple sets $D_1, \dots D_n$ where D_{i-1} lies in the interior of D_i and every trajectory starting in D_i reaches D_{i-1} in finite time and does not leave the set for $i \in 1, \dots n$. Another proof presented in [92] uses a homogeneous

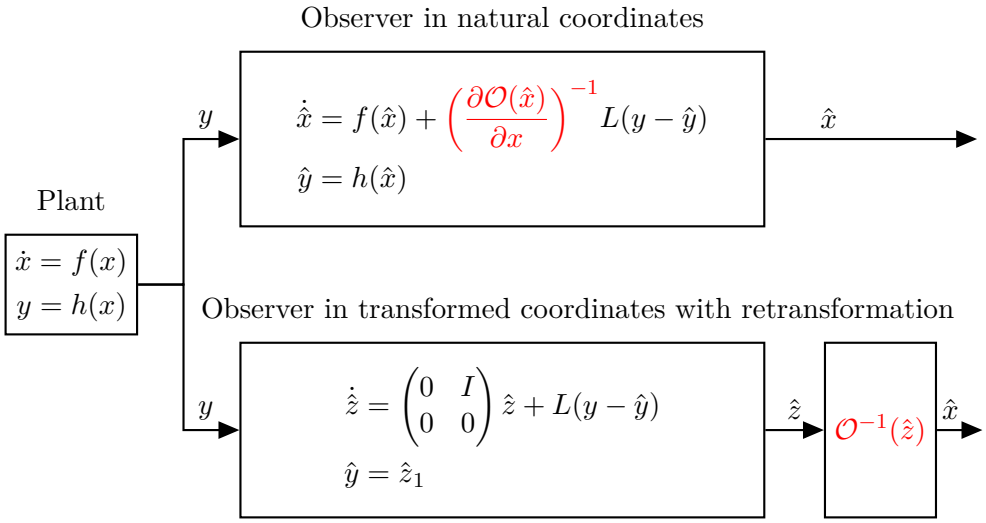


Figure 2.2: Two observer schemes using the transformation $\mathcal{O}$ into canonical form.

Lyapunov function to prove the stability of the differentiator. A set of parameters leading to an asymptotically stable observer is given by $\lambda = [12,\ 10,\ 7,\ 5,\ 3,\ 2,\ 1.5,\ 1.1]$. The remaining parameter L can be used as a tuning parameter to generate the desired balance between performance and robustness.

We continue by introducing a method that can be used to design an observer for any nonlinear system using these introduced differentiators.

2.3 Observer for Nonlinear Systems: Observer Design Using Observability Canonical Form

In this section we introduce the main idea of the designed observers in this work. We consider nonlinear autonomous systems of the form

$$\dot{x} = f(x) \tag{2.17a}$$
$$y = h(x) \tag{2.17b}$$

with the state $x \in \mathbb{R}^n$, the $n-1$ times differentiable vector field $f : \mathbb{R}^n \times \mathcal{U} \to \mathbb{R}^n$, the output $y \in \mathbb{R}^p$ and the corresponding n times differentiable output function $h : \mathbb{R}^n \to \mathbb{R}^p$. Further, let the observability mapping $\mathcal{O} : \mathbb{R}^n \to \mathbb{R}^n$ with $\mathcal{O}(x) = [h(x), \mathcal{L}_f h(x) \dots \mathcal{L}_f^{(n-1)} h(x)]$ be bijective. Then we can use the transformation $z = \mathcal{O}(x)$ to obtain the system in canonical form

$$\dot{z} = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} z + \begin{pmatrix} 0 \\ \mathcal{L}_f^{(n)} h(x)|_{x=\mathcal{O}^{-1}(z)} \end{pmatrix}. \tag{2.18}$$

Because the system (2.18) has the same form as (2.12), estimating z can be achieved using any differentiator discussed in Section 2.2 of the form

$$\dot{\hat{z}} = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \hat{z} + L(y - \hat{y})\,. \tag{2.19}$$

The desired estimate of $\hat{x}$ can then be obtained using the inverse transformation

$$\hat{x} = \mathcal{O}^{-1}(\hat{z}). \tag{2.20}$$

However, instead of finding the inverse of $\mathcal{O}$ it is also possible to design the observer in its equivalent form in the original coordinates. By differentiating Equation (2.20) we obtain with $\frac{\partial \mathcal{O}^{-1}(z)}{\partial z} = \left(\frac{\partial \mathcal{O}(x)}{\partial x}\right)^{-1}$ that

$$\dot{\hat{x}} = f(\hat{x}) + \left(\frac{\partial \mathcal{O}(\hat{x})}{\partial x}\right)^{-1} L(y - \hat{y})\,. \tag{2.21}$$

This design has the benefit that it uses the inverse of the Jacobian of $\mathcal{O}(x)$ instead of the inverse of $\mathcal{O}(x)$. We call observers resulting from (2.19) and (2.20) *observer in transformed coordinates*, as they estimate the transformed states using the dynamic equations and then use the inverse transformation to obtain the original coordinates. We call the observer (2.21) the *observer in natural coordinates* as it uses the dynamics in the original states to obtain the estimate.

We formalise this approach by establishing the requirements necessary which have to be fulfilled such that converging transformed $\hat{z}$-coordinates lead to converging natural $\hat{x}$-coordinates.

Proposition 2.12. *Let the diffeomorphism $\mathcal{O} : x \mapsto \mathcal{O}(x) = z$ be a transformation from x to z coordinates. Further, let $\hat{z}$ be a converging estimate of z coordinates, i.e*

$$\lim_{t\to\infty} \|z(t) - \hat{z}(t)\| = 0 \tag{2.22}$$

and let any of the following assumptions hold

1. *$t \to \mathcal{O}^{-1}(z(t))$ is bounded and $\mathcal{O}^{-1}$ is globally Lipschitz*
2. *There exists T, $\bar{\varepsilon} > 0$ such that the set $\{f^{-1}(z(t) + h) \mid t > T, \|h\| < \varepsilon\}$ is bounded*
3. *$z(t)$ is bounded for all t .*

Then the original coordinates converge as well, i.e. for $\hat{x} = \mathcal{O}^{-1}(\hat{z})$ holds

$$\lim_{t\to\infty} \|x(t) - \hat{x}(t)\| = 0\,.$$

Proof. Due to the mean value theorem [93], the following holds

$$\begin{aligned}\|x(t) - \hat{x}(t)\| &= \|\mathcal{O}^{-1}(z(t)) - \mathcal{O}^{-1}(\hat{z}(t))\| \\ &= \sup_{\tau\in[0,1]} \left\| \frac{\partial \mathcal{O}^{-1}\Big(z(t) + \tau\big(z(t) - \hat{z}(t)\big)\Big)}{\partial z} \right\| \|z(t) - \hat{z}(t)\| .\end{aligned}$$

Thus, for the limit holds with (2.22) and any of the imposed assumptions

$$\lim_{t\to\infty} \|x(t) - \hat{x}(t)\| = \lim_{t\to\infty} \sup_{\tau\in[0,1]} \left\| \frac{\partial \mathcal{O}^{-1}\Big(z(t) + \tau\big(z(t) - \hat{z}(t)\big)\Big)}{\partial z} \right\| \|z(t) - \hat{z}(t)\| = 0$$

because the first factor is bounded and the second factor converges to zero. □

This allows to state the two observers in transformed and natural coordinates to estimate the desired states.

Proposition 2.13. *Consider the system*

$$\begin{aligned}\dot{x} &= f(x) \\ y &= h(x)\end{aligned}$$

with the observability mapping $\mathcal{O} : \mathbb{R}^n \to \mathbb{R}^n$ *with* $\mathcal{O}(x) = [h(x), \mathcal{L}_f h(x) \dots \mathcal{L}_f^{(n-1)} h(x)]$. *Then both the observer in transformed coordinates*

$$\dot{\hat{z}} = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \hat{z} + L(y - \hat{y}) \tag{2.23a}$$

$$\hat{y} = z_1 \tag{2.23b}$$

$$\hat{x} = \mathcal{O}^{-1}(\hat{z}) \tag{2.23c}$$

as well as the observer in natural coordinates

$$\dot{\hat{x}} = f(\hat{x}) + \left(\frac{\partial \mathcal{O}(\hat{x})}{\partial x} \right)^{-1} L(y - \hat{y}) \tag{2.24a}$$

$$\hat{y} = h(x) \tag{2.24b}$$

converge to the desired states, i.e.

$$\lim_{t\to\infty} \|x(t) - \hat{x}(t)\| = 0$$

if any of the conditions from Proposition 2.12 are fulfilled.

Proof. As established in Section 2.2 the observer in transformed coordinates (2.23) fulfils $\lim_{t\to\infty} z(t) - \hat{z}(t) = 0$. The observer in natural coordinates (2.24) has the transformed

dynamics

$$\dot{\hat{z}} = \begin{pmatrix} 0 & I \\ 0 & 0 \end{pmatrix} \hat{z} + \begin{pmatrix} 0 \\ \mathcal{L}_f^{(n)} h(\hat{x})|_{\hat{x}=\mathcal{O}^{-1}(\hat{z})} \end{pmatrix} + L(y - \hat{y}) \tag{2.25a}$$

$$y = z_1 \tag{2.25b}$$

which leads to $\lim_{t\to\infty} z(t) - \hat{z}(t) = 0$. Thus, Proposition 2.12 leads to the claim. □

Figure (2.2) illustrates the two different possibilities to construct the observer using the transformation $\mathcal{O}$. While for the observer in natural coordinates the inverse of the Jacobian of the transformation τ with respect to the states is necessary, the observer in transformed coordinates needs the inverse transformation. These are marked in red as they can be complex to calculate. Note that these two proposed observers are not equivalent despite being based on the same idea. This is due to the fact that the nonlinear term $\mathcal{L}_f^{(n)} h(\hat{x})|_{\hat{x}=\mathcal{O}^{-1}(\hat{z})}$ present in (2.25) is not existent in (2.23) as it requires the inverse mapping to be calculated.

An extension of the systems (2.17) with inputs such that the dynamics have the form $\dot{x} = f(x) + g(x)u$ can easily be achieved. In view of the observer design there are two possible augmentations. The first uses the observability mapping described by the derivatives of the output as before. This has the disadvantage that derivatives of the input up to the order $n - r$ are required where r is the relative degree of the system as defined in Appendix A.1. The second option is that the observability mapping is chosen as the transformation into Byrnes-Isidori form [67]. In this case no derivatives of the input are necessary but the detectability property (2.11) is required to allow a straightforward observer design as discussed beforehand. For a relative degree $n = r$ these two options are identical.

This finishes the general concepts of observability used in this work. In the next chapter we discuss the fundamentals of attitude determination and control. Naturally, this incorporates the observability properties of the attitude dynamics which are defined on a non-euclidean domain.

3 Spacecraft Attitude Determination and Control

This chapter gives an overview of the topic of spacecraft attitude determination. It gives the fundamental mathematical background for orbit propagation and attitude estimation and control that is essential for the topics investigated in this work.

3.1 Reference Frames

In attitude analysis there are a number of special reference frames that are of interest. They usually facilitate or even enable the formulation of the underlying tasks for attitude determination and control. In this section we give the four most important frames that are used throughout this work and are also used in [41]. The first three are illustrated in Figure 3.1.

3.1.1 Earth-Centered Inertial (ECI) Frame

The most common reference frame is the Earth-Centered Inertial (ECI) frame. It is described by its three normalized axes $(x^{\mathrm{I}}, y^{\mathrm{I}}, z^{\mathrm{I}})$ with

x^{I} - points in direction of vernal equinox, this is the intersection between the equator and the Sun's apparent orbit during spring,

y^{I} - completes right handed coordinate system,

z^{I} - parallel to the mean Earth's rotation axis and towards the North Pole.

Even though neither the polar axis nor the ecliptic plane is actually fixed, it is reasonable due to the very slow change compared to Earth and spacecraft rotation, to assume this frame to be constant over the course of time.

For all following reference frames the rotation matrix to ECI coordinates is given by their coordinate axes. Each coordinate system will be denoted with a superscript which will be used for the notation of their coordinate axes and rotation matrices. We denote $A^{\mathrm{S},I} = \begin{bmatrix} x^{\mathrm{S}} & y^{\mathrm{S}} & z^{\mathrm{S}} \end{bmatrix}^{\top}$ as the transformation from ECI coordinates to the coordinates with superscript S, i.e. for r^{I} in ECI coordinates the transformed vector r^{S} is calculated via $r^{S} = A^{\mathrm{S},I} r^{I}$. Additionally, due to the geometry of the rotation matrices we can use the notation $A^{\mathrm{S_1,S_2}} = A^{\mathrm{S_1,I}} A^{\mathrm{I,S_2}}$.

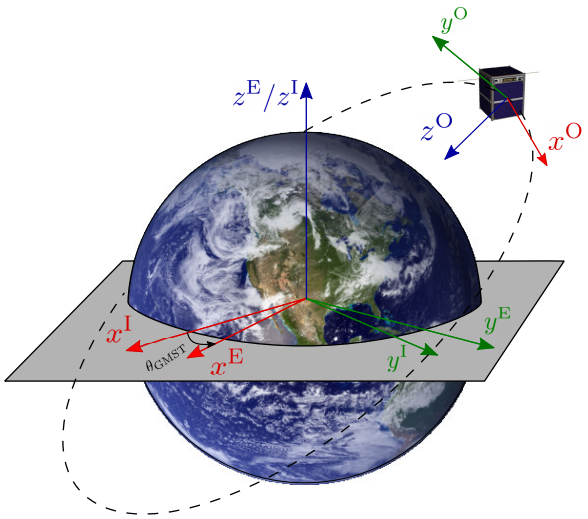


Figure 3.1: Illustration of the ECI, ECEF and orbit frame. Courtesy: NASA.

3.1.2 Earth-Centered/Earth-Fixed (ECEF) Frame

The Earth-Centered/Earth-Fixed (ECEF) is defined by the following normalised axes

x^{E} - points to the prime meridian,

y^{E} - completes right handed coordinate system,

z^{E} - parallel to the mean Earth's rotation axis and towards the North Pole.

The ECEF rotates around the z^{E} axis. Therefore, the transformation from ECI coordinates to ECEF coordinates is described by

$$A^{\mathrm{E,I}}(t) = \begin{pmatrix} \cos\left(\theta_{\mathrm{GMST}}(t)\right) & \sin\left(\theta_{\mathrm{GMST}}(t)\right) & 0 \\ -\sin\left(\theta_{\mathrm{GMST}}(t)\right) & \cos\left(\theta_{\mathrm{GMST}}(t)\right) & 0 \\ 0 & 0 & 1 \end{pmatrix} \tag{3.1}$$

where $\theta_{\mathrm{GMST}}(t)$ is the Greenwich Mean Sidereal Time (GMST) angle. The angle $\theta_{\mathrm{GMST}}(t)$ can be determined by

$$\theta_{\mathrm{GMST}}(t) = 2\pi(0.7790572732640 + 1.00273781191135448(JD + \frac{t}{86400} - 2451545.0))$$

where JD is the initial Julian Date [94].

3.1.3 Orbit Frame

The orbit frame is defined for a spacecraft in an elliptical orbit with position $r^{\mathrm{I}}(t)$ and velocity $v^{\mathrm{I}}(t)$ in inertial coordinates by

x^{O} - completes right handed coordinate system (for circular orbits in direction of velocity),

y^{O} - normal to orbit plane in direction of negative angular momentum,

z^{O} - points to geocentric nadir.

We omit the time argument on the right hand side and obtain the transformation from ECI coordinates to orbit coordinates as

$$A^{\mathrm{O,I}}(t) = \left[\frac{r^{\mathrm{I}} \times (r^{\mathrm{I}} \times v^{\mathrm{I}})}{\|r^{\mathrm{I}} \times (r^{\mathrm{I}} \times v^{\mathrm{I}})\|} \quad -\frac{r^{\mathrm{I}} \times v^{\mathrm{I}}}{\|r^{\mathrm{I}} \times v^{\mathrm{I}}\|} \quad -\frac{r^{\mathrm{I}}}{\|r^{\mathrm{I}}\|} \right]^{\top} . \tag{3.2}$$

Note that the order of the axes may differ throughout the literature used. Additionally, the axis y^{O} may also be defined not in direction of the angular momentum but in the direction of the Sun and may therefore lead to discontinuities throughout the year.

3.1.4 Spacecraft Body Frame

The origin of the spacecraft body frame is usually defined by the mass distribution and corresponding centre of mass. The orientation of the spacecraft body frame with respect to the orbit frame is usually subject to control. Vectors in body frame are indicated by the superscript $\cdot^{B}$.

3.2 Orbital Dynamics

The orbital dynamics are fundamental for attitude estimation as they allow the description of the position and velocity of the spacecraft as an ordinary differential equation. In this section we derive the gravity potential to describe this ordinary differential equation and give the formulation in its simplest forms. We use this to introduce the Kepler elements that can be used to transform the position and velocity states into six elements that evolve in linear fashion with slower dynamics.

3.2.1 Gravitation Field

In this section we discuss the dynamics of an object at the position $r = [r_1 \; r_2 \; r_3]^{\top} \in \mathbb{R}^3$ which is solely under the influence of the Earth's gravity. The acceleration $\ddot{r}$ is described by the gravity potential U such that

$$\ddot{r} = \nabla U \, . \tag{3.3}$$

Details about the derivation and the description of the gravity potential can be found in [95] and [41]. It is dependent on the spherical harmonics coefficients C_n^m and S_n^m of the order $n, m \in \mathbb{N}$ which are obtained by fitting gravity measurements. Using coefficients of high order improves the accuracy of the model but also lead to more complex gravity descriptions. In this work we only use zonal coefficients $J_n = -C_n^0$ which are sufficient to describe the two

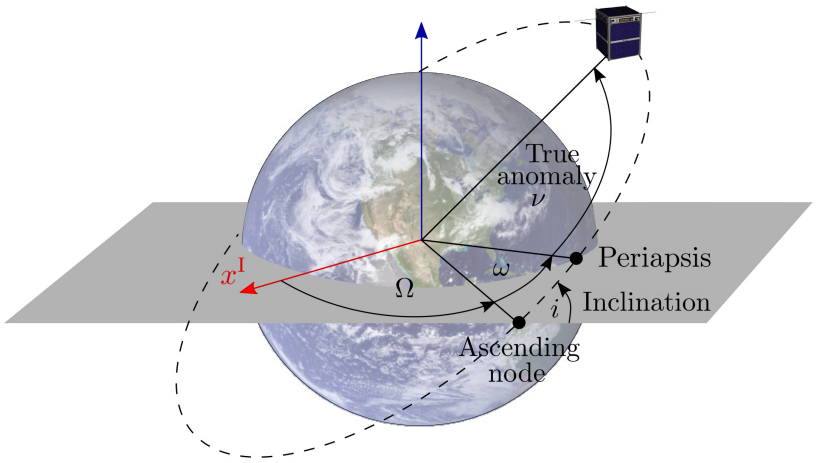


Figure 3.2: Illustration of the classical orbital elements. Courtesy: NASA.

simplest gravity fields. The simplest form of gravity assumes the Earth to be spherically symmetric and the mass to be uniformly distributed. Then, Equation (3.3) has the form

$$\ddot{r} = -\frac{GM_{\oplus}}{\|r\|^3} r \tag{3.4}$$

where G is Newton's gravitational constant and $M_{\oplus}$ the mass of the Earth. Most of the gravity effects due to Earth's shape are described by the second zonal coefficient J_2. Then, Equation (3.3) has the form

$$\ddot{r} = -\frac{GM_{\oplus}}{\|r\|^3} r - \frac{3}{2} J_2 \left(\frac{GM_{\oplus}}{r^2}\right) \left(\frac{R_{\oplus}}{r}\right)^2 \begin{bmatrix} \left(1 - 5\left(\frac{r_3}{\|r\|}\right)^2\right) \frac{r_1}{\|r\|} \\ \left(1 - 5\left(\frac{r_3}{\|r\|}\right)^2\right) \frac{r_2}{\|r\|} \\ \left(3 - 5\left(\frac{r_3}{\|r\|}\right)^2\right) \frac{r_3}{\|r\|} \end{bmatrix} \tag{3.5}$$

where $R_{\oplus}$ denotes the Earth equatorial radius. The gravitation fields of higher order are defined in the same manner and can be found e.g. in [41]. For our purposes, the considered time frame is within the region of few orbits, therefore the usage of the gravity field of first order is sufficient. The usage of these simple gravity fields motivates the introduction of a transformation from position and velocity to another set of six variables for the description of the spacecraft with a simpler dynamics.

3.2.2 Orbital Elements

The elliptical orbit of a spacecraft can be uniquely described with respect to the ECI frame by a set of six orbital elements. A bijective transformation between the position r and velocity v in ECI and the six orbital elements exists. Details about this transformation and its derivation can be found in [95]. Throughout this work we use these six parameters for a more compact and meaningful orbit description. The six parameters can be divided

into three categories. The *ellipse parameters* describe the form of the elliptical orbit. The *orientation parameters* determine the orientation of the orbit while the *time parameter*, defines the satellite's position in the orbit at a point of time. The six orbital elements $(a, \varepsilon, i, \Omega, \omega, M_0)$ are defined as the following:

Ellipse parameters

1. a - Semi-major axis of the elliptical orbit. The distance from the centre of the orbit to its periapsis.
2. ε - Eccentricity of the elliptical orbit. The closer the eccentricity of the orbit is to zero, the more circular is the orbit .

Orientation parameters

3. i - Inclination of the orbit. This parameter describes the angle between the equatorial $(x^{\mathrm{I}}, y^{\mathrm{I}})$-plane and the orbit $(x^{\mathrm{O}}, z^{\mathrm{O}})$-plane .
4. Ω - Longitude of the ascending node of the orbit. Describes the angle between the ascending node and the vernal equinox. The ascending node is obtained as the intersection of equatorial $(x^{\mathrm{I}}, y^{\mathrm{I}})$-plane and orbit $(x^{\mathrm{O}}, z^{\mathrm{O}})$-plane. The sign of this direction is defined by the rotation direction of the orbit. The vernal equinox is identical to the x^{I}-axis.
5. ω - Argument of the periapsis. Describes the angle between the ascending node and the periapsis. The ascending node is calculated as before and the vector pointing towards the periapsis is obtained as $\frac{v\times(r\times v)}{GM_{\oplus}} - \frac{r}{\|r\|}$.

Time parameter

6. M_0 - Mean anomaly at epoch t_0. The angle at epoch from the periapsis to the spacecraft if it would move with a constant velocity in a circular orbit.

When a point mass gravity field is assumed, these six orbital elements are constant over time. In order to obtain the corresponding position and velocity of the spacecraft, and additional variable, the time since epoch t, is required. This can be avoided by using a time variant mean anomaly M instead of M_0 to reduce the required amount of variables. Instead of the mean anomaly, the *true anomaly* can also be used which is the angle between the spacecraft and the periapsis. The elements are visualised in Figure 3.2. The fact that these orbital elements are constant is extremely helpful for visualisation as well as time invariant descriptions of key variables such as shadow time.

However, by introducing additional gravity perturbations as in (3.5) the orbital elements change with time. The derivatives with respect to time of each of these six elements can be found in [96]. The calculated rates can be divided into short period oscillations, long period oscillations and secular drifts. As for long term oscillations the secular drift is of

great interest. The mean orbital elements $\bar{a}, \bar{\varepsilon}, \bar{i}, \bar{\Omega}, \bar{\omega}, \bar{M}_0$ are defined and the mean effect of the second zonal coefficient expressed by their rates as in [41] has the form

$$\frac{\mathrm{d}\bar{a}}{\mathrm{d}t} = 0 \tag{3.6a}$$

$$\frac{\mathrm{d}\bar{\varepsilon}}{\mathrm{d}t} = 0 \tag{3.6b}$$

$$\frac{\mathrm{d}\bar{i}}{\mathrm{d}t} = 0 \tag{3.6c}$$

$$\frac{\mathrm{d}\bar{\Omega}}{\mathrm{d}t} = -\frac{3}{2} J_2 \left(\frac{R_\oplus}{\bar{a}}\right)^2 (1-\bar{\varepsilon}^2)^{-2} \cos\bar{i} \sqrt{\frac{GM_\oplus}{\bar{a}^3}} \tag{3.6d}$$

$$\frac{\mathrm{d}\bar{\omega}}{\mathrm{d}t} = \frac{3}{2} J_2 \left(\frac{R_\oplus}{\bar{a}}\right)^2 (1-\bar{\varepsilon}^2)^{-2} \left(2 - \frac{5}{2}\sin^2\bar{i}\right) \sqrt{\frac{GM_\oplus}{\bar{a}^3}} \tag{3.6e}$$

$$\frac{\mathrm{d}\bar{M}_0}{\mathrm{d}t} = \frac{3}{2} J_2 \left(\frac{R_\oplus}{\bar{a}}\right)^2 (1-\bar{\varepsilon}^2)^{-\frac{3}{2}} \left(1 - \frac{3}{2}\sin^2\bar{i}\right) \sqrt{\frac{GM_\oplus}{\bar{a}^3}}\,. \tag{3.6f}$$

These still comparatively simple dynamics allow the introduction of an orbit that is used for a high number of spacecraft missions.

3.2.3 Sun-Synchronous Orbits

A spacecraft on a Sun-synchronous orbit rotates with the movement of the Sun in order to obtain constant sunlight over the course of an orbit [41]. In this section we discuss the orbital elements that usually describe a Sun-synchronous orbit.

Typical Sun-synchronous orbits are about $600-800\mathrm{km}$ in altitude h_{alt} which allows a direct calculation of semi-major axis a as

$$a = R_\oplus + h_{\mathrm{alt}}\,. \tag{3.7}$$

Most Sun-synchronous orbits are designed to be circular which makes the eccentricity identical to zero, i.e. $\varepsilon = 0$.

The main idea of the design of a Sun-synchronous orbit lies within the choice of the inclination i. It is chosen such that the orbit rotates with the movement of the Sun. This is achieved by assuming that the Sun rotation is uniform. This allows establishing the desired condition for a Sun-synchronous orbit

$$\dot{\Omega}(t) \approx \frac{2\pi}{1\,\mathrm{y}} \tag{3.8}$$

where y denotes the sidereal year. The longitude of the ascending node Ω is constant if a homogeneous gravity field is assumed. Thus, Equation (3.8) can only be fulfilled if a gravity field of higher order than point gravity (3.4) is assumed. Using the gravity field of first order we consider Equation (3.6d) describing the average change Ω. Combining this with Equation

(3.8) enables the calculation of the average inclination necessary for a Sun-synchronous orbit with average semi-major axis a and eccentricity ε as

$$i = \arccos\left(-\frac{2}{3}\frac{2\pi}{1\,\mathrm{y}}\frac{1}{J_2\sqrt{GM_\oplus}}\frac{a^{7/2}}{R_\oplus^2}(1-\varepsilon^2)^2\right).$$

The choice of this inclination is the main factor that defines a Sun-synchronous orbit.

Instead of the longitude of the ascending node Ω the local time of the descending node is given which is approximately constant. The local solar time of the descending node can be calculated by finding the position of the node in ECI coordinates by rotating the vector pointing to vernal equinox around the z axis

$$r^{\mathrm{I}} = \begin{pmatrix} \cos(\Omega+\pi) & -\sin(\Omega+\pi) & 0 \\ \sin(\Omega+\pi) & \cos(\Omega+\pi) & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}. \tag{3.9}$$

Then, this position is transformed into ECEF coordinates using the transformation introduced in (3.1) and

$$r^{\mathrm{E}} = A^{\mathrm{E,I}} r^{\mathrm{I}}. \tag{3.10}$$

In order to calculate the longitude of the node we obtain the angle between this position in ECEF and the angle between the vector pointing towards zero longitude, i.e.

$$\theta = \arccos\left(\frac{r^E x^E}{\|r^E\|\|x^E\|}\right). \tag{3.11}$$

Then, we can calculate the Julian Date of the descending node JD^{LTDN} by transforming the longitude and adding it to the current Julian date JD^{c} with

$$JD^{\mathrm{LTDN}} = JD^{\mathrm{c}} + \frac{\theta}{2\pi}\,\mathrm{d} \tag{3.12}$$

where d denotes the unit of a Julian day. Finally the hour of the ascending node can be obtained as

$$h^{\mathrm{LTDN}} = f^{JD,h}(JD^{\mathrm{LTDN}}). \tag{3.13}$$

where $f^{JD,h}$ denotes the function extracting the current hour from a Julian Date as given in e.g. [41]. Note that this time is constant except of slight variations due to the short periods oscillations as mentioned before.

The argument of periapsis ω does not change the trajectory of the orbiting spacecraft as the orbit is circular. Therefore, it is usually put at the ascending node, i.e. $\omega = 0$.

Similarly, the mean anomaly M_0 at epoch t_0 influences only the position of the spacecraft on the already fully defined orbit. Therefore, if not stated differently this parameter is also chosen to zero, i.e. $M_0 = 0$.

This allows to describe a Sun-synchronous orbit using only two additional parameters, namely the attitude of the orbit and the time of the descending node. This is a compact specification of such orbits without discussing the elements in detail.

This finishes the topic of orbital dynamics of this section in which we have discussed the fundamentals required to describe the three translational degrees of freedom of a spacecraft. In the next section we discuss the basics of the remaining three rotational degrees of freedom.

3.3 Attitude Dynamics and Control

Attitude is commonly referred to as the orientation of a body in the three dimensional space. A lot of different work exists on the topic of attitude modelling, control and estimation e.g. [40, 41, 44, 97–104] to name but a few. We restrict ourselves to the essential parts of these works that are required to deal with the topic of attitude estimation. There are a number of different spaces with various different properties that allow an identification with attitudes. Commonly, the 3D rotation group $\mathcal{SO}_3$ is identified with the set of 3×3 orthogonal matrices with determinate equal to one. These matrices are called rotation matrices and have the disadvantage of describing a three dimensional space with nine parameters. This problem is tackled by the introduction of various other parameters such as quaternion, rotation vector and Euler angle to name but a few. For aerospace, quaternions are common as they use only four variables and have only a single constraint namely the quaternion being normalised. They have no singularities and admit a continuous evolution. This comes at the price of having every attitude represented by exactly two quaternions with opposite sign. An overview of the relations between the individual attitudes as well as identities used in this work can be found in Appendix A.2 and A.3. We will mostly rely on quaternions as the attitude description. Therefore, we introduce only their dynamics and fundamental control laws as well as observers using this description.

The quaternion dynamics have the form as given in [41]

$$\dot{q} = \frac{1}{2}\Omega(\omega)q = \frac{1}{2}\Xi(q)\omega \tag{3.14}$$

where $q \in \mathbb{S}_3$ denotes the quaternion and $\omega \in \mathbb{R}^3$ the angular velocity and the matrices $\Xi(q)$ and $\Omega(\omega)$ have the form

$$\Xi(q) = \begin{pmatrix} q_4 I_3 + [q_{1:3}\times] \\ -q_{1:3}^\top \end{pmatrix}, \quad \Omega(\omega) = \begin{pmatrix} -[\omega\times] & \omega \\ -\omega^\top & 0 \end{pmatrix}$$

where $[\omega\times]$ denotes the cross product matrix

$$[\omega\times] := \begin{pmatrix} 0 & -\omega_3 & \omega_2 \\ \omega_3 & 0 & -\omega_1 \\ -\omega_2 & \omega_1 & 0 \end{pmatrix}.$$

Some useful properties for these matrices used throughout this work can be found in Appendix A.3.

While these dynamics (3.14) are nonlinear if considered as a function of (q, ω), they are linear if considered to be a function of either q or ω. This facilitates the calculation of their linearisation as carried out later in Section 5.2.1.. Furthermore, these dynamics inherently guarantee that the norm of the quaternion does not change, i.e. $\frac{\mathrm{d}}{\mathrm{d}t}\|q(t)\| = 0$ which makes the norm of the solution at every time step have the same norm as the initial quaternion. Additionally, it can be readily verified that the analytical solution of (3.14) has the form

$$q(t) = \mathrm{e}^{\frac{1}{2}\Omega\left(\int_0^t \omega(\tau)\mathrm{d}\tau\right)} q(0) \tag{3.15}$$

where $q(0)$ denotes the initial state of q. This requires the calculation of a matrix exponential as well as the explicit integration of the angular velocity.

The dynamics of the latter is governed by Euler's rotational equation and has the form

$$\dot{\omega} = J^{-1}(-\omega \times J\omega + u) \tag{3.16}$$

where $J \in \mathbb{R}^{3\times3}$ denotes the inertia matrix and u the torque responsible for the motion of the body. In the spacecraft context this is usually provided in a direct or indirect fashion by thrusters, reaction wheels or magnetic torquers. In this work the actuation plays only a subordinate role which is why we assume to have direct control over the variable u.

3.3.1 Attitude Control

The problem of attitude control therefore deals with the problem of driving the system

$$\begin{aligned}\dot{q} &= \frac{1}{2}\Xi(q)\omega \\ \dot{\omega} &= J^{-1}(-\omega \times J\omega + u)\end{aligned}$$

to the desired quaternion q_d or follow the desired quaternion and angular velocity trajectory $q_\mathrm{d}(\cdot), \omega_\mathrm{d}(\cdot)$. We consider the attitude tracking control law given in [41] as

$$u = [(A^\mathrm{e}\omega_\mathrm{d})\times]JA^\mathrm{e}\omega_\mathrm{d} + JA^\mathrm{e}\dot{\omega}_\mathrm{d} - k_\mathrm{p}q^\mathrm{e} - k_\mathrm{d}\omega^\mathrm{e} \tag{3.17}$$

with the error quaternion, error attitude matrix and error angular velocity defined by

$$q^\mathrm{e} = q \otimes q_\mathrm{d}^{-1} \tag{3.18}$$
$$A^\mathrm{e} = A(q)A(q_\mathrm{d})^\top \tag{3.19}$$
$$\omega^\mathrm{e} = \omega - A^\mathrm{e}\omega_\mathrm{d} \tag{3.20}$$

and the gains $k_\mathrm{p}, k_\mathrm{d} \in \mathbb{R}$. This control law is in its essence a PD control law with a nonlinear compensation term. It requires to define a desired quaternion trajectory q_d with its corresponding angular velocity ω_d and its derivative $\dot{\omega}_\mathrm{d}$. In this work we use the attitude control law (3.17) to follow a given trajectory. Accuracy and quality due to the choice of

the used control law are hereby not discussed as the focus of this work lies on the accuracy and quality of the estimation of the attitude that are used in the control law. This problem of attitude estimation is discussed in the next section.

3.3.2 Attitude Estimation and Observers

We differentiate between the two main attitude estimation problems occurring. The first is a filtering problem which refines an already existing estimate of the attitude under the usage of the attitude dynamics and angular velocity to obtain a better estimate. The second poses the question of how the attitude can be estimated using a number of vector measurements. This is divided into static and dynamic estimation methods. Naturally, in order to obtain a good attitude estimation, either a dynamical estimation method is used or the static estimation method is combined with a filtering method to obtain the best results. This section shall not give a detailed analysis of the existing methods but merely state the underlying problems and the methods used throughout this work. For more details the reader is referred to the literature (e.g. [41] or [45]).

3.3.2.1 Attitude Filtering Using a Nonlinear Observer

We consider the system dynamics of the attitude q^{real} with

$$\begin{aligned}\dot{q}^{\text{real}} &= \frac{1}{2}\Xi^{\top}(q^{\text{real}})\omega \\ \hat{q} &= q^{\text{real}} + \nu\end{aligned}$$

where $\hat{q}$ is an estimate obtained through measurements or estimation methods. This estimate can be described as the sum of the real quaternion and a disturbance ν. We want to design an observer

$$\dot{q} = f^{\text{obs}}(q, \hat{q}, \omega)$$

that allows its estimate $q \in \mathbb{S}_3$ to be closer to the real attitude q^{real} than the estimate $\hat{q}$. The observer uses the vector field f^{obs} which combines an already existing estimate $\hat{q}$ with the angular velocity ω and the current estimate q. Note that this is an observability problem in the form of Chapter 2. In the spirit of this section it is straight forward to design an observer as the system is already in normal form and every state has a corresponding output. However, due to the structure of the quaternion space there are more elegant solutions to incorporate the error than in the additive fashion conventionally done. This is derived in the following. We consider the quaternion dynamics (3.14) in DAE form

$$2\Xi^{\top}(q)\dot{q} = \omega \tag{3.21}$$

$$q^{\top}q = 1\,. \tag{3.22}$$

It is natural to design an observer that retains the algebraic constraint (3.22) to allow an interpretation of the quaternion as an attitude. Thus, we vary only the right hand side of

Equation (3.21) and consider observers of the form

$$2\Xi^\top(q)\dot{q} = \omega + f^{\text{cor}}(q, \hat{q}, \omega) \tag{3.23a}$$
$$q^\top q = 1 \tag{3.23b}$$

with the correction term f^{cor}. By multiplying Equation (3.23a) with $\Xi(q)$ from the left hand side we obtain back the ODE formulation of the dynamics

$$\dot{q} = \frac{1}{2}\Xi(q)\omega + \frac{1}{2}\Xi(q)f^{\text{cor}}(q, \hat{q}, \omega)\,.$$

By using the vector part of the error quaternion as defined in (A.23) as the correction term we obtain the desired observer

$$\dot{\hat{q}} = \frac{1}{2}\Xi(\hat{q})\omega + k\frac{1}{2}\Xi(\hat{q})\Xi^\top(\hat{q})q \tag{3.24}$$

with the tuning parameter $k \in \mathbb{R}^+$. Its stability can be shown using the Lyapunov function $V = \frac{1}{2}q^\top\Xi(\hat{q})\Xi^\top(\hat{q})q$. A proof for the closed loop system can be found in [105] and its augmentation on tracking problems in [106].

3.3.2.2 Attitude Estimation Using Vector Measurements

It is common for the attitude estimation to consider $m \in \mathbb{N}$ vectors in reference frame $r_i \in \mathbb{S}_2$ and their corresponding vector measurements in body frame $b_i \in \mathbb{S}_i$ that fulfil the equation

$$Ar_i = b_i \qquad \forall i \in \{1, \ldots, m\} \tag{3.25}$$

where $A \in \mathcal{SO}_3$ denotes the attitude matrix which transforms the reference frame to the body frame. The task is then to obtain an estimate of the attitude A based on these vectors. This can be cast in the observability problem formulation of the previous section by formulating the system as

$$\dot{q} = \frac{1}{2}\Xi(q)\omega \tag{3.26a}$$
$$b_1 = A(q)r_1 \tag{3.26b}$$
$$\vdots \tag{3.26c}$$
$$b_m = A(q)r_m\,. \tag{3.26d}$$

This problem may of course be augmented with additional state or parameter dynamics. Transformation into normal form as proposed in the previous chapter using the outputs is difficult as there are potentially more outputs than states and the resulting dynamics are highly linear. Simpler solutions are widely known in literature and give satisfying results. In [41] the solutions for this problem observation problem are divided into two classes, static and dynamic methods. Static methods only solve the algebraic equation system (3.26b)-(3.26d) in every time step without considering the attitude dynamics. This requires usually

to employ an additional filter to remove emerging noise. Dynamic methods directly incorporate the filter into the observer to estimate the attitude.

Static Estimation: Wahba's Problem

The attitude estimation problem can be formulated in the sense of Wahba [107] as the optimisation problem derived from Equation (3.26b)-(3.26d) as

$$\min_{A \in \mathcal{SO}_3} \frac{1}{2} \sum_{i=1}^{m} a_i \|b_i - A r_i\|^2 \tag{3.27}$$

whose solution is the desired attitude matrix A. Various different algorithms exist to solve this problem, e.g. Davenport's q Method, the Quaternion Estimator QUEST, or the Singular Value Decomposition Method to name but a few. Details of these algorithms can be found in e.g. [41]. We denote with the function f^{Wahba} any of these algorithms such that the resulting quaternion $f^{\text{Wahba}}(a, r, b) \in \mathbb{S}_3$ solves the optimisation problem (3.27) for $A = A(q)$. One method which has proven to give good results is the Singular Value Decomposition Method which uses the matrix

$$B^{\text{Wahba}} = \sum_{i=1}^{N} a_i b_i r_i^\top$$

and its singular value decomposition $B^{\text{Wahba}} = U \Sigma V^\top$ to define the function f^{Wahba} as

$$f^{\text{Wahba}}(a, r, b) = q\Big(U \operatorname{diag}\left(1, 1, \det(U), \det(V)\right) V^\top\Big)$$

where we denote here by $q(\cdot)$ the transformation from attitude matrices to quaternions as given in Appendix A.2.

Dynamic Estimation: Multiplicative Extended Kalman Filter

One very common method to solve the observation problem posed by the system (3.26) is to employ a multiplicative extended Kalman filter. It uses the quaternions $\hat{q}$ as a global attitude representation while using the three component vector $\delta\vartheta$ as the local representation of attitude errors. The real quaternion q can be obtained in a multiplicative fashion using the estimate and the quaternion error δq as

$$q = \delta q(\delta\vartheta) \otimes \hat{q}\,.$$

The usage of these two attitudes representations allows minimal dimensionality of the covariance matrix because $\hat{q}$ is always a unit quaternion and avoids singularities because $\delta\vartheta$ represents only small attitude errors. A detailed discussion and illustration of the indicated advantages can be found in [41]. The algorithm used to estimate the attitude is given in Table 3.1. It is identical to the one presented in [41, Chapter 6.2]. The algorithm is easily

Table 3.1: Extended Kalman filter for attitude estimation.

Initialise	$\hat{q}(0)=\hat{q}_0, \quad P(0) = P_0$
Gain	$K_k = P_k^- H_k^\top(x_k^-)\Big(H_k(x_k^-)P_k^- H_k^\top(x_k^-) + R_k\Big)^{-1}$
	$H_k(\hat{q}^-) = \Big[[A(\hat{q})r_1\times]^\top \quad \cdots \quad [A(\hat{q})r_N\times]^\top\Big]^\top$
Update	$P_k^+ = (I - K_k H_k(\hat{x}_k^-))P_k^-$
	$h_k(\hat{q}^-) = \Big[(A(\hat{q}^-)r_1)^\top \quad \cdots \quad (A(\hat{q}^-)r_N)^\top\Big]^\top$
	$\delta\vartheta_k^+ = K_k\Big(y_k - h_k(\hat{q}^-)\Big)$
	$\hat{q}^* = q_k^- + \frac{1}{2}\Xi(\hat{q}_k^-)\delta\hat{\vartheta}_k^+$
	$\hat{q}_k^+ = \frac{\hat{q}^*}{\|\hat{q}^*\|}$
Propagation	$\dot{\hat{q}} = \frac{1}{2}\Xi(\hat{q})\omega$
	$\dot{P} = [\omega\times]P + P[\omega\times]^\top + Q$

extended to estimate other states or parameters such as biases, scale factors or misalignments simultaneously. The algorithm consist of three steps, a measurement update, a state vector reset and a propagation to the next time. This is a slight deviation from the Kalman filter algorithm commonly used in literature and introduced in [10] as the order of the measurement and propagation step is reversed and the state $\delta\hat{\vartheta}$ is set to zero in every time step. Further, it shall be pointed out that the correction is implemented in a discrete manner in the local variables $\delta\hat{\vartheta}$ while the propagation is performed in the global coordinates $\hat{q}$ using the continuous description. Furthermore, the quaternion is also normalised in every step to ensure its interpretation as an attitude. Note that the angular velocity ω is also incorporated in the quaternion dynamics but that their respective dynamics is not used in this algorithm.

This completes the basics on attitude estimation and control used throughout this work. We continue by introducing the remaining thermal dynamics that allow to formulate our system of interest and the problem investigated in this work.

4 Spacecraft Thermal Modelling

In order to estimate the attitude of a spacecraft using temperature data, it is necessary to know the underlying dynamics of a temperature sensor. These dynamics are discussed and introduced in this chapter. Further, their characteristics are analysed which allows to determine the required accuracies and sampling rate of the temperature sensors.

4.1 Thermal Model

A detailed derivation of the thermal dynamics can be found e.g. in [108]. Assuming a isothermal surface with area A_s, temperature T and thermal capacitance C, the temperature dynamics are obtained using the principle of conservation of energy as

$$C\frac{\mathrm{d}T}{\mathrm{d}t} = Q_\mathrm{d} - Q_\mathrm{c} - Q_\mathrm{r}\,,$$

with Q_d the dissipated heat, Q_c the heat conduction to neighbours and Q_r the heat radiation to neighbours. Let there be n neighbours and each neighbour j has a temperature T_j. Then Fourier's law as in [109] gives the heat conduction as

$$Q_\mathrm{c} = \sum_{j=1}^{n} k_j(T - T_j)$$

with the thermal conductivity between the body and its neighbour k_j. The heat radiation reads

$$Q_\mathrm{r} = \sum_{j=1}^{n} A\sigma\mathfrak{F}_j(T^4 - T_j^4)$$

with σ being the Stefan–Boltzmann constant and $\mathfrak{F}_j$ the exchange factor between the surfaces. In view of the major heat being produced by Earth and Sun, we can recast the equation to obtain

$$C\frac{\mathrm{d}T}{\mathrm{d}t} = Q_\mathrm{sun} + Q_\mathrm{alb} + Q_\mathrm{planet} - Q_\mathrm{ds} + Q_\mathrm{d} - \sum_{j=1}^{n} k_j(T - T_j) - \sum_{j=1}^{n} A_\mathrm{s}\sigma\mathfrak{F}_j(T^4 - T_j^4) \tag{4.1}$$

where Q^sun is heat produced by the solar radiation, Q^alb by the Earth's albedo and Q^planet due to the infra-red radiation of the Earth. The term Q_ds describes the radiation of the surface to deep space which is assumed to have 0 K and Q_d is the dissipated heat.

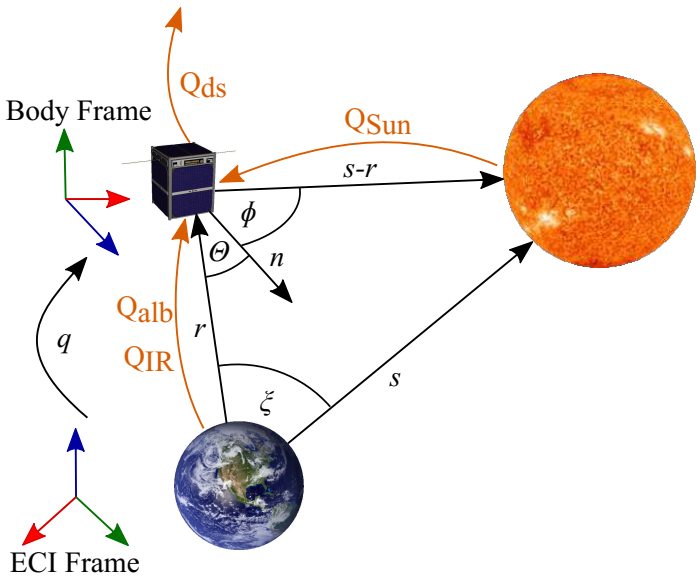


Figure 4.1: Illustration of the considered system. Courtesy: NASA.

Typically, a mathematical model representing the geometry of the system is developed and divided into nodes. The dynamics of each node are described by (4.1) as in [110]. We consider a cuboid spacecraft with six surfaces, which individual temperatures T depend only on the influence of the radiation to space and from Earth and Sun and dissipated heat, i.e. Equation (4.1) simplifies to

$$C\frac{\mathrm{d}T}{\mathrm{d}t} = Q_{\mathrm{sun}} + Q_{\mathrm{alb}} + Q_{\mathrm{planet}} - Q_{\mathrm{ds}} + Q_{\mathrm{d}}\,. \tag{4.2}$$

An illustration of the acting irradiations and their corresponding angles which will be discussed in the next section can be found in Figure 4.1. In the following we derive formulae for the heat inputs. First, we describe each input using functions of the Earth position $r \in \mathbb{R}^3$, the Sun position $s \in \mathbb{R}^3$ and the normal direction of the surface $n^{\mathrm{I}} \in \mathbb{R}^3$ given in a common frame. We choose this common frame to be the initial frame, however the use of any other frame is possible as well. Then we give the explicit formula for determining the heat flows of the stated variables.

Solar Heat

The solar radiation is the main factor influencing temperature changes of the spacecraft. A solar constant G_{s0} is defined as in [110] which gives the mean solar irradiance acting on a unit area perpendicular to the solar rays in a distance of 1 au. As the amount of irradiance crossing spherical surfaces with different radii is constant, the solar irradiation G_s scales with distance as

$$G_s(d) = G_{s0}\frac{1}{d^2}$$

where d is the distance in astronomical units.

Further, we introduce the normal solar angle ϕ which is defined between the normal of the spacecraft surface n^{I} and the vector pointing to the Sun $s - r$ as

$$\phi := \angle(s - r, n^{\mathrm{I}})$$

where the angle function $\angle$ is defined in Appendix A.2. Because the Sun is extremely far away from the Earth if compared to the spacecraft, i.e. $\|s\| >> \|r\|$ this is almost identical to the angle between Sun and normal, i.e.

$$\phi \approx \angle(s, n^{\mathrm{I}}) .$$

The normal in inertial coordinates changes with the attitude of the spacecraft, which makes this angle substantial for the desired attitude reconstruction.

Finally, we introduce the shadow function which gives the occultation of the satellite due to the Earth. As the distance between Earth and Sun is significantly higher than the difference of their radii, it is sufficient to assume cylindrical shadows instead of conics as illustrated in Figure 4.2. Thus, the shadow can be calculated as

$$\nu(r, s) = \begin{cases} 0 & \text{if } \angle(r, s) > \frac{\pi}{2} \wedge \|r - r^\top s \frac{s}{\|s\|^2}\| < r_\oplus \\ 1 & \text{otherwise} \end{cases} . \tag{4.3}$$

There exist more elaborate methods to model the shadow by dividing the Earth's shadow into umbra and penumbra. The shadow coefficient $\nu \in (0, 1)$ in penumbra is then determined by the overlapping of two circular disks. This is out of the scope for this work as this is not necessary for low Earth orbits and a detailed derivation can be found in [95].

This allows to model the solar irradiation as

$$Q^{\mathrm{sun}}(r, s, n^{\mathrm{I}}) = \begin{cases} \alpha_{\mathrm{s}} G_s \left(\frac{\|s-r\|}{1\,\mathrm{au}}\right) A_{\mathrm{s}} \cos\left(\phi(n^{\mathrm{I}}, s)\right) \nu(r, s) & \text{if } 0 < \phi(n^{\mathrm{I}}, r, s) < \frac{\pi}{2} \\ 0 & \text{if } \frac{\pi}{2} < \phi(n^{\mathrm{I}}, r, s) < \pi \end{cases}$$

where α_{s} denotes the solar absorptance of the surface. Note that usage of $\cos\left(\phi(n, r, s)\right)$ instead of a more complicated form factor as required for the upcoming albedo and infrared irradiation. This is due to the fact that the Sun is considered as a point heat source and not a spheric one because the radius of the Sun is extremely small in comparison to the distance between spacecraft and Sun.

Albedo

As albedo irradiation we denote the part of the solar radiation which is reflected by the Earth or scattered by the planet surface and atmosphere.

This makes two angles that influence this irradiation. This first, the solar zenith angle ξ describes the angle between the spacecraft vector r and the Sun vector s. This angle describes

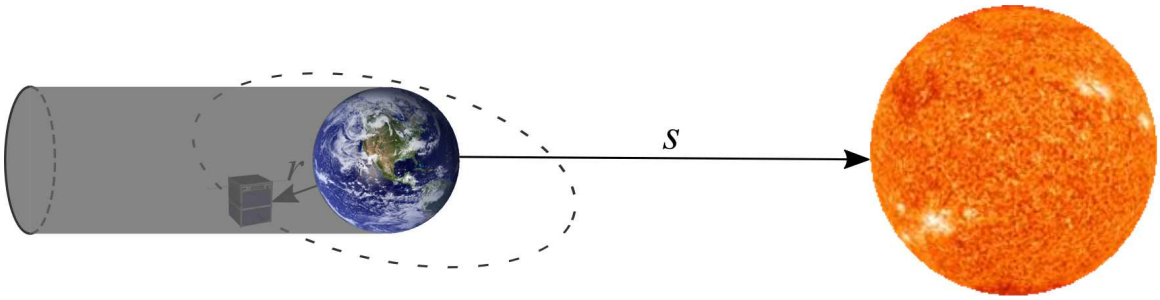


Figure 4.2: Cylindrical Shadow Model. Courtesy: NASA.

the portion of the illuminated planet which is seen by the spacecraft and is defined by

$$\xi(s,r) = \angle(s,r)\,.$$

It is independent of the attitude of the spacecraft and only dependent on the orbit and the time of the year.

The second angle is the angle between the between the normal of the plate n^{I} and spacecraft r

$$\theta(r,n^{\mathrm{I}}) = \angle(r,n^{\mathrm{I}})\,.$$

In comparison to the Sun, the Earth should not be considered as a point heat source but as a spherical heat source. Therefore, the influence of this normal Earth angle needs to be modelled by a form factor as introduced in [111]. This has the form

$$F(\theta,r) = \begin{cases} \frac{r_\oplus^2}{\|r\|^2}\cos(\theta) & \text{if } \theta < \frac{\pi}{2} - \arcsin\left(\frac{r_\oplus}{\|r\|}\right) \\ F_2 & \text{if } \frac{\pi}{2} - \arcsin\left(\frac{r_\oplus}{\|r\|}\right) < \theta < \frac{\pi}{2} + \arcsin\left(\frac{r_\oplus}{\|r\|}\right) \\ 0 & \text{if } \theta > \frac{\pi}{2} + \arcsin\left(\frac{r_\oplus}{\|r\|}\right) \end{cases} \tag{4.4}$$

where $r_\oplus$ denotes the radius of the Earth and

$$\begin{aligned} F_2 = \frac{1}{2} - \frac{1}{\pi}\arcsin\left(\frac{\sqrt{H^2-1}}{H\sin(\theta)}\right) \\ + \frac{1}{\pi H^2}\Big(\cos(\theta)\arccos\big(-\sqrt{H^2-1}\cot(\theta)\big) - \sqrt{H^2-1}\sqrt{1-H^2\cos(\theta)^2}\Big) \end{aligned}$$

where $H = \frac{\|r\|}{r_\oplus}$. Note that for $\theta < \frac{\pi}{2} - \arcsin\left(\frac{r_\oplus}{\|r\|}\right)$ the relation $F(\theta,r) = \frac{1}{G_{\mathrm{s0}}} G_{\mathrm{s}}\left(\frac{r}{r_\oplus}\right)\cos(\theta)$ confirms with $\lim_{\frac{r_\oplus}{\|r\|}\to 0}\arcsin\left(\frac{r_\oplus}{\|r\|}\right) = 0$ the simplifications made for the solar irradiation if the radius $r_\oplus$ is much higher than the distance $\|r\|$.

This leads to the model for the albedo irradiations as a slight variation of [112] and [110] as

$$Q^{\mathrm{alb}}(r,s,n^{\mathrm{I}}) = \begin{cases} \rho_{\mathrm{alb}}\alpha_{\mathrm{s}}G_s(d)A_{\mathrm{s}}F(\theta(r,n^{\mathrm{I}}),r)\cos\left(\xi(s,r)\right) & \text{if } 0 < \xi(s,r) < \frac{\pi}{2} \\ 0 & \text{if } \frac{\pi}{2} < \xi(s,r) < \pi \end{cases} \tag{4.5}$$

where $\rho_{\mathrm{alb}} \in [0,1]$ is the albedo coefficient. This albedo coefficient is assumed to be constant over time. This is a common assumptions as the coefficient is almost periodic with a period of one orbit and usually the mean value is chosen. We use this assumption for this chapter and Chapter 5 as it facilitates the analytical considerations, but we drop it in Chapter 6 and define the coefficient as a function of the latitude of the spacecraft.

Planetary Radiation

As planetary radiation we denote the thermal radiation which is emitted by the planet as infra-red radiation. The emitted radiation can be calculated as the absorbed solar radiation of the planet minus the radiation emitted via albedo. Then, by assuming the planet to be a black body we obtain the planetary infra-red thermal heat acting on a surface of a spacecraft as in [110] as

$$Q^{\mathrm{planet}}(r,n^{\mathrm{I}}) = \varepsilon_{\mathrm{e}}A_{\mathrm{s}}I_{\mathrm{IR}}F(\theta(r,n^{\mathrm{I}}),r)\,, \tag{4.6}$$

where I_{IR} denotes Earth emitted infra red flux. The infra red flux is directly correlated to the albedo coefficient. The same assumptions as for this coefficient are use throughout this work for the emitted infra-red flux.

Deep Space Radiation

The deep space irradiation is the irradiation emitted of the spacecraft via radiation into deep space. As deep space is assumed to have approximately 0K , the radiation has the form

$$Q_{\mathrm{ds}}(T) = A_{\mathrm{s}}\sigma\varepsilon_{\mathrm{e}}T^4\,,\,. \tag{4.7}$$

This is the only irradiation proposed which actually depends on the current temperature T.

Thermal Dynamics

We use these four previously introduced irradiation in view to give a compact model description in view of (4.2). We consider the variables as function of temperature T, attitude represented as quaternions q, angular velocity ω and time t as these are considered as the states for our system. Other dependency for example on the normal vector or the parameters are not explicitly stated in the notation. However, they play a vital role in Chapter 6 and the variable definitions will be adjusted there. The final model introduced in this section has small simplifications in comparison to the variables proposed in the previous

section. The time dependency of the Sun s and the solar irradiation G_s is neglected due to their slow time variance. The time variance of r must not be neglected for low Earth orbits as it influences the temperature evolution significantly for the considered time scale. We introduce the variables and parameters α, β, γ, δ of the form

$$\alpha(t) = \frac{1}{C}\alpha_\mathrm{s} G_s A_\mathrm{s} \nu(r(t)) \in \{0, \bar{\alpha}\} \qquad \beta(t) = \frac{1}{C}\rho_\mathrm{alb}\alpha_\mathrm{s} G_s A_\mathrm{s} \cos\left(\xi(t)\right) \in [-\bar{\beta}, \bar{\beta}] \tag{4.8a}$$

$$\gamma = \frac{1}{C}\varepsilon_\mathrm{e} A_\mathrm{s} I_\mathrm{IR}, \qquad \delta = \frac{1}{C}\varepsilon_\mathrm{e} A_\mathrm{s} \sigma_\mathrm{s} \tag{4.8b}$$

to allow a compact model representation. The variable α incorporates some of the spacecraft and material specific parameters and the shadow function. Therefore this function is either $\bar{\alpha} = \frac{1}{C}\alpha_\mathrm{s} G_s A_\mathrm{s}$ or zero, depended if the spacecraft is in the shadow or not. The second variable β incorporates some of the introduced parameters and the solar zenith angle which varies between 0 and π. Therefore β can be any value between zero and $\bar{\beta} = \frac{1}{C}\rho_\mathrm{alb}\alpha_\mathrm{s} G_s A_\mathrm{s}$. The remaining parameters γ and δ are constant and depend on the mission. Note that all parameters may change over the course of a mission due to ageing. However, as we are only interested in time periods of multiple orbits we assume that the parameters are not affected by ageing.

The angle between the normal of the spacecraft surface and Sun $\phi(q)$, the angle between normal and spacecraft position $\theta(q,t)$ and the solar zenith angle $\xi(q,t)$ are defined using the definition of the previous section and the current attitude q. Because the normal vector n is only known in body coordinates we use $n := A(q)n^\mathrm{I}$ to define the three angles

$$\phi(q) = \arccos\left(\frac{s^\top}{\|s\|} A(q)^\top n\right) \in [0, \pi] \qquad \theta(q,t) = \arccos\left(-\frac{r^\top(t)}{\|r(t)\|} A(q)^\top n\right) \in [0, \pi] \tag{4.9a}$$

$$\xi(q,t) = \arccos\left(\frac{r^\top(t)}{\|r(t)\|}\frac{s}{\|s\|}\right) \in [0, \pi] \tag{4.9b}$$

with the rotation matrix

$$A(q) = \|q\|^{-2}\Big((q_4^2 - \|q_{1:3}\|^2)I_3 + 2q_{1:3}q_{1:3}^\top - 2q_4[q_{1:3}\times]\Big)$$

where $q = \begin{bmatrix} q_1 & q_2 & q_3 & q_4 \end{bmatrix}^\top = \begin{bmatrix} q_{1:3} & q_4 \end{bmatrix}^\top$ is the quaternion describing the attitude of the body frame of the spacecraft with respect to the world ECI frame. We write the solar irradiation, the albedo irradiation, the infrared irradiation and the deep space irradiation introduced in the previous section in a more compact form using the maximum function

$$Q_\mathrm{sun}(q,t) = \max\Big(\alpha(t)\cos(\phi(q)), 0\Big) \qquad Q_\mathrm{alb}(q,t) = \max\Big(\beta(t)F(\theta(t,q), r(t)), 0\Big) \tag{4.10a}$$

$$Q_\mathrm{IR}(q,t) = \gamma F(\theta(t,q), r(t)) \qquad Q_\mathrm{ds}(T) = \delta T^4\,. \tag{4.10b}$$

We finish this section by considering the dissipative heat Q_d as a function of time, and introduce the temperature dynamics that will be used for the remainder of this work as a

basis for the analysis

$$\dot{T} = Q_{\text{sun}}(q,t) + Q_{\text{alb}}(q,t) + Q_{\text{IR}}(q,t) - Q_{\text{ds}}(T) + \frac{Q_{\text{d}}(t)}{C}\,. \tag{4.11}$$

4.2 Characteristics of the Thermal System

In this section we analyse the dynamics (4.11) introduced in the previous section. We discuss potential model and measurement errors and determine the time constant based on a first order linearisation for a set of parameters given in Appendix A.4. We review the most common attitude dynamics that influence the temperature evolution. This is put into perspective with the sampling time and measurements and model errors to determine possible estimation accuracies.

We start by discussing potential model uncertainties of the system. The model (4.11) is proposed in view of its simplicity to allow an analytical design of an observer. This leads to a number of disturbances and uncertainties that act on (4.11) when compared to the real system. All these disturbances can be modelled introducing a new variable $\phi_T(T,q,t)$ which is added to the right side of the dynamics. This disturbance usually consists of uncertainties induced by estimation errors of the variables $\alpha, \beta, \gamma, \delta$ and unmodelled heat flows in the dissipative heat Q_{ds}. The uncertainty errors are due to estimation errors of the variables $\alpha_{\text{s}}, \varepsilon_{\text{e}}, A_{\text{s}}$, C_{s}, ρ_{alb} and I_{IR}. These errors may vary with the state, over the course of the orbit and also over the course of the lifetime of the mission. In order to compensate these last mentioned degradation effects, the model parameters should be adjusted periodically. The exact domains of these uncertainties and disturbance are not investigated in this work. Based on the measurements discussed in Chapter 6 a 10 % margin for every parameter appears sufficient to make first simple robustness studies.

The second class of errors in this context are the measurement errors which are induced due to the imperfectness of the temperature sensors. Herby, the accuracy of the measurements varies dependent on the sensor used. Generally, the temperature signals for spacecraft in the considered orbits are in the range between $-100\,^{\circ}\text{C}$ to $100\,^{\circ}\text{C}$. For commonly available sensors with an accuracy of $0.1\,^{\circ}\text{C}$ [113–115] this leads to errors of about 0.1 % in the considered temperature ranges. Other works such as [56] use sensors with an even higher accuracy of up to $0.001\,^{\circ}\text{C}$ for the temperature based attitude estimation. The sensors used to obtain the measured data in Chapter 6 have an accuracy of $0.1\,^{\circ}\text{C}$.

In order to obtain a first impression of the time constants associated with the thermal dynamics, we linearise the system around a steady state temperature T^* with the error temperature $\Delta T = T - T^*$. All other heat flows are considered to be constant to obtain an approximation of the temperature eigendynamics. The linearisation has the form

$$\Delta\dot{T} = -4\delta T^{*3}\Delta T + \mathcal{O}(\Delta T^2)\,. \tag{4.12}$$

This yields the time constant $T_{\text{th}} = \frac{1}{-4\delta T^{*3}}$ which describes as a first rough estimation the temperature dynamics. This time constant depends on the considered steady state temperature and is shown in Figure 4.3 for the parameters given in Appendix A.4. Clearly,

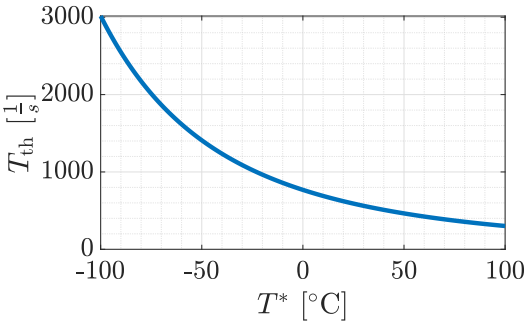


Figure 4.3: Time Constant changing with the considered steady state temperature

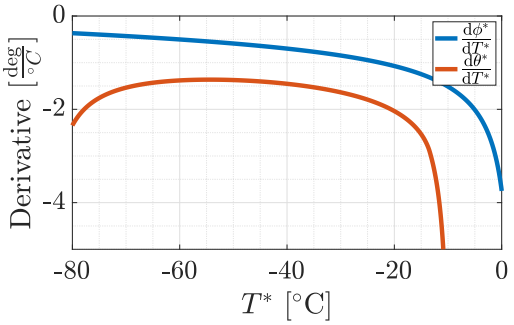


Figure 4.4: Derivative of the steady state angles

the time constant of the system falls with rising temperature. While the time constant for small temperatures is as high as 3000 $\frac{1}{\text{s}}$, it becomes around 300 $\frac{1}{\text{s}}$ for high temperatures. The latter corresponds to a cutoff frequency of approximately $f_{\text{co}} = 3\,\text{mHz}$ and to a convergence time of approximately 23 min until the temperature is within a 1 % interval of the steady state temperature.

This needs to be put into perspective with the underlying attitude dynamics which determine the steady state temperature T^*. First it needs to be pointed out, that due to the orbit motion the orientation of the uncontrolled spacecraft is not fixed with respect to Sun and Earth. Therefore, even for an uncontrolled spacecraft, the temperature usually does not remain in steady state due to the time variance of its position $r(t)$. For orbits considered in this work, the orbit period is about 90 minutes which corresponds to an angular velocity of about $-1.2 \cdot 10^{-3}\,\frac{\text{rad}}{\text{s}}$. This leads to the first attitude manoeuvre considered in this work, namely the three axis stabilisation. For this manoeuvre the spacecraft rotation matches the induced rotation due to the orbit to maintain the same orientation with respect to the Earth. In other words, the body frame of the spacecraft is identical to the orbit frame defined in Section 3.1. The second slow manoeuvre considered in this work is a rotation as employed in Chapter 6. This kind of manoeuvre achieves a 50 degrees change of orientation in around 13.3 minutes. This corresponds to an average angular velocity of $-1.1 \cdot 10^{-3}\,\frac{\text{rad}}{\text{s}}$. On the other side, extremely agile slew manoeuvres are currently subject to research as shown in [116, 117]. For these manoeuvres high torque wheels or control moment gyroscopes are used that aim for a 30 degrees rotation within 10 seconds. This is equivalent to an average angular velocity of $5.24 \cdot 10^{-2}\,\frac{\text{rad}}{\text{s}}$ which is about ten times faster than the slow manoeuvres considered before.

For these kind of manoeuvres the sampling rate is of major importance. For fast manoeuvres the sampling rate of the control must be sufficiently fast to achieve the desired performance. For the design of the observer, the sampling rate of the temperature signal is important to achieve a correct estimate. The first factor to determine this sampling rate is the time constant describing the system. With this time constant obtained in (4.12) and

the Nyquist–Shannon sampling theorem a required sampling frequency of at least

$$f_{\text{Sampling}} = 2f_{\text{co}} \approx 6\,\text{mHz} \tag{4.13}$$

is required to avoid aliasing effects. Note that this frequency is directly proportional to the area A, time constant C and emissivity ε introduced in Section 4.1. Dependent on the size and material properties of the spacecraft these parameters may vary significantly. This can be seen for example by the fact that the cutoff frequency obtained from the measurement data in Section 6.1 are about 10 times higher compared to the considered parameters in Appendix (A.4). Nevertheless, common sensors such as [113–115] have sample frequency ranges from 20 Hz to 1 Hz which is more than sufficient with respect to these slow eigendynamics. The issues related to the sampling time are therefore only occurring for fast manoeuvres in which a slow sampling rate of 1 Hz for the control may not be sufficient to obtain the desired performance.

With these informations it is possible to obtain a first approximation of the potential quality of the attitude estimate based on temperature data with steady state calculations. In order to make this easier to interpret, we consider the angles of the normal to Earth and Sun ϕ and θ as the representation of the attitude. Note that these two angles are simple to obtain for a known attitude q and can be transformed with some additional information into the complete attitude q as discussed later in Chapter 6. The steady state temperature obtained from (4.11) as a function of the angles between normal and Sun ϕ as well as normal and Earth θ has the form

$$T^*(\phi^*, \theta^*) = \left(\frac{1}{\delta} \left(\max\left(\alpha \cos(\phi^*), 0 \right) + \max\left(\beta F(\theta^*, r), 0 \right) + \gamma F(\theta^*, r) + \frac{Q_{\text{d}}}{C} \right) \right)^{\frac{1}{4}} . \tag{4.14}$$

If we assume that only solar irradiation or infrared irradiation is acting, we can reformulate this equation for the angle ϕ^* and θ^*

$$\phi^* = \arccos\left(\frac{\delta T^{*4} - \frac{Q_{\text{d}}}{C}}{\alpha} \right), \qquad \theta^* = F^{-1}\left(\frac{\delta T^{*4} - \frac{Q_{\text{d}}}{C}}{\gamma} \right). \tag{4.15a}$$

This allows to calculate the derivative $\frac{\mathrm{d}\phi^*}{\mathrm{d}T^*}$ and $\frac{\mathrm{d}\theta^*}{\mathrm{d}T^*}$, respectively. Both of these derivatives are displayed as a function of the steady state temperature in Figure 4.4. It can be seen that the derivative of the solar angle is smaller than the one from the Earth angle. Both functions exhibit a singularity at the temperature that is obtained if the spacecraft is pointing towards the heat source due to the influence of arccosine. An additional singularity exists for the derivative of the Earth angle for vanishing infrared irradiation. Nevertheless, both derivatives yield a finite mean value on a bounded interval, which is approximately $-1.0\,\frac{\text{deg}}{^\circ\text{C}}$ for the solar angle and $-2.0\,\frac{\text{deg}}{^\circ\text{C}}$ for the Earth angle. Thus, by only considering measurements errors the estimation accuracy for the location of the solar irradiation source is two times better than for the infrared irradiation source at least for steady state calculations. In combination with the measurement errors discussed beforehand an accuracy of 0.1 deg is possible for the steady state calculations. However as mentioned before, the temperature is rarely in steady state but follows changes in irradiation either introduced by the orbit or

the commanded manoeuvres of the spacecraft. In this case, it is more difficult to obtain an approximation of the quality of the attitude estimation. This attitude estimation now relies on an approximation of the derivative of the temperature. The quality of this approximation relies on the sampling time, the measurement error, the temperature dynamics as well as the considered manoeuvre. Every approach to obtain an estimation of the possible accuracies relies on assumptions on the quality of estimation of $\dot{T}$ and thereby requires almost the complete design of the observer itself and is therefore not discussed in detail here. This topic is treated in detail in Chapter (6).

5 Observer Based Attitude Estimation with a Single Temperature Measurement

In this chapter we are interested in estimating the attitude using only a single temperature sensor. This section shall show what is possible from an algorithmic point of view when a single perfect sensor is considered and discuss possible limitations and challenges. Instead of merging multiple temperature signals, the proposed methods focus on using only a single signal and uses higher order derivatives of the obtained temperature signal. We start the chapter using the model introduced in the previous section to give a mathematically sound problem definition. In order to do so we define the dynamics of the previous section in a compact manner by introducing the vector fields

$$f_T(T,q,t) = \max\Big(\alpha(t)\cos(\phi(q)),0\Big) + \max\Big(\beta(t)F(\theta(t,q),0\Big) + \gamma F(\theta(t,q),t) - \delta T^4 \tag{5.1a}$$

$$f_q(q,\omega) = \frac{1}{2}\Omega(\omega)q \tag{5.1b}$$

$$f_\omega(\omega) = J^{-1}(-\omega \times J\omega) \tag{5.1c}$$

and define the complete state and the total dynamics by

$$\begin{aligned} x &= \begin{bmatrix} T & q & \omega & t \end{bmatrix}^\top \\ f(x) &= \begin{bmatrix} f_T(T,q,t) & f_q(q,\omega) & f_\omega(\omega) & 1 \end{bmatrix}^\top \\ g &= \begin{bmatrix} 0 & 0 & J^{-1} & 0 \end{bmatrix}^\top. \end{aligned}$$

Note that we neglect the dissipative heat Q^{d} and assume a disturbance free sensor as well as model for this section. This allows to focus on the main algorithmic challenges and properties.

5.1 Problem Definition

We are interested in the observability properties of the thermal system (5.2) if we use only a single temperature measurement and additional angular velocity measurements. This means

we want to analyse the system

$$\dot{T} = f_T(T, q, t) \tag{5.2a}$$
$$\dot{q} = f_q(q, \omega) \tag{5.2b}$$
$$\dot{\omega} = f_\omega(\omega) + J^{-1}u\,. \tag{5.2c}$$
$$\dot{t} = 1 \tag{5.2d}$$

with the output

$$y = h(x) = \begin{bmatrix} T & q^\top q - 1 & \omega & t \end{bmatrix}^\top \tag{5.3}$$

in view of its observability properties and design an observer which estimates the complete state x and therefore in particular the attitude q. The six dimensional output y incorporates the temperature T, the angular velocity ω and the (measurable) time t. It is well known that the attitude space is of dimension three and that the quaternion representation uses four variables to allow a singularity free representation. In order to make sure that the quaternion space is also only of order three despite being described by four variables it is restricted to unit quaternions. Thus, every quaternion q is an element of the unit sphere $\mathbb{S}_3$ and this is accounted for by adding a virtual output $q^\top q - 1$ which is constantly zero (not needed to be measured). In the observer design this variable will be omitted later on because the proposed observer structure inherently guarantees that this quaternion constraint is fulfilled. We start by grouping the states into meaningful configurations to analyse their resulting observability properties.

State Space Separation

The complete state space $\mathbb{X}$ is governed by the domains of temperature $\mathbb{R}^+$, quaternions $\mathcal{Q} = \mathbb{S}_3$, angular velocity $\mathcal{W} = \mathbb{R}^3$ and time $\mathcal{T} = \mathbb{R}$, i.e. $\mathbb{X} = \mathbb{R}^+ \times \mathcal{Q} \times \mathcal{W} \times \mathcal{T}$. We isolate the points where the vector field f_T is not differentiable to obtain a family of smooth systems, the domains $\mathbb{X}_1, \ldots, \mathbb{X}_6$ of which form a partition of the state space $\mathbb{X}$. The functions ν, ξ, ϕ and θ define the points where the vector field is not differentiable. We distinguish between ν, ξ which are functions of time t and θ, ϕ which are functions of the attitude q. Thus, we use ν and ξ to divide the space of time into subspaces. Note that a spacecraft in solar eclipse has a solar zenith angle larger than 90 degrees, i.e. if $\nu(r(t)) = 0$ then $\xi(t) \geq \frac{\pi}{2}$. With this in mind we divide the orbit into three regions with respect to time. A region in which the spacecraft is in solar eclipse, another one when the spacecraft is sunlit, but no albedo radiation is acting, and finally the region when the spacecraft is sunlit and under the influence of albedo radiation. These three time sets are defined by

$$\mathcal{T}_{\text{ec}} = \left\{ t \in \mathbb{R} \mid \nu(t) = 0 \right\}$$
$$\mathcal{T}_{\text{noalb}} = \left\{ t \in \mathbb{R} \mid \nu(t) = 1 \wedge \xi(t) \geq \frac{\pi}{2} \right\}$$
$$\mathcal{T}_{\text{sun}} = \left\{ t \in \mathbb{R} \mid \xi(t) < \frac{\pi}{2} \right\}$$

and partition the complete time domain $\mathcal{T}$. Note that, in our proposed model, time t acts only on the position of the spacecraft $r(t)$, which is periodic. Thus, in the observability analysis it is sufficient to consider only a single period, i.e. we use $\mathcal{T} = [0, \frac{2\pi}{\omega_0})$. Similarly, we divide the quaternion space $\mathcal{Q}$ into subspaces using $\phi(q)$ and $\theta(q,t)$ such that

$$\mathcal{Q}_{\text{sun}} = \left\{ q \in \mathbb{S}_3 \mid \phi(q) < \frac{\pi}{2} \right\},$$
$$\mathcal{Q}_{\text{earth}}(t) = \left\{ q \in \mathbb{S}_3 \mid \theta(q,t) < \frac{\pi}{2} + \arcsin\left(\frac{r_\oplus}{\|r(t)\|} \right) \right\}$$

where $r_\oplus$ is the mean Earth radius. Attitudes within $\mathcal{Q}_{\text{sun}}$ allow the surface to be sunlit. Attitudes within $\mathcal{Q}_{\text{earth}}(t)$ allow the irradiation from Earth to act on the surface. The second set varies with the position of the spacecraft and as such with time t. These variations lead to six different smooth models which piecewisely describe the dynamics. Table 5.1 gives an overview of the partitioning of the state space and the corresponding dynamics. An entry of 1 in the table means that the state is part of the set of the corresponding column while 0 defines the state to be in the complement of the set. The fourth column describes the resulting smooth right hand side for each state x where the irradiation terms are only noted when they are not identical to zero. This allows to define the six subsets of the state space $\mathbb{X}_1, \dots \mathbb{X}_6$ as follows

$$\mathbb{X}_1 = \bigcup_{t \in \mathcal{T}_{\text{sun}}} \mathbb{R}^+ \times Q_{\text{sun}} \cap Q_{\text{earth}}(t) \times \mathcal{W} \times \{t\}$$
$$\mathbb{X}_2 = \bigcup_{t \in \mathcal{T}_{\text{noalb}}} \mathbb{R}^+ \times Q_{\text{sun}} \cap Q_{\text{earth}}(t) \times \mathcal{W} \times \{t\}$$
$$\mathbb{X}_3 = \bigcup_{t \in \mathcal{T}_{\text{sun}} \cup \mathcal{T}_{\text{noalb}}} \mathbb{R}^+ \times Q_{\text{sun}} \cap Q^c_{\text{earth}}(t) \times \mathcal{W} \times \{t\}$$
$$\mathbb{X}_4 = \Big(\bigcup_{t \in \mathcal{T}_{\text{noalb}} \cup \mathcal{T}_{\text{ec}}} \mathbb{R}^+ \times Q^c_{\text{sun}} \cap Q_{\text{earth}}(t) \times \mathcal{W} \times \{t\} \Big) \cup \Big(\bigcup_{t \in \mathcal{T}_{\text{ec}}} \mathbb{R}^+ \times Q_{\text{sun}} \cap Q_{\text{earth}}(t) \times \mathcal{W} \times \{t\} \Big)$$
$$\mathbb{X}_5 = \bigcup_{t \in \mathcal{T}_{\text{sun}}} \mathbb{R}^+ \times Q^c_{\text{sun}} \cap Q_{\text{earth}}(t) \times \mathcal{W} \times \{t\}$$
$$\mathbb{X}_6 = \Big(\bigcup_{t \in \mathcal{T}} \mathbb{R}^+ \times Q_{\text{sun}} \cap Q^c_{\text{earth}}(t) \times \mathcal{W} \times \{t\} \Big) \cup \Big(\bigcup_{t \in \mathcal{T}_{\text{ec}}} \mathbb{R}^+ \times Q^c_{\text{sun}} \cap Q^c_{\text{earth}}(t) \times \mathcal{W} \times \{t\} \Big)$$

where Q^c_{sun} denotes the complement of the set Q_{sun}. Each set $\mathbb{X}_i$ is denoted as Case i for $i \in \{1, \dots, 6\}$ as displayed in the final column of Table 5.1. Each case has a superposition of different irradiations governing the temperature evolution.

Finally, we shall add some small notes on the influence of the inertia matrix J and the normal vector n. The inertia matrix J induces a cross coupling term in the equation of motion. This cross coupling term vanishes if the satellite has an inertia matrix which is a multiple of the identity matrix. In this case the angular velocities along each axis are not coupled, i.e. $f_\omega = J^{-1}u$. This may alter the observability properties. Generally however, the satellite is not completely symmetric which is why we consider only non-trivial inertia matrices. The normal vector n of the surface determines the magnitude of all three irradiations. However, this vector occurs only in combination with the attitude matrix $A(q)$. Thus, we assume the

Table 5.1: Overview of piecewise continuous dynamics dependent on time and attitude

$\mathcal{T}$	$\mathcal{Q}_{\text{sun}}$	$\mathcal{Q}_{\text{earth}}(t)$	$f(x)$	Case
$\mathcal{T}_{\text{sun}}$	1	1	$Q_{\text{sun}} + Q_{\text{alb}} + Q_{\text{IR}} - Q_{\text{ds}}$	1
$\mathcal{T}_{\text{noalb}}$	1	1	$Q_{\text{sun}} + Q_{\text{IR}} - Q_{\text{ds}}$	2
$\mathcal{T}_{\text{sun}}$	1	0	$Q_{\text{sun}} - Q_{\text{ds}}$	3
$\mathcal{T}_{\text{noalb}}$	1	0	$Q_{\text{sun}} - Q_{\text{ds}}$	3
$\mathcal{T}_{\text{noalb}}$	0	1	$Q_{\text{IR}} - Q_{\text{ds}}$	4
$\mathcal{T}_{\text{ec}}$	1	1	$Q_{\text{IR}} - Q_{\text{ds}}$	4
$\mathcal{T}_{\text{ec}}$	0	1	$Q_{\text{IR}} - Q_{\text{ds}}$	4
$\mathcal{T}_{\text{sun}}$	0	1	$Q_{\text{alb}} + Q_{\text{IR}} - Q_{\text{ds}}$	5
$\mathcal{T}_{\text{sun}}$	0	0	$-Q_{\text{ds}}$	6
$\mathcal{T}_{\text{noalb}}$	0	0	$-Q_{\text{ds}}$	6
$\mathcal{T}_{\text{ec}}$	1	0	$-Q_{\text{ds}}$	6
$\mathcal{T}_{\text{ec}}$	0	0	$-Q_{\text{ds}}$	6

normal vector to be $n = e_3$ where e_3 denotes the third unity vector in $\mathbb{R}^3$. Any other vector $\tilde{n}$ can be achieved by the choice of a suitable attitude $\tilde{q}$ such that $A(\tilde{q})e_3 = \tilde{n}$.

5.2 Observability Analysis

In order to assess if it is even possible to build an observer which estimates the attitude using only a single temperature measurement, we first have to analyse the observability of the system. This is done first by linearising the system and evaluating the resulting observability matrices. Then, we continue to analyse the observability of the nonlinear system with the methods from Chapter 2.

5.2.1 Analysis of the Linearised System

In order to obtain first insights of the observability properties, we linearise the system (5.2). The linearised system has the form

$$
\begin{aligned}
\dot{T} &= f_T(T^*, q^*, t^*) + \frac{\partial f_T(T^*, q^*, t^*)}{\partial T}(T - T^*) + \frac{\partial f_T(T^*, q^*, t^*)}{\partial q}(q - q^*) + \frac{\partial f_T(T^*, q^*, t^*)}{\partial t}(t - t^*) \\
\dot{q} &= \frac{1}{2}\Omega(\omega^*)q^* + \frac{1}{2}\Omega(\omega^*)(q - q^*) + \frac{1}{2}\Xi(q^*)(\omega - \omega^*) \\
\dot{\omega} &= -J^{-1}(\omega^* \times J\omega^*) - J^{-1}([\omega^*\times]J - [J\omega^*]\times)(\omega - \omega^*) + J^{-1}u \\
\dot{t} &= 1
\end{aligned}
$$

where the analytical expressions for the derivatives of f_T can be found in Appendix A.5. To facilitate the analysis and allow for an equilibrium, we neglect the dependency on

time, i.e. $\frac{\partial f_T(T,q^*,t^*)}{\partial t} = 0$. This allows to analyse the (T, q, ω) dynamics for an equilibrium point $(T^*, q^*, \omega^*, u^*)$. Such an equilibrium always fulfils $u^* = \omega^* = 0$. Then with $\frac{\partial f_T^*}{\partial T} := \frac{\partial f_T(T^*,q^*,t^*)}{\partial T}$ and $\frac{\partial f_T^*}{\partial q} := \frac{\partial f_T(T^*,q^*,t^*)}{\partial q}$ the system takes the form

$$\begin{bmatrix} \dot{T} \\ \dot{q} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} \frac{\partial f_T^*}{\partial T} & \frac{\partial f_T^*}{\partial q} & 0 \\ 0 & 0 & \frac{1}{2}\Xi(q^*) \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} T \\ q \\ \omega \end{bmatrix}.$$

To determine the observability of the system we consider the output T and the observability matrix for the output $y_1 = T$

$$\mathrm{d}\mathcal{O} = \begin{bmatrix} 1 & 0 & 0 \\ \frac{\partial f_T^*}{\partial q} & \frac{\partial f_T^*}{\partial q} & 0 \\ \left(\frac{\partial f_T^*}{\partial q}\right)^2 & \frac{\partial f_T^*}{\partial T}\frac{\partial f_T^*}{\partial q} & \frac{\partial f_T^*}{\partial q}\frac{1}{2}\Xi(q^*) \\ \left(\frac{\partial f_T^*}{\partial q}\right)^3 & \left(\frac{\partial f_T^*}{\partial T}\right)^2\frac{\partial f_T^*}{\partial q} & \frac{\partial f_T^*}{\partial T}\frac{\partial f_T^*}{\partial q}\frac{1}{2}\Xi(q^*) \\ \left(\frac{\partial f_T^*}{\partial q}\right)^4 & \left(\frac{\partial f_T^*}{\partial T}\right)^3\frac{\partial f_T^*}{\partial q} & \left(\frac{\partial f_T^*}{\partial T}\right)^2\frac{\partial f_T^*}{\partial q}\frac{1}{2}\Xi(q^*) \\ \vdots & \vdots & \vdots \end{bmatrix}.$$

The rank of this matrix is three if T^* and $\frac{\partial f_T(q^*,T^*)}{\partial q}$ are non zero elements. For the output of interest $y = h(x)$ this leads to the observability matrix with rank seven. This suggests that a minimum of three temperature sensors is required to fully determine the attitude of the spacecraft. Note however, that the eigenvalues of the quaternion and angular velocity dynamics have real part zero and therefore every equilibrium point is not hyperbolic. Then, Hartman–Grobman theorem [118] does not guarantee the existence of a diffeomorphism between the linearisation and the original dynamics. Additionally, the time dependence of the system was neglected in this analysis which may be justified for geostationary but not low Earth orbits. Therefore, the shown observability analysis for the linearisation does not forbid that the nonlinear system (5.2) + (5.3) is weakly observable as will be shown in the next section.

5.2.2 Analysis of the Nonlinear System

We analyse observability in the spirit of Chapter 2 and for $u = 0$ define the specific observability mapping

$$\mathcal{O}(x) = \left[h_1(x), \mathcal{L}_f h_1(x), \mathcal{L}_f^2 h_1(x), \mathcal{L}_f^3 h_1(x), q^\top q - 1, \omega, t\right]^\top \tag{5.4}$$

where $h_1(x)$ denotes the first entry of $h(x)$, the temperature measurement. The mapping also consists of the first three derivatives with respect to time of the temperature measurements.

The Jacobian of this observability mapping has the form

$$\begin{aligned}\mathrm{d}\mathcal{O}(x) &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ \frac{\partial}{\partial T}\mathcal{L}_f h_1(x) & \frac{\partial}{\partial q}\mathcal{L}_f h_1(x) & 0 & \frac{\partial}{\partial t}\mathcal{L}_f h_1(x) \\ \frac{\partial}{\partial T}\mathcal{L}_f^2 h_1(x) & \frac{\partial}{\partial q}\mathcal{L}_f^2 h_1(x) & \frac{\partial}{\partial \omega}\mathcal{L}_f^2 h_1(x) & \frac{\partial}{\partial t}\mathcal{L}_f^2 h_1(x) \\ \frac{\partial}{\partial T}\mathcal{L}_f^3 h_1(x) & \frac{\partial}{\partial q}\mathcal{L}_f^3 h_1(x) & \frac{\partial}{\partial \omega}\mathcal{L}_f^3 h_1(x) & \frac{\partial}{\partial t}\mathcal{L}_f^3 h_1(x) \\ 0 & 2q^\top & 0 & 0 \\ 0 & 0 & I_3 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix} \\ &=: \Big[\mathrm{d}\mathcal{O}_T(x)\ \mathrm{d}\mathcal{O}_q(x)\ \mathrm{d}\mathcal{O}_\omega(x)\ \mathrm{d}\mathcal{O}_t(x)\Big] \in \mathbb{R}^{9\times 9}\,. \end{aligned} \tag{5.5}$$

The first three rows are implemented as symbolic expressions while the fourth row of this matrix is calculated numerically. If the Jacobian has full rank, the system is called *locally weakly observable* at x. For brevity of exposition, we write *observability* if we refer to local weak observability in this section. It is straightforward to see that the rank of the Jacobian is five plus the rank of $\mathrm{d}\mathcal{O}_q(x)$, with a total maximum of nine. Thus, we determine the rank of $\mathrm{d}\mathcal{O}_q(x)$ to get the rank of the full Jacobian. The rank of $\mathrm{d}\mathcal{O}_q(x)$ is at least one, because $2q^\top$ cannot be identical to zero.

Remark 5. Using even higher derivatives of T may yield additional information for the reconstructions of the states and could theoretically let replace the measurements of the angular velocity. However, the added value by these derivatives might be little due to its bad numerical conditioning. Using additional derivatives of $h_2(x)$, $h_3(x)$, $h_4(x)$, i.e. $q^\top q - 1$, ω or t, does not yield further information.

Remark 6. The observability is analysed for input $u = 0$. At this point it is not clear if u is an universal input or how altering u may change the observability properties. But since the input is occurring only in the third derivative of T, the influence is limited.

Remark 7. There are numerous other ways to analyse the observability of a system, e.g. using a linearisation or the explicit solution of the system, as suggested in [52]. We advocate the proposed approach because it avoids any linearisation or numerical integration, not suitable for non-euclidean spaces as for example quaternion spaces.

Determining the precise set of states x for which the matrix $\mathrm{d}\mathcal{O}_q(x)$ has full rank is not trivial. We relegate feasible analytical proofs to the appendix. In the following, we shall distinguish between three ways to validate the proposed statements. In an analytical proof the statement is analytically simplified until its evidence is clear. Proofs are carried out symbolically, if computer algebra is used to verify and simplify the given mathematical expressions. Finally, if the preceding methods are too costly, the statement is validated numerically. Then the expression is calculated for a number of uniformly distributed states approximating the state space. Table A.1 contains the parameters for the symbolical and numerical analysis.

The observability analysis is split into two parts. The first part covers the special cases, i.e. the parameter configurations in manifolds of small dimensions that admit no observability. In the second part the observability for the remaining configurations is proven and qualitatively evaluated.

5.2.2.1 Non-Observable Configurations

In the previous section we have established the relevant six cases. These cases differ in the right hand side, i.e by the irradiations that are acting on the spacecraft as can be seen in Table 5.1. The orbit described by r is fitted based on real data in which $t = 0$ is in the middle of the eclipse phase. Naturally, the different cases have different observability properties. In this section, for every case we determine the subsets of the domains $\mathbb{X}_1, \ldots, \mathbb{X}_6$ that contain the non-observable states. These are characterised by the Jacobian of the observability mapping not having full rank. As this corresponds to a singular value of zero, the sets of non-observable states are characterised by

$$\mathbb{X}_i^{\mathrm{no}} = \{x \in \mathbb{X}_i \mid \sigma_{\min}(\mathrm{d}\mathcal{O}(x)) = 0\} \tag{5.6}$$

for $i \in \{1, \ldots 6\}$ where $\sigma_{\min}(\cdot)$ maps a matrix to its smallest singular value. We start with the non-observable states for cases that can be stated analytically. For better readability, we relegate the technical derivation to Appendix A.5. Afterwards, we illustrate the remaining non-observable states that can only be found numerically.

Case 1 and Case 2

These are the cases when solar and infrared irradiation are acting. The cases differ in whether albedo irradiation is acting or not. Observability in these two cases is lost if the spacecraft is not rotating, i.e. $\omega = 0$. Then f_q and f_ω become identical to zero. The detailed proof can be found in Appendix A.5.1. Whenever $\omega = a\,n$ with $a \in \mathbb{R}$, the rotation axis aligns with the surface normal. In this case, observability is lost as shown in Appendix A.5.4.

Case 3

In this case only solar irradiation is acting. Thus, the dynamics are not explicitly dependent on the state t. This renders all states non-observable as shown in Appendix A.5.5 and lets the temperature measurement contain even less information for $\omega = 0$ as shown in Appendix A.5.2. Furthermore, Sun pointing also yields less usable information from the temperature data as shown in Appendix A.5.3.

Case 4 and 5

In these cases, infrared irradiation is acting on the spacecraft. They differ in the albedo irradiation. As in Case 1 and 2, observability is lost if the spacecraft is not rotating or the

Table 5.2: Analytical unobservable configurations.

Case	Condition	rank(d$\mathcal{O}$)
1,2	$\omega = 0$	8
1,2	$\omega = a\,n$	8
3	$\omega \neq 0$	8
3	$\omega = 0$	7
3	$A(q)^\top n = \frac{s}{\|s\|}$	7
4,5	$\omega = 0$	8
4,5	$A(q)^\top n = -\frac{r}{\|r\|}$	8
4,5	$\omega = a\,n$	8
6		6

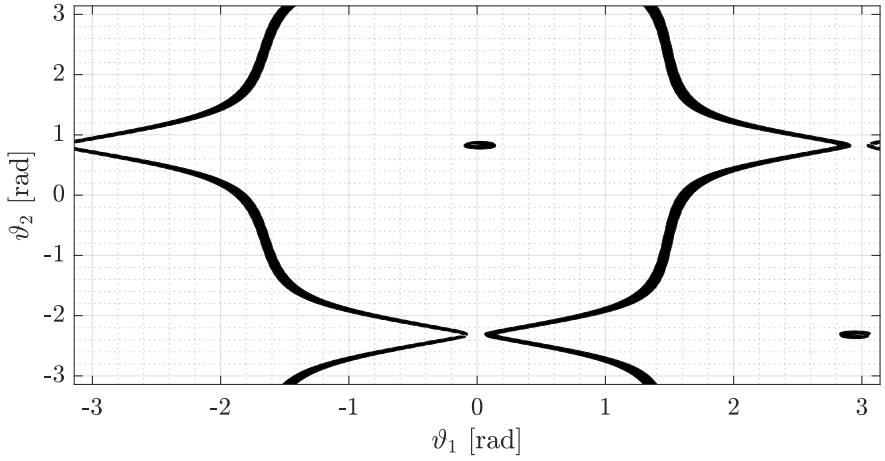


Figure 5.1: Illustration of the non-observable points.

rotation axis aligns with the surface normal, i.e. $\omega = 0$ or $\omega = a\,n$. Additionally, observability is lost if the spacecraft is Earth pointing, i.e. $A(q)n = \frac{r}{\|r\|}$ as shown in Appendix A.5.3.

Case 6

Here no irradiation is acting on the spacecraft, so no information can be gained from the temperature sensor to estimate the attitude as shown in Appendix A.5.6.

The discussed special cases in which the system is not weakly observable are summarised in Table 5.2.

There are also other non-observable states that cannot be given in a simple analytical expression. These points can be found by the use of extensive gridding or finding all solutions

of the optimisation problem

$$\min_{x} \quad \sigma_{\min}(\mathrm{d}\mathcal{O}(x)), \tag{5.7}$$

to be solved for multiple initial values. The numerical analysis has shown that the set of non-observable states is of at least one dimension less than the state space. More precisely, this means that for every state (T, ω, t) considered, exists a subset of $\mathcal{Q}$ of dimension two which contains all the non-observable points. We illustrate this with Case 4. The non-observable points are displayed in ϑ_1, ϑ_2 coordinates which result from a transformation discussed in detail in Section 5.3.3. In short, the section introduces for every q the angles

$$\begin{aligned}
\theta &= \arccos(r^\top A(q)^\top n) \\
\vartheta_1 &= \operatorname{atan2}((r \times n^{\theta,0})^\top n^q, {n^{\theta,0}}^\top n^q - (r^\top n^{\theta,0})^2) \\
\vartheta_2 &= \begin{cases} +2\arccos(\bar{q}_4), & \text{if } \frac{\bar{q}_{1:3}}{\|\bar{q}_{1:3}\|} = n^q \\ -2\arccos(\bar{q}_4), & \text{else} \end{cases}
\end{aligned}$$

with the two rotation axes $n^q = A(q)^\top n$, $n^{\theta,0} = A\left(\frac{e_i \times r}{\|e_i \times r\|}, \theta\right)^\top r$ and the quaternion which describes the rotation around the n^q vector as $\bar{q} = q^{-1}(r, \vartheta_1) \otimes q^{-1}\left(\frac{e_i \times r}{\|e_i \times r\|}, \theta\right) \otimes q^{-1}(v, \phi) \otimes q$ with the rotation axis $v = \frac{n \times r}{\|n \times r\|}$ and angle $\phi = \arccos(n^\top r)$. This allows to illustrate all non-observable points for the state $(T, \theta, \omega^\top, t) = (292, \frac{\pi}{2}, 0.0058, 0.0058, 0.0058, 100)$ in Fig. 5.1. The black lines are overapproximations of the non-observable points. They are obtained using detailed gridding which leads to different thicknesses of the lines. It can be seen that the resulting set separates the ϑ_1-ϑ_2 space into four connected spaces. The set is symmetric with respect to a point close to $(0.07, 0.84)$ and some neighbourhoods of the points $(0, -2)$ and $(2.9, 0.84)$ are not part of the set and separate the set. Interesting are also the sets at $(0.07, 0.84)$ and $(3, -2)$ which are both shaped similar to an annulus with non-trivial inner area. This shows that the set of non-observable points is not simple-shaped. Even though the majority of the points are observable, as illustrated in the next section, the set $\mathbb{X}_i^{\mathrm{no}}$ plays an important role in observer design.

5.2.2.2 Observable Configurations

We calculate the singular values of matrix $\mathrm{d}\mathcal{O}_q(x)$ to analyse if the state x is observable. If four of the singular values are not identical to zero, then the Jacobian has full rank and the system is observable. The analysis is carried out for a number of uniformly distributed states (T, q, ω, t). The temperature T takes values between 0 and 500 Kelvin and each individual component of the angular velocity ω_i is between 0 and 0.1. The quaternions are obtained by points that are uniformly distributed within a unit sphere and then projected on its boundary. This way of determining the states allows to cover the state space in an unbiased way. If the system is observable for all points generated with this method, this indicates that the system is observable in practical applications. However, this method generates only points of manifolds that have the same dimension as the state space. In particular, most of the special cases discussed in the previous section hold for states in manifolds of

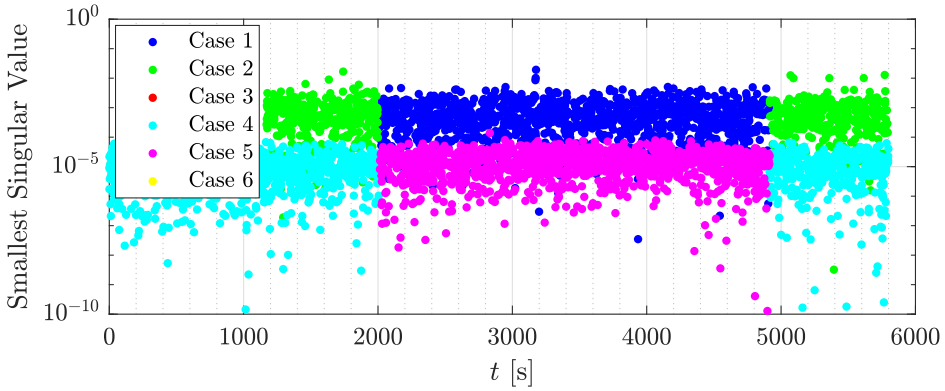


Figure 5.2: Smallest singular value of $\mathrm{d}\mathcal{O}_q$.

smaller dimension. Thus, points of these special cases may not occur in this analysis. Fig. 5.2 shows the smallest singular value in a semi-logarithmic scale over time. We notice that this singular value differs significantly between the six cases. The remaining cases are delimited horizontally due to the time dependency of the cases. Cases 1 and 2, and 4 and 5 have pairwise singular values with similar mean. Cases 1 and 2 have consistently the highest singular values.

Table 5.3 displays the mean of the calculated singular values. The singular values confirm the result of the previous section. Cases 3 and 6 are never observable and as such have always a singular value identical to zero. Cases 1, 2, 4 and 5 do not have a singular value identical to zero but the values are close if the specified point is near the unobservable manifold specified in the previous section. The highest singular value is identical to two in all cases. This is due to the term $2q^{\top}$ in $\mathrm{d}\mathcal{O}_q$ and the quaternion to be of norm one. We can see in all cases that the mean of singular values differs about one to two decades to the next singular value. Furthermore, the mean singular values of Cases 1–3 are larger than of the ones of Cases 4–6. In particular, the smallest singular value is more than a decade smaller. This can be ascribed to the solar irradiation which is not acting in Cases 4–6. The results reveal the best observability properties when the solar irradiation and at least one time variant irradiation is acting on the spacecraft, i.e. Case 1 or 2.

All in all, observability of this system is a delicate matter. While the majority of the states is observable there are some connected manifolds that contain unobservable states. These manifolds may become a problem if the to be observed trajectory crosses it or the observer crosses it during its transient phase. These issues will be illustrated and discussed in the next section where we design a nonlinear observer in transformed coordinates.

5.3 Observer in Transformed Coordinates

The previous section has shown for which state configurations the system (5.2) is locally observable by determining the Jacobian of the observability mapping (5.4). Consequently, this observability mapping can be used in order to transform the system into a normal form

Table 5.3: Mean of singular values.

Case	Largest SV	Second SV	Third SV	Smallest SV
1	2	1.0e-1	1.3e-2	7.6e-4
2	2	1.3e-1	2.6e-2	8.3e-4
3	2	1.3e-1	1.5e-2	0
4	2	0.5e-1	0.6e-2	0.1e-4
5	2	0.6e-1	0.8e-2	0.2e-4
6	2	0	0	0

as discussed in Section 2.3. The observability mapping $\mathcal{O}$ with its domain and image has the form

$$\begin{aligned}\mathcal{O} &: \mathbb{R}^+ \times \mathbb{S}_3 \times \mathbb{R}^3 \times \mathbb{R}^+ \to \mathbb{R}^+ \times \mathbb{R} \times \mathbb{R} \times \mathbb{R} \times \{0\} \times \mathbb{R}^3 \times \mathbb{R}^+ \\ &(T, q, \omega, t) = x \mapsto z = \left[T\ \mathcal{L}_f^1 h_1(x)\ \mathcal{L}_f^2 h_1(x)\ \mathcal{L}_f^3 h_1(x) + \mathcal{L}_g \mathcal{L}_f^2 h_1(x) u\ q^\top q - 1\ \omega\ t\right]^\top\end{aligned} \tag{5.8}$$

where $z_{1:4}$ are the temperature and its first three derivatives, z_5 is the quaternion constraint, $z_{6:8}$ the angular velocity and z_9 the time. The resulting dynamics have the form

$$\dot{z} = \begin{bmatrix} 0_{3,1} & I_{3,8} \\ 0_{1,1} & 0_{1,8} \\ 0_{1,1} & 0_{1,8} \\ 0_{3,1} & 0_{3,8} \\ 0_{1,1} & 0_{1,7} \end{bmatrix} z + \begin{bmatrix} 0_{3,1} \\ \varphi_4(z) \\ \varphi_5(z) \\ \varphi_{6:8}(z_{6:8}) \\ \varphi_9(z_9) \end{bmatrix} + \begin{bmatrix} 0_{3,3} \\ \phi_5(z) \\ 0_{3,3} \\ I_{3,1} \\ 0_{3,3} \end{bmatrix} u \tag{5.9a}$$

$$y = \begin{bmatrix} z_1 \\ z_5 \\ z_{6:8} \\ z_9 \end{bmatrix} \tag{5.9b}$$

where

$$\varphi_4(z) = \mathcal{L}_f^{(4)} h_1(x)|_{x=\mathcal{O}^{-1}(z)} \tag{5.10}$$

$$\varphi_5(z) = 2q|_{x=\tau^{-1}(z)} \tag{5.11}$$

$$\varphi_{6:8}(z) = f_\omega(z_{6:8}) \tag{5.12}$$

$$\varphi_9(z) = 1 \tag{5.13}$$

$$\phi_5(z) = \mathcal{L}_g \mathcal{L}_f^{(3)} h_1(x)|_{x=\mathcal{O}^{-1}(z)} \tag{5.14}$$

and $\mathcal{O}^{-1}$ denotes the local inverse of $\mathcal{O}$ at the point x where x is a state for which the local inverse exists as analysed in the previous section. No global inverse can exist as there are non-observable states as shown in the previous section. We are interested in obtaining more insight into the inverse of $\mathcal{O}$. The system (5.9) has six outputs, each forming a differentiator subsystem where the unknown nonlinearity does only occur in the final derivative. The nonlinearities φ_3, φ_4 are known and functions of the output while the nonlinearities φ_1, φ_2

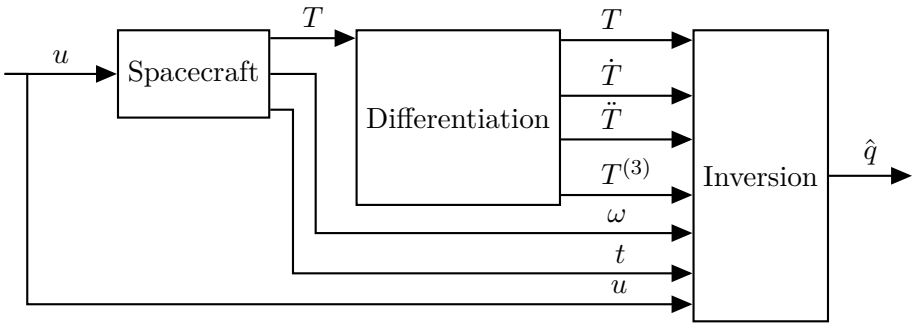


Figure 5.3: Attitude Reconstruction Using Differentiation and Inversion.

and φ_5 can only be numerically approximated and are functions of the approximated states (not only the measured output). This makes an observer of the form

$$\dot{\hat{z}} = \begin{bmatrix} 0_{3,1} & I_{3,8} \\ 0_{1,1} & 0_{1,8} \\ 0_{1,1} & 0_{1,8} \\ 0_{3,1} & 0_{3,8} \\ 0_{1,1} & 0_{1,7} \end{bmatrix} \hat{z} + \begin{bmatrix} 0 \\ 0 \\ 0 \\ \varphi_{6:8}(\hat{z}_{6:8}) \\ \varphi_9(\hat{z}_9) \end{bmatrix} + \begin{bmatrix} L_{1:3}(z_1 - \hat{z}_1) \\ L_4(z_1 - \hat{z}_1) \\ L_5(z_5 - \hat{z}_5) \\ L_{6:8}(z_{6:8} - \hat{z}_{6:8}) \\ L_9(z_9 - \hat{z}_9) \end{bmatrix} + \begin{bmatrix} 0 \\ \phi_5(\hat{z}) \\ 0 \\ I_{3,1} \\ 0 \end{bmatrix} u \tag{5.15a}$$

$$\hat{y} = \begin{bmatrix} \hat{z}_1 \\ \hat{z}_5 \\ \hat{z}_{6:8} \\ \hat{z}_9 \end{bmatrix} \tag{5.15b}$$

a canonical choice for the estimation of z where L can be a matrix or a nonlinear function. Choosing $L_{5:9}$ is a simple task in view of stability. The main task is choosing $L_{1:4}$ which can be considered as finding a suitable differentiation algorithm to differentiate z_1. This topic is discussed in detail in Section 2.2 and its application can be found in the next section. Finally, when a good estimate of z is found the transformation into the original coordinates must be performed using

$$\hat{x} = \mathcal{O}^{-1}(\hat{z}). \tag{5.16}$$

This is discussed in Section 5.3.2. These two steps are illustrated in the block diagram 5.3.

5.3.1 Differentiator Design

In this section, we discuss the established differentiation methods discussed in Section 2.2 and compare their results for our system of interest

$$\dot{z}_{1:4} = \begin{pmatrix} 0_{3,1} & I_{3,8} \\ 0_{1,1} & 0_{1,8} \end{pmatrix} z_{1_4} + \begin{pmatrix} 0 \\ \tilde{\varphi}_4(t) \end{pmatrix} \tag{5.17}$$

resulting from (5.9) where z and u on the right-hand side have been replaced by their solutions which are functions of time, i.e. $\tilde{\varphi}_4(t) = \varphi_4(z(t)) + \phi_4(z(t)) + u(t)$.

We simulate the three presented differentiation methods, linear differentiator, high gain differentiator and sliding mode differentiator for a temperature signal produced by a satellite fulfilling equation (5.17) with the initial conditions $T(0) = 220$, $q(0) = l^{r,n}(\frac{\pi}{2}, 0, 0)$ and $\omega(0) = 0.01\frac{[1\ 1\ 1]}{\|[1\ 1\ 1]\|}$.

The linear differentiator 2.2.2 is chosen to be of order $n = 3$. Its parameters $(a_0 \dots a_{n-1})$ are the ones of a Bessel filter with cutoff frequency $\omega_0 = 0.6$.

The High Gain differentiator 2.2.3 is chosen to be of order $n = 4$. Its states are the estimated derivatives of the temperature of interest. Its parameters $(\alpha_0 \dots \alpha_{n-1})$ are also of the same form as for a Bessel Filter.

The Sliding Mode differentiator proposed in Section 2.2.4 is very sensitive to the choice of the integrator as well as the step size. Good results for these simulations are obtained using a simple Euler integration and a step size of 1.5×10^{-5}. Note, however, that commercial temperature sensors are at the moment not able to realise such a small step size as this would require much higher sensor accuracy than currently possible. Furthermore, the sliding mode differentiator exhibits chattering if the relation between function value and step size is too small. This chattering can be attenuated in lower derivatives by using a differentiator of higher order than necessary. Thus, the sliding mode differentiator is chosen to be of order $n = 6$. Its parameters are the result of the optimisation

$$\int_{t_0}^{t_{\text{end}}} |T^{(3)} - \hat{T}^{(3)}| \mathrm{d}t$$

where the desired settling time is chosen to be $t_0 = 20$ and the simulated time is $t_{\text{end}} = 40$. The optimisation yields the parameters $L = 0.89146$ and $\lambda = [56\ 1.4\ 0.78\ 0.35\ 0.12\ 0.005]$. Note that these parameters differ greatly from the parameters recommended in the literature due to being optimised for this specific scenario. Further note, that the parameters of all three differentiators are chosen to have a similar settling time.

Figure (5.4) presents the simulated derivatives. It can be seen that all three differentiators give reasonable estimations. The linear differentiator admits the highest delay for estimating T. Both linear differentiators converge to the analytical derivatives within a very small error bound. The first two derivatives of the sliding mode differentiator have a very small stationary error as well. However, the estimate of the third derivative chatters around the real value despite the parameters being the result of an optimisation to reduce this chattering. This makes the sliding mode differentiator the least appropriate choice to estimate these derivatives of very small magnitude.

5.3.2 The Inversion Problem

In this section, we consider the problem of finding the inverse $\mathcal{O}^{-1}$ of Equation (5.8) or equivalently, we want to find a way to obtain the estimated states $(\hat{T}, \hat{q}, \hat{\omega}, \hat{t})$ through the estimated derivatives of the temperature. The measured temperature is described by $z_1 = y_1$.

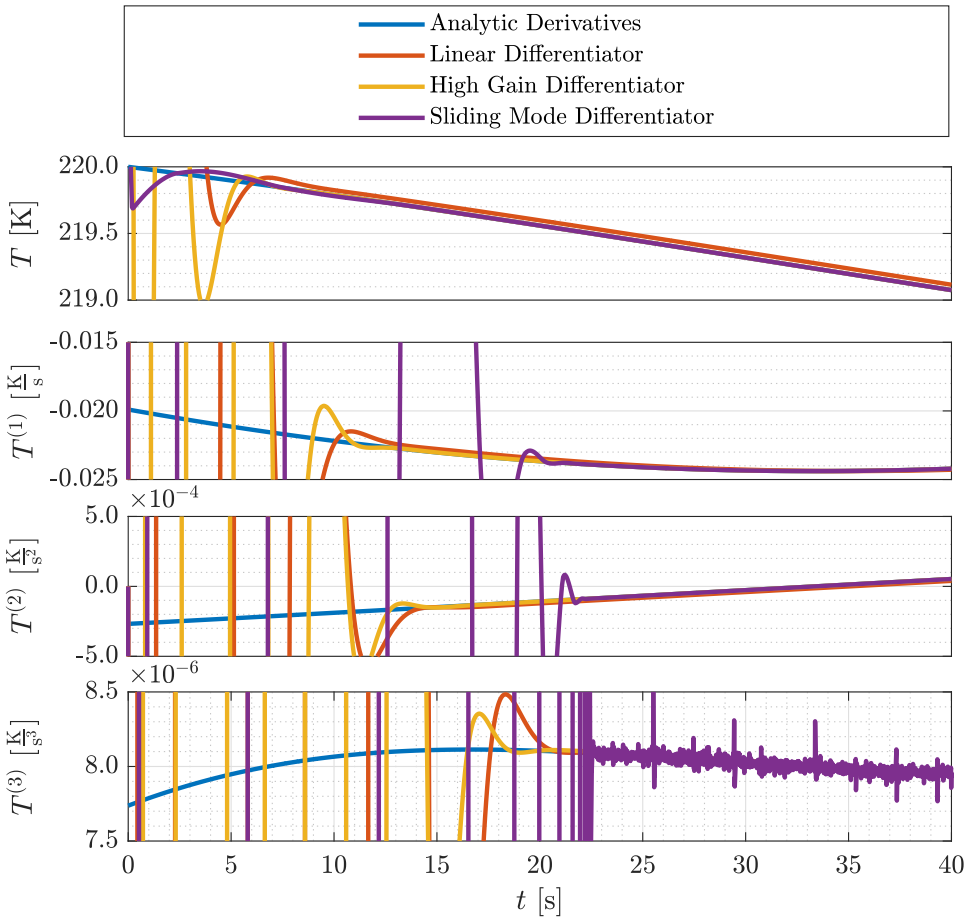


Figure 5.4: Temperature signal differentiated with three different differentiators.

The output $z_5 = y_2$ is the quaternion constraint. The output $z_{6:8} = y_3$ is the angular velocity and $z_9 = y_4$ is the time t. We use this to define the functions $L_1^{T,\omega,t}(q) := \mathcal{L}_f^{(1)}h(x)$, $L_2^{T,\omega,t}(q) := \mathcal{L}_f^{(2)}h(x)$ and $L_3^{T,\omega,t}(q) := \mathcal{L}_f^{(3)}h(x)$ with $T = y_1$, $\omega = y_3$ and $t = y_4$. Then we can remove the trivial parts of Equation (5.8) and obtain the function of interest

$$\begin{aligned}\mathcal{O}_{2:5}^{T,\omega,t} : \mathbb{S}_3 &\to \mathbb{R}\times\mathbb{R}\times\mathbb{R}\times\{0\}\\ q &\to (z_2, z_3, z_4, z_5)\end{aligned}$$

where the image is defined by the the nonlinear algebraic equation system

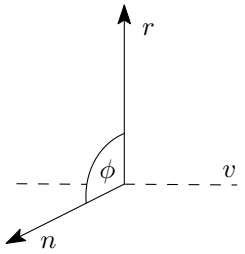
$$z_2 = L_1^{T,\omega,t}(q) \tag{5.18a}$$
$$z_3 = L_2^{T,\omega,t}(q) \tag{5.18b}$$
$$z_4 = L_3^{T,\omega,t}(q) \tag{5.18c}$$
$$z_5 = q^\top q - 1\,. \tag{5.18d}$$

The full rank of the Jacobian of the observability mapping guarantees by the implicit function theorem the existence and local uniqueness of q, i.e. there exists a vicinity U of (z_2, z_3, z_4, z_5) such that there exists a unique continuously differentiable function $\mathcal{O}_{2:5}^{-1}$ such that Equation (5.18) is fulfilled with $q = \mathcal{O}_{2:5}^{-1}(z_2, z_3, z_4, z_5)$. However, this does not guarantee the global uniqueness of q nor that this q is found by an algorithm for arbitrary initial values. Furthermore, due to the non-euclidean structure of the quaternion space, it is desirable to introduce an intermediate transformation l_1 which guarantees an euclidean domain. This transformation is introduced in the following section by a new set of angles $(\theta, \vartheta_1, \vartheta_2)$ where the angle θ describes the influence of one irradiation on the temperature.

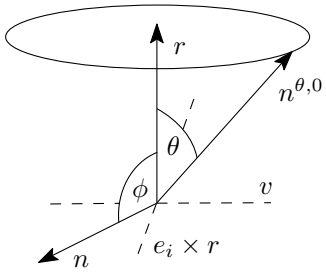
5.3.3 Irradiation Angle Transformation

We derive the new attitude representation via the two mappings $l_1^{r,n} : q \mapsto (\theta, \vartheta_1, \vartheta_2)$ and $l_2^{r,n} : (\theta, \vartheta_1, \vartheta_2) \mapsto q$. The superscripts shall suggest that these mappings change dependent on the vectors r and n. We define the mappings for normalised $r, n \in \mathbb{S}_2$. Throughout this section the vectors r, n are always normalised. For not normalised $\bar{r}, \bar{n} \in \mathbb{R}^3$, the mapping $l_1^{\bar{r},\bar{n}}$ is defined by the mapping $l_1^{r,n}$ with $r = \frac{\bar{r}}{\|\bar{r}\|}$ and $n = \frac{\bar{n}}{\|\bar{n}\|}$. As motivated in the previous section, this representation shall incorporate θ as its first variable, which is the angle between two normalised vectors $r, n \in \mathbb{S}_2$ that form a linear space of dimension two. Since the attitude space is of dimension three, it is natural to define two additional angles ϑ_1, ϑ_2 for allowing to describe every attitude by the three angles $(\theta, \vartheta_1, \vartheta_2)$. For a reference attitude representation, we use quaternions. Thus, in order to show that the three angles $(\theta, \vartheta_1, \vartheta_2)$ are a valid attitude representation, we need to show that for every quaternion q there exists one unique angle triple $(\theta, \vartheta_1, \vartheta_2)$. Then the corresponding mapping $l_1^{r,n}$ is bijective and an inverse mapping $l_2^{r,n}$ exists. Additionally, these mappings need to fulfil certain continuity conditions to make the attitude representation usable.

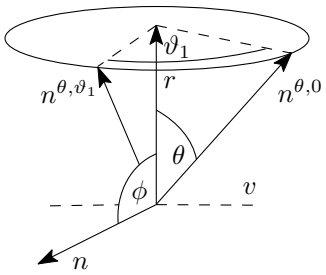
We start by deriving $l_2^{r,n} : (\theta, \vartheta_1, \vartheta_2) \mapsto q$, i.e. the function that maps every angle set $(\theta, \vartheta_1, \vartheta_2)$ onto its quaternion $q \in \mathbb{S}_3$. These steps are illustrated in Figure 5.5 and define



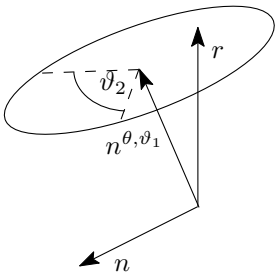
Step 1: Rotate with ϕ around v.
This rotates n on r.



Step 2: Rotate with θ around $e_i \times r$.
This rotates r on $n^{\theta,0}$.



Step 3: Rotate with ϑ_1 around r.
This rotates $n^{\theta,0}$ on n^{θ,ϑ_1}



Step 4: Rotate with ϑ_2 around n^{θ,ϑ_1}.
The result is the desired attitude

Figure 5.5: Level sets of the three angles.

the function $l_2^{r,n}$. First, the rotation axis $v := \frac{n \times r}{\|n \times r\|}$ and the angle $\phi := \arccos(n^\top r)$ are used to rotate n on r, i.e.

$$r = A^\top(v, \phi)\, n\,. \tag{5.19}$$

Then, a rotation is carried out to achieve the desired angle θ. Define the vector

$$n^{\theta,0} := A^\top\left(e_i \times r, \theta\right) A^\top(v, \phi)\, n \tag{5.20}$$

where e_i is the first unit vector if $r \notin \operatorname{span}\{e_1\}$, otherwise it is the second unit vector. By definition, it is clear that $r^\top n^{\theta,0} = \cos(\theta)$. The vector n^{θ,ϑ_1} is defined by a rotation of ϑ_1 around the r axis, hence

$$n^{\theta,\vartheta_1} := A^\top(r, \vartheta_1)\, n^{\theta,0}\,. \tag{5.21}$$

Finally, rotation of ϑ_2 around n^{θ,ϑ_1} gives the desired attitude q, i.e.

$$q := q(v, \phi) \otimes q\left(e_i \times r, \theta\right) \otimes q(r, \vartheta_1) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2)\,. \tag{5.22}$$

Definition 5.1. *For two vectors $r, n \in \mathbb{S}_2$ the attitude defined by the angles θ, ϑ_1 and ϑ_2 is represented as a quaternion defined by the mapping $l_2^{r,n}$ which has the form*

$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) := q(v, \phi) \otimes q(e_i \times r, \theta) \otimes q(r, \vartheta_1) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2)$$

with the rotation axis and rotation angle $v := \frac{n \times r}{\|n \times r\|}$, $\phi := \arccos(n^\top r)$ *and rotation axis* $n^{\theta,\vartheta_1} := A^\top(r, \vartheta_1) A^\top(e_i \times r, \theta)\, r$.

Note that we have so far deliberately not given the domain and the image of $l_2^{r,n}$. Clearly, the image is a subset of the quaternion space because the definition of $l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$ is merely a result of a quaternion product which retains the norm condition. The domain is obtained when deriving the inverse transformation. Recall that the original motivation for these three angles is that the first angle θ is the angle between r and n. This property and a helpful equation for the rotation axis n^{θ,ϑ_1} are shown in the following lemma.

Lemma 5.2. *The quaternion* $l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$ *described by the rotation resulting from the three angles* $\theta, \vartheta_1, \vartheta_2$ *leads to an angle* θ *between* r *and* n*, i.e.*

$$r^\top A^\top(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))\, n = \cos(\theta) \tag{5.23}$$

where

$$A^\top(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))\, n = n^{\theta,\vartheta_1}.$$

Proof. We obtain with $q = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$ that

$$\begin{aligned}
A^\top(q)\, n &\overset{(A.20)}{=} A^\top(n^{\theta,\vartheta_1}, \vartheta_2) A^\top(r, \vartheta_1) A^\top(e_i \times r, \theta) A^\top(v, \phi)\, n \\
&\overset{(5.19)}{=} A^\top(n^{\theta,\vartheta_1}, \vartheta_2) A^\top(r, \vartheta_1) A^\top(e_i \times r, \theta)\; r \\
&\overset{(5.20)}{=} A^\top(n^{\theta,\vartheta_1}, \vartheta_2) A^\top(r, \vartheta_1)\, n^{\theta,0} \\
&\overset{(5.21)}{=} A^\top(n^{\theta,\vartheta_1}, \vartheta_2)\, n^{\theta,\vartheta_1} \\
&= n^{\theta,\vartheta_1}\,.
\end{aligned}$$

By using the definition of n^{θ,ϑ_1} we obtain

$$\begin{aligned}
r^\top A^\top(q) n = r^\top n^{\theta,\vartheta_1} &\overset{(5.21)}{=} r^\top A^\top(r, \vartheta_1) n^{\theta,0} \\
&= r^\top n^{\theta,0} \overset{(5.20)}{=} \cos(\theta)
\end{aligned}$$

which finishes the proof. □

It remains to show that for every attitude q, there is one unique angle set $(\theta, \vartheta_1, \vartheta_2)$. We do so by introducing $l_1^{r,n}$ and show that it is the inverse mapping of $l_2^{r,n}$.

For a quaternion q we define the angle θ along (5.23) as

$$\theta := \arccos(r^\top A^\top(q)\, n)\,.$$

Consequently, the domain of θ is chosen to be $[0, \pi]$. It is clear that for $n^q := A^\top(q)\, n$ the equations

$$r^\top n^{\theta,0} = r^\top n^q \tag{5.24a}$$

$$n^{\theta,0\top} n^{\theta,0} = n^{q\top} n^q \tag{5.24b}$$

hold. Now we calculate the angle ϑ_1 rotating $n^{\theta,0}$ on n^q around the rotation axis r, i.e. $n^q = A^\top(r, \vartheta_1)\, n^{\theta,0}$. Note that this equation is equivalent to (5.21). Since (5.24) holds, the angle ϑ_1 can be obtained using Rodrigues' formula

$$\vartheta_1 := \operatorname{atan2}\left((r \times n^{\theta,0})^\top n^q, n^{\theta,0\top} n^q - (r^\top n^{\theta,0})^2\right).$$

Then ϑ_1 is obtained using the atan2 function which maps the two arguments to an angle in $(-\pi, \pi]$ and defines the domain of ϑ_1. Note that for $\theta = 0$ and $\theta = \pi$, both arguments are identical to zero which is where atan2 is undefined. This singularity will be discussed in detail in the next section. For our application, we define atan2 to be identical to zero to allow a definition on complete $[0, \pi]$. For calculating ϑ_2, we use (5.22) and define

$$\bar{q} := q^{-1}(r, \vartheta_1) \otimes q^{-1}\left(e_i \times r, \theta\right) \otimes q^{-1}(v, \phi) \otimes q\,.$$

It can be seen that this quaternion describes a rotation around the $\pm n^q$ axis since

$$\begin{aligned} A^\top(\bar{q})\, n^q &\overset{(A.20)}{=} A^\top(q) A^\top(q^{-1}(v,\phi)) A^\top\left(q^{-1}(e_i \times r, \theta)\right) n^{\theta,0} \\ &\overset{(5.20)}{=} A^\top(q) A^\top(q^{-1}(v,\phi))\, r \\ &\overset{(5.19)}{=} A^\top(q)\, n \\ &= n^q\,. \end{aligned}$$

This makes the first three entries of the quaternion describe the n^q axis, i.e. $\frac{\bar{q}_{1:3}}{\|\bar{q}_{1:3}\|} = \pm n^q$ or the unit quaternion and $\bar{q}_{1:3} = 0$. The angle ϑ_2 results from

$$\vartheta_2 := \begin{cases} +2\arccos(\bar{q}_4), & \text{if } \frac{\bar{q}_{1:3}}{\|\bar{q}_{1:3}\|} \in \{0, n^q\} \\ -2\arccos(\bar{q}_4), & \text{if } \frac{\bar{q}_{1:3}}{\|\bar{q}_{1:3}\|} = -n^q \end{cases}.$$

Clearly, the result of this definition is an angle $\vartheta_2 \in (-2\pi, 2\pi]$ which is chosen to be the domain of ϑ_2. These steps define the mapping $l_1^{r,n}$.

Definition 5.3. *For two vectors $r, n \in \mathbb{S}_2$ the attitude given by q is transformed into the three angles using the mapping $l_1^{r,n} : \mathbb{S}_3 \to [0, \pi] \times (-\pi, \pi] \times (-2\pi, 2\pi]$ where $l_1^{r,n}(q) = (\theta, \vartheta_1, \vartheta_2)$*

is defined by

$$\begin{aligned}\theta &:= \arccos(r^\top A^\top(q)n)\\ \vartheta_1 &:= \operatorname{atan2}((r \times n^{\theta,0})^\top n^q, {n^{\theta,0}}^\top n^q - (r^\top n^{\theta,0})^2)\\ \vartheta_2 &:= \begin{cases} +2\arccos(\bar{q}_4), & \text{if } \frac{\bar{q}_{1:3}}{\|\bar{q}_{1:3}\|} = n^q \\ -2\arccos(\bar{q}_4), & \text{otherwise} \end{cases}\end{aligned}$$

with the rotation axes $n^q := A(q)^\top n$, $n^{\theta,0} := A^\top(e_i \times r, \theta)\, r$ *and the quaternion describing the rotation around the* n^q *axis as* $\bar{q} := q^{-1}(r, \vartheta_1) \otimes q^{-1}(e_i \times r, \theta) \otimes q^{-1}(v, \phi) \otimes q$ *with rotation axis and angle* $v := \frac{n \times r}{\|n \times r\|}, \phi := \arccos(n^\top r)$.

This allows to state the main theorem of the section.

Theorem 5.4. *Consider the mappings*

$$\begin{aligned} l_1^{r,n} : \mathbb{S}_3 &\to [0, \pi] \times (-\pi, \pi] \times (-2\pi, 2\pi] \\ q &\mapsto l_1^{r,n}(q) \end{aligned}$$

and

$$\begin{aligned} l_2^{r,n} : [0, \pi] \times (-\pi, \pi] \times (-2\pi, 2\pi] &\to \mathbb{S}_3 \\ (\theta, \vartheta_1, \vartheta_2) &\mapsto l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) \end{aligned}$$

as defined in Definitions 5.3 and 5.1. Then $l_1^{r,n}$ *is the inverse of* $l_2^{r,n}$ *for* $\theta \notin \{0, \pi\}$ *and vice versa. The first argument of* $l_2^{r,n}$ *describes the cosine of the angle between* r *and* n*, i.e.*

$$r^\top A^\top(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))\, n = \cos(\theta)\,. \tag{5.25}$$

Proof. By the transparent way of defining the functions $l_1^{r,n}$ and $l_2^{r,n}$ it can be seen that they are corresponding inverses for θ not identical to zero or π. □

5.3.4 Properties of the Irradiation Angle Transformation

In this section we discuss some of the properties of the attitude representation and their mappings.

5.3.4.1 Underlying System of Equations

First, we summarise the three equations that are the basis for the two mappings.

Remark 8. The attitude representation is based on the three equations

$$\begin{aligned} r^\top A^\top(q)\, n &= \cos(\theta) \\ A^\top(q)\, n &= A^\top(r,\vartheta_1) A^\top(e_i \times r, \theta)\, r \\ q &= q(v,\phi) \otimes q\,(e_i \times r, \theta) \otimes q(r,\vartheta_1) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2) \end{aligned}$$

with the rotation axis and angles as in Definition 5.3.

5.3.4.2 Singularities

It is well known that the attitude space is of dimension three. Attitude representations that use only three variables come with the disadvantage of having singularities while representations with more than three variables have additional constraints. The proposed attitude representation has only three variables, but a singularity occurs when the angle θ is identical to zero or π. In this case, ϑ_1 and ϑ_2 have the same rotation axis, namely r.

Proposition 5.5. *For $\theta = 0$ and $\theta = \pi$ it holds*

$$l_2^{r,n}(0, \vartheta_1 - \vartheta_2, \vartheta_2) = l_2^{r,n}(0, \vartheta_1, 0) \tag{5.26}$$
$$l_2^{r,n}(\pi, \vartheta_1 + \vartheta_2, \vartheta_2) = l_2^{r,n}(\pi, \vartheta_1, 0)\,. \tag{5.27}$$

For $l_1^{r,n}$, we obtain

$$\begin{aligned} l_1^{r,n}(l_2^{r,n}(0, \vartheta_1 - \vartheta_2, \vartheta_2)) &= (0, 0, \vartheta_1) \\ l_1^{r,n}(l_2^{r,n}(\pi, \vartheta_1 + \vartheta_2, \vartheta_2)) &= (\pi, 0, -\vartheta_1)\,. \end{aligned}$$

Proof. For $\theta = 0$ the rotation axis n^{θ,ϑ_1} is identical to r. Thus, we obtain

$$\begin{aligned} l_2^{r,n}(0, \vartheta_1 - \vartheta_2, \vartheta_2) &\overset{(5.22)}{=} q(v,\phi) \otimes q(r, \vartheta_1 - \vartheta_2) \otimes q(r, \vartheta_2) \\ &= q(v,\phi) \otimes q(r, \vartheta_1) \otimes q(r, 0) \\ &= l_2^{r,n}(0, \vartheta_1, 0)\,. \end{aligned}$$

The second and third equation follow from the fact that $n^q = n^{\theta,0} = r$ and that we have defined atan2 to be zero for two arguments identical to zero. The same calculations can be made for $\theta = \pi$ and $-r$. □

Note that these singularities make $l_1^{r,n}$ not surjective and $l_2^{r,n}$ not injective which is why invertibility was only stated for the sets without these singularities in the previous section.

5.3.4.3 Augmentation of the Mappings

The mappings $l_1^{r,n}$ and $l_2^{r,n}$ can be augmented easily for $r \in \operatorname{span}\{n\}$. In this case, we define any vector which is perpendicular to r as $v = e_i \times r$ with any unity vector $e_i \notin \operatorname{span}\{n\}$ and proceed as usual.

For reconstruction purposes, it might be beneficial to expand the domain of $l_2^{r,n}$. This can avoid discontinuities in the attitude representation. In this case, the periodicity of the functions allows distinguishing equivalent attitude representations.

Proposition 5.6. *If we extend the domain of l_2 to $\mathbb{R} \setminus \pi\mathbb{N} \times \mathbb{R} \times \mathbb{R}$, we obtain the negative quaternions as*

$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = -l_2^{r,n}(\theta + 2\pi, \vartheta_1, \vartheta_2) \tag{5.28a}$$
$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = -l_2^{r,n}(\theta, \vartheta_1 - 2\pi, \vartheta_2) \tag{5.28b}$$
$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = -l_2^{r,n}(\theta, \vartheta_1, \vartheta_2 - 2\pi) \tag{5.28c}$$

while also the identities

$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = l_2^{r,n}(-\theta, \vartheta_1 + \pi, \vartheta_2 - \pi) \tag{5.29a}$$
$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = l_2^{r,n}(\theta, \vartheta_1 - 2\pi, \vartheta_2 - 2\pi) \tag{5.29b}$$
$$l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2 - 4\pi) \tag{5.29c}$$

hold.

Proof. Equation (5.28a) holds because the definition $n^{\theta+2\pi,\vartheta_1} := A^\top(r, \vartheta_1)A^\top(e_i \times r, \theta + 2\pi)r$ gives

$$n^{\theta,\vartheta_1} = A^\top(r, \vartheta_1)A^\top(e_i \times r, \theta)r = A^\top(r, \vartheta_1)A^\top(e_i \times r, \theta + 2\pi)r = n^{\theta+2\pi,\vartheta_1}.$$

Then the identity

$$\begin{aligned} l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) &= q(v, \phi) \otimes q(e_i \times r, \theta) \otimes q(r, \vartheta_1) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2) \\ &= -q(v, \phi) \otimes q(e_i \times r, \theta + 2\pi) \otimes q(r, \vartheta_1) \otimes q(n^{\theta+2\pi,\vartheta_1}, \vartheta_2) \\ &= -l_2^{r,n}(\theta + 2\pi, \vartheta_1, \vartheta_2) \end{aligned}$$

follows.

In order to prove the identity (5.29a), we define $n^{-\theta,\vartheta_1+\pi} := A^\top(r, \pi + \vartheta_1)A^\top(e_i \times r, -\theta)r$ and obtain the identity

$$\begin{aligned} n^{-\theta,\vartheta_1+\pi} &= A^\top(r, \pi + \vartheta_1)A^\top(e_i \times r, -\theta)r \\ &= A^\top(r, \vartheta_1)A^\top(r, \pi)A^\top(e_i \times r, -\theta)r \\ &= A^\top(r, \vartheta_1)A^\top(e_i \times r, \theta)r = n^{\theta,\vartheta_1}\,. \end{aligned}$$

Further the following identity can be verified

$$q(e_i \times r, \theta) \otimes q(r, \vartheta_1) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2) = q(e_i \times r, -\theta) \otimes q(r, \vartheta_1 + \pi) \otimes q(n^{\theta,\vartheta_1}, \vartheta_2 - \pi)\,.$$

Consequently we obtain Equation (5.29a) via

$$\begin{aligned} l_2^{r,n}(\theta,\vartheta_1,\vartheta_2) &= q(v,\phi)\otimes q(e_i\times r,\theta)\otimes q(r,\vartheta_1)\otimes q(n^{\theta,\vartheta_1},\vartheta_2) \\ &= q(v,\phi)\otimes q(e_i\times r,-\theta)\otimes q(r,\vartheta_1+\pi)\otimes q(n^{-\theta,\vartheta_1+\pi},\vartheta_2-\pi) \\ &= l_2^{r,n}(-\theta,\vartheta_1+\pi,\vartheta_2-\pi)\,. \end{aligned}$$

With the identities $q(n^{\theta,\vartheta_1},\vartheta_2)=-q(n^{\theta,\vartheta_1},\vartheta_2-2\pi)$, $q(r,\vartheta_1)=-q(r,\vartheta_1-2\pi)$ and (5.22) we obtain (5.28b) and (5.28c). Then, by using the identities

$$\begin{aligned} -l_2^{r,n}(\theta,\vartheta_1,\vartheta_2-2\pi) &= l_2^{r,n}(\theta,\vartheta_1-2\pi,\vartheta_2-2\pi) \\ -l_2^{r,n}(\theta,\vartheta_1,\vartheta_2-2\pi) &= l_2^{r,n}(\theta,\vartheta_1,\vartheta_2-4\pi) \end{aligned}$$

we obtain (5.29b) and (5.29c).

□

On the other hand, the domain of $l_2^{r,n}$ for ϑ_2 is chosen in the definition of the mappings to have the form $(-2\pi,2\pi]$. This is done to achieve the bijectivity on the quaternion space. However, restricting this domain to $(-\pi,\pi]$ is sufficient to represent all orientations because $\pm q$ describe the same orientation and (5.28c) allows the transformation into the reduced domain.

5.3.4.4 Smoothness

The smoothness properties of $l_1^{r,n}$ and $l_2^{r,n}$ are inherited from their defining functions. The function $l_1^{r,n}$ is not continuous on its complete domain. The points of discontinuity are defined by the following subset

$$\begin{aligned} \mathcal{Q}^{\mathrm{dc}} = \Big\{ & q\in\mathbb{S}_3 \mid \exists\,\theta\in[0,\pi],\ \vartheta_1\in(-\pi,\pi],\ \vartheta_2\in(-2\pi,2\pi] : \\ & q\in\{l_1^{r,n}(\bar\theta,\vartheta_1,\vartheta_2), l_1^{r,n}(\theta,\pi,\vartheta_2), l_1^{r,n}(\theta,\vartheta_1,2\pi)\} \text{for } \bar\theta\in\{0,\pi\}\Big\}\,. \end{aligned}$$

We summarise the smoothness properties as follows.

Proposition 5.7. *The function $l_2^{r,n}$ is smooth everywhere while $l_1^{r,n}$ is only smooth on $\mathbb{S}_3\setminus\mathcal{Q}^{\mathrm{dc}}$.*

Proof. The definition of $l_2^{r,n}$ is the result of a composition of smooth functions. Thus $l_2^{r,n}$ is smooth [119]. Also the smoothness of $l_1^{r,n}$ is the result of the individually composed functions. Possible discontinuities are induced by atan2 and the piecewise definition of ϑ_2. Indeed the discontinuities due to atan2 at $\theta\in\{0,2\pi\}$ can be shown using two simple sequences and the property (5.27). The same can be done for the singularities at $\vartheta_1=\pi$ and $\vartheta_2=2\pi$ and the properties (5.29). □

5.3.4.5 Dynamics

Considering the three angles as functions of time we may define their dynamics by using either of the functions $l_1^{r,n}$, $l_2^{r,n}$ and the quaternion dynamics from [41], i.e.

$$\dot{q} = \frac{1}{2}\Xi^\top(q)\,\omega\,. \tag{5.30}$$

By differentiating $(\theta, \vartheta_1, \vartheta_2) = l_1^{r,n}(q)$ with respect to time we obtain

$$\begin{bmatrix}\dot{\theta}\\ \dot{\vartheta}_1\\ \dot{\vartheta}_2\end{bmatrix} = \frac{\partial l_1^{r,n}(l_2^{r,n}(\theta,\vartheta_1,\vartheta_2))}{\partial q}\frac{1}{2}\Xi^\top(l_2^{r,n}(\theta,\vartheta_1,\vartheta_2))\,\omega + \frac{\partial l_1^{r,n}(l_2^{r,n}(\theta,\vartheta_1,\vartheta_2))}{\partial r}\,\dot{r}\,,$$

defined for all $(\theta, \vartheta_1, \vartheta_2)$ with $l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) \in \mathbb{S}_3 \setminus \mathcal{Q}^{\mathrm{dc}}$. Note that we also assume r to be time-varying because it is identified in the further course to the position of the spacecraft. These dynamics have the disadvantage that they are not defined for ϑ_1 or ϑ_2 at their boundaries. This can be avoided using $l_2^{r,n}$ instead.

Proposition 5.8. *For the functions of time $(r, \theta, \vartheta_1, \vartheta_2)$ with $q = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$ and q obeying the quaternion dynamics* (5.30)*, the dynamics of $(\theta, \vartheta_1, \vartheta_2)$ have the form*

$$\begin{bmatrix}\dot{\theta}\\ \dot{\vartheta}_1\\ \dot{\vartheta}_2\end{bmatrix} = \frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}^{+}\frac{1}{2}\Xi^\top(l_2^{r,n}(\theta,\vartheta_1,\vartheta_2))\,\omega - \frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}^{+}\frac{l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial r}\,\dot{r} \tag{5.31}$$

where $A^+ := (A^\top A)^{-1}A^\top$ denotes the Moore–Penrose inverse of a matrix A with full rank. The domain of the dynamics is $(0, \pi) \times (-\pi, \pi] \times (-2\pi, 2\pi]$.

Proof. Differentiating $q = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$ yields

$$\dot{q} = \frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}\begin{bmatrix}\dot{\theta}\\ \dot{\vartheta}_1\\ \dot{\vartheta}_2\end{bmatrix} + \frac{l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial r}\,\dot{r}\,. \tag{5.32}$$

With $\partial l_2^{r,n} := \frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}$ the product of the Jacobians is

$$\partial l_2^{r,n\top}\partial l_2^{r,n} = \begin{pmatrix}0.25 & 0 & 0\\ 0 & 0.25 & 0.25\cos\theta\\ 0 & 0.25\cos\theta & 0.25\end{pmatrix}. \tag{5.33}$$

This was validated in a numerical fashion. The expression for $\partial l_2^{r,n}$ itself is rather complicated and computed using computer algebra. The matrix (5.33) has full rank for $\theta \in (0, \pi)$. It does not have full rank at the singularities $\theta \in \{0, \pi\}$. Consequently, the Moore–Penrose inverse

$\frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}^+$ exists on $(0,\pi)\times(-\pi,\pi]\times(-2\pi,2\pi]$. Multiplying $\frac{\partial l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)}^+$ from the right to (5.32) and inserting the quaternion dynamics (5.30) leads to (5.31). □

5.3.4.6 Level Sets

In order to define distances between two orientations described by the proposed attitude representation, we use the angle defined on the quaternion space.

Definition 5.9. *For two sets of angles $(\theta,\vartheta_1,\vartheta_2)$ and $(\bar\theta,\bar\vartheta_1,\bar\vartheta_2)$ we define the angle ϕ using the mapping into quaternions as*

$$\phi(\theta,\vartheta_1,\vartheta_2,\bar\theta,\bar\vartheta_1,\bar\vartheta_2) := 2\arccos\left(e_4^\top\begin{bmatrix}\Xi^\top(q)\\ q\end{bmatrix}\bar q\right)$$

with $q = l_2^{r,n}(\theta,\vartheta_1,\vartheta_2)$ and $\bar q = l_2^{r,n}(\bar\theta,\bar\vartheta_1,\bar\vartheta_2)$.

Note that this metric considers the quaternions q and $-q$ as two different attitudes with distance 2π. It allows to define sublevel sets for a given attitude.

Definition 5.10. *For a reference attitude $(\bar\theta,\bar\vartheta_1,\bar\vartheta_2)$ the sublevel set for an angle $c\in[0,2\pi]$ is defined by*

$$\Omega^{q,c} = \{(\theta,\vartheta_1,\vartheta_2)\mid \phi(\theta,\vartheta_1,\vartheta_2,\bar\theta,\bar\vartheta_1,\bar\vartheta_2) < c\}\,.$$

These sublevel sets allow identifying angle tuples that represent similar attitudes despite exhibiting a large euclidean distance in their individual variables.

Figure 5.6 shows these sublevel sets for 27 points which are uniformly spread in $[0,\frac{\pi}{2}]\times[-\pi,\pi]\times[0,2\pi]$. It can be seen that the form of the level sets changes significantly in θ direction. For θ close to $\frac{\pi}{2}$ the level sets have an ellipsoidal form. For θ converging to zero, the level sets become cylindrical. This is due to the singularity at $\theta=0$ as discussed in Proposition 5.5. In this case only the sum of ϑ_1 and ϑ_2 are of interest and not the individual angles themselves. In ϑ_1 and ϑ_2 direction the form and size of the level set stay of similar magnitude. Furthermore, for points close to the boundary of the state space the level sets are cut off and continue at the opposite side of the space in the fashion stated in Proposition 5.6.

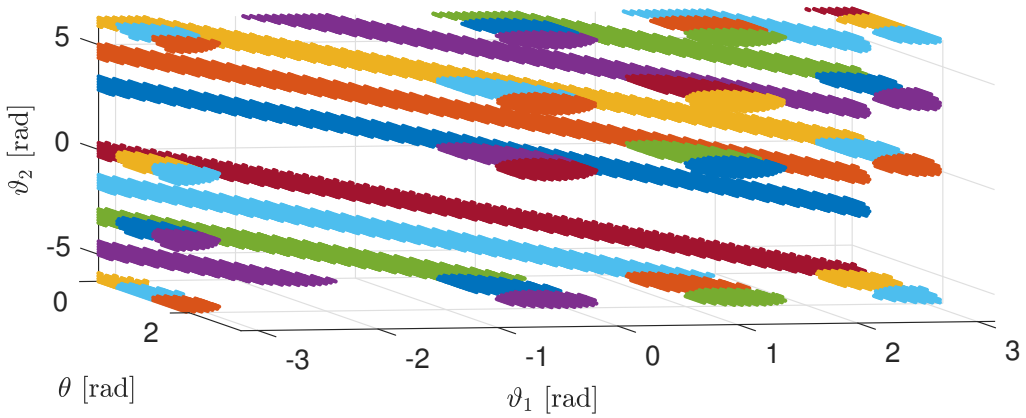


Figure 5.6: Level sets of the three angles.

5.3.5 Inversion Utilizing the Irradiation Angle Transformation

The in the previous section introduced transformation allows to write system (5.18) in the form

$$z_2 = L_1^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)) \tag{5.34a}$$
$$z_3 = L_2^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)) \tag{5.34b}$$
$$z_4 = L_3^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)) \tag{5.34c}$$
$$z_5 = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)^\top l_2^{r,n}(\theta, \vartheta_1, \vartheta_2) - 1\,. \tag{5.34d}$$

As presented in (5.18a), z_5 is identical to zero. Thus, the final Equation (5.34d) is fulfilled for any angle because the image of $l_2^{n,r}$ is the unit sphere. Thus, such a transformation simplifies the original nonlinear system of equations of order four to a system of equations of order three with the θ variable acting linearly in the first equation.

In this chapter, we restrict ourselves to Case 4 where only infrared irradiation is acting. In this special case, the angle θ can be directly determined from Equation (5.34a). If $F_{\text{form}}(\theta) = F_{\text{form},1}$ the angle can even by analytically calculated as Equation (5.34a) takes the form

$$z_2 = \gamma \frac{r_\oplus^2}{\|r\|^2} \cos(\theta) - \delta T^4$$

which allows to directly calculate the angle θ via

$$\theta = \arccos\left(\frac{1}{\gamma} \frac{\|r\|^2}{r_\oplus^2}(z_2 + \delta T^4)\right). \tag{5.35}$$

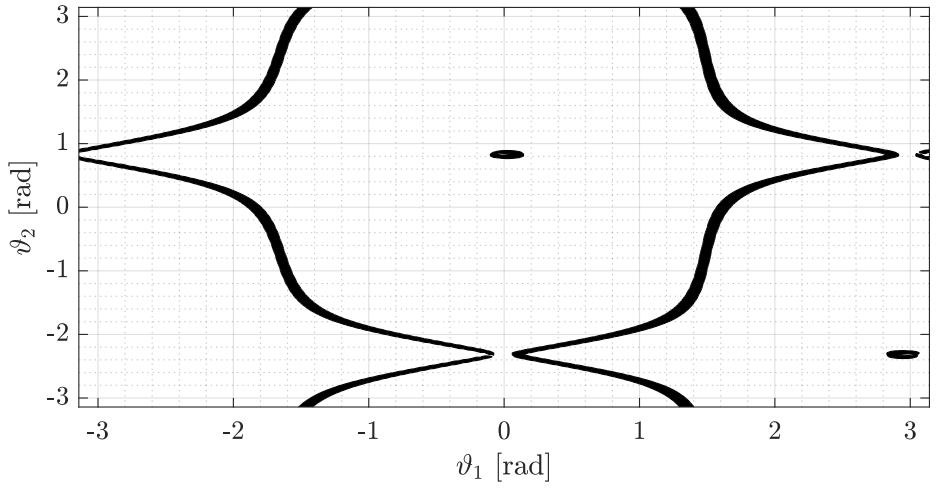


Figure 5.7: Illustration of the Non-observable points.

This simplifies the system (5.34) to the second order system

$$z_3 = L_2^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)) \tag{5.36a}$$

$$z_4 = L_3^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))\,. \tag{5.36b}$$

with the definitions $H_3(\vartheta_1, \vartheta_2) = L_2^{T,\omega,t}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))$ and $H_4(\vartheta_1, \vartheta_2) = L_3^{T,\omega,t,u}(l_2^{r,n}(\theta, \vartheta_1, \vartheta_2))$. As this is a highly nonlinear system of equations it is not possible to give an analytical expression for ϑ_1, ϑ_2 as a function of z. On top of that, this expression varies with the other states T, ω, t, θ and the parameters. In the next section, we formulate the nonlinear equation into an optimisation problem and discuss its solution for the parameters and states shown in Table A.1 and A.2.

5.3.6 Illustration of the Observability Mapping and its Singularities

Dependent on the transient behaviour of the observer, it might occur that the observation crosses the set of unobservable points. The might lead to a number of different phenomena such as discontinuities or chattering and complicates the convergence of the observer. Thus, it is important to identify the singularities in order to be able to avoid them. They can be easily identified using the method proposed in Section 5.2.2 which suggests calculating the observability matrix and determining if some of the singular values are close to zero or solving the optimisation problem (5.7). Figure 5.7 shows an overestimation of the set of non-observable points for $\theta = \frac{\pi}{2}$ obtained with the first method for a fine grid of points. It can be seen that this set is highly nonlinear and separates the ϑ_1-ϑ_2 space into four connected spaces. The set is symmetric with respect to a point close to $(0.07, 0.84)$ and some neighbourhoods of the points $(0, -2)$ and $(2.9, 0.84)$ are not part of the set and separate the set. Interesting are also the sets at $(0.07, 0.84)$ and $(3, -2)$ which are both shaped similar to an annulus with a non-trivial inner area.

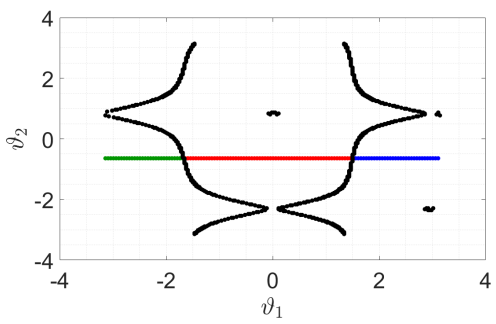
As shown in the previous section, nonlinearities only arise from the reduced observability mapping

$$H_{3,4} : (-\pi, \pi] \times (-\pi, \pi] \to \mathbb{R} \times \mathbb{R}$$
$$(\vartheta_1, \vartheta_2) \mapsto (z_3, z_4)$$

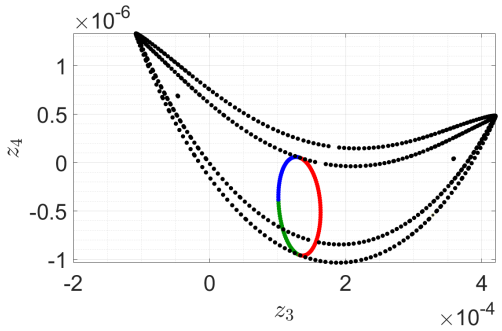
as defined in Equation (5.36). In the following, we denote the domain of $H_{3,4}$ as the ϑ-space and the image of $H_{3,4}$ as the z-space. Figure 5.8 illustrates the mapping of four different sets from the ϑ to the z space. The left-hand side shows the ϑ-plane and the set that is mapped to the z-plane which can be found on the right-hand side. The black points in the ϑ-plane are the points for which the Jacobian of the observability mapping becomes singular. It can be seen that these points divide the plane into roughly three parts. All points in the left part depicted in green colour, all parts in the middle are in red and all points in the right part are shown as blue points. The corresponding points in the z-plane have the same colour. It can be seen that the singular points are mapped to a closed curve displayed in black on the right-hand side. The left boundary in the ϑ-plane maps onto the outer boundary of the set. The right boundary maps onto the inner boundary. In this analysis we ignore the two separate annuli. [1]

Figure 5.8a and 5.8b illustrate the mapping of a set with constant ϑ_2 value to the z-space. It can be seen that the straight line is mapped to an ellipse. The points touch only the lower outer boundary and the upper inner boundary in the z-space. The other two boundaries are either intersected or not touched at all. Figure 5.8c and 5.8d map a set consisting of multiple sets with constant ϑ_2 value to the z-plane. It can be seen that each set with constant ϑ_2 value is mapped to an ellipse and that the ϑ_2 component determines the position of the ellipse. It can be seen that these ellipses gradually fill the space between the boundaries. Starting for $\vartheta_2 = 1.7$, the ellipse fill the z-plane with rising ϑ_2 from right to left up to $\vartheta_2 = 0.9$ when the following ellipses fill the z-space from left to right but the ellipses now touch the lower inner and upper outer boundary. At this point, it becomes clear that the mapping is not bijective. All points in the middle of the two inner boundaries in the z-plane have at least four corresponding points in the ϑ-plane while all remaining points have at least two, this will also be evident in the next section.

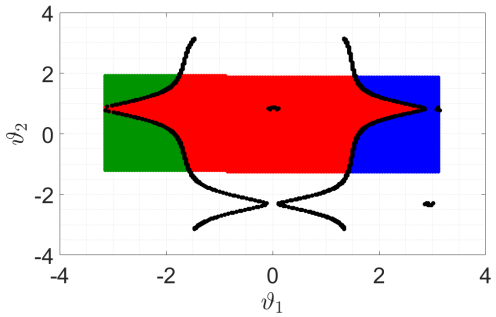
Similar illustrations can be made with Figure 5.8e, 5.8f, 5.8g and 5.8h. Here, sets with constant ϑ_1 component are mapped to the z-space. It can be seen that the mapping of such a single set has a similar form to the mapping of the singularity points. For the displayed sets, it can be seen that the space between the inner boundaries in the z-space becomes gradually filled for rising ϑ_1. When the inner boundary is reached, the direction is reversed and the mapped sets reach back into the inner.



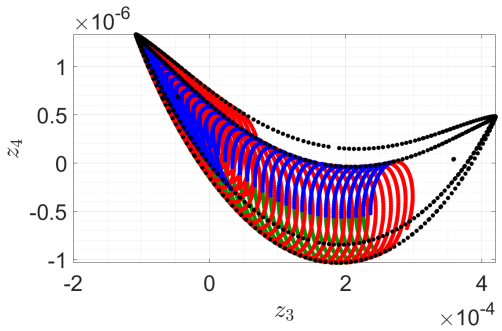
(a) Points with constant ϑ_2



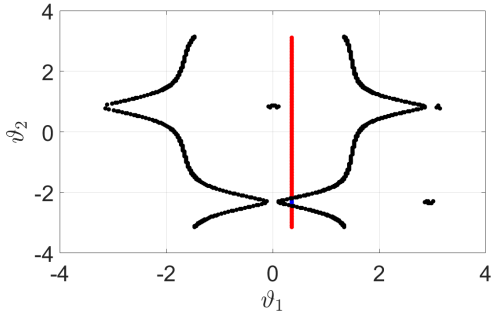
(b) Image of points with constant ϑ_2



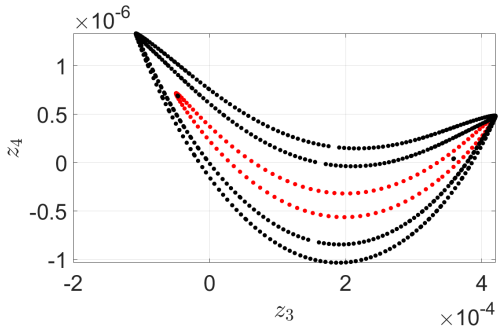
(c) Multiple sets. Each set has constant ϑ_2.



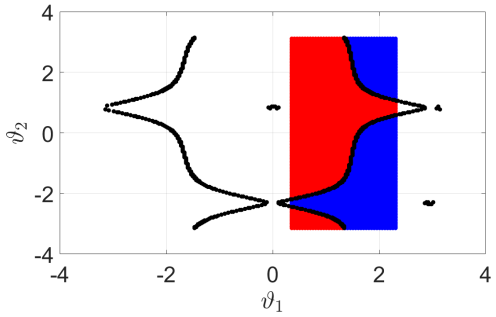
(d) Image of the sets with constant ϑ_2



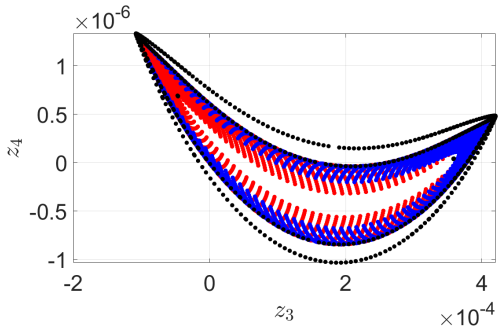
(e) Points with constant ϑ_1



(f) Image of points with constant ϑ_1



(g) Multiple sets. Each set has constant ϑ_1.



(h) Image of the sets with constant ϑ_1

Figure 5.8: Illustration of the observability mapping $H_{3,4}$. The left column displays the $(\vartheta_1, \vartheta_2)$ space in radians with the coloured points to be mapped by $H_{3,4}$. Green points are left, red points are in the middle and blue points are right of the two boundaries. The right column displays (z_3, z_4) space with the mapped points and corresponding colours.

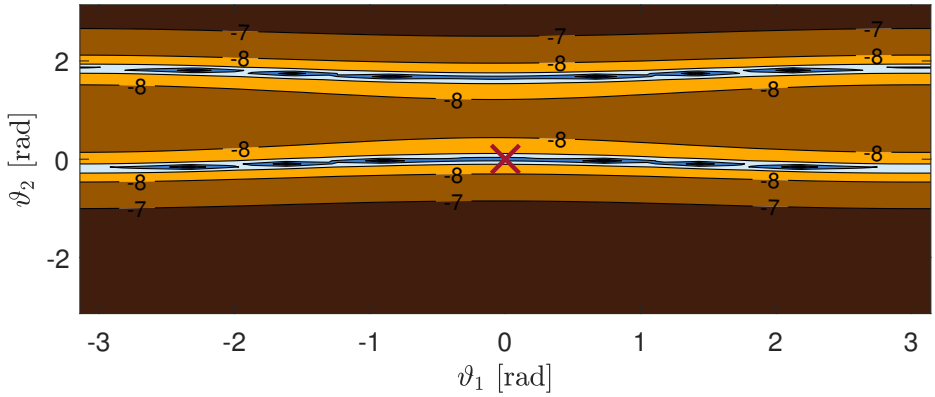


Figure 5.9: Contour plot of the image of the logarithm of the cost function of optimisation problem (5.37) for outputs created of the corresponding states $(\vartheta_1, \vartheta_2) = 0$.

5.3.6.1 Formulation of the Optimisation Problem

The inversion problem (5.36) can be reformulated into an optimisation problem of the form

$$\min_{\vartheta_1,\vartheta_2} \quad \left\| \begin{bmatrix} H_3(\vartheta_1,\vartheta_2) - z_3 \\ H_4(\vartheta_1,\vartheta_2) - z_4 \end{bmatrix} \right\|_2 . \tag{5.37a}$$

$$\text{s.t.} \quad [\vartheta_1, \vartheta_2] \in [-\pi, \pi] \times [-\pi, \pi] \tag{5.37b}$$

It is straightforward to see that every solution of (5.36) is a solution of (5.37) and vice versa. Plotting the image of the cost function (5.37) allows getting an estimate of the difficulty of this optimisation problem. Figure 5.9 shows this image in the $(\vartheta_1, \vartheta_2)$-space for $[z_3, z_4] = [H_3(0,0), H_4(0,0)]$ and the given parameters and states. . It can be seen that most of the level sets are approximately parallel to the ϑ_1 axis. However, there are two valleys around which the level sets are locally decreasing. One contains the expected optimum while the other has a similar shape and seemingly admits multiple optima as well. Furthermore, the cost function varies barely for varying ϑ_1. This is coherent with the results of Section 5.2.2, as one singular value is of the magnitude of 10^{-4}. Values along the corresponding eigendirection vary only slowly. This makes the determination of a solution difficult because all values along this manifold are optimal results under certain tolerances which will make the optimisation stop if a point $(\vartheta_1, 0)$ with $\vartheta_1 \in (-\pi, \pi)$ is reached. In order to avoid this issue, we need to transform the cost function to remove the slowly varying manifold. A simple linear transformation is proposed by considering the Taylor Expansion of first order at the point $(\vartheta_1^*, \vartheta_2^*)$

$$H_{3:4}(\vartheta_1, \vartheta_2) \approx H_{3:4}(\vartheta_1^*, \vartheta_2^*) + \frac{\partial H(\vartheta_1^*, \vartheta_2^*)}{\partial(\vartheta_1, \vartheta_2)} \left(\begin{bmatrix} \vartheta_1 \\ \vartheta_2 \end{bmatrix} - \begin{bmatrix} \vartheta_1^* \\ \vartheta_2^* \end{bmatrix} \right).$$

[1]Note that we mean by outer boundary the boundary of the set described by the closed curve and all its inner points. The inner boundary is the complement of the closed curve without the image of the annuli.

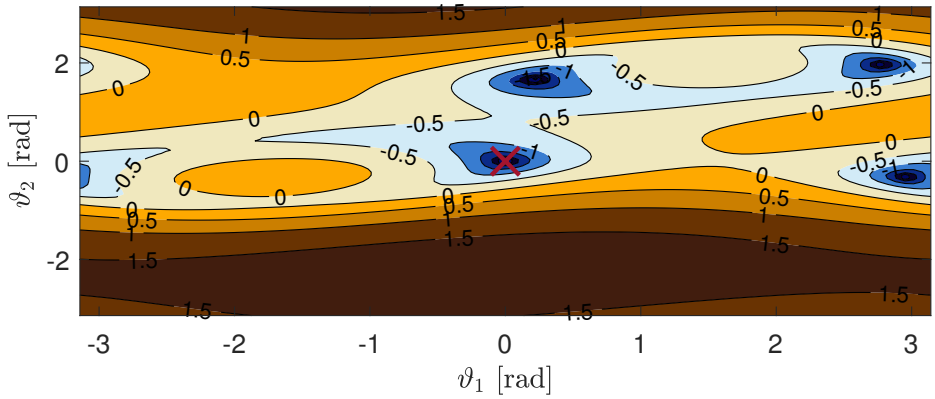


Figure 5.10: Contour plot of the image of the logarithm of the cost function of optimisation problem (5.38) for outputs created of the corresponding states $(\vartheta_1, \vartheta_2) = 0$.

Using the Jacobian allows to shape the optimisation problem (5.37) to obtain a better conditioned problem in the form

$$\min_{\vartheta_1,\vartheta_2} \quad \left\| \left[\frac{\partial H(\vartheta_1,\vartheta_2)}{\partial(\vartheta_1,\vartheta_2)} \right]^{-1} \begin{bmatrix} H_3(\vartheta_1,\vartheta_2) - z_3 \\ H_4(\vartheta_1,\vartheta_2) - z_4 \end{bmatrix} \right\|_2 \tag{5.38a}$$

$$\text{s.t.} \quad [\vartheta_1, \vartheta_2] \in D_1 \times D_2 \,. \tag{5.38b}$$

The image of this cost function is shown in Figure 5.10. It can be seen that the image admits a smooth behaviour in the domain. The form of the level sets have changed significantly through the transformation. The attitude $(\vartheta_1, \vartheta_2) = 0$ used to determine the outputs now describes a minimum with near-circular level sets in its vicinity. This is a desirable property in respect of numerically solving the optimisation problem. Additionally, it can be seen that there are multiple other minima. Some of them are due to the boundedness of the domain and can be found at its boundary. They can clearly be identified as non-global minima. Others, that can be found in the inner of the domain exhibit a cost function value in the same range as for $(\vartheta_1, \vartheta_2) = 0$. The difference between these minima and their region of attraction is discussed in the next section.

Remark 9. In (5.38) the inverse of the Jacobian matrix is multiplied by the measurement values. As the eigenvalues are of the Jacobian matrix are of the order 10^{-6} a multiplication of its inverse assumes a sufficiently high accuracy of these measurement values.

Remark 10. Note that the domain of ϑ_2 in the optimisation problem is chosen to be $[-\pi, \pi]$ instead of $[-2\pi, 2\pi]$. The latter allows a continuous trajectory in the quaternion space but allows two solutions of the optimisation problem representing the same attitude.

Remark 11. In the sections where we use numerical methods to identify solutions we consider to points to be equal if their euclidean distance is in a certain tolerance.

5.3.6.2 Local Minima and their Region of Attraction

As pointed out in the previous section, there exist multiple points that solve the optimisation problem locally. This makes the solution of the optimisation problem depending on the usage of the initial value and the optimisation method. Local minima can be avoided using multiple starting values. Global minima exist if the transformation τ is not injective and can only be distinguished from the real solution by incorporating additional constraints. In this section, we identify the minima the solver is converging to and determine its region of attraction.

We solve the optimisation problem (5.38) for a number of initial values uniformly covering the ϑ-state space $[-\pi, \pi] \times [-\pi, \pi]$. If the optimisation converges, the result $\bar{\vartheta}$ is a local minimum which allows us to define as in [120] the function

$$\mathcal{LS} : [-\pi, \pi] \times [-\pi, \pi] \to \mathrm{im}(\mathcal{LS}) \tag{5.39a}$$

$$\vartheta \mapsto \bar{\vartheta} \tag{5.39b}$$

mapping every initial value ϑ to the optimum $\mathcal{LS}(\vartheta)$ the solver is converging to. Note that $\mathrm{im}(\mathcal{LS})$ denotes the image of $\mathcal{LS}$ which is a set of finite points resulting from the optimisations. It is clear that every minimum $\bar{\vartheta} \in \mathrm{im}(\mathcal{LS})$ is a global minimum if the cost function of $\bar{\vartheta}$ is equal to zero. Otherwise, the minimum is local. In Table 5.4 the calculated minima and its cost can be seen. It can be seen that there are four global minima and four other local minima that can be clearly identified as such by a higher value of the cost function. We define the basin of attraction of a minimum $\Omega(\bar{\vartheta})$ as the pre-image of $\bar{\vartheta}$ under $\mathcal{LS}$, i.e.

$$\Omega(\bar{\vartheta}) := \{\vartheta \in [-\pi, \pi] \times [-\pi, \pi] \mid \mathcal{LS}(\vartheta) = \bar{\vartheta}\}\,.$$

This basin is displayed for the individual minima $\bar{\vartheta} \in \mathrm{im}(\mathcal{LS})$ in Figure 5.11. The minima are marked by black crosses. In red, the union of the initial values that do not converge to a global minimum can be seen. The green region displays the basin of attraction of $\bar{\vartheta} = [0, 0]^\top$ which was used to construct the optimisation problem. The remaining colours correspond to the other three global minima. It can be seen that every minimum has a basin of attraction of non-negligible size. These regions of attraction are mostly connected and clearly distinguishable. However, there are some a points in each basin that make the smallest convex sets in the basins very conservative. It can be concluded by the size and the form of the sets that a classical optimisation method allows finding the desired minimum, granted that the initial value is in its region of attraction. If this can not be ensured then a population-based optimisation method can find all global minima and reconstruct the attitudes that lead to the measured output and its derivatives. In the next section, we analyse the evolution of the global minima over the course of varying attitudes.

Table 5.4: Determined minima $\bar{\vartheta} \in \mathcal{LS}(\vartheta)$ and their rounded costs.

Minimum	$\begin{bmatrix}0\\0\end{bmatrix}$	$\begin{bmatrix}2.955\\-0.32\end{bmatrix}$	$\begin{bmatrix}2.77\\1.96\end{bmatrix}$	$\begin{bmatrix}0.208\\1.66\end{bmatrix}$	$\begin{bmatrix}-\pi\\-0.33\end{bmatrix}$	$\begin{bmatrix}-\pi\\2.2\end{bmatrix}$	$\begin{bmatrix}-2.23\\-\pi\end{bmatrix}$	$\begin{bmatrix}-\pi\\-\pi\end{bmatrix}$
Costs	2×10^{-8}	2×10^{-8}	2×10^{-8}	2×10^{-8}	3×10^{-2}	0.1	5	8

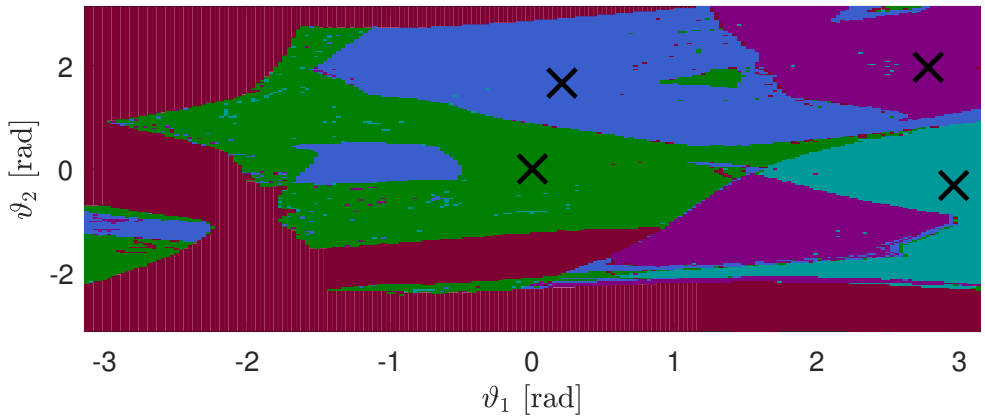


Figure 5.11: Illustration of the basin of attraction of the four global minima.

5.3.6.3 Optimisation Problem for Different Attitudes

As pointed out in the previous section, the analysis of the contour plot allows evaluating the difficulty of the optimisation problem by identifying local minima and a rough approximation of their basin of attraction. Thus, we consider the contour plot for attitudes distributed along a grid as displayed in Figure 5.12. Each plot shows the contour plot of the image of the logarithm of the cost function of optimisation problem (5.38) for outputs (z_3, z_4) determined by (5.36) with the angles $(\vartheta_1^*, \vartheta_2^*)$ displayed as the red marker. It can be observed that all plots admit multiple global minima. It can be seen that for every ϑ_2 value exists regions $[\underline{\vartheta}_1, \overline{\vartheta}_2]$ with $\underline{\vartheta}_1, \overline{\vartheta}_2 \in \mathbb{R}$ that lack any minima with small cost function value. Except that, few more general regularities can be derived from these contour plots as the global minima vary in their number as well as form of region of attraction in highly nonlinear fashion.

Remark 12. The image of the cost function varies greatly with the states q and ω.

Remark 13. Parameter variations have shown that a high angular velocities allow a better reconstruction of the ϑ_2 coordinate while lower angular velocities allow a better ϑ_1 reconstruction. The best results are obtained for angular velocities in the range of the angular velocity of the circulating spacecraft $\frac{r \times v}{\|r\|^2}$.

5.3.7 Simulation

The differentiation and inversion are combined in the algorithm illustrated in the block diagram 5.13. For this simulation, the derivatives of the temperature measurement are analytically generated and the inversion problem is solved in every time step in order to determine the attitude.

As we have seen in the previous section for every attitude might exist multiple other attitudes that lead to the same temperature and its derivatives. Thus, the choice of the initial state for the optimisation problem is very important as this will lead to different solutions. In

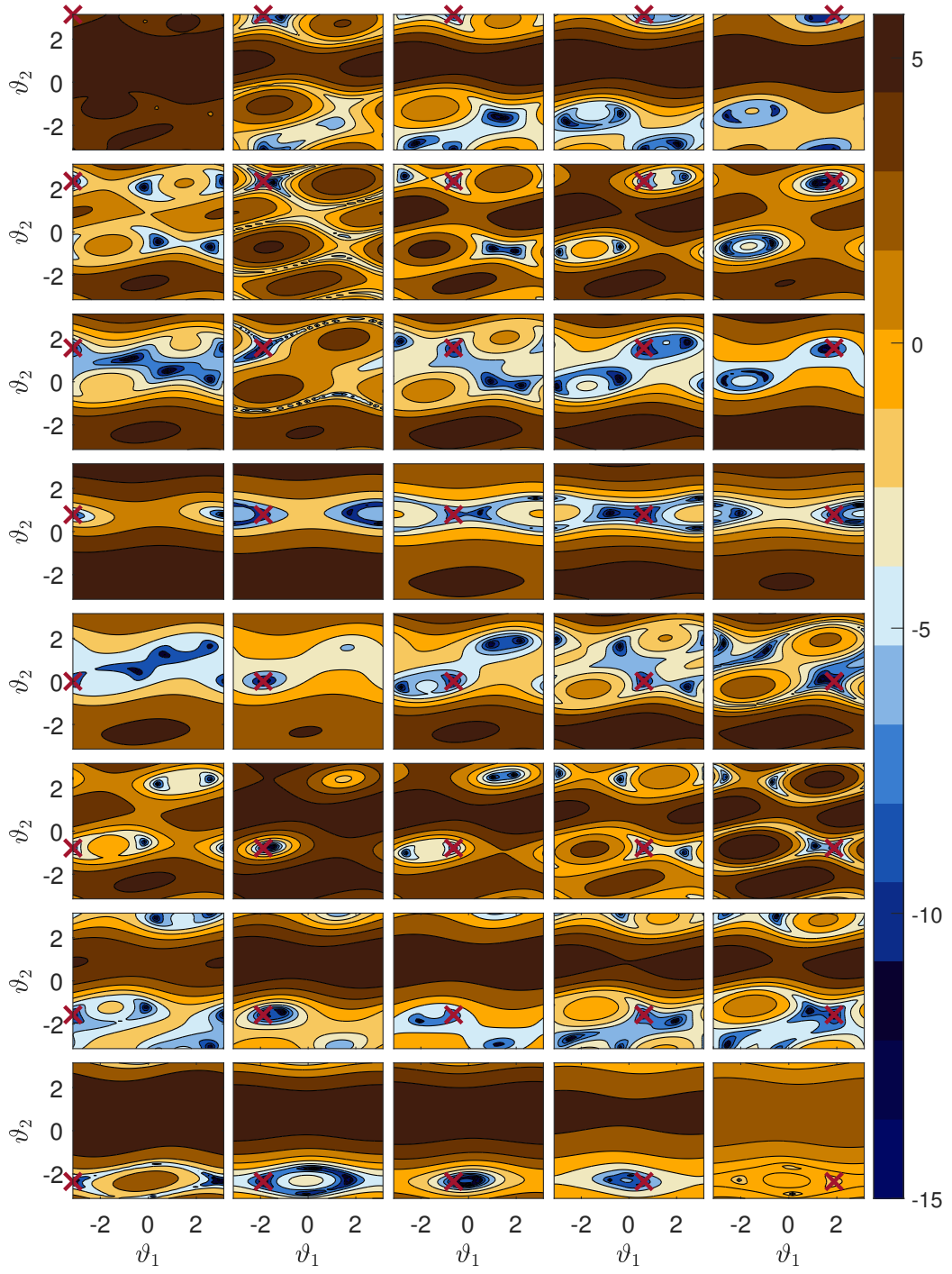


Figure 5.12: Contour plot of the image of the logarithm of the cost function of optimisation problem (5.38) for (z_3, z_4) determined by (5.36) with the angles $(\vartheta_1, \vartheta_2)$ displayed as the red marker. Angles are given in radians.

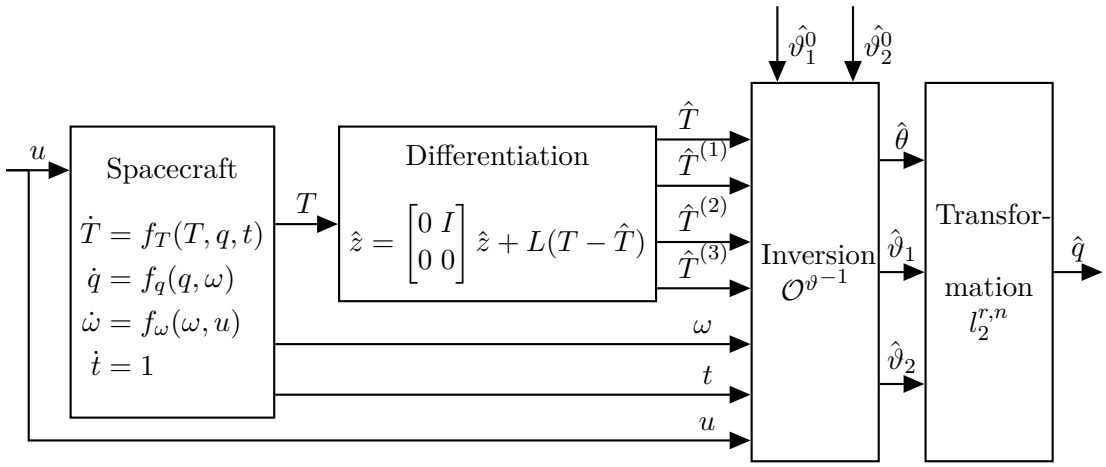


Figure 5.13: Attitude Reconstruction Using Differentiation and an inversion with transformation.

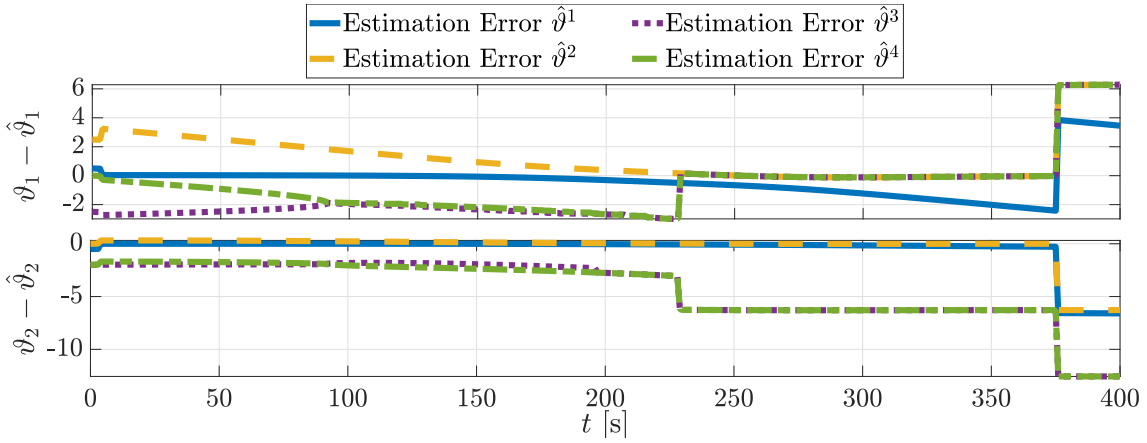


Figure 5.14: Inversion result for four different initial states.

order to obtain continuous solutions, we use the attitude of the last step as the initial value of the next step. Still, the problem of having appropriate initial values in the first place persists. This problem is illustrated in the following simulation where four different initial values are used. These initial values are all in the basin of attractivity of another one of the four existing global minima. The error of the estimated angles is displayed in Figure 5.10. The initial values are $\hat{\vartheta}^1(0) = (-0.5, 0.5)$ and $\hat{\vartheta}^2(0) = (-2.5, 0)$ and $\hat{\vartheta}^3(0) = (2.5, 2)$ and $\hat{\vartheta}^4(0) = (0, 2)$. Because we consider for this simulation Case 4, the angle θ can be immediately and accurately calculated from $\dot{T}$ and is assumed to be known. Thus, only the remaining irradiation angles ϑ_1 and ϑ_2 are of interest and are reconstructed by the optimisation. In order to allow for a continuous estimation the domain $D_1(t)$ and $D_2(t)$ of ϑ_1 and ϑ_2 are moved with the last estimate, i.e. $D_1(t_{k+1}) = \vartheta_1(t_k) + [-\pi, \pi]$, $D_2(t_{k+1}) = \vartheta_2(t_k) + [-4\pi, 4\pi]$. These are the domains used in the optimisation.

The real and estimated angles can be found in Figure 5.14. All four estimations admit

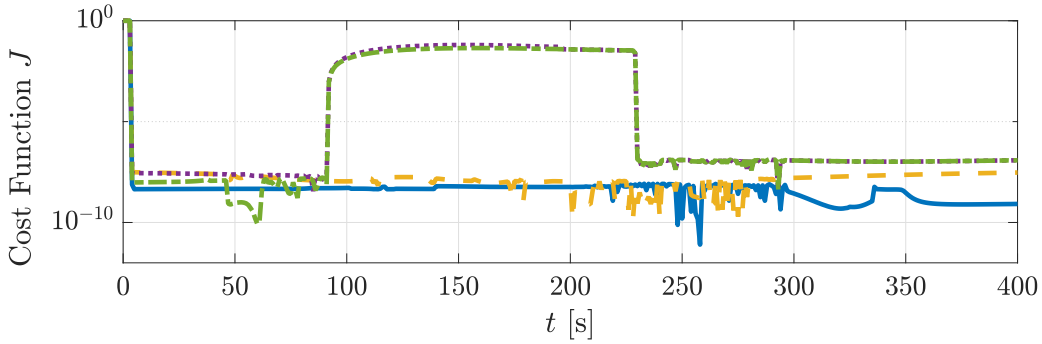


Figure 5.15: Value of the cost function for the four estimations.

different behaviour. The estimate $\hat{\vartheta}^1$ with its initial value close to the real initial value converges to the real states and the estimate $\hat{\vartheta}^1$ follows the real states until 205s. Then the estimated and real state differs for the remainder of the simulation.

The other three estimates $\hat{\vartheta}^2$, $\hat{\vartheta}^3$, $\hat{\vartheta}^4$ all start at different initial values and follow different estimates. However, at 205s they start following the real state ϑ. Note the discontinuity of the real state at 376s. This is due to the boundedness of the domain of ϑ to avoid ambiguities. The estimates are kept within a moving domain to avoid this discontinuity. It can however be easily re-transformed into the real domain using Lemma 5.6.

In order to evaluate whether all estimates really yield the same outputs, the cost function is displayed in Figure 5.15. All costs below 10^{-7} suggest a good approximation which admits really the outputs expected. Thus, ϑ^1, ϑ^2 can both be seen as attitude trajectories that lead to the same output. For ϑ^3 and ϑ^4 it can be seen that the optimisation results between 94s and 225s are not satisfactory. In the remaining time, however, the estimated attitudes achieve the measured outputs as well.

These observations reveal an important issue. It is not possible to reconstruct the correct attitude only from the measured output T and its derivatives. In the first 100s there are four different attitude trajectories that lead to the same outputs. And it cannot be determined which one is the real one solely by the output.

Consequently, the initial states are either undistinguishable or one of the reconstructed trajectories is not a solution of the dynamics (5.2) with $q = l_2^{r,n}(\theta, \vartheta_1, \vartheta_2)$. In order to verify whether the dynamic equations are fulfilled, we determine the derivative of q using the difference quotient as well as the right-hand side of (5.2) and determine the difference between them

$$\varepsilon^q := \frac{q(t+h) - q(t)}{h} - \frac{1}{2}\Omega(\omega)q\,. \tag{5.40}$$

This variable is close to zero for a solution q which solves (5.2). It is displayed in Figure (5.15) for the real and estimated states and a step size $h = 1$. The black dashed line corresponds to the error of the real quaternion trajectory. This error is dependent on the step size and converges to zero for small step sizes. It can be seen that for ϑ and all regions

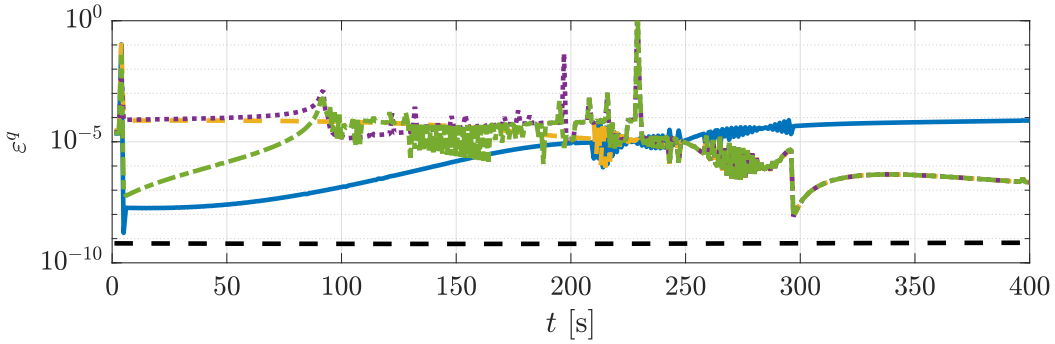


Figure 5.16: Quaternion Derivative Error as defined in (5.40) for the real quaternions and the four estimates. The black dashed line corresponds to the error of the real quaternion trajectory.

where the estimate is identical to ϑ the error is very small and exhibits almost no dynamics. In the remaining time, the error varies and is of high magnitude. This shows that despite the ambiguity of the reconstruction algorithm a correct estimate can be identified using the information of a time interval of the estimation. It can be used to validate a posteriori which estimate is the correct one or it can be incorporated into the optimisation problem. Note that a similar result can be achieved if the fourth derivative of the temperature is considered. However, this approach is not recommended as the numerical accuracy of the sensors is generally not sufficient.

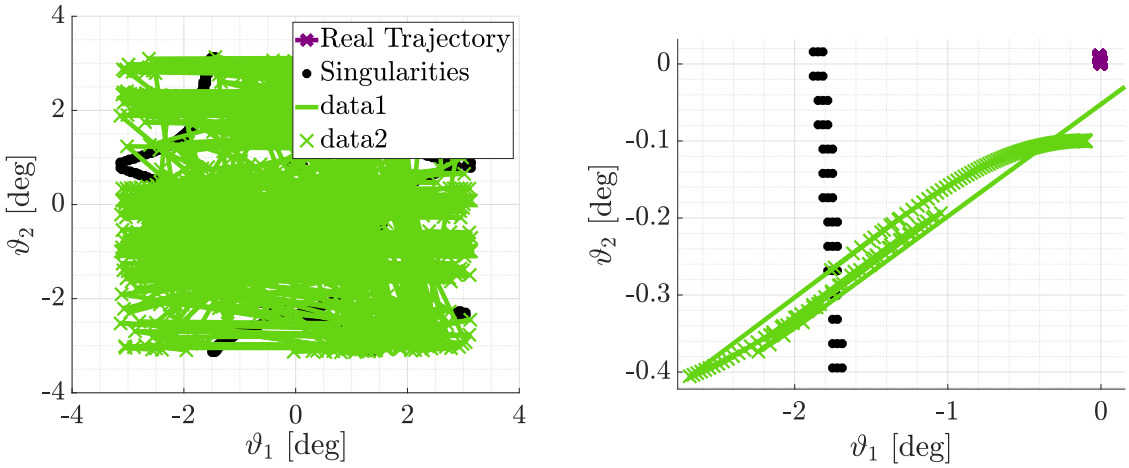
5.4 Observer in Natural Coordinates

In this section, we discuss the observer obtained in the previous section transformed into natural coordinates as proposed in Section 2.3. This has the advantage that the optimisation problem is avoided and instead only the inversion of a matrix is required. However, other problems that are of no concern in the previous design occur and are discussed in detail. The observer in natural coordinates has the form

$$\dot{\hat{x}} = f(\hat{x}) + \left(\frac{\partial \mathcal{O}(\hat{x})}{\partial x}\right)^{-1} L(y - \hat{y}) \tag{5.41a}$$

$$\hat{y} = h(\hat{x})\,. \tag{5.41b}$$

We show the result of a straightforward application for the parameters and states displayed in Table A.1 and A.2 for the initial value $(\vartheta_1, \vartheta_2) = (-0.1, -0.1)$. Further, the coefficients of the high gain observer were chosen such that the poles are at $[-1, -4, -5, -6]$ and the parameter ε is equal to 1. Figure 5.17a shows the evolution of the real and estimated trajectory in $(\vartheta_1, \vartheta_2)$-space for a simulation time of 2s and a step size of 10^{-3}. Note that the step size is chosen to be smaller than in the other simulations to illustrate the issue. It can be seen that the estimate diverges despite having an initial value close to the real value. The reason for this behaviour can be seen in Figure 5.17b where the simulation time is reduced



(a) Complete ϑ-space

(b) Subset of the ϑ-space

Figure 5.17: Simulation results using the common observer in natural coordinates. The estimated trajectory crosses the singularities and becomes unstable.

to 0.2s and only a smaller part of the $(\vartheta_1, \vartheta_2)$-space is shown. In the transient phase the estimation moves towards the boundary indicating the points where the Jacobian matrix of the observability mapping is singular. If the trajectory comes too close to this boundary, the estimate becomes discontinuous and obtains a new value which is heavily reliant on the step size. After this jump, the trajectory moves back into the direction of the boundary where a discontinuous jump occurs again. At this point, it is not to be expected any more that the estimate converges as the boundary describing the singularity appears to be attractive in some sense. Consequently, in order to make the observer design (5.41) work, it needs to be ensured that neither the real trajectory nor the estimate crosses the singularities. One method to achieve this is by using an invariant observer design.

5.4.1 Invariant Observer Design/ Image Extension of a Diffeomorphism

As motivated in the previous section we want to adjust the observer design

$$\dot{\hat{x}} = f(\hat{x}) - \frac{\partial \bar{\mathcal{O}}(\hat{x})}{\partial \hat{x}}^{-1} L(\hat{y} - y) \tag{5.42}$$

for a non-bijective observability mapping $\bar{\mathcal{O}} : \mathbb{R}^n \mapsto \mathbb{R}^n$ such that the evolution of $\hat{x}$ is only on a subset $E \subsetneq \mathbb{R}^n$ where E is chosen such that $\mathcal{O} : E \to \mathcal{O}(E)$ with $\mathcal{O}(x) = \bar{\mathcal{O}}(x)$ is bijective. On the one hand, this problem formulation can be considered as the task of finding an *invariant observer* in the sense that $\hat{x}(0)$ is in E as well as $\hat{x}(t)$ for all t. On the other hand, the term *invariant observer* is already discussed and used by [121, 122] with a slightly different meaning. It describes an observer that has the same invariance with respect to a transformation as the original system, i.e. using specific kinds of transformation on the system does not change the dynamics. Applications of this idea can be found e.g. in [123] and [124] where the attitude of a flying rigid body is estimated. This is done by using an

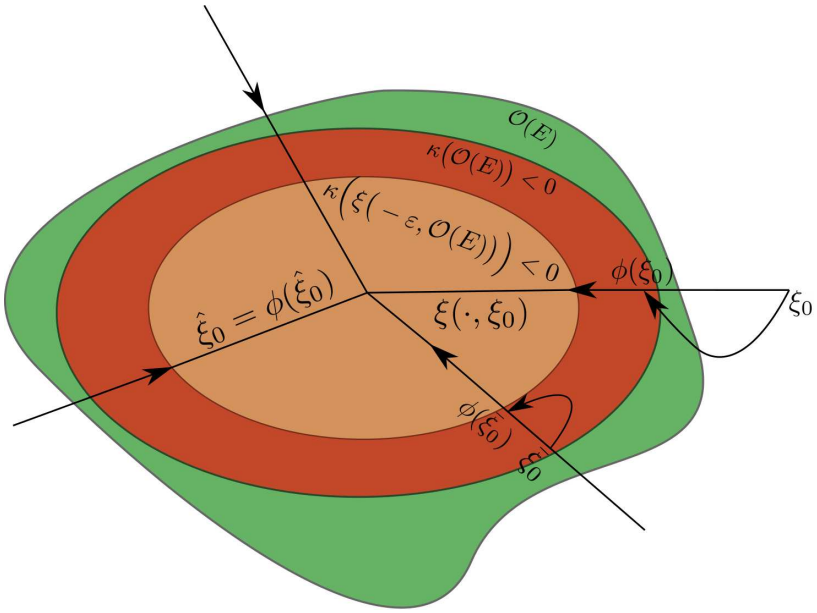


Figure 5.18: Illustration of the invariant observer idea. The green set illustrates the set $\mathcal{O}(E)$, the set denoted by $\kappa(\mathcal{O}(E))$ illustrates the largest set M which is a subset of $\mathcal{O}(E)$ and can be described by κ. The orange set $\kappa(\xi(-\varepsilon, \mathcal{O}(E))) < 0$ denotes the set on which ϕ is identical to the identity function.

observer which retains the invariance of the original system to be unaffected by a constant velocity translation in the Earth-fixed frame and constant rotation of the body-fixed frame. This kind of invariance is different from the one we are trying to achieve in this section.

We solve this issue by using the approach suggested in [19, 125] using the solution to the problem *image extension of a diffeomorphism*. This means that we extend the diffeomorphism $\mathcal{O} : E \to \mathcal{O}(E)$ using a bijective function $\phi : \mathbb{R}^n \to \mathcal{O}(E)$. The idea is that the dynamics of the transformed coordinates $\xi = \phi^{-1}(\mathcal{O}(x))$ are the same as of the transformed coordinates $z = \mathcal{O}(x)$, which are stable by design. All states $z \notin \mathcal{O}(E)$ for which $\mathcal{O}$ is not defined are projected into $\mathcal{O}(E)$ using ϕ which leads to $x = \mathcal{O}^{-1}(\phi(\xi)) \in E$. Therefore, the new extended coordinate transformation $T = \phi^{-1} \circ \mathcal{O}$ maps every x to a point in $\mathbb{R}^n$ while the usual transformation $\mathcal{O}$ maps every point x only to a subset of $\mathbb{R}^n$, namely $\mathcal{O}(E)$. Thus, the inverse transformation $T^{-1} = \mathcal{O}^{-1} \circ \phi$ is defined for every state in $\mathbb{R}^n$ and the issues illustrated in the previous section can not occur.

Thus, we can use the observer in natural coordinates as

$$\dot{x} = \frac{\partial \mathcal{O}(x)}{\partial x}^{-1} \frac{\partial \phi(\xi)}{\partial \xi} \dot{\xi}|_{\xi = \phi^{-1}(\mathcal{O}(x))} .$$

Note that for $\phi(\xi) = \xi$ the dynamics are identical to the dynamics of the original observer (5.42) except for a nonlinear term that is usually neglected. Therefore, this method can be considered as an extension of the observer (5.42).

Design of the observer

The choice of ϕ is not trivial and [19] proposed a generic way to design ϕ. Hereby, the following definition proposed in [19] is introduced.

Definition 5.11. *An open subset M of $\mathbb{R}^n$ is said to fulfil Property $\mathfrak{C}$ if there exists a C^1 function $\kappa : \mathbb{R}^n \to \mathbb{R}$, a bounded C^1 vector field χ, and a closed set K_0 contained in M such that*

1. *$M = \{x \in \mathbb{R}^n \mid \kappa(\xi) < 0\}$*
2. *K_0 is globally attractive for χ*
3. *the following transversality property is fulfilled for all ξ with $\kappa(\xi) = 0$:*

$$\frac{\partial \kappa}{\partial \xi}\chi(\xi) < 0\,.$$

If M fulfils the property $\mathfrak{C}$, then ϕ exists and can be constructed as shown in [19]. The idea is illustrated in Figure 5.18. First, it shall be emphasised that the design needs to be made in the ξ domain and not in the x domain which might be more natural, i.e. knowledge about the set $\mathcal{O}(E)$ is required. Then the largest set $\kappa(\mathcal{O}(E))$ and a subset of it needs to be found for which it can be ensured that the real trajectory does not leave these sets. This subset is depicted in Figure 5.18 in orange while the set $\kappa(\mathcal{O}(E))$ without this subset is depicted in red. Then, every point outside of the red and orange sets is mapped into the red part. Note that this red part is designed to be arbitrarily small. Every point in the red subset is mapped closer to the orange subset. Finally, every point in the orange subset is mapped onto itself.

Note that in general, this analytical construction requires the solution of a differential equation and an inversion problem which might be both nonlinear dependent on the choice of κ and χ. Furthermore, choosing appropriate κ and χ is difficult in its own right. In the next theorem, we propose specific κ and χ to simplify the design process and calculation of the corresponding functions ϕ and ϕ^{-1}.

Theorem 5.12. *Let $P \in \mathbb{R}^{n\times n}$ be a positive definite symmetric matrix, $d \in \mathbb{R}^n$ be a vector, $\varepsilon \in \mathbb{R}^+$ be a positive real number and $\kappa : \mathbb{R}^n \to \mathbb{R}$ with*

$$\kappa(\xi) := \xi^\top P\xi - 2d^\top P\xi + d^\top Pd - c \tag{5.43}$$

be the function defining the set

$$M := \{\xi \in \mathbb{R}^n \mid \kappa(\xi) < 0\}. \tag{5.44}$$

Further, let the flow $\xi : \mathbb{R}^n \times \mathbb{R} \to \mathbb{R}$ *and the function dilating the time* $\nu : \mathbb{R} \to (0, \infty)$ *be*

$$\xi(\xi_0, t) = e^{-t}(\xi_0 - d) + d \tag{5.45}$$

$$\nu(t) = \begin{cases} \frac{(t+\varepsilon)^2}{2\varepsilon+t} & \text{for } t > -\varepsilon \\ 0 & t \leq -\varepsilon \end{cases} . \tag{5.46}$$

Combining these with the function $\mathfrak{t} : \mathbb{R}^n \to \mathbb{R}$ *with*

$$\mathfrak{t}(\xi_0) = -\frac{1}{2} \ln \frac{c}{(\xi_0 - d)^\top P (\xi_0 - d)} \tag{5.47}$$

which maps an initial state to its time needed to reach the boundary ∂M*, allows to define the* C^1*-diffeomorphism* $\phi : \mathbb{R}^n \to M$ *with*

$$\phi(\xi_0) = \begin{cases} \xi\Big(\xi_0, \nu\big(\mathfrak{t}(\xi_0)\big)\Big) & \text{for } \mathfrak{t}(\xi_0) > -\varepsilon \\ \xi_0 & \text{for } \mathfrak{t}(\xi_0) \leq -\varepsilon \end{cases} . \tag{5.48}$$

Then its inverse $\phi^{-1} : M \to \mathbb{R}^n$ *has the form*

$$\phi^{-1}(y) = \xi(y, -\nu(\mathfrak{s}(y))) \tag{5.49}$$

where the function $\mathfrak{s} : M \to (-2\varepsilon, \infty)$ *is defined by*

$$\mathfrak{s}(y) = \frac{-\varepsilon^2}{\mathfrak{t}(y)} - 2\varepsilon \,. \tag{5.50}$$

Proof. We give a constructive proof in the sense of [19]. We show the relations between the given functions and show that M fulfils the property $\mathfrak{C}$ of M. The set M is defined in (5.44) according to Condition 1. Its boundary has the form

$$\partial M = \{x \in \mathbb{R}^n \mid \kappa(\xi) = 0\} \,. \tag{5.51}$$

Using the vector field $\chi(\xi) := -\xi + d$ and the differential equation $\dot{\xi} = \chi(\xi)$ leads to the solution defined in (5.45). We define the closed set $K_0 := \{0\}$ which is contained in M and attractive as required in Condition 2. Mapping the solution (5.45) with the vector field κ from equation (5.43) yields

$$\begin{aligned} \kappa(\xi(\xi_0, t)) &= (e^{-t}(\xi_0 - d) + d)^\top P (e^{-t}(\xi_0 - d) + d) - 2d^\top P (e^{-t}(\xi_0 - d) + d) + d^\top P d - c \\ &= e^{-t}(\xi_0 - d)^\top P e^{-t}(\xi_0 - d) - c = e^{-2t}(\xi_0 - d)^\top P (\xi_0 - d) - c \,. \end{aligned}$$

The time $\mathfrak{t}(\xi_0)$ is the time required to reach the boundary of M as illustrated in Figure 5.19,

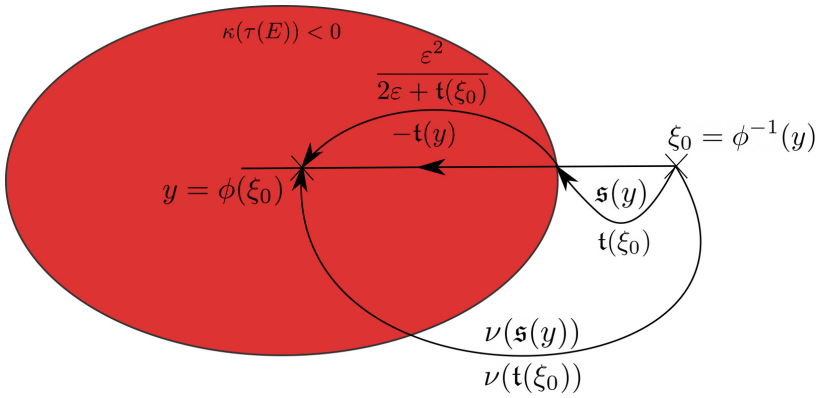


Figure 5.19: Illustration of the mappings along the flow with $\mathfrak{t}$ and $\mathfrak{s}$. With $\mathfrak{t}(\xi_0)$ the point ξ_0 is mapped to the boundary of $\kappa(E)$ and with $\nu(\mathfrak{t}(\xi_0))$ it is mapped further inside $\kappa(E)$. The mapping $\mathfrak{s}$ gives the same time as $\mathfrak{t}$ but for the final point y and not the initial point ξ_0.

i.e.

$$\kappa(\xi(\xi_0, \mathfrak{t}(\xi_0))) = e^{-2\mathfrak{t}(\xi_0)}(\xi_0 - d)^\top P(\xi_0 - d) - c \tag{5.52a}$$

$$\overset{(5.47)}{=} \frac{c}{(\xi_0 - d)^\top P(\xi_0 - d)}(\xi_0 - d)^\top P(\xi_0 - d) - c \tag{5.52b}$$

$$= 0\,. \tag{5.52c}$$

Note that $\mathfrak{t}$ exists because Condition 3 is fulfilled. The function $\nu : \mathbb{R} \to (0, \infty)$ maps every time t that is bigger than $-\varepsilon$ in a bijective manner to a higher time, i.e.

$$\nu(\mathfrak{t}) - \mathfrak{t} = \frac{(\mathfrak{t} + \varepsilon)^2}{2\varepsilon + \mathfrak{t}} - \mathfrak{t} = \frac{\mathfrak{t}^2 + 2\mathfrak{t}\varepsilon + \varepsilon^2 - 2\mathfrak{s}\varepsilon - \mathfrak{t}^2}{2\varepsilon + \mathfrak{t}} = \frac{\varepsilon^2}{2\varepsilon + \mathfrak{t}} > 0\,. \tag{5.53}$$

This expression can also be used to divide the time $\nu(\mathfrak{t}(\xi_0))$ into the time that is required to reach that boundary $\mathfrak{t}(\xi_0)$ plus the time used to map the point further into the inner of M, i.e.

$$\nu(\mathfrak{t}(\xi_0)) = \mathfrak{t}(\xi_0) + \frac{\varepsilon^2}{2\varepsilon + \mathfrak{t}(\xi_0)}\,. \tag{5.54}$$

This is also illustrated in Figure 5.19.

Note that the definitions of M and $\mathfrak{s}$ ensure that $\mathfrak{t}(y) < 0$ and $\mathfrak{s}(y) > -2\varepsilon$ for all $y \in M$. Then with (5.53) and (5.50) the following identity can be established

$$\nu(\mathfrak{s}) - \mathfrak{s} = \frac{\varepsilon^2}{2\varepsilon + \mathfrak{s}} = -\mathfrak{t}\,. \tag{5.55}$$

Further, because ξ is the solution of an ODE with a unique solution it is

$$\begin{aligned}\kappa\Big(\xi\Big(\xi\big(y,\mathfrak{t}(y)-\mathfrak{s}(y)\big),\mathfrak{s}(y)\Big)\Big) &= \kappa\Big(\xi\big(y,\mathfrak{t}(y)-\mathfrak{s}(y)+\mathfrak{s}(y)\big)\Big)\\ &= \kappa\Big(\xi\big(y,\mathfrak{t}(y)\big)\Big)\\ &= 0\,.\end{aligned}$$

Combining this result with the uniqueness of the solution ξ and the equation (5.52) leads to the identity

$$\mathfrak{t}\Big(\xi\big(y,\mathfrak{t}(y)-\mathfrak{s}(y)\big)\Big) = \mathfrak{s}(y)\,. \tag{5.56}$$

It remains to show that ϕ^{-1} is the inverse of ϕ. First, we consider all $y \in M$ with $\mathfrak{s}(y) > -\varepsilon$. This leads to

$$\begin{aligned}\phi(\phi^{-1}(y)) = \phi\Big(\xi\Big(y,-\nu\big(\mathfrak{s}(y)\big)\Big)\Big) &= \xi\Big(\xi\Big(y,-\nu\big(\mathfrak{s}(y)\big)\Big),\nu\Big(\mathfrak{t}\big(\xi(y,-\nu(\mathfrak{s}(y)))\big)\Big)\Big)\\ &\overset{(5.55)}{=} \xi\Big(\xi\Big(y,-\nu\big(\mathfrak{s}(y)\big)\Big),\nu\Big(\mathfrak{t}\Big(\xi\big(y,\mathfrak{t}(y)-\mathfrak{s}(y)\big)\Big)\Big)\Big)\\ &\overset{(5.56)}{=} \xi\Big(\xi\Big(y,-\nu\big(\mathfrak{s}(y)\big)\Big),\nu\big(\mathfrak{s}(y)\big)\Big)\\ &= y\,.\end{aligned}$$

For y such that $\mathfrak{s}(y) \leq -\varepsilon$ it is clear with Equation (5.50) that $\mathfrak{t}(y) \leq -\varepsilon$ and therefore

$$\phi(\phi^{-1}(y)) = \phi\Big(\xi\Big(y,-\nu\big(\mathfrak{s}(y)\big)\Big)\Big) = \phi\Big(\xi\big(y,0\big)\Big) = \phi(y) = \xi(y,0) = y\,.$$

This shows that ϕ^{-1} is the right inverse of ϕ.

It remains to show that ϕ^{-1} is also the left inverse. Using that ξ is the solution of an ODE with a unique solution leads for $\mathfrak{t}(\xi_0) > -\varepsilon$ to

$$\mathfrak{t}\Big(\xi\Big(\xi_0,\mathfrak{t}(\xi_0)+\frac{\varepsilon^2}{2\varepsilon+\mathfrak{t}(\xi_0)}\Big)\Big) = \mathfrak{t}\Big(\xi\Big(\xi(\xi_0,\mathfrak{t}(\xi_0)),\frac{\varepsilon^2}{2\varepsilon+\mathfrak{t}(\xi_0)}\Big)\Big) \tag{5.57a}$$

$$= -\frac{\varepsilon^2}{2\varepsilon+\mathfrak{t}(\xi_0)} \tag{5.57b}$$

where the last equality holds because $\xi(\xi_0,\mathfrak{t}(\xi_0))$ is at the boundary of M, i.e $\kappa\Big(\xi\big(\xi_0,\mathfrak{t}(\xi_0)\big)\Big) = 0$.

This allows to obtain the following equality for $\mathfrak{s}$:

$$\mathfrak{s}\Big(\xi\Big(\xi_0, \nu(\mathfrak{t}(\xi_0))\Big)\Big) = \mathfrak{s}\Big(\xi\Big(\xi_0, \mathfrak{t}(\xi_0) + \frac{\varepsilon^2}{2\varepsilon + \mathfrak{t}(\xi_0)}\Big)\Big) \tag{5.58a}$$

$$\overset{(5.50)}{=} \frac{-\varepsilon^2}{\mathfrak{t}\Big(\xi\Big(\xi_0, \mathfrak{t}(\xi_0) + \frac{\varepsilon^2}{2\varepsilon+\mathfrak{t}(\xi_0)}\Big)\Big)} - 2\varepsilon \tag{5.58b}$$

$$\overset{(5.57)}{=} \frac{-\varepsilon^2}{-\frac{\varepsilon^2}{2\varepsilon+\mathfrak{t}(\xi_0)}} - 2\varepsilon \tag{5.58c}$$

$$= \mathfrak{t}(\xi_0)\,. \tag{5.58d}$$

Then it is clear that for ξ_0 such that $\mathfrak{t}(\xi_0) > -\varepsilon$ holds

$$\begin{aligned}
\phi^{-1}(\phi(\xi_0)) &= \phi^{-1}\Big(\xi(\xi_0, \nu(\mathfrak{t}(\xi_0)))\Big) \\
&= \xi\Big(\xi\Big(\xi_0, \nu(\mathfrak{t}(\xi_0))\Big), -\nu\Big(\mathfrak{s}(\xi(\xi_0, \nu(\mathfrak{t}(\xi_0))))\Big)\Big) \\
&\overset{(5.58)}{=} \xi\Big(\xi(\xi_0, \nu(\mathfrak{t}(\xi_0))), -\nu\big(\mathfrak{t}(\xi_0)\big)\Big) \\
&= \xi_0
\end{aligned}$$

For ξ_0 such that $\mathfrak{t}(\xi_0) \leq -\varepsilon$ it is clear with Equation (5.50) that $\mathfrak{s}(\xi_0) \leq -\varepsilon$ and consequently

$$\begin{aligned}
\phi^{-1}(\phi(\xi_0)) &= \phi^{-1}(\xi_0) \\
&= \xi\Big(\xi_0, -\nu\big(\mathfrak{s}(\xi_0)\big)\Big) \\
&= \xi(\xi_0, 0) \\
&= \xi_0
\end{aligned}$$

This shows that ϕ^{-1} is the inverse of ϕ. □

The presented theorem gives a method with specific choices to shape the invariant set and to define the flow. This reduces the design effort to the choice of a suitable matrix positive definite matrix P, a vector d and a small constant ε. However, the more generalised approach proposed in [19] gives more design freedom at the cost of solving differential as well as nonlinear equations as stated in the next remark.

Remark 14. The approach can be adapted to use different κ, χ or ν. A different choice of these functions leads to the following steps that need to be carried out.

1. Find ξ by calculating the solution of the ordinary differential equation

$$\dot{\xi} = \chi(\xi)\,. \tag{5.59}$$

2. Find $\mathfrak{t}$ by solving the nonlinear equation

$$\kappa(\xi(\xi_0, \mathfrak{t}(\xi_0))) = 0 \tag{5.60}$$

3. Find $\mathfrak{s}$ by solving the nonlinear equation

$$\nu(\mathfrak{s}) - \mathfrak{s} = -\mathfrak{t}. \tag{5.61}$$

A small discrepancy to the original statement is pointed out in the next remark.

Remark 15. The used χ is not bounded as assumed in the original statement. Boundedness is a sufficient assumption that shall guarantee the desired behaviour of the solution. The solution (5.45) has this desired behaviour despite the absence of this property.

With the definition of this image extending function ϕ, it is possible to state the main theorem of the section, which allows to use an observer in transformed coordinates to state an observer in natural coordinates which does not suffer from the singularity problems in the transient phase as described in the introduction.

Theorem 5.13. *Let $E \subset \mathbb{R}^n$ be an open set, and $K \subset E$ be compact. Further, let the solutions of the system*

$$\dot{x} = f(x) \tag{5.62a}$$
$$y = h(x) \tag{5.62b}$$

be $x(t) \in K^\circ$ for all $t \in \mathbb{R}$. Suppose there exists a vector field F and a diffeomorphism $\mathcal{O} : E \to \mathcal{O}(E)$ such that the system

$$\dot{\hat{z}} = F(\hat{z}, y) \tag{5.63a}$$
$$\hat{x} = \mathcal{O}^{-1}(z) \tag{5.63b}$$

is an observer of the system (5.62). Further let $\phi : \mathbb{R}^n \to M$ with $M \subset \mathcal{O}(E)$ be a $\mathcal{C}^1$-diffeomorphism with $\phi(x) = x$ for $x \in K$. Then the system

$$\dot{\hat{x}} = \frac{\partial \mathcal{O}(\hat{x})}{\partial x}^{-1} \frac{\partial \phi(\phi^{-1}(\mathcal{O}(\hat{x})))}{\partial \xi} F(\phi^{-1}(\mathcal{O}(\hat{x})), y) \tag{5.64}$$

is an observer of the system (5.62) and $\hat{x}(t) \in E$ for all t.

Proof. Consider the transformation $T = \phi^{-1} \circ \mathcal{O}$ and the transformed state $\hat{\xi} = T(\hat{x})$. Then,

using (5.64) the resulting dynamics have the form

$$\begin{aligned}\dot{\hat{\xi}} &= \frac{\partial \phi^{-1}(\mathcal{O}(\hat{x}))}{\partial z}\frac{\partial \mathcal{O}(\hat{x})}{\partial x}\dot{\hat{x}} \\ &= \frac{\partial \phi(\phi^{-1}(\mathcal{O}(\hat{x})))}{\partial \xi}^{-1}\frac{\partial \mathcal{O}(\hat{x})}{\partial x}\frac{\partial \mathcal{O}(\hat{x})}{\partial x}^{-1}\frac{\partial \phi(\phi^{-1}(\mathcal{O}(\hat{x})))}{\partial \xi}F(\phi^{-1}(\mathcal{O}(\hat{x})), y) \\ &= F(\hat{\xi}, y)\,.\end{aligned}$$

These dynamics have the same form as (5.63). Because this system is an observer of (5.62) it holds

$$\lim_{t\to\infty} \mathcal{O}^{-1}\Big(\hat{\xi}(t)\Big) - x(t) = 0\,. \tag{5.65}$$

Further, due to this convergence and since $x(t) \in K^\circ$ and K is compact there exists $T \in \mathbb{R}$ such that $\mathcal{O}^{-1}(\hat{\xi}(t)) \in K$ for all $t \geq T$. Consequently, it is $\hat{\xi}(t) \in \mathcal{O}(K)$ for all $t \geq T$. Then with $\hat{x} = T^{-1}(\hat{\xi}) = \mathcal{O}^{-1}(\phi(\hat{\xi})) = \mathcal{O}^{-1}(\hat{\xi})$ for $\hat{\xi} \in \mathcal{O}(K)$ and (5.65) it is shown that the system is an observer of system (5.62). The identity $\hat{x} = \mathcal{O}^{-1}(\phi(\hat{\xi}))$ ensures by the definition of ϕ and $\mathcal{O}$ that $\hat{x}(t) \in E$ for all t. □

While this design is a very helpful tool as also highlighted in the next section, some of its problems that might occur in practical implementation are mentioned here. Firstly, conventional approaches define $\mathcal{O}$ as the derivatives of the output to obtain the system (5.63). In this case the expression $\mathcal{L}_f^n h(x)|_{x=\mathcal{O}^{-1}(\xi)}$ is needed to define F. For the observer in natural coordinates this takes the form $\mathcal{L}_f^n h(x)|_{x=\mathcal{O}^{-1}(\phi(\mathcal{O}(\hat{x})))}$. Usually the function $\mathcal{O}^{-1}$ is not known. Thus this term may be either neglected or the approximation $\mathcal{O}^{-1}(\phi(\mathcal{O}(\hat{x}))) \approx \hat{x}$ may be used. The latter approximation is exact for $\hat{x} \in K$ and the quality at the complement depends heavily on $\mathcal{O}$. Secondly, it is very natural that from a theoretic point of view it is desired to choose ε as small as possible as this leads to a smaller transition set. However, in practice a smaller choice of ε leads to a smaller transition set which requires a higher numerical accuracy for the calculations. Finally, it shall be emphasised that the proposed method does not solve the problem that occurs if the real trajectory crosses one of the singularity boundaries. If this happens the main challenge is to find the E in which the current trajectory lies.

We continue by applying the proposed method to our attitude estimation problem.

5.4.2 Simulation

In this section we apply the proposed theory to our system of interest

$$\dot{x} = f(x)$$

with $x = [T, \theta, \vartheta_1, \vartheta_2, \omega, t]$ and $f(x) = [f_T(T, l_2^{-r,n}(\theta, \vartheta_1, \vartheta_2)), f_{\theta,\vartheta_1,\vartheta_2}(\theta, \vartheta_1, \vartheta_2, \omega), f_\omega(\omega), 1]$ where f_T, f_ω are defined as in Equation (5.1) and $f_{\theta,\vartheta_1,\vartheta_2}$ is defined as the right hand side of Equation (5.31). We consider the system to be in Case 4, i.e. only infrared irradiation

is acting on the surface and in particular the angle between Sun and normal $\phi(s,n)$ is not of interest and we can define ϕ as in the previous section and have no ambiguities. The observability mapping $\mathcal{O}^{\vartheta}$ has the form

$$\mathcal{O}^{\vartheta}(x) = V^{\top}(4)\mathcal{O}(T, l_2^{-r,n}(\theta,\vartheta_1,\vartheta_2),\omega,t)$$

$$\frac{\partial \mathcal{O}^{\vartheta}(x)}{\partial x} = V^{\top}(4)\frac{\partial \mathcal{O}(T, l_2^{-r,n}(\theta,\vartheta_1,\vartheta_2),\omega,t)}{\partial(T,q,\omega,t)} \begin{bmatrix} 1 & 0_{1,3} & 0_{1,3} & 0_{1,1} \\ 0_{4,1} & \frac{l_2^{-r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial(\theta,\vartheta_1,\vartheta_2)} & 0_{4,3} & -\frac{l_2^{-r,n}(\theta,\vartheta_1,\vartheta_2)}{\partial r}\dot{r} \\ 0_{3,1} & 0_{3,3} & I_{3,3} & 0_{3,1} \\ 0_{1,1} & 0_{1,3} & 0_{1,3} & 1 \end{bmatrix}$$

with $\mathcal{O}$ defined as in (5.8) and $V^{\top}(4)$ defined as in (A.49) which removes the quaternion constraint of the mappings.

The observer is designed as in Equation (5.64) but only ϑ_1, ϑ_2 are estimated by the proposed observer design. It has the form

$$\dot{\hat{x}} = \frac{\partial \mathcal{O}^{\vartheta}(\hat{x})}{\partial x}^{-1} M^{\text{obs}} \left(\begin{bmatrix} 0_{3,1} & I_{3,3} & 0_{3,3} & 0_{3,1} \\ 0_{1,1} & 0_{1,3} & 0_{1,3} & 0_{1,1} \\ 0_{3,1} & 0_{3,3} & 0_{3,3} & 0_{3,1} \\ 0_{1,1} & 0_{1,3} & 0_{1,3} & 0_{1,1} \end{bmatrix} \xi + \begin{bmatrix} 0_{3,1} \\ 0_{1,1} \\ -J^{-1}(\omega \times J\omega) \\ 1 \end{bmatrix} + L \begin{bmatrix} T-\hat{T} \\ w-\hat{w} \\ t-\hat{t} \end{bmatrix} \right) \tag{5.66}$$

where the state ξ is identical to the state z of the common observer design apart from the third and fourth state which are shaped by the design to avoid singularities, i.e.

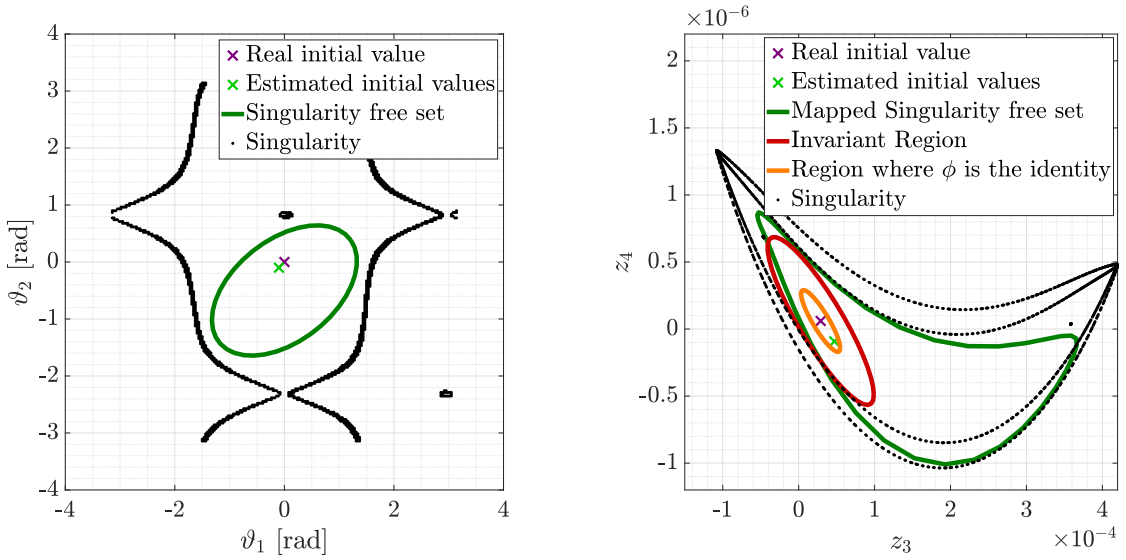
$$\xi = \left[V^{\perp}(1,2)^{\top}\mathcal{O}^{\vartheta}(\hat{x})\ \phi^{-1}\left(V^{\perp}(3,4)^{\top}\mathcal{O}^{\vartheta}(\hat{x})\right)\ V^{\perp}(5,6,7)^{\top}\mathcal{O}^{\vartheta}(\hat{x})\ V^{\perp}(8)^{\top}\mathcal{O}^{\vartheta}(\hat{x}) \right] \tag{5.67}$$

with the function ϕ defined as in Theorem 5.12 in Equation 5.48 with the parameters $\varepsilon = 1$, $P = \begin{pmatrix} 0.9\cdot 10^{-4} & 8.7\cdot 10^{-3} \\ 8.7\cdot 10^{-3} & 1.1 \end{pmatrix} \cdot 10^{13}$, $d = \left(0.3\ 4\cdot 10^{-4}\right)\cdot 10^{-4}$ and $c = 1$. The matrix $V^{\perp}(1,2)^{\top}$ keeps the first two entries of the multiplied vector. The matrix M^{obs} is the identity matrix but for one block that contains the derivative of ϕ, i.e.

$$M^{\text{obs}} = \begin{bmatrix} I_2 & 0_{2,2} & 0_{2,3}\ 0_{1,1} \\ 0_{2,2} & \frac{\partial \phi\left(\phi^{-1}\left(V^{\perp}(3,4)^{\top}\mathcal{O}^{\vartheta}(\hat{x})\right)\right)}{\partial \xi} & 0_{2,3}\ 0_{1,1} \\ 0_{3,2} & 0_{3,2} & I_{3,3}\ 0_{3,1} \\ 0_{1,2} & 0_{1,2} & 0_{1,3}\ I_{1,1} \end{bmatrix}.$$

The observer matrix L is defined to be of high gain with the parameters stated in the introduction of Section 5.3. It has the form

$$L = \begin{bmatrix} 16 & 0_{1,4} \\ 89 & 0_{1,4} \\ 194 & 0_{1,4} \\ 120 & 0_{1,4} \\ 0_{4,1} & I_{4,4} \end{bmatrix}.$$

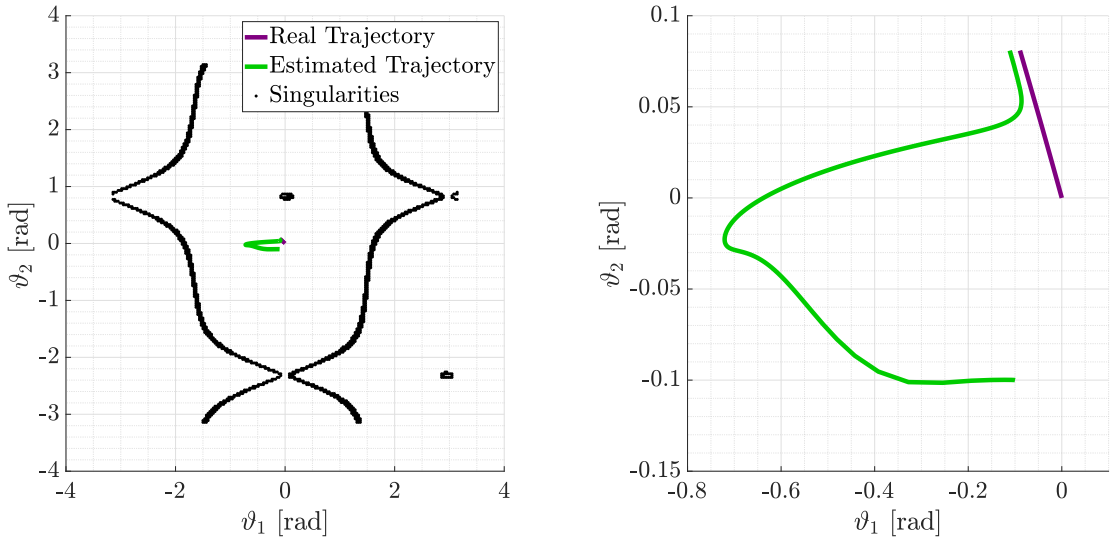


(a) ϑ-space. (b) z-space.

Figure 5.20: Left Figure displays ϑ space with an exemplary choice of a set without singularities. The right figure shows the z-space with the projected singularity free set and the choice of the designed invariant region. The space between the red and the orange ellipse is the layer in which all points in the ξ-plane which are not in the z-plane are projected into.

In order to evaluate the results, we distinguish between the three coordinates introduced in the previous section. The natural coordinates x are the coordinates of ϑ_1, ϑ_2 which we are actually interested in. We denote their domain E as the ϑ-space. The coordinates z result from the transformation of x using the observability mapping $\mathcal{O}^{\vartheta}$. The observer is designed such that these coordinates are always inside the desired set $\mathcal{O}(E)$. We denote this domain as the z-space. The final coordinates ξ are defined using the inverse of ϕ on z. These coordinates are defined on the whole of $\mathbb{R}^n$ and its domain is denoted as the ξ-space. The z-space is a subset of the ξ-space which is why we show the z and ξ coordinates in the same figure.

The first difficulty when designing the observer is finding an appropriate subset $M \subset \mathcal{O}(E)$. Ideally, E or an appropriate subset of E is known that allows a straightforward description of M. Unfortunately, this task can be quite difficult as illustrated in Figures 5.20. The left figure shows the ϑ-space with the singular points and the real and estimated initial values. The designed set must contain at least both of these values and their evolution over time. A simple choice of this set is displayed by the ellipse. The right plot shows the transformation of this set into the z-space. It can be seen that the transformed set is not free of singularity points any more and that the resulting set cannot be easily described by a quadratic function. Instead, we design the quadratic set M directly in the z-space as displayed. From this set M, a layer needs to be defined using the parameter ε in which all points outside will be projected into. Naturally, the estimate and the real trajectory need to be inside M without this layer to obtain an interpretable result. Theoretically, the layer can be made arbitrarily



(a) Complete ϑ-space.

(b) Subset of the ϑ-space.

Figure 5.21: Estimated and real solution in natural coordinates.

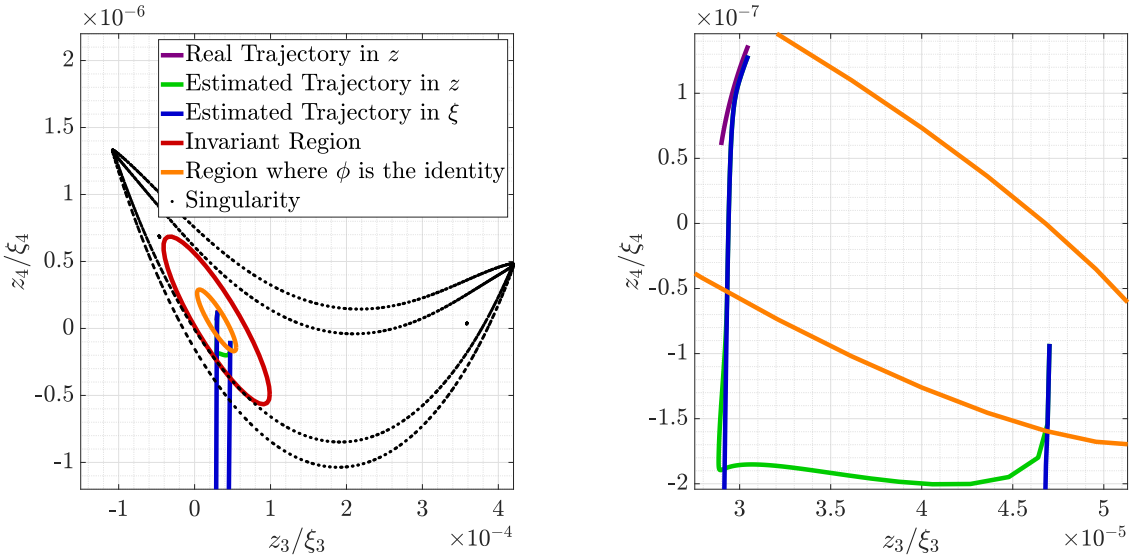
small while guaranteeing the validity of the algorithm. In practice however, a layer that is too small leads to numerical issues, thus we have chosen a conservative layer suited for this scenario.

The real and estimated trajectories in natural coordinates are shown in Figure (5.21). It can be seen that in comparison to the original observer design displayed in Figure (5.17), this time the estimated trajectory does not cross the singularity boundary. On the contrary, due to the design the commonly occurring peaking phenomenon for high gain observers is attenuated. It can be seen that after some time, the estimate follows the real trajectory closely with a small steady state error. This is a significant improvement over the original observer design and deems the invariant observer design to be valuable in these kind of scenarios.

In order to deepen the understanding of the proposed observer, Figure (5.22) shows the real and estimated states in z as well as ξ coordinates. As expected, the ξ trajectory exhibits the standard peaking phenomenon as expected for a high gain observer. It leaves the z-space and returns after a while. The z trajectory however stays always within the invariant region. If the ξ trajectory is outside of the invariant region, the z trajectory is inside the layer of the invariant region. When the ξ trajectory returns inside the invariant set it converges to the z trajectory until it is outside the layer and in the invariant set. Then z and ξ are identical.

5.5 Conclusions

Observer design for the attitude estimation based on a single temperature and angular velocity measurements is a difficult task. The state space needs to be divided to identify the



(a) Complete z-space displaying the real and estimated solution.

(b) Subset of the ϑ-space displaying the real and estimated solution.

Figure 5.22: Estimated and real solution in z and ξ coordinates

locally weakly observable points. We analyse two observer designs, one in transformed and one in natural coordinates.

The observer in transformed coordinates relies on a novel angle transformation and estimates the attitude based on an approximation of the measured temperature and its first three order derivatives. This method relies on an optimisation which leads due the non-global observability to ambiguities which can only be avoided using a-priori information.

The second observer design in natural coordinates uses the inverse of the Jacobian of the observability mapping instead of an optimisation. This method requires an additional function which ensures that the estimate does not cross the manifold of singularities. It avoids the time-consuming issue of optimisation but requires a smaller step size as the proposed invariant observer design is numerically more sensitive than the first design.

In the current state, for both observer designs, the required accuracy of the sensors as well as the computational power is very high which suggests that using additional temperature sensors as analysed in the next chapter should be favoured.

6 Observer Based Attitude Estimation with Twelve Temperature Measurements

In this chapter we alter the problem introduced in the previous chapter by introducing eleven additional temperature measurements. In detail, we consider a spacecraft with twelve sensors pointing along the body axes. Along each body axis there are are two sensors with different physical parameters. This allows the usage of a single temperature derivative, instead of three as in the previous chapter, to obtain an attitude estimation. The algorithm proposed in this chapter will be validated on real data acquired from Coarse Earth Sun Sensors (CESS) discussed in 1.1. Much of the notations and algorithms proposed in this section are chosen to allow a simple identification with the considered GRACE data. Thus we start by introducing the reference data set and make slight modifications to the model proposed in Section 4 to allow a proper representation of the data after fitting the models. Then an observer is developed in transformed coordinates and its pendant in natural coordinates is analysed.

6.1 GRACE Mission Data Sets

The reference mission used throughout this chapter is the Gravity Recovery and Climate Experiment (GRACE) [126–129]. The mission launched in march 2002 consists of two twin satellites which took detailed measurements to improve the model of the Earth's gravity field . The author kindly acknowledges the organisation unit Mission Operations (MIB-LEM OP) of the German Aerospace Center for providing the data sets used in this section.

The data sets considered consist of the time t when a measurement is obtained, the corresponding position and velocity of the spacecraft r and v, the current attitude of the spacecraft q and twelve temperature measurements T. In this section we present relevant parts of the received data to provide an overview of the interesting dynamics.

Table 6.1: Initial spacecraft position, spacecraft velocity and Sun position [all SI units].

Variable	Value	Variable	Value
r_0^{SM}	$\begin{bmatrix}-1.9265 & -1.8703 & -6.2987\end{bmatrix}^\top \cdot 10^6$	v_0^{SM}	$\begin{bmatrix}-4.79 & -5.11 & 2.98\end{bmatrix}^\top \cdot 10^3$
r_0^{Slew}	$\begin{bmatrix}3.4998 & -4.8428 & 3.2395\end{bmatrix}^\top \cdot 10^6$	v_0^{Slew}	$\begin{bmatrix}2.27 & -2.89 & -6.72\end{bmatrix}^\top \cdot 10^3$
s^{SM}	$\begin{bmatrix}1.4779 & 0.1984 & 0.0860\end{bmatrix}^\top \cdot 10^{11}$	s^{Slew}	$\begin{bmatrix}1.4029 & 0.4844 & 0.2100\end{bmatrix}^\top \cdot 10^{11}$

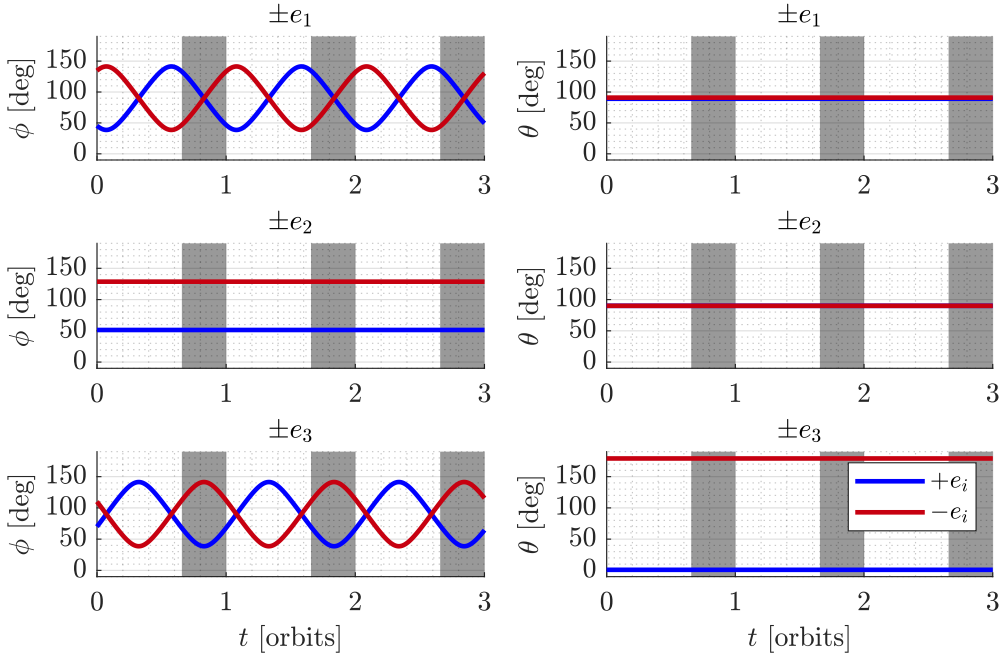


Figure 6.1: Angles of the spacecraft for the first date. The periods with black background illustrate the shadow phases.

Data from two different dates, namely 2012/03/29 00:00:00 UT and 2014/04/11 00:00:00 UT is considered. In the following we denote all data from the first date by a superscript $\cdot^{\mathrm{SM}}$ while data from the second date is denoted by $\cdot^{\mathrm{Slew}}$. Both data series have a sample rate of 1 s. This allows a direct derivation of the corresponding time vectors t^{SM} and t^{Slew} in seconds.

The corresponding orbit is a low Earth orbit as it has a semi major axis of approximately 6800 km, i.e. an altitude of about 450 km. This leads to a period of the orbit T^{Orbit} of approximately 93.0 minutes. Further, the orbit is almost circular with an eccentricity of 0.003 and it is a polar orbit with an inclination of about 89 degrees. For the longitude of the ascending node (LTAN) holds $\Omega^{\mathrm{SM}} = -131.78$ and $\Omega^{\mathrm{Slew}} = 126.39$. This shows that the considered orbit is not Sun-synchronous as the LTAN is very different for two times of the year that are very close together. Consequently, the eclipse phases of this orbit vary over the years. For the first date, the eclipse phase is approximately a third of the orbit duration long, while at the second date no eclipse occurs in the following orbits. The initial positions and velocities of the spacecraft and the position of the Sun in ECI can be found in Table 6.1.

The attitude of the spacecraft is defined by the mode the spacecraft is in. In science mode the attitude is described by the Orbit frame defined in Section 3.1.3 with an additional -0.92 degree pitch. The second mode we consider in this work is the slew mode, in which the spacecraft performs a slew manoeuvre over the course of 107.2 minutes. During the slew manoeuvre the spacecraft rotates 90 degrees around its body z-axis and remains in the new

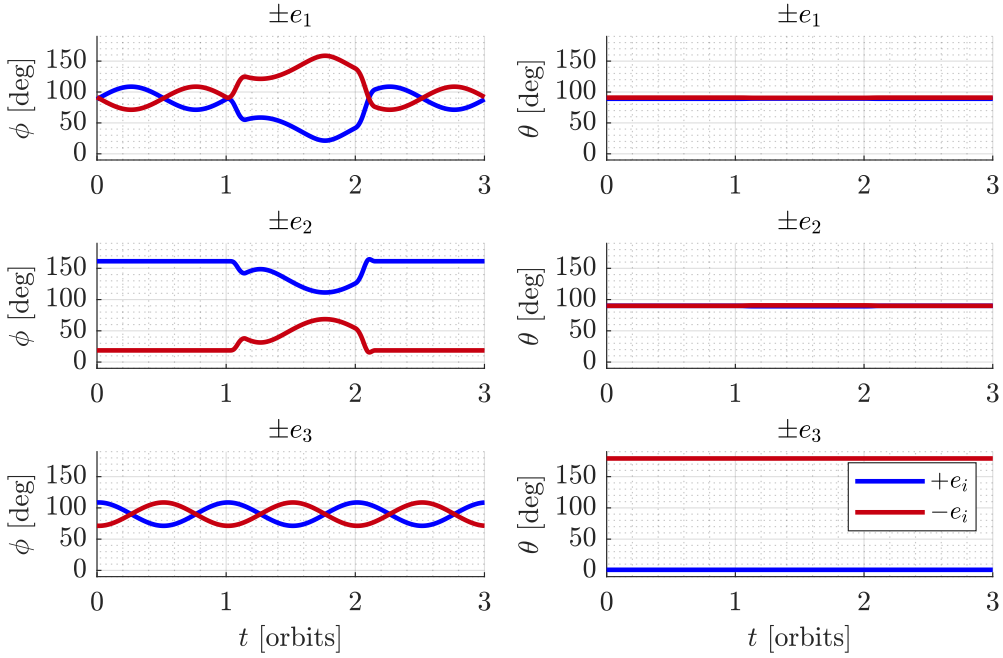


Figure 6.2: Angles of the spacecraft for the second date.

attitude for 78.8 minutes and then rotates back −90 degrees. For the measurements from the first date, the spacecraft is always in science mode. For measurements from the second date a time frame can be identified in which the slew mode occurs. In the remaining time the spacecraft is in science mode. The attitude for these two time series is illustrated by the means of their angles to Sun and Earth in Figure 6.1 and 6.2. The time is normalised such that the x-axis displays the time passed in orbits, i.e. $t^{\text{Orbit}} = \frac{t-t_0}{T^{\text{Orbit}}}$ where t_0 is chosen such that the eclipse phase ends at this point of time and T^{Orbit} is the orbit period. Figure 6.1 shows the angles for the first date. It consists of two columns, each consisting of three plots. Every plot displays two angles for normals in opposing unit directions. Note that there is always a simple connection between two angles for normals pointing into opposing direction. One results from the other by subtracting the angle from 180 degrees. The left column displays the ϕ angles, the right column displays the θ angles. In the first plot in the left column we can see the angle to the Sun for the normal vector pointing in and oppose the velocity direction. It can be seen that both angles can be approximately defined by a sinus function with a mean of 90 degrees and an amplitude of approximately 50 degrees. Their phase is shifted with exactly 180 degrees. The second plot in the left column shows the angles for the two normal vectors pointing perpendicular to the orbit. The angles are constant over the course of multiple orbits, due to the position of the Sun varying very slowly in this time frame. One angle is about 130 degrees while the other one is approximately 50 degrees. The last side is commonly referred to as the side pointing towards the Sun. The third plot in the left column shows the angles to the Sun of the zenith and nadir pointing side. These have a similar behaviour as the sides pointing into velocity direction. In the second column the angles between the vector to Earth and each surface normal can be seen.

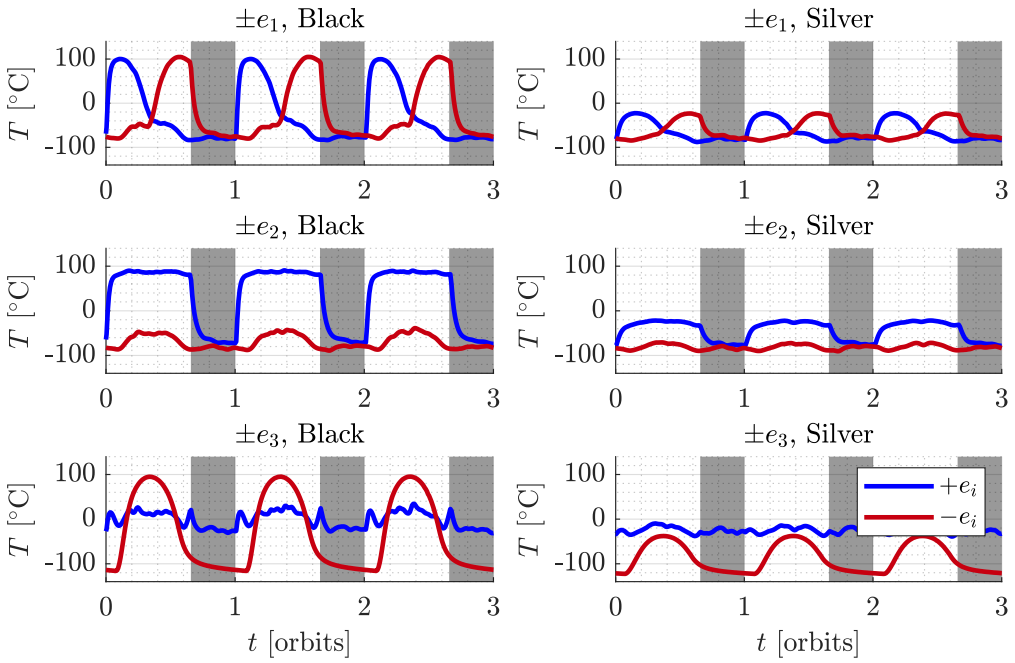


Figure 6.3: Temperature measurements $T_n^{\mathrm{SM},j}$ of the spacecraft for the first date. The periods with black background illustrate the shadow phases.

These are all constant because the spacecraft is controlled to follow the orbit frame 3.1. Consequently, the angles of the sides pointing nadir and zenith have approximately 0 and 180 degrees, respectively. The angles of the other four sides are all about 90 degrees. Note that small deviations to these values can be seen because of the additional -0.92 degrees pitch.

Figure 6.2 shows the angles in the same manner for the measurements from the second date. The time is normalised such that the x axis displays the time passed in orbits where the initial time t_0 is chosen such that the slew manoeuvre starts exactly one orbit later. In science mode, the qualitative behaviour is identical but the angles pointing to Sun differ in their numerical values. For the normals in and oppose velocity direction as well as in zenith and nadir direction, the amplitude is only about 20 degrees. The normal pointing towards the Sun has now an angle of about 20 degrees while the one pointing away from it has about 160 degrees. The angles to Earth are basically identical to the first date. The solar angles become interesting for the normals in and oppose to velocity direction during the slew manoeuvre. The side along $-e_1$ makes a quick motion towards the Sun, then points gradually away and then back to it until the angle reaches its minimum. This behaviour is mirrored by the side pointing into $+e_1$ direction. The sides pointing perpendicular to the orbit endure a qualitatively similar motion.

The temperatures are obtained from the coarse Earth sun sensors (CESS). Every sensor consists of two surfaces, one reflective silver surface and one black surface leading to two temperature measurements per sensor. These surfaces have different thermal coefficients α

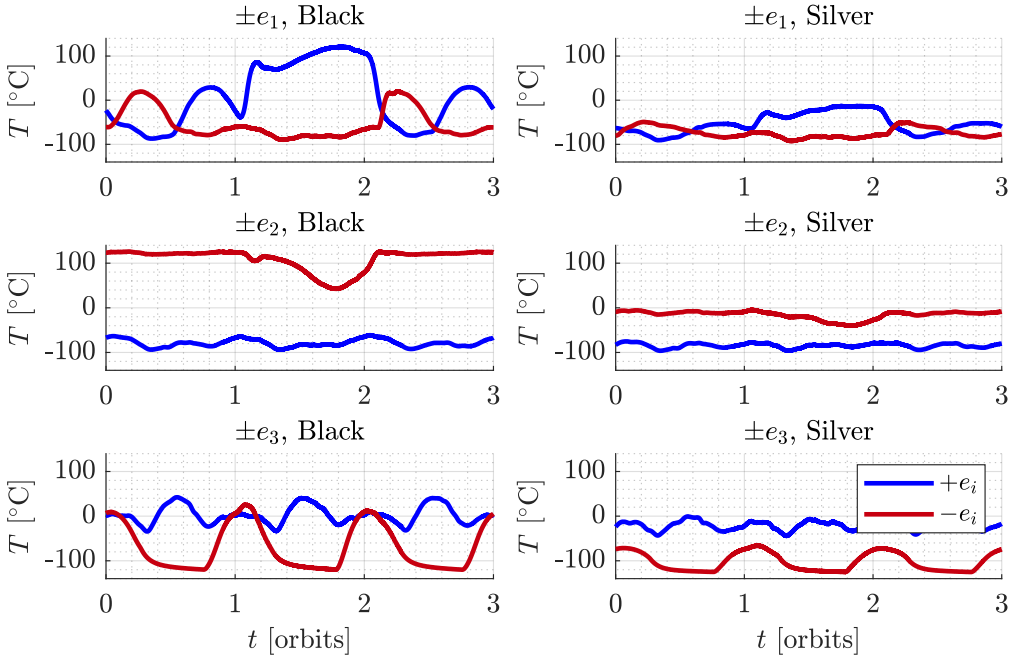


Figure 6.4: Temperature measurements $T_n^{\text{Slew},j}$ of the spacecraft for the second date.

and ε. A temperature measurement of a black surface will be denoted by the superscript $\cdot^{\text{B}}$ while one of a silver surface will be denoted by $\cdot^{\text{S}}$. One sensor is located at each side of the spacecraft, i.e. $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. We denote a temperature measurement of the location n with a subscript $\cdot_n$. In total there are 24 different temperature trajectories $T_n^{i,j}$ with $i \in \{\text{SM}, \text{Slew}\}$, $j \in \{\text{B}, \text{S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$.

Figure 6.3 shows the twelve temperature measurements over the course of three orbits for the first measurements series. Each column consist of three plots, each containing two temperature measurements with opposing normals. The left column shows the temperature for the black sensors and the right column shows the temperature for the silver sensors. Three orbits are displayed and it can be seen that the measurements are almost periodic with a period of one orbit. The first plot in the left column shows the temperature of the black sensor pointing in and oppose to the velocity direction of the spacecraft. Both sensors show similar qualitative behaviour. Under the influence of solar irradiation the temperature rises with the solar angle to its maximum of around 100 degrees Celsius and then drops back to approximately -75 degrees Celsius when no solar irradiations is acting any more. This behaviour results mostly from the sinusoidal evolution of the angles. Note that the eclipse starts after two thirds of the orbit and therefore no solar irradiation is acting in this phase despite some of the angles to the Sun being smaller than 90 degrees. The second plot shows the temperature of the black sensor pointing towards and oppose the orbit normal. The temperature of the sensor with the normal pointing toward the Sun is mostly constant except for the dynamic change of value when the spacecraft enters and leaves eclipse. This suggests that the influence of the solar irradiation is dominant on this side. For the oppose pointing side, no solar irradiation is acting. Here the temperature evolution

is dominated by the albedo and the infrared irradiation. The albedo irradiation occurring outside of the solar eclipse leads to the rise in temperature at the beginning of each orbit. Oscillations occur because of varying albedo and infrared irradiation factors. These depend on the current cloud coverage and the terrain the spacecraft passes. The third plot on the left column shows the evolution of the measured temperature from the sensors pointing towards zenith and nadir. For the nadir pointing sensor e_3 it can be seen that the influence of solar irradiation is sparse and only occurring at the very beginning and near the end of the sunlit phase. This leads to the clearly distinguishable peaks in these time periods. Over the course of the complete orbit, infrared as well as albedo irradiation is acting on the sensor if the spacecraft is not in eclipse. This leads to a highly time variant temperature signal due to the changing surface of the Earth. A more distinct temperature evolution can be observed by the sensor pointing towards zenith $-e_3$. Here, only solar irradiation is acting and the sinusoidal influence of the solar angle can be seen. The temperature evolutions for the silver sensors are qualitatively almost identical. The main difference lies in the amplitude of the temperatures as the influence of the solar irradiation on the temperature is diminished.

Figure 6.4 shows the evolution of the temperature from the second date. The first orbit is in science mode and the measurements have a similar behaviour as from the first date. The main difference is that during the second date, no solar eclipse occurs. Therefore, no significant temperature change for the normal pointing towards the Sun $-e_3$ can be witnessed. From the second orbit on, the slew manoeuvre starts. Here, the temperature evolution follows roughly the evolution of the solar angles ϕ during the slew manoeuvre. Afterwards, the temperature returns to its periodic behaviour in science mode.

This completes all the data provided from the GRACE mission for this work. We will use this data in the upcoming sections to fit the proposed model and validate the results.

6.2 Model Augmentation and Parameter Identification

In this section we adapt the model from Section 4 and identify the parameters required to estimate the dynamics of the measured temperature data. The temperature dynamics of a single of these twelve sensors is governed by Equation (4.11). It is required to adjust the model by defining the terms I_{IR} and ρ_{alb} as functions of the position $r(t)$ rather than as constants in Equation (4.11). The necessity of this will be verified in the upcoming section. We use the Tables 1 and 2 from [130] to implement a position dependent infrared and albedo coefficient. This leads to the new dynamics

$$\begin{aligned}\dot{T} = &\max\left(\frac{1}{C}\alpha_{\mathrm{s}} G_{\mathrm{s}} A_{\mathrm{s}} \nu(r(t)) \cos(\phi_n(q)), 0\right) + \max\left(\frac{1}{C}\rho_{\mathrm{alb}}(r(t)) G_{\mathrm{s}} A_{\mathrm{s}} \cos(\xi(t)) F(\theta_n(t,q), r(t)), 0\right) \\ &+ \max\left(\frac{1}{C}\alpha_{\mathrm{s}}\varepsilon_{\mathrm{e}} A_{\mathrm{s}} I_{\mathrm{IR}}(r(t)) F(\theta_n(t,q), r(t)), 0\right) - \frac{1}{C}\varepsilon_{\mathrm{e}} A_{\mathrm{s}} \sigma T^4 + \frac{Q^{\mathrm{d}}}{C}\end{aligned} \tag{6.1}$$

where ϕ_n and θ_n are defined as in Equation (4.9). The additional subscript $\cdot_n$ shall highlight the dependence on the normal vector n. In this section, we consider the attitude q of the system to be known and use it with the measured temperatures and position to identify the

other variables depending on the spacecraft and its thermal topology. More precisely, we determine the optimal numerical values of the five parameters C, α_s, ε_e, A_s and Q^d based on the temperature data.

In order to reduce the amount of parameters to be estimated, it is important to determine which parameters are redundant in the system dynamics sense. It can readily verified that the parameter tupel $(C, \alpha_\mathrm{s}, \varepsilon_\mathrm{e}, A_\mathrm{s}, Q^\mathrm{d})$ and $(abC, a\alpha_\mathrm{s}, a\varepsilon_\mathrm{e}, bA_\mathrm{s}, abQ^\mathrm{d})$ yield the same dynamics for $a, b \in \mathbb{R} \setminus \{0\}$. Consequently, two of these parameters can be arbitrarily chosen for the estimation process. We choose A_s and ε_e to allow an easier comparison of the dynamics and the thermal properties. The values are set to be $A_\mathrm{s} = 1$ and $\varepsilon_\mathrm{e} = 0.8$.

We denote the right hand side of (6.1) by $f(T, q, r, n, \alpha_\mathrm{s}, C, Q^\mathrm{d})$ which explicitly incorporates all states, variables and remaining parameters of the system. To reduce the amount of indices used we write α for α_s and Q for Q^d for the remainder of this section. This is unambiguous as the definition of α in Equation (4.8) is not used in this section. For the first and the second date the spacecraft follows two different position trajectories which we denote by r^SM and r^Slew. Further, the attitudes are different for the two dates and thus we have the two orientations q^SM and q^Slew. The spacecraft has six sensors, each normal n oriented along one of the six unit vectors. Sensor with black and silver coating differ in the parameters that describe the absorption and the emissivity, which we will denote by α^B, α^S and ε^B, ε^S. In order to drop some of the argument to allow a more compact notation, we define the dynamics of $\hat{T}_n^{i,j}$ as $\dot{\hat{T}}_n^{i,j} = f_n^{i,j}(\hat{T}_n^{i,j}, t, c_n^{i,j})$ with $f_n^{i,j}(\hat{T}_n^{i,j}, t, c_n^{i,j}) := f(\hat{T}_n^{i,j}, q^i, r^i, n, \alpha_n^{i,j}, C_n^{i,j}, Q_n^{i,j})$ with the parameter vector $c_n^{i,j} = (\alpha_n^{i,j}, C_n^{i,j}, Q_n^{i,j})$ for $i \in \{\mathrm{SM}, \mathrm{Slew}\}$, $j \in \{\mathrm{B}, \mathrm{S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$.

6.2.1 Parameter Identification

In this section, we discuss the parameters obtained for the model (6.1). In order to allow an unbiased fit, we divide both of the measured trajectories $T_n^{i,j}$ into a part that is used for fitting and another part that is used for validation purposes. As the trajectories are periodical with the period T^orbit, it is sufficient to not fit the complete time series, but only a part that incorporates at least a single orbit. We use an optimisation horizon t_end of about two orbits to identify the parameters. This allows in the second step to use the remaining data to validate the fit. In particular, for the data set of the second date, the slew manoeuvre can be used to evaluate the estimations.

In order to use all the available data, it is sensible establishing the physical relations between the parameters describing the 24 temperature measurements. In the following we establish some physical constraints that can be woven into the optimisation problem. First, ideally each set of the twelve temperature sensors should have similar physical properties, i.e. the absorption α should be the same for all directions n and both dates, this can be written as

$$\alpha_n^{i,j} \approx \alpha_{\bar{n}}^{\bar{i},j} \qquad \forall\, i, \bar{i} \in \{\mathrm{SM}, \mathrm{Slew}\},\ j \in \{\mathrm{B}, \mathrm{S}\},\ n, \bar{n} \in \{\pm e_1, \pm e_2, \pm e_3\}\,. \tag{6.2}$$

However, in reality the sensors are subject to degradation and possibly slightly different manufacturing. The time constant C should be similar for two surfaces pointing into the

same direction, i.e.

$$C_n^{i,j} \approx C_n^{\bar{i},\bar{j}} \qquad \forall\, i,\bar{i} \in \{\mathrm{SM},\mathrm{Slew}\},\ j,\bar{j} \in \{\mathrm{B},\mathrm{S}\},\ n \in \{\pm e_1, \pm e_2, \pm e_3\}\,. \tag{6.3}$$

Again, this approximation might be impaired by different degradation and manufacturing. Finally, the dissipative heat may vary heavily with the side n as it is dependent on how good the sensor is isolated from the heat produced by the electrical components. However, it should be consistent for configurations with the same normal, i.e.

$$Q_n^{i,j} \approx Q_n^{\bar{i},\bar{j}} \qquad \forall\, i,\bar{i} \in \{\mathrm{SM},\mathrm{Slew}\},\ j,\bar{j} \in \{\mathrm{B},\mathrm{S}\},\ n \in \{\pm e_1, \pm e_2, \pm e_3\}\,. \tag{6.4}$$

These equalities may not hold if the operating condition of the spacecraft has changed.

All in all, it might be sensible to introduce these established physical constraints into the optimisation problem, but they may also be relaxed or even neglected to obtain better results. The full optimisation problem generated from Equation (6.1) with all the physical constraints has the form

$$\min_{c_n^{i,j}} \quad \sum_{\substack{i\in\{\mathrm{SM},\mathrm{Slew}\}\\ j\in\{\mathrm{B},\mathrm{S}\}\\ n\in\{\pm e_1,\pm e_2,\pm e_3\}}} \int_0^{t_\mathrm{end}} |T_n^{i,j}(t) - \hat{T}_n^{i,j}(t)|\mathrm{d}t \tag{6.5a}$$

$$\text{s. t.} \quad \dot{\hat{T}}_n^{i,j} = f_n^{i,j}(\hat{T}_n^{i,j}, t, c_n^{i,j}) \tag{6.5b}$$

$$\alpha_n^{i,j} = \alpha_{\bar{n}}^{\bar{i},j} \tag{6.5c}$$

$$C_n^{i,j} = C_n^{\bar{i},\bar{j}} \tag{6.5d}$$

$$Q_n^{i,j} = Q_n^{\bar{i},\bar{j}} \tag{6.5e}$$

$$c_n^{i,j} = (\alpha_n^{i,j}, C_n^{i,j}, Q_n^{i,j}) \subset [0,1] \times \mathbb{R}^+ \times \mathbb{R}_0^+ \tag{6.5f}$$

$$\text{for} \quad i,\bar{i} \in \{\mathrm{SM},\mathrm{Slew}\},\ j,\bar{j} \in \{\mathrm{B},\mathrm{S}\},\ n,\bar{n} \in \{\pm e_1, \pm e_2, \pm e_3\} \tag{6.5g}$$

where $T_n^{i,j}$ denotes the measured temperature and $\hat{T}_n^{i,j}$ the estimate generated by the model (6.5b). Equation (6.5a) is the cost function that minimises the difference of the estimate and the real value over a desired time interval $[0, t_\mathrm{end}]$. Eqs. (6.5c) to (6.5e) contain the physical constraints that are included as equalities. Finally, the inequalities in (6.5f) determine the domain of the parameters. In detail, the solar absorption α takes always values between 0 and 1, the time constant C is always bigger than zero, while the dissipated heat is always greater or equal to zero.

This optimisation problem has a total of 72 optimisation parameters. It is is constrained by 24 differential equations. For the equality constraints we neglect the linear dependent equations and obtain 22 linear independent equations for the absorption α and 18 linear independent equations for each C and Q. Due to the simple form of the algebraic constraints, this optimisation problem is equivalent to an optimisation problem with 14 optimisation parameters constrained by 24 differential equations.

Another option is to relax the physical constraints and incorporate them with weights into

the cost function, i.e.

$$\begin{aligned}
\min_{c_n^{i,j}} \quad & \sum_{\substack{i\in\{\mathrm{SM,Slew}\}\\ j\in\{\mathrm{B,S}\}\\ n\in\{\pm e_1,\pm e_2,\pm e_3\}}} \int_0^{t_{\mathrm{end}}} |T_n^{i,j}(t) - \hat{T}_n^{i,j}(t)|\mathrm{d}t + \sum_{\substack{k=1\\ i,\bar{i}\in\{\mathrm{SM,Slew}\}\\ j\in\{\mathrm{B,S}\}\\ n,\bar{n}\in\{\pm e_1,\pm e_2,\pm e_3\}\\ n\neq\bar{n}}}^{20} w_k^{\alpha} |\alpha_n^{i,j} - \alpha_{\bar{n}}^{\bar{i},j}| \\
& + \sum_{\substack{k=1\\ i,\bar{i}\in\{\mathrm{SM,Slew}\}\\ j,\bar{j}\in\{\mathrm{B,S}\}\\ n\in\{\pm e_1,\pm e_2,\pm e_3\}\\ i\neq\bar{i},\ j\neq\bar{j}}}^{18} w_k^{C} |C_n^{i,j} - C_n^{\bar{i},\bar{j}}| + \sum_{\substack{k=1\\ i,\bar{i}\in\{\mathrm{SM,Slew}\}\\ j,\bar{j}\in\{\mathrm{B,S}\}\\ n\in\{\pm e_1,\pm e_2,\pm e_3\}\\ i\neq\bar{i},\ j\neq\bar{j}}}^{18} w_k^{Q} |Q_n^{i,j} - Q_n^{\bar{i},\bar{j}}| \\
\text{s. t.} \quad & \dot{\hat{T}}_n^{i,j} = f_n^{i,j}(\hat{T}_n^{i,j}, t, c_n^{i,j}) \qquad i \in \{\mathrm{SM,Slew}\},\ j \in \{\mathrm{B,S}\},\ n \in \{\pm e_1, \pm e_2, \pm e_3\} \\
& c_n^{i,j} \subset [0,1] \times \mathbb{R}^+ \times \mathbb{R}_0^+ \qquad i \in \{\mathrm{SM,Slew}\},\ j \in \{\mathrm{B,S}\},\ n \in \{\pm e_1, \pm e_2, \pm e_3\}\,.
\end{aligned}$$

For this system the optimisation has to be performed for the complete set of 72 optimisation parameters. The design parameters w^{α}, w^{C}, w^{Q} allow to weigh the former equality constraints in relation to the best possible temperature estimation.

Finally, it was chosen to use the weights $w^{\alpha} = w^{C} = w^{Q} = 0$ to simply minimise the estimation error. Then the optimisation problem can be divided into 24 different optimisation problems, each of the form

$$\min_{c_n^{i,j}} \quad \int_0^{t_{\mathrm{end}}} |T_n^{i,j}(t) - \hat{T}_n^{i,j}(t)|\mathrm{d}t \tag{6.7a}$$

$$\text{s. t.} \quad \dot{\hat{T}}_n^{i,j} = f_n^{i,j}(\hat{T}_n^{i,j}, t, c_n^{i,j}) \tag{6.7b}$$

$$(\alpha_n^{i,j}, C_n^{i,j}, Q_n^{i,j}) \subset [0,1] \times \mathbb{R}^+ \times \mathbb{R}_0^+ \tag{6.7c}$$

for $i \in \{\mathrm{SM,Slew}\}$, $j \in \{\mathrm{B,S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. Every problem has 3 optimisation parameters (6.7c) and only a single dynamic equation (6.7b). This optimisation problem was chosen to be used as it is the simplest to solve and yields models that give the best temperature estimation possible as it incorporates all possible estimation parameters. The potential drawback is that the remaining model may give worse results for other attitudes which would be of course devastating for the observer design. However, the resulting quality of the model will be evaluated using the physical constraints and the slew manoeuvre which where not used in the design as benchmarks.

In order to solve the optimisation problem (6.7) we use the software environment MOPS (Multi-Objective Parameter Synthesis), an optimisation tool of DLR-SR presented in [131]. Tables 6.2 to 6.5 show the parameters resulting from fitting the data by solving optimisation problem (6.7) for the 24 configurations where $T = T_n^{i,j}$ with $i \in \{\mathrm{SM,Slew}\}$, $j \in \{\mathrm{B,S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. Each table shows the parameters α, C and Q^d for six of the specified configurations. The first row specifies the normal vector n of the considered sensor. Some entries in Tables 6.4 and 6.5 have different colours. An entry with an orange colour indicates that the entry deviates significantly from the value that would be enforced by the

Table 6.2: Science Mode, Black Sensor.

Variable	Unit	e_1	$-e_1$	e_2	$-e_2$	e_3	$-e_3$
$\alpha_n^{\mathrm{SM,B}}$		0.72	0.78	0.76	0.72	0.71	0.73
$C_n^{\mathrm{SM,B}}$	$\frac{\mathrm{K}}{\mathrm{s}}$	472	543	542	485	645	493
$Q_n^{d,\mathrm{SM,B}}$	W	11.28	12.58	18.43	7.24	0	28.07

Table 6.3: Science Mode, Silver Sensor.

	e_1	$-e_1$	e_2	$-e_2$	e_3	$-e_3$
$\alpha_n^{\mathrm{SM,S}}$	0.13	0.12	0.12	0.18	0.14	0.12
$C_n^{\mathrm{SM,S}}$	634	431	501	660	743	437
$Q_n^{d,\mathrm{SM,S}}$	10.54	11.69	12.33	6.53	0	24.77

physical constraints. An entry in a red colour means the same, but also that the system dynamics are not sensitive to this parameter, i.e. changing this parameter makes the error $\int_0^{t_{\mathrm{end}}} T(t) - \hat{T}(t)\mathrm{d}t$ only slightly higher. Thus, these entries are replaced by values that have a lower residuum considering the physical constraints as discussed in the next section. Table 6.2 also specifies the units of the variables, these are omitted in the other tables.

6.2.2 Model Verification

We verify the results of the determined model by two means. First, we compare the obtained parameter to parameters from literature and evaluate if the physical constraints from the previous section are fulfilled. Secondly, we use the slew manoeuvre of the second data set which was not incorporated into the fit to verify the quality of the model.

In [132] values for the thermal absorption and emission of different coatings are provided. The thermal values for black coatings are $\alpha^B \in [0.9, 0.97]$ and $\varepsilon^B \in [0.72, 0.91]$. For reflecting silver coatings the thermal values are $\alpha^S \in [0.07, 0.22]$ and $\varepsilon^S \in [0.4, 0.87]$. Due to degradation these values may change over time. Ideally, as stated in (6.2) the absorption values are independent of their surface normal for both, black and silver coatings, i.e. $\alpha_n^{i,j} = \alpha_{\bar{n}}^{i,j}$ with $i \in \{\mathrm{SM}, \mathrm{Slew}\}$, $j \in \{\mathrm{B}, \mathrm{S}\}$ and $n, \bar{n} \in \{\pm e_1, \pm e_2, \pm e_3\}$. Note that we drop the equality for values of the two different data sets due to possible degradation. Further note, in our case the degradation of α and ε are both covered in the change of α. In science mode, the estimated α values have a mean of 0.73 and a small standard deviation of 0.02

Table 6.4: Slew Mode, Black Sensor.

	e_1	$-e_1$	e_2	$-e_2$	e_3	$-e_3$
$\alpha_n^{\mathrm{Slew,B}}$	0.70	0.57	0.53	0.81	0.50	0.57
$C_n^{\mathrm{Slew,B}}$	570	392	361	2873/443	495	398
$Q_n^{d,\mathrm{Slew,B}}$	15.57	12.73	8.37	0	24.42	26.10

Table 6.5: Slew Mode, Silver Sensor.

	e_1	$-e_1$	e_2	$-e_2$	e_3	$-e_3$
$\alpha_n^{\text{Slew,S}}$	0.11	0.13	0.06/0.12	0.11	0.14	0.13
$C_n^{\text{Slew,S}}$	451	484	334	636	463	340
$Q_n^{d,\text{Slew,S}}$	12.92	11.64	6.62	29.19	3.03	23.24

being only roughly 3%. Their ratio $\frac{\alpha}{\varepsilon}$ being 0.91 is smaller than in the literature which is most likely caused by the degradation. This allows a high confidence level in the estimate of α due to its consistency. The same can be said for the estimate for the silver coating parameter in science mode. The high value of $\alpha_{-e_2}^{\text{Slew},B}$ is due to the lower sensitivity of α in this configuration.
In slew mode, the standard deviation between the α values of the black sensor is comparatively high. Also the degradation of approximately 0.2 over two years appears to be relatively high. Nevertheless, it appears to be the best way to keep these parameters despite these physical discrepancies. Adjusting these parameters to obtain a smaller standard deviation and degradation would lead to discrepancies in the dynamics. All slew mode parameters of the silver coated sensors, but $\alpha_3^{\text{Slew},S}$ are within an appropriate range. The parameter $\alpha_{e_2}^{\text{Slew},S}$ is off due to its low sensitivity and is replaced by the mean of the other $\alpha_i^{\text{Slew},S}$ values for $i \in \{\pm e_1, -e_2, \pm e_3\}$.

As stated in (6.3), the time constant C should be similar for the individual surfaces, i.e. $C_n^{i,j} = C_n^{\bar{i},\bar{j}}$ with $i, \bar{i} \in \{\text{SM}, \text{Slew}\}$, $j, \bar{j} \in \{\text{B}, \text{S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. While the deviations between the individual values appear significant, the sensitivity of the dynamics for these deviations is not very high. This suggests that almost all of the values can be kept. In particular the parameter $C_{-e_2}^{\text{Slew},B}$ has a low sensitivity in this scenario which is why its estimate is so high. The value is replaced by the mean value of the remaining C parameters as this allows for a better fit in other orientations while not impairing the fit too heavily.

The dissipated heat should be consistent for sensors with the same normal n as stated in (6.4), i.e. ideally $Q_n^{i,j} = Q_n^{\bar{i},\bar{j}}$ with $i, \bar{i} \in \{\text{SM}, \text{Slew}\}$, $j, \bar{j} \in \{\text{B}, \text{S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. Indeed the dissipated heat varies significantly dependent on the normal n which suggests that a more elaborate model for the internal heat flows might be beneficial in future work. For the same normal, the heat flows are mostly consistent. The most noticeable discrepancies are $Q_{e_3}^{\text{Slew,B}}$ and $Q_{-e_2}^{\text{Slew,S}}$. They appear to compensate small α value. All in all, the comparison to literature and to the physical constraints suggests that the obtained parameters are sensible in combination with the proposed model. However, some additional information might be required in future applications about the internal heat flow to allow a prediction of the dissipated heat.

The second step uses the remaining data points to verify the model and its parameters. First, we verify that using a position dependent infrared and albedo coefficient in model (6.1) has enough benefits to compensate the added complexity. This can be seen in Figure 6.5 where the temperature of the silver nadir pointing surface in science mode is considered. As it is pointing nadir, almost no solar irradiation is acting on the surface. The temperature evolution is mostly determined by the influence of albedo and infrared irradiation. We have

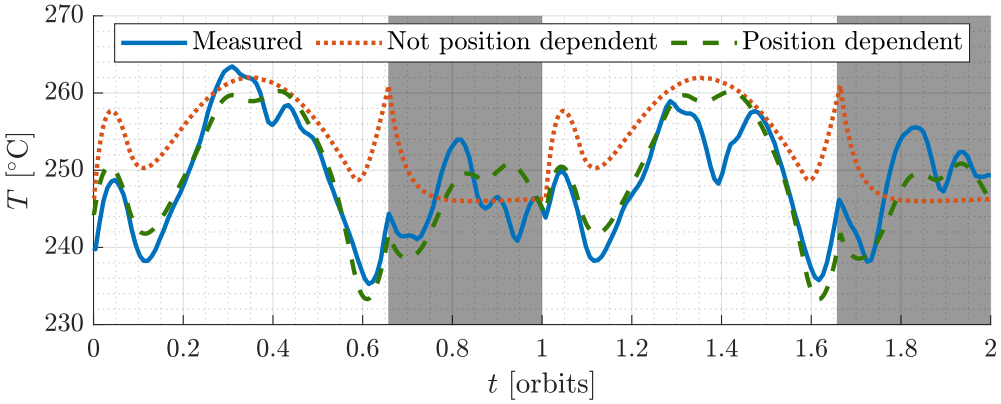
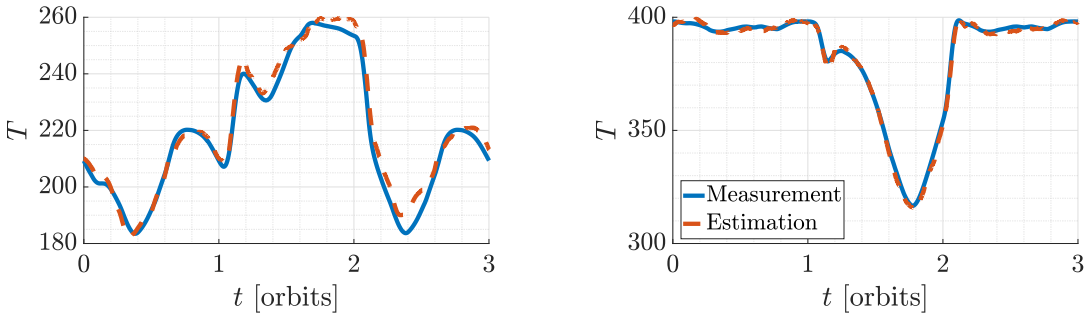


Figure 6.5: Comparison of measurement and estimates based on models (4.11) (dotted red line) and (6.1) (dashed green line). The infrared variable I_{IR} and albedo coefficient ρ are either constants or functions of the position.

fitted two models using the optimisation (6.7). The red dotted line shows the result of fitting model (4.11) which does not contain position dependent infrared and albedo coefficients. The green dashed line shows the fit of model (6.1) where the infrared and albedo coefficient are functions of the position. It is clear that the first model incorporates the most distinct dynamics, namely the peaks resulting from the short time with solar irradiation and the influence of the Earth albedo during the Sun phase, however many other dynamics that are seemingly part of the real temperature evolution are missing. This leads to an overall insufficient fit. However, incorporating the position dependence improves the fit significantly, as the estimate now exhibits most of the prominent dynamics. Nevertheless, due to the simplicity of the model there are still errors remaining which allow for improvement if more accurate descriptions of the two coefficients were available.

In the following, we verify the fit of the models showing some of the most interesting configurations and their fits. We show the simulations from the second date ($\cdot^{\mathrm{Slew}}$) for 30 000 s with the slew manoeuvre between 11 810 s and 18 244 s. Note that only data between 0 s and 10 000 s was used to fit the model and the remaining time can now be used to give an unbiased validation. Figure 6.6 shows two plots that compare the measurement with the estimates. Figure 6.6a shows the temperature of the silver coated surface pointing into the direction of the first unity vector for three orbits of the second date. This surface is under the influence of all three irradiations or a combination of them. This makes the temperature vary a lot which is why there is a maximum temperature difference of about 40 degrees Celsius over an orbit when no slew manoeuvre was performed. It can be seen that as expected the estimate follows the measurement closely outside the slew manoeuvre, because this data has been used to fit the model. During the slew manoeuvre the temperature rises significantly because the surface is rotated towards the Sun. It can be seen that the estimate shows the same behaviour but gives slightly bigger errors. Figure 6.6b shows the temperature of the black coated surface pointing into the direction of the negative second unity vector for three orbits of the second date. During the nominal state the temperature varies only with small amplitudes. Thus, there are almost no dynamics during this phase which



(a) Silver, Slew, e_1

(b) Black, Slew, $-e_2$

Figure 6.6: Comparison of the temperature measurement and estimate from the fitted models. Figure 6.6a shows temperature of the silver coated surface pointing into the direction of the first unity vector starting from the second date. Figure 6.6b shows temperature of the black coated surface pointing into the direction of the negative second unity vector starting from the second date.

is why the dynamics were insensitive to the parameter $C_{-e_2}^{\text{Slew,B}}$ as discussed in the previous section. By the decreasing temperature during slew mode it can be seen that the surface is rotated away from the Sun. The estimate follows this behaviour also rather well which confirms that replacing $C_{-e_2}^{\text{Slew,B}}$ with a more physical parameter was appropriate.

In order to give a compact evaluation of all the fits we calculate the coefficient of determination denoted by R^2-value defined as

$$R^2 = 1 - \frac{\sqrt{\sum_{k=1}^{N} |y_k - \hat{y}_k|^2}}{\sqrt{\sum_{k=1}^{N} |y_k - \hat{y}^{\text{av}}|^2}}$$

with the measurement values $y_k = T_n^{i,j}(t_k)$, the values obtained from the model $\hat{y}_k = \hat{T}_n^{i,j}(t_k)$ with the average value $\hat{y}^{\text{av}} = \frac{1}{N}\sum_{k=1}^{N} \hat{T}_n^{i,j}(t_k)$ for the measurement times $(t_k)_{k\in\mathbb{N}} = (1, \dots, N)$ and the configurations $i \in \{\text{SM}, \text{Slew}\}$, $j \in \{\text{B}, \text{S}\}$ and $n \in \{\pm e_1, \pm e_2, \pm e_3\}$. An R^2 value of 1 means that the estimation is perfect. A value of 0 is obtained if the model is giving the average value of the measured time series at every point of time. A negative value means that the model is even worse than that. Table 6.6 gives the R^2 values for the 24 different configurations. The values in orange are the ones below 0.95, which is where the most improvements can be made. The worst of these configurations, silver coated in science mode and $n = e_3$, was already discussed in the review of Figure 6.5 and should still be sufficient to allow a sensible attitude reconstruction. Generally speaking, the configurations without solar irradiation admit a worse fit as they rely on accurate knowledge of the albedo and infrared coefficient which is not ensured at that moment. This can be seen by the fact that the surfaces that are nadir pointing have the lowest R^2 values. In summary, all values are very close to one which suggests that the fitted model is appropriate. We use this model in the next sections to design the observer.

Table 6.6: R^2 values for all 24 configurations.

	e_1	$-e_1$	e_2	$-e_2$	e_3	$-e_3$
Black, Science Mode	0.997	0.999	0.999	0.967	0.946	0.999
Silver, Science Mode	0.991	0.995	0.996	0.869	0.788	0.994
Black, Slew Mode	0.992	0.986	0.918	0.996	0.916	0.991
Silver, Slew Mode	0.979	0.964	0.848	0.970	0.913	0.983

6.3 Problem Definition

In this section we use the model introduced in the previous Section 6.2 to formulate the observability problem. We introduce a compact notation for the individual states and parameters by denoting

$$T := [T_1, \dots, T_{12}] := [T_{n_1}^{\mathrm{B},j}, T_{n_2}^{\mathrm{B},j}, T_{n_3}^{\mathrm{B},j}, T_{n_4}^{\mathrm{B},j}, T_{n_5}^{\mathrm{B},j}, T_{n_6}^{\mathrm{B},j}, T_{n_1}^{\mathrm{S},j}, T_{n_2}^{\mathrm{S},j}, T_{n_3}^{\mathrm{S},j}, T_{n_4}^{\mathrm{S},j}, T_{n_5}^{\mathrm{B},j}, T_{n_6}^{\mathrm{S},j}]$$

for $j \in \{\mathrm{SM}, \mathrm{Slew}\}$. Note that we omit the indication of j in the temperature vector to simplify the readability In the following sections, it will be made explicit in the description if data from the first or second date is considered which is why no indication of j is required. Accordingly, we define as in Equation (4.8) for each date the parameters α_i, β_i, γ_i and δ_i for each sensor $i \in \{1, \dots 12\}$ with the parameters determined in the previous Section 6.2. Note that α in this section does not denote α_s but α defined in Equation (4.8). We apply the same straightforward notation for the angle vector ϕ and θ. Note that with these definitions it is $\phi_i = \phi_{i+6}$ and $\theta_i = \theta_{i+6}$ for $i \in \{1, \dots 6\}$.

This allows to state the temperature dynamics as in Equation (6.1) for every sensor as

$$\dot{T}_i = \max\left(\alpha_i(t)\cos(\phi_i(q)), 0\right) + \max\left(\beta_i(t)F(t, \theta_i(t,q), 0)\right) + \gamma_i F(t, \theta_i(t,q)) - \delta_i T_i^4 + \frac{Q_i}{C_i}$$

for $i \in \{1, \dots 12\}$. We introduce the remaining notation in the same manner as in Chapter 5 and obtain the compact system

$$\begin{aligned}
\dot{T}_1 &= f_{T_1}(T_1, q, t) && \text{(6.8a)}\\
&\vdots && \text{(6.8b)}\\
\dot{T}_{12} &= f_{T_{12}}(T_{12}, q, t) && \text{(6.8c)}\\
\dot{q} &= f_q(q, \omega) && \text{(6.8d)}\\
\dot{\omega} &= f_\omega(\omega) + J^{-1}u\,. && \text{(6.8e)}\\
\dot{t} &= 1\,. && \text{(6.8f)}\\
y &= h(x) = \left[T_1 \dots T_{12}\ q^\top q - 1\ \omega\ t\right]^\top\,. && \text{(6.8g)}
\end{aligned}$$

The objective is to analyse the observability of the system (6.8) and design an observer. Note that the main difference to the problem definition in Chapter 5 is that the output now

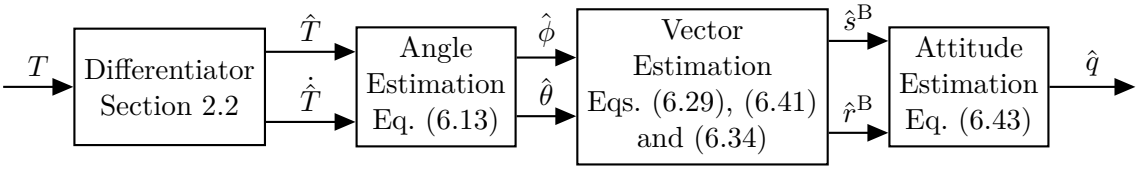


Figure 6.7: Block diagram of the attitude estimation based on temperature.

contains twelve temperature signals instead of a single one. This facilitates the problem in the sense that higher order derivatives of the same temperature signal are not required.

6.4 Observer in Transformed Coordinates

In this section the observer in transformed coordinates consisting of a differentiator and an algebraic equation is derived. Further, we evaluate the performance of the proposed observer on the experimental temperature data. Figure 6.7 shows a block diagram which presents the idea of the algorithm proposed in the upcoming sections. First, the temperature signal needs to be differentiated to obtain an estimate of the temperature as well as its derivative. These two signals shall then be used in combination with their dynamics presented in Eqs. (6.8a) to (6.8c) to estimate the angles to Sun and Earth. These angles shall then be used to estimate the vectors of Sun and Earth in body coordinates. With these, the actual desired attitude can be estimated. The differentiator design obtained extensive attention in Section 2.2 and got applied in Section 5.3.1. The application of this differentiator to the current problem is straightforward and not further discussed. The angle estimation will be discussed in detail in the next sections. It turns out that this requires merely a linear inversion. The vector estimation consists of two parts, first finding the measurements that contain useful information and second using these measurements to estimate the vectors. These will turn out to be merely the result of multiple matrix multiplications. Finally, the real attitude estimation is a standard problem for which various different algorithms exist and will be discussed. The upcoming sections also solve the question of the observability of the system by naturally solving the underlying inversion problem.

6.4.1 Estimation of the Angle to Sun and Earth

The first step uses each pair of temperature sensors with the same normal vector but different coating in order to determine the angle between the normal vector and Sun vector as well as the angle between normal vector and Earth vector. We start by splitting the temperature vectors into T_i with $i \in \{1, \dots, 6\}$ for black sensors and T_j with $j \in \{7, \dots, 12\}$ for silver sensors. In the following we always consider a tupel (i, j) with $i \in \{1, \dots, 6\}$ and $j = i + 6$.

The temperature dynamics have the form

$$\dot{T}_i = \max\Big(\alpha_i(t)\cos(\phi_i(q)), 0\Big) + \max\Big(\beta_i(t)F(t,\theta_i(t,q),0)\Big) + \gamma_i F(t,\theta_i(t,q)) - \delta_i T_i^4 + \frac{Q_i}{C_i} \tag{6.9a}$$

$$\dot{T}_j = \max\Big(\alpha_j(t)\cos(\phi_j(q)), 0\Big) + \max\Big(\beta_j(t)F(t,\theta_j(t,q),0)\Big) + \gamma_j F(t,\theta_j(t,q)) - \delta_j T_j^4 + \frac{Q_j}{C_j}\,. \tag{6.9b}$$

Because i and j were chosen such that they have the same angles, i.e. $\phi_i = \phi_j$ and $\theta_i = \theta_j$, we can write the system (6.9) in matrix form as

$$\begin{bmatrix}\dot{T}_i\\ \dot{T}_j\end{bmatrix} = \begin{bmatrix}\alpha_i(t) & \max(\beta_i(t),0)+\gamma_i\\ \alpha_j(t) & \max(\beta_j(t),0)+\gamma_j\end{bmatrix}\begin{bmatrix}\max\Big(\cos(\phi_i(q)),0\Big)\\ F(t,\theta_i(t,q))\end{bmatrix} - \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C_i}\\ \delta_j T_j^4 - \frac{Q_j}{C_j}\end{bmatrix}. \tag{6.10}$$

For the spacecraft in solar eclipse, i.e $t \in \mathcal{T}_{\text{ec}}$ as defined in (5.1) this is equivalent to

$$\begin{bmatrix}\gamma_i\\ \gamma_j\end{bmatrix} F(t,\theta(t,q)) = \left(\begin{bmatrix}\dot{T}_i\\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C_i}\\ \delta_j T_j^4 - \frac{Q_j}{C_j}\end{bmatrix}\right). \tag{6.11}$$

In this case it is not possible to obtain an estimate for ϕ.

For $t \in \mathcal{T}_{\text{alb}} \cup \mathcal{T}_{\text{sun}}$ and $\alpha_i(\max(\beta_j(t),0)+\gamma_j) - \alpha_j(\max(\beta_i(t),0)+\gamma_i) \neq 0$ we can solve Equation (6.10) to obtain

$$\begin{bmatrix}\max\Big(\cos(\phi_i(q)),0\Big)\\ F(t,\theta_i(t,q))\end{bmatrix} = \begin{bmatrix}\alpha_i(t) & \max(\beta_i(t),0)+\gamma_i\\ \alpha_j(t) & \max(\beta_j(t),0)+\gamma_j\end{bmatrix}^{-1}\left(\begin{bmatrix}\dot{T}_i\\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C}\\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix}\right). \tag{6.12}$$

We introduce the augmented inverse function of the cosine $\overline{\arccos} : \mathbb{R} \to [0,\pi]$ with

$$\overline{\arccos}(y) = \begin{cases}\pi & \text{if } y < -1\\ \arccos(y) & \text{if } -1 \le y \le 1\\ 0 & \text{if } y > 1\end{cases}.$$

Further we denote by F^{-1} the inverse of the form factor for a fixed time t, i.e. the inverse is defined by $F^{-1} : \mathbb{R} \to [0, \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right)]$ with $F^{-1}(F(t,\theta)) = \theta$. This inverse can be realised by a simple Newton method due to the monotonicity of F. We augment this inverse to obtain

$$\overline{F}^{-1}(y) = \begin{cases}\frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) & \text{if } y \le 0\\ F^{-1}(y) & \text{if } 0 \le y \le \frac{r_\oplus^2}{\|r\|^2}\\ 0 & \text{if } y > \frac{r_\oplus^2}{\|r\|^2}\end{cases}.$$

This allows to define a method to estimate the angles ϕ and θ for a sunlit spacecraft.

Proposition 6.1. *Consider two temperature sensors which dynamics obey Equation* (6.9) *with the angles* $\phi \in [0, \pi]$ *and* $\theta \in [0, \pi]$ *as defined in Equation* (4.9a)*. Let the spacecraft be sunlit, i.e.* $t \in \mathcal{T}_{\text{alb}} \cup \mathcal{T}_{\text{sun}}$ *and* $\alpha_i(\max(\beta_j(t), 0) + \gamma_j) - \alpha_j(\max(\beta_i(t), 0) + \gamma_i) \neq 0$ *. Then the estimate for the angle between the normal and the Sun and Earth is defined to be*

$$\hat{\phi}_i = \overline{\arccos}\left(\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\alpha_i(t) & \max(\beta_i(t), 0) + \gamma_i \\ \alpha_j(t) & \max(\beta_j(t), 0) + \gamma_j\end{bmatrix}^{-1}\left(\begin{bmatrix}\dot{T}_i \\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C} \\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix}\right)\right) \tag{6.13a}$$

$$\hat{\theta}_i = \overline{F}^{-1}\left(\begin{bmatrix}0 & 1\end{bmatrix}\begin{bmatrix}\alpha_i(t) & \max(\beta_i(t), 0) + \gamma_i \\ \alpha_j(t) & \max(\beta_j(t), 0) + \gamma_j\end{bmatrix}^{-1}\left(\begin{bmatrix}\dot{T}_i \\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C} \\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix}\right)\right). \tag{6.13b}$$

For the estimate of the solar angle holds

$$\phi_i(q) \leq \frac{\pi}{2} \Rightarrow \hat{\phi}_i = \phi_i(q) \tag{6.14a}$$

$$\phi_i(q) > \frac{\pi}{2} \Rightarrow \hat{\phi}_i = \frac{\pi}{2}. \tag{6.14b}$$

For the estimate of the angle between normal and Earth holds

$$\theta_i(q) \leq \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow \hat{\theta}_i = \theta_i(q) \tag{6.15a}$$

$$\theta_i(q) > \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow \hat{\theta}_i = \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right). \tag{6.15b}$$

Proof. From $0 \leq \phi_i(q) < \frac{\pi}{2}$ it follows

$$\hat{\phi}_i = \overline{\arccos}\left(\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\alpha_i(t) & \max(\beta_i(t), 0) + \gamma_i \\ \alpha_j(t) & \max(\beta_j(t), 0) + \gamma_j\end{bmatrix}^{-1}\left(\begin{bmatrix}\dot{T}_i \\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C} \\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix}\right)\right) \tag{6.16}$$

$$\overset{(6.12)}{=} \overline{\arccos}(\max(\cos(\phi_i(q), 0)) \tag{6.17}$$

$$\overset{\phi_i(q) \in [0, \frac{\pi}{2}]}{=} \overline{\arccos}(\cos(\phi_i(q)) \tag{6.18}$$

$$= \phi_i(q) \tag{6.19}$$

and the claim (6.14a) is shown. For $\frac{\pi}{2} < \phi < \pi$ it is

$$\hat{\phi}_i = \overline{\arccos}\left(\begin{bmatrix}1 & 0\end{bmatrix}\begin{bmatrix}\alpha_i(t) & \max(\beta_i(t), 0) + \gamma_i \\ \alpha_j(t) & \max(\beta_j(t), 0) + \gamma_j\end{bmatrix}^{-1}\left(\begin{bmatrix}\dot{T}_i \\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C} \\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix}\right)\right) \tag{6.20}$$

$$\overset{(6.12)}{=} \overline{\arccos}(\max(\cos(\phi_i(q), 0)) \tag{6.21}$$

$$\overset{\phi_i(q) \in (\frac{\pi}{2}, \pi]}{=} \overline{\arccos}(0) \tag{6.22}$$

$$= \frac{\pi}{2} \tag{6.23}$$

and the claim (6.14b) is shown. The claims (6.15a) and (6.15b) can be shown in analogue fashion. □

The proposition has shown that we have to distinguish between the two cases whether the side is pointing into direction of the heat source or not. Equation (6.14a) ensures that for a sensor pointing into the direction of the Sun, i.e. $\phi \leq \frac{\pi}{2}$, the estimate is accurate. However, for a sensor pointing away from the Sun as in Equation (6.14b), i.e. $\phi \in (\frac{\pi}{2}, \pi]$ the estimated angle is always $\frac{\pi}{2}$. In this case, this estimate is generally not useful to estimate the attitude. Similar results are obtained for the angle to the Earth but the critical angle is higher, due to the Earth being considered as a spherical heat source and not a point heat source.

For a spacecraft in eclipse we define the estimates in a similar manner. However, the angle between normal and Sun cannot be estimated. However, the angle between normal and Earth can be estimated more accurately using the Moore-Penrose inverse $\cdot^+$ of $\begin{bmatrix}\gamma_i \ \gamma_j\end{bmatrix}^\top$ as defined in Appendix A.1.

Proposition 6.2. *Consider two temperature sensors which dynamics obey Equation* (6.9) *with the angles* $\phi \in [0, \pi]$ *and* $\theta \in [0, \pi]$ *as defined in Equation* (4.9a)*. Let the spacecraft be in solar eclipse, i.e.* $t \in \mathcal{T}_{\text{ec}}$ *with either* $\gamma_i \neq 0$ *or* $\gamma_j \neq 0$ *or both. Then an estimate for the angle between the normal and the Earth is obtained using*

$$\hat{\theta} = \overline{F}^{-1} \left(\begin{bmatrix}\gamma_i \\ \gamma_j\end{bmatrix}^+ \left(\begin{bmatrix}\dot{T}_i \\ \dot{T}_j\end{bmatrix} + \begin{bmatrix}\delta_i T_i^4 - \frac{Q_i}{C} \\ \delta_j T_j^4 - \frac{Q_j}{C}\end{bmatrix} \right) \right) . \tag{6.24a}$$

For the estimate of the angle between normal and Earth holds

$$\theta \leq \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow \hat{\theta} = \theta \tag{6.25a}$$

$$\theta > \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow \hat{\theta} = \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) . \tag{6.25b}$$

Proof. Analogue to Proposition 6.1 using the solution of Equation (6.11). □

Clearly, in the case of solar eclipse the single angle will not be sufficient to estimate the attitude. Thus, we focus on the sunlit case and use the 12 calculated angles to obtain an estimate of the Sun vector and the Earth vector in body coordinates.

6.4.2 Estimation of the Sun and Earth Vector in Body Frame

Naturally, from the twelve calculated angles not all of them are of use, e.g. the estimate of the angle to the Sun of a plate pointing away from the Sun will always be $\frac{\pi}{2}$ even though the real angle may be higher. Thus, there are two main tasks in this section. The first is to find the angles that contain actual information to determine the desired vectors. The second task is to use these angles to determine the vectors. We start by solving the second task as this shows the information that is required to calculate the angles.

Calculation of Sun and Earth Vector in Body Frame from Solar Angles

We obtain with the definition of ϕ and θ according to Equation (4.9a) the system of equations

$$\frac{s^\top}{\|s\|} A(q)^\top n_i = \cos(\phi_i) \quad i \in \{1, \dots 6\} \tag{6.26a}$$

$$-\frac{r^\top}{\|r\|} A(q)^\top n_i = \cos(\theta_i) \quad i \in \{1, \dots 6\} \tag{6.26b}$$

with $(n_1, \dots, n_6) = (e_1, -e_1, e_2, -e_2, e_3, -e_3)$. Note that this equation uses that the position vectors s and r are given in ECI frame while the normal vector n_i is given in body frame and that $A(q)$ denotes the transformation matrix from ECI to body frame. We denote the matrix of normal vectors by $N = [n_1, \dots, n_6] \in \mathbb{R}^{3\times 6}$, the vector of cosines of the angle between Sun and normal by $b_\phi = [\cos(\phi_1), \dots \cos(\phi_6)]^\top \in \mathbb{R}^6$ and the vector of the cosines of the angles between normal and Earth as $b_\theta = [\cos(\theta_1), \dots \cos(\theta_6)]^\top \in \mathbb{R}^6$. With $b_{\hat{\phi}}$ and $b_{\hat{\theta}}$ we denote the vectors b_ϕ and b_θ with $\phi = \hat{\phi}$ and $\theta = \hat{\theta}$, respectively. Then Equation (6.26a) can be written as

$$\frac{s^\top}{\|s\|} A(q)^\top N = b_\phi^\top .$$

Transposing both sides leads to

$$N^\top A(q) \frac{s}{\|s\|} = b_\phi . \tag{6.27}$$

We aim to replace b_ϕ by our estimate $b_{\hat{\phi}}$ and formulate a conventional attitude estimation problem in the sense of Wahba [41]. Therefore, we need to solve the system for $A(q)s \in \mathbb{R}^3$. Clearly, this is an overdetermined linear system of equations with six equations and three variables. For an imperfect model and measurements the system has no solution but needs to be considered as an optimisation problem. However, generally not all of the estimates are sensible for the reconstruction as discussed in the previous section. Thus, we multiply a positive semi-definite diagonal weight matrix $W_\phi = [w_{\phi,1}, \dots w_{\phi,6}]^\top \in \mathbb{R}^{6\times 6}$ to the system which decides which of the measurements should be taken into account. The weight matrix is chosen such that $W_\phi N^\top$ has full rank 3. The choice of its individual entries will be discussed in the next section. We now show how to estimate the vectors in body frame and show that the estimate is correct for a perfect model and measurements.

Proposition 6.3. *Consider two temperature sensors which dynamics obey Equation* (6.9) *and the assumptions and angle estimations as in* (6.13)*. Further let W_ϕ and W_θ be positive semi-definite diagonal matrices such that $W_\phi N^\top$ and $W_\theta N^\top$ have full rank and fulfil*

$$\phi_i > \frac{\pi}{2} \Rightarrow w_{\phi,i} = 0 \tag{6.28a}$$

$$\theta_i > \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow w_{\theta,i} = 0 \tag{6.28b}$$

for $i \in \{1, \dots 6\}$.

Then the estimated vectors

$$\hat{s}^{\mathrm{B}} := (W_\phi N^\top)^+ W_\phi b_{\hat{\phi}} \|s\| \tag{6.29a}$$

$$\hat{r}^{\mathrm{B}} := -(W_\theta N^\top)^+ W_\theta b_{\hat{\theta}} \|r\| \tag{6.29b}$$

are the Earth and the Sun vector in body coordinates, i.e.

$$\hat{s}^{\mathrm{B}} = A(q)s \tag{6.30a}$$

$$\hat{r}^{\mathrm{B}} = A(q)r\,. \tag{6.30b}$$

Proof. The solution of (6.27) exists by the definition of ϕ. Because N has full rank, the solution of (6.27) is unique. Multiplying the weight matrix W_ϕ to both sides in (6.27)

$$W_\phi N^\top A(q) \frac{s}{\|s\|} = W_\phi b_\phi\,. \tag{6.31}$$

Condition (6.28a) ensures with Equation (6.14a) and the diagonal form of W_ϕ that

$$e_i^\top W_\phi b_\phi = w_{\phi,i}^\top b_\phi = w_{\phi,i}^\top b_{\hat{\phi}} = e_i^\top W_\phi b_{\hat{\phi}}$$

for all $i \in \{1, \dots 6\}$. The equality in the middle holds because $w_{\phi,i}$ is either a multiple of the i-th unit vector with $\cos(\phi_i) = \cos(\hat{\phi}_i)$ or identical to zero. Consequently, the equation $W_\phi b_\phi = W_\phi b_{\hat{\phi}}$ holds. Multiplying $NW_\phi^\top$ from the left side to obtain a quadratic matrix yields

$$NW_\phi^\top W_\phi N^\top A(q) \frac{s}{\|s\|} = NW_\phi^\top W_\phi b_{\hat{\phi}}\,.$$

The inverse of $NW_\phi^\top W_\phi N^\top$ exists because $W_\phi N^\top$ has full rank. Multiplying with the inverse and $\|s\|$ from the left side leads to Equation (6.29) and (6.30). The same arguments can be made for θ and $A(q)\frac{-r}{\|r\|}$. □

This proposition in Equation (6.28) has provided the requirements that need to be met by the weight matrix and has shown in (6.29) and (6.30) that the resulting estimation of the vectors in body frame is accurate. Note that for an imperfect model and measurements, the estimation is still optimal as stated in the following formulation.

Remark 16. The estimates of the solar and the Earth vector in body frame $\hat{s}^{\mathrm{B}}$, $\hat{r}^{\mathrm{B}}$ obtained from (6.29) are the best possible solutions in that sense, that they are the least square

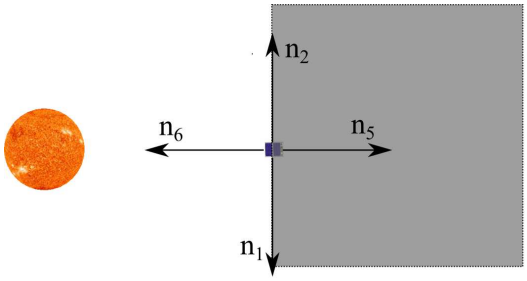


Figure 6.8: Illustration of the planar projection of the Sun, spacecraft and its normals.

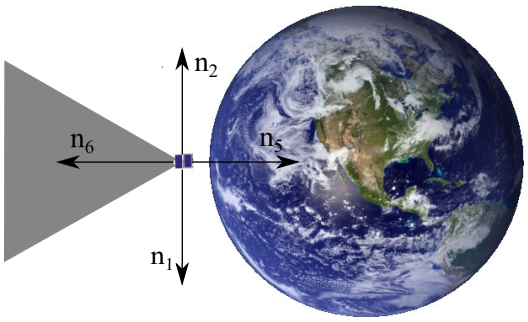


Figure 6.9: Illustration of the planar projection of the Earth, spacecraft and its normals.

solutions which solve the optimisation problem

$$\min_{\hat{s}^{\mathrm{B}},\hat{r}^{\mathrm{B}}} \quad e_1^\top e_1 + e_2^\top e_2 \tag{6.32a}$$

$$\text{s. t.} \quad e_1 = W_\phi N^\top \frac{\hat{s}^{\mathrm{B}}}{\|s\|} - W b_{\hat{\phi}} \tag{6.32b}$$

$$e_2 = W_\theta N^\top \frac{\hat{r}^{\mathrm{B}}}{\|r\|} - W b_{\hat{\theta}} \tag{6.32c}$$

resulting from (6.31).

Proposition (6.3) assumes that enough of the twelve measurements are useful to estimate the Earth and Sun vector in body frame i.e. the existence of the weight matrices W_ϕ, W_θ to fulfil the rank condition. This is not always ensured,e.g. if the spacecraft is in the Earth's shadow, i.e $t \in \mathcal{T}^{\mathrm{ecl}}$. In this case, no matrix W_ϕ exists which ensures that $W_\phi N^\top$ has full rank and Equation (6.28) is fulfilled. Then it is not possible to give a sensible estimate of the Sun vector in body frame $A(q)s$ using temperature data. For all other cases it can be shown that there exists such a weight matrix and that the choice of the weight matrix influences the quality of the solution. This is interrogated in the following and completes this section.

Choice of the Weight Matrices

As established in the previous section, the choice of the weight matrix means to choose the estimated angles that are of use for the attitude estimation. The choice of this matrix varies with the configuration of the sensor up to the point that the matrices do not necessarily exist if the configuration is chosen poorly. In our case, the sensors are pointing along each of the unit vectors which ensures the existence of the weight matrices. There is basically no design freedom for the choice of the weight matrix W_ϕ as illustrated in Figure 6.8 which shows the planar projection of the Sun, spacecraft and its normals. The Sun is depicted to be very small, instead of displaying the large distance between Sun and spacecraft. Thus, the Sun can be considered a point heat source and the spacecraft surfaces as infinitesimal

small. This leads to the black half-plane which describes the normals of the surfaces for which there is no exposure to solar irradiation. In the depicted case three normal vectors are inside the plane and one is outside. In all other cases every rotation leads to exactly two normal vectors inside the plane and two outside. With the additional vector for the third dimension not displayed in the illustration this leads to exactly three normal vectors outside the half-plane which are the minimum of information required to estimate the Sun position. More design freedom exists for the choice of W_θ as illustrated in Figure 6.9 by the planar projection of the Earth, spacecraft and its normals. The Earth is considered as a spherical heat source and the spacecraft surfaces as infinitesimal small. The black cone with an angle of about 60 degrees describes the normals for which no infrared and albedo irradiation acts on the surface. It can be seen that for all spacecraft rotations at most one normal vector lies inside the cone. Therefore most pairs of positive and negative unit vectors contain redundant information that can be used in an optimal manner to estimate the position of the Earth. A mathematically more profound and detailed description of these properties and design for both weight matrices is proposed in the following.

We start with the weight matrix W_ϕ for the Sun vector estimation. Naturally, we want to weigh any measurement that gives information about the location of the Sun as also used in Equation (6.14a) and (6.28). Therefore, the choice of the six normal vectors leads to a matrix of rank three which is defined in the following way

$$w_{\phi,i}(\phi) = \begin{cases} e_i & \text{if } \phi_i \leq \frac{\pi}{2} \\ 0 & \text{otherwise} \end{cases} \tag{6.33}$$

where e_i is the i-th six dimensional unit vector. This matrix is uniquely defined and there is no real way to optimise it. Changing the weights that are identical to zero would lead to the usage of undesired information while changing the non-zero weights does not change the solution as they weigh independent variables. Only in the special case $\phi_i = \phi_{i+1} = \frac{\pi}{2}$ would changing the weights to $w_{\phi,i}(\phi) = a_i e_i$ and $w_{\phi,i+1}(\phi) = (1 - a_i)e_{i+1}$ with the parameter $a_i \in [0, 1]$ be sensible. The more accurate measurement should obtain a higher weight.

As we do not know the exact angle ϕ but only its estimate $\hat{\phi}$ which is generally subject to model uncertainties and measurement errors, it is not advisable to use equation (6.33) with $\phi = \hat{\phi}$ to define the weight matrix. Instead, we make sure that only one of the two in the opposite direction pointing sides is used as formulated in the following proposition.

Proposition 6.4. *Consider the assumptions and angle estimates as in* (6.13)*. Then, the choice of the weights as*

$$w_{\phi,i}(\hat{\phi}_i) = \begin{cases} e_i & \textit{if } \min(\hat{\phi}_i, \hat{\phi}_{i+1}) = \hat{\phi}_i \textit{ or } \hat{\phi}_i = \hat{\phi}_{i+1} \textit{ for } i \in \{1, 3, 5\} \\ e_i & \textit{if } \min(\hat{\phi}_{i-1}, \hat{\phi}_i) = \hat{\phi}_i \textit{ or } \hat{\phi}_{i-1} = \hat{\phi}_i \textit{ for } i \in \{2, 4, 6\} \\ 0 & \textit{otherwise} \end{cases} . \tag{6.34}$$

ensures that the matrix $W_\phi N^\top$ has full rank and it is

$$\phi_i > \frac{\pi}{2} \Rightarrow w_{\phi,i} = 0 \tag{6.35}$$

Proof. Due to the structure of N and the definition of ϕ we can establish the connection between ϕ_{i+1} and ϕ_i as

$$\phi_{i+1} = \pi - \phi_i \tag{6.36}$$

for $i = \{1, 3, 5\}$. This shows that the following statement holds

$$\left(\phi_i < \frac{\pi}{2} \wedge \phi_{i+1} > \frac{\pi}{2}\right) \vee \left(\phi_i > \frac{\pi}{2} \wedge \phi_{i+1} < \frac{\pi}{2}\right) \vee \left(\phi_i = \frac{\pi}{2} \wedge \phi_{i+1} = \frac{\pi}{2}\right).$$

Consequently, for the estimation holds according to Proposition 6.1

$$\left(\hat{\phi}_i < \frac{\pi}{2} \wedge \hat{\phi}_{i+1} = \frac{\pi}{2}\right) \vee \left(\hat{\phi}_i = \frac{\pi}{2} \wedge \hat{\phi}_{i+1} < \frac{\pi}{2}\right) \vee \left(\hat{\phi}_i = \frac{\pi}{2} \wedge \hat{\phi}_{i+1} = \frac{\pi}{2}\right). \tag{6.37}$$

The definition of W_ϕ leads to

$$\begin{aligned}\left(W_\phi N^\top e_i = e_i \wedge W_\phi N^\top e_{i+1} = 0\right) \vee \left(W_\phi N^\top e_i = 0 \wedge W_\phi N^\top e_{i+1} = -e_i\right) \\ \vee \left(W_\phi N^\top e_i = e_i \wedge W_\phi N^\top e_{i+1} = -e_i\right)\end{aligned}$$

for $i \in \{1, 3, 5\}$. In either case, the matrix $W_\phi N$ has full rank as its rows consist of all three unit vectors.

It remains to prove the claim (6.35). Suppose there is $i \in \{1, 3, 5\}$ with $\phi_i > \frac{\pi}{2}$. Then $\hat{\phi}_i = \frac{\pi}{2}$ and $\hat{\phi}_{i+1} < \frac{\pi}{2}$ and consequently the definition of the weight matrix (6.34) yields $w_{\phi,i} = 0$. The same argument can be made for $\phi_i > \frac{\pi}{2}$ with $i \in \{2, 4, 6\}$. □

This proposition has given a suitable choice of W_ϕ to ensure the requirements of the weight matrix W_ϕ imposed in Proposition 6.3.

For W_θ the choice of the weights is less straightforward because the influence of θ is propagated by the form factor F and not the cosine. The critical angle $\theta_{\text{crit}} = \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right)$ introduced in (6.28b) is approximately $\frac{5}{6}\pi$ for the considered orbit. Consequently, at least five out of the six sensors provide information that can be used to estimate $A(q)r$ at all times. For a perfect model and measurements it is simple to determine whether a measurement should be incorporated. If $\hat{\theta}_i \geq \theta_{\text{crit}}$ and $\pi - \hat{\theta}_{i+1} \leq \theta_{\text{crit}}$ the angle should not be incorporated into the design for $i \in \{1, 3, 5\}$. A simple robust approach always uses five of the six measurements for the attitude estimation.

Proposition 6.5. *Consider the assumptions and angle estimates as in* (6.13)*. Suppose that* $\theta_{\text{crit}} > \frac{3}{4}\pi$ *and find the index* k *of the highest angle which fulfils*

$$\hat{\theta}_k = \max_{i=1\ldots6} \hat{\theta}_i \,. \tag{6.38}$$

Choose the weight matrix as

$$w_{\theta,i}^{\mathrm{no}}(\hat{\theta}_i) = \begin{cases} \alpha_{\frac{i+1}{2}} e_i & \text{if } i \in \{1,3,5\} \setminus \{k\} \\ (1-\alpha_{\frac{i}{2}}) e_i & \text{if } i \in \{2,4,6\} \setminus \{k\} \\ 0 & \text{if } i = k\,. \end{cases} \tag{6.39}$$

for some design parameters $\alpha_1, \alpha_2, \alpha_3 \in [0,1]$. *Then the matrix* $W_\theta N^\top$ *has full rank and*

$$\theta_i > \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow w_{\theta,i} = 0$$

holds.

Proof. The full rank of the matrix $W_\theta N^\top$ is clear. Since $\theta_{\mathrm{crit}} > \frac{3}{4}\pi$ there is at most one θ_i such that $\theta_i > \theta_{\mathrm{crit}}$ because $\min_{i,j\in\{1,\dots 6\}, j\neq i} \angle(\theta_i, \theta_j) = \frac{\pi}{2}$. If there is k such that $\theta_k > \theta_{\mathrm{crit}}$, then it is $\hat{\theta}_k = \theta_{\mathrm{crit}}$ and consequently $\hat{\theta}_k = \max_{i=1\dots 6} \hat{\theta}_i$. Because all the other estimates are exact we obtain $\hat{\theta}_k^{\mathrm{no}} = \theta_k > \theta_{\mathrm{crit}}$ and therefore $w_{\theta,k}(\hat{\theta}_k) = 0$. □

This Proposition has given a choice of the weight matrix W_θ that ensures the requirements for Proposition (6.38). As shown in the proposition, the k-th measurement is never required to estimate the attitude, as the remaining five measurements are sufficient to generate a valid estimate of q. However, this measurement can improve the estimation if $\theta_k < \theta_{\mathrm{crit}}$. Thus, we proceed with the proposed algorithm, i.e. we use Proposition 6.1, 6.3 and Proposition 6.7, which will be introduced in the next chapter, to define an estimate of the attitude $\hat{q}^{\mathrm{no}}$ that does not use the k-th measurement. Then, we can obtain a second estimate of θ_k as

$$\hat{\theta}_k^{\mathrm{no}} = \arccos\left(\frac{r^\top}{\|r\|} A(\hat{q}^{\mathrm{no}})^\top n_k\right). \tag{6.40}$$

It is sensible to include the k-th measurement if the estimate of this angle is smaller than the critical angle.

Proposition 6.6. *Consider the assumptions and angle estimates as in* (6.13). *Suppose that* $\theta_{\mathrm{crit}} > \frac{3}{4}\pi$. *Define* k *as in* (6.38) *and use* W_θ *as defined in* (6.39) *to obtain a first estimate of the attitude* $\hat{q}^{\mathrm{no}}$. *Then, with the second estimate from the* k*-th angle* $\hat{\theta}_k^{\mathrm{no}}$ *defined by* (6.40) *we choose the new weight matrix as*

$$w_{\theta,i}(\hat{\theta}_i) = \begin{cases} \alpha_{\frac{i+1}{2}} e_i & \text{if } i \in \{1,3,5\} \setminus \{k\} \vee (i \in \{1,3,5\} \wedge \hat{\theta}_k^{\mathrm{no}} \leq \theta_{\mathrm{crit}}) \\ (1-\alpha_{\frac{i}{2}}) e_i & \text{if } i \in \{2,4,6\} \setminus \{k\} \vee (i \in \{2,4,6\} \wedge \hat{\theta}_k^{\mathrm{no}} \leq \theta_{\mathrm{crit}}) \\ 0 & \text{if } i = k \wedge \hat{\theta}_k^{\mathrm{no}} > \theta_{\mathrm{crit}}\,. \end{cases} \tag{6.41}$$

which ensures that the matrix $W_\theta N^\top$ *has full rank and*

$$\theta_i > \frac{\pi}{2} + \sin^{-1}\left(\frac{r_\oplus}{\|r\|}\right) \Rightarrow w_{\theta,i} = 0$$

holds.

Proof. Identical to the proof of Proposition 6.5. □

This Proposition has given a choice of the weight matrix W_θ that ensures the requirements for Proposition (6.38). Of course, choosing the weight matrix W_θ according to (6.5) is sufficient and straightforward. But the approach that uses the weight matrix defined in Proposition (6.6) allows to use all sensible measurements an gives therefore in general a more robust attitude estimation. The next section now shows how to obtain the attitude estimation from the two body vectors.

6.4.3 Attitude Estimation from the Vectors in Body Frame

Obtaining the attitude from two vectors in body frame is a standard problem and sufficiently analysed in the literature [41, 44]. In this section we merely show how we transform our problem into the standard problem and state how this is commonly solved.

From Equation (6.27) follows that the equation

$$A(q)\frac{s}{\|s\|} = \frac{s^{\mathrm{B}}}{\|s^{\mathrm{B}}\|} \tag{6.42a}$$

$$A(q)\frac{r}{\|r\|} = \frac{r^{\mathrm{B}}}{\|r^{\mathrm{B}}\|} \tag{6.42b}$$

holds with

$$s^{\mathrm{B}} = (N^\top)^+ b_\phi \|s\|$$
$$r^{\mathrm{B}} = -(N^\top)^+ b_\theta \|r\|$$

where $\|s^{\mathrm{B}}\| = \|s\|$ and $\|r^{\mathrm{B}}\| = \|r\|$. Estimating the attitude matrix of these systems of equations is a standard problem as discussed in Section 3.3.2. We simply denote by f^{Wahba} as introduced in Section 3.3.2 the function that maps the estimated vectors in initial and body coordinates to the desired attitude q.

Proposition 6.7. *Consider the assumptions and estimates from Proposition* (6.3)*. Then the estimate*

$$\hat{q} = f^{\mathrm{Wahba}}\left(\begin{bmatrix}1 & 1\end{bmatrix}^\top, \begin{bmatrix}\frac{s}{\|s\|} & \frac{r}{\|r\|}\end{bmatrix}^\top, \begin{bmatrix}\frac{\hat{s}^{\mathrm{B}}}{\|\hat{s}^{\mathrm{B}}\|} & \frac{\hat{r}^{\mathrm{B}}}{\|\hat{r}^{\mathrm{B}}\|}\end{bmatrix}^\top\right) \tag{6.43}$$

represents the current attitude, i.e.

$$A(q) = A(\hat{q})\,. \tag{6.44}$$

Proof. Due to Proposition (6.3) it is $\hat{r}^{\mathrm{B}} = r^{\mathrm{B}}$ and $\hat{s}^{\mathrm{B}} = s^{\mathrm{B}}$. Consequently, $\hat{q}$ is a solution of Equation (6.42) which leads directly to the claim. □

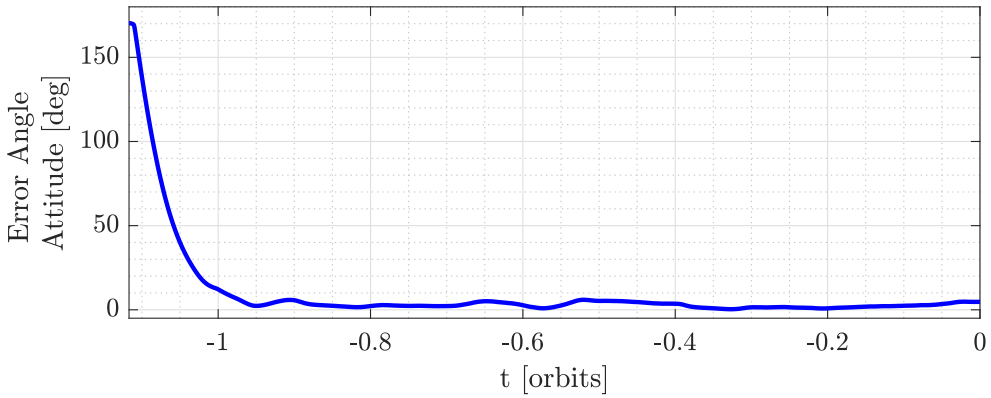


Figure 6.10: Evolution of the error angle of the attitude in the transient phase.

At this point, the attitude estimation can be refined by adding additional vector measurements such as from the magnetic field or the gravity to the optimisation problem. Consequently, adding a single of these vectors can even ensure during eclipse a fine estimation of the attitude.

All in all the proposed algorithm appears to be very suitable to solve the posed problem. The method uses only the measurements of the temperatures T and not of the angular velocity ω and neither of the virtual measurements $q^\top q - 1$ or t. The temperature dynamics described in Eqs. (6.8a) to (6.8c) are used to calculate the estimations of the angle. Neither the quaternion, the angular velocity dynamics in Eqs. (6.8d) and (6.8e) nor the input u is explicitly used in this algorithm. This makes this algorithm simple to implement and a straightforward choice. Improvements can be made by adding the mentioned quantities into the differentiator design or constructing an additional observer using the dynamics, the input and the estimated attitude to refine the estimation. In the next section, we focus on the attitude estimation solely dependent on the temperature measurements and verify them with real data from GRACE.

6.4.4 Validation of the Algorithm Based on Real Data

In this section we use the proposed observer from the previous section to estimate the attitude q solely based on the GRACE measurements discussed in Section 6.1. We implement the algorithm proposed in the previous section and use the nonlinear filter discussed in Section 3.3.2 to smooth the estimation. For the position r we use the direct measurements.

We normalise the time with respect to the orbit period as in Section 6.1. The normalised time is defined such that the slew manoeuvre begins after exactly one orbit. Thus, the simulation starts at a normalised time of around -1.1 with the initial attitude $q_0 = [-0.88\ 0.44\ 0.15\ 0.085]^\top$ and the initial guess $\hat{q}_0 = [0\ 0\ 0\ 1]^\top$, i.e. the initial error is around 170 degrees. The algorithm guarantees through the convergence of the differentiator the convergence to the real attitude with errors induced by the imperfectness of the model. As displayed in Figure 6.10 the transient phase is finished after approximately 0.15 orbits.

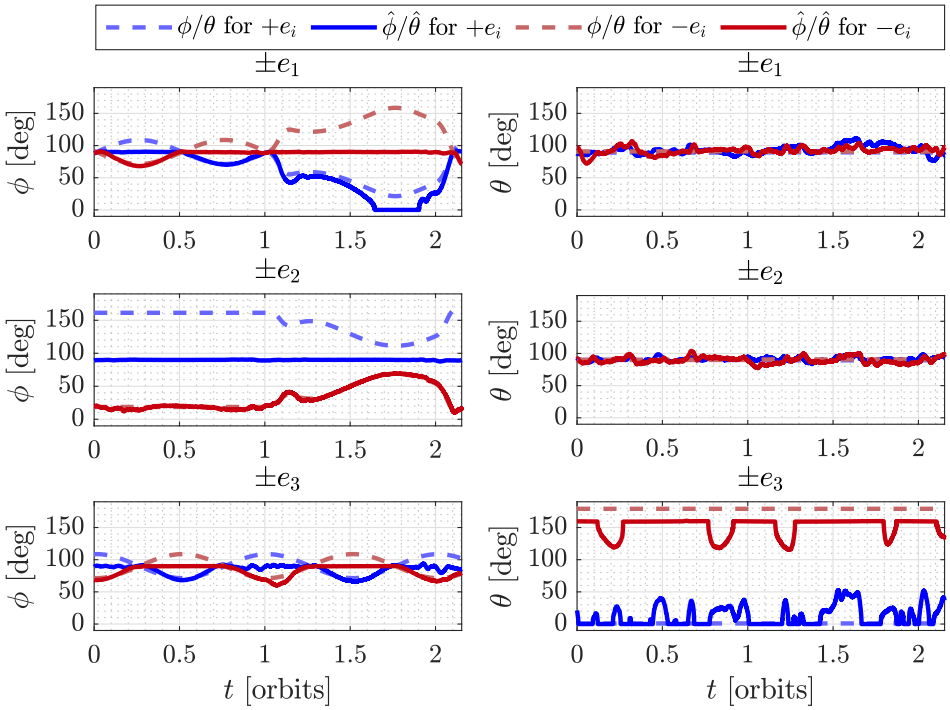


Figure 6.11: Estimated angles $\hat{\phi}$ and $\hat{\theta}$ compared with the real angles ϕ and θ for the normal $n = \pm e_i$.

This is a solid convergence time for the herein considered initial values and time frame. Faster convergence can also be achieved by adjusting the differentiator gains and the filter but this is out of the scope of this section. Here the focus lies on the stationary behaviour when the error due to model faults becomes evident. Therefore, we show in the following the estimates between orbit 0 and 2.

Figure 6.11 shows all twelve estimated angles $\hat{\phi}$ and $\hat{\vartheta}$ over the number of orbits. Each column consist of three plots, each containing the real angles displayed by a dotted line and the estimated angles indicated by the solid lines. Each of the three plots corresponds to two normals pointing into opposite directions. The left column shows the angles to the Sun ϕ and the right column shows the angles to the Earth θ. The first plot in the left column shows the estimates of the solar angle for the sensors pointing in and oppose the velocity direction. During the first orbit in science mode the angles below 90 degrees are estimated pretty accurately. As expected, angles over 90 degrees are estimated to be 90 degrees by the algorithm. During the slew manoeuvre the estimation error becomes more visible as the solar irradiation appears to be estimated to be stronger than it actually is. This suggests that lower values in the model for the parameter α_s may improve this estimation. The second plot in the left column displays the angle of the sensors perpendicular to the orbit plane. It can be seen that the estimates are very good even during the slew manoeuvre. The third plot in the left column shows the estimates for the nadir and zenith pointing sensors. Here the estimation shows some small errors during the science mode as well as the slew mode. In particular, the large error during the start of the slew manoeuvre is due to some

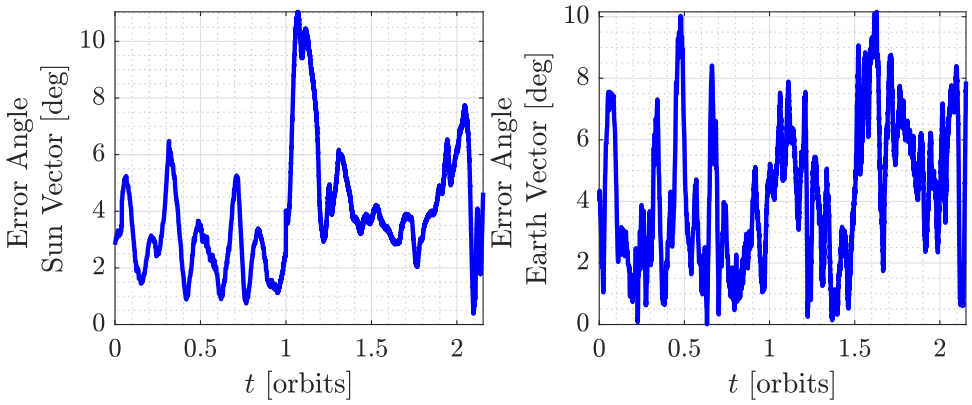


Figure 6.12: Angle error between the estimated vector and the real vectors.

dynamics that have not been incorporated in the model. For the second column the first and second plots show the angle to Earth of the sensors pointing in direction of the velocity and perpendicular to the orbit plane. These errors are higher than the ones from the solar angle but they appear to oscillate around the correct mean value. The third plot in the second column shows the estimates for the angles to $\pm e_3$. The normal $-e_3$ is pointing towards zenith and can therefore not be correctly estimated and leads instead mostly to an estimation of the maximal angle. The normal $+e_3$ is pointing towards nadir, the estimation shows some significant errors for this normal as the time variance of the infrared and albedo irradiation is only roughly incorporated in the model. All in all the estimated angles show good results with error sources that have already been mentioned and discussed in Section 6.2.

Figure 6.12 shows the error angle between estimated and real Sun and Earth vector. Table 6.7 shows the mean and the standard deviation of the error angles. It can be seen that the error angle between the Sun vector and its estimate in science mode is of low dynamics in comparison to the error angle of the Earth vector. Additionally, during science mode the solar vector is more accurately estimated than the Earth vector. Remember that the model was fitted to the data during science mode which is why it is expected to have better results in science mode than in slew mode. Indeed, during slew mode the estimate of both vectors becomes less accurate. For the solar vector, a significant spike occurs at the start of the manoeuvre as discussed due to the high error in the estimate of the ϕ angle in e_1 direction. The Earth vector estimate admits larger errors during the second phase of the orbit.

Figure 6.13 and Table 6.7 show the error of the estimated attitude in comparison to the real attitude. Note that these results were obtained after the observer (3.24) with the gain $k = 0.01$, so the high oscillations that could be seen in the vector estimation in Figure 6.12 are not existent any more. Again, it is evident that the results are better during the science mode than during the slew mode. All in all, the mean of the total error is about 4 degrees which makes the method a viable option for attitude estimation.

The proposed algorithm is very well structured in the different tasks of differentiation, angle estimation, vector estimation and attitude estimation. This gives numerous possibilities where extensions can be made to the proposed algorithm in order to potentially improve

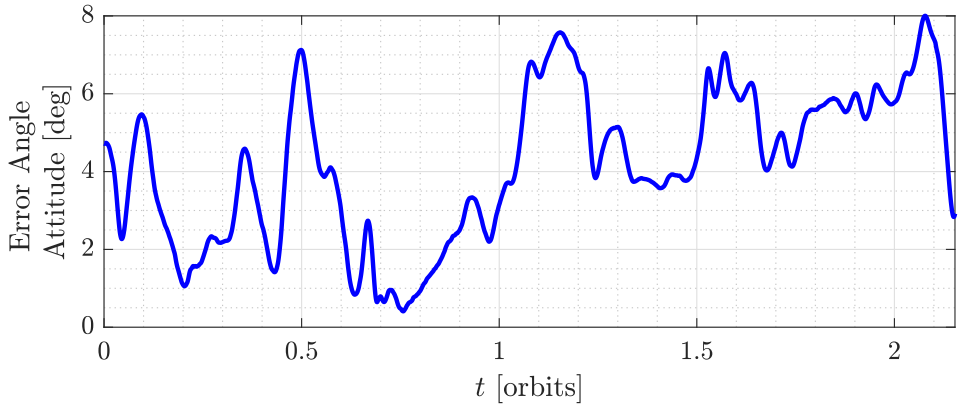


Figure 6.13: Error between the real and the estimated attitude.

Table 6.7: Mean and Standard deviation of the error angles of the estimated vectors and attitude over the course of two orbits. The science mode (SM) denotes the time frame until the first orbit and the slew mode (Slew) denotes the time frame from the first orbit when the slew manoeuvre starts to the end of the slew manoeuvre. Total denotes the mean and the standard deviation for the complete time frame.

	Mean in deg			Standard Deviation in degree		
Error Angle	SM	Slew	Total	SM	Slew	Total
$\angle(s^{\mathrm{B}}, \hat{s}^{\mathrm{B}})$	2.8	4.82	3.80	1.26	2.15	2.03
$\angle(r^{\mathrm{B}}, \hat{r}^{\mathrm{B}})$	3.36	4.92	4.14	2.15	2.19	2.31
$\angle(q, \hat{q})$	2.71	5.27	3.99	1.57	1.10	1.86

the results. For example, additional filters and observers may be introduced between the individual steps to smooth the results in a different manner. Also interesting is the idea of casting the proposed observer in transformed coordinates into natural coordinates. This avoids the differentiation task and allows the observer design directly in a set of desired coordinates.

6.5 Observer in Natural Coordinates

In this section we use the observer proposed in the previous section to design the corresponding observer in natural coordinates in the sense of Section 2.3. While the theoretic application is straightforward and the singularity problems from Section 5.4 do not occur, some other issues arise that are discussed in this section. As proposed in Section 2.3, the natural observer can be obtained by simply augmenting the dynamics with the Jacobian of the observability mapping proposed in the previous section. However, it is desirable to avoid incorporating quaternions in the state as this requires some form of normalisation. Furthermore, the observability mapping discussed in the previous section is divided into some meaningful intermediate steps. It is beneficial to estimate one of the intermediate

variables and use the remaining equations to determine the actual attitude. The choice of the intermediate variable is not trivial and discussed in this section. Secondly, the actual observer design suffers from a similar problem as in the previous section, namely how to detect and combine the outputs containing actual attitude information. This task is not trivial and turns out to be more challenging than before. This leads to two observer designs which performance will then be evaluated in simulations using the experimental data.

6.5.1 Choice of the Observer States

As mentioned in the introduction, choosing q as the explicit states in the observer is not optimal due to most observers being designed with an Euclidean space in mind. Consequently, we analyse the temperature dynamics used in the previous section

$$\dot{T}_i = \alpha_i(t)\max\Big(\cos(\phi_i(q)),0\Big) + \max\Big(\beta_i(t)F(t,\theta_i(t,q),0)\Big) + \gamma_i F(t,\theta_i(t,q)) - \delta_i T_i^4 + \frac{Q_i}{C_i}$$

in order to find more suitable states for the observer design. A natural choice is to use the angles ϕ and θ as these are also estimated in Equation (6.13). Another intuitive choice for the states is the cosine of the angle ϕ, its maximum, as well as the form factor F as these are directly used in the temperature dynamics. Using these variables as states requires to define the mapping that transforms these states into the attitude q. Thus, we denote for each combination of these states (φ,ϑ) for $\varphi \in \{\max(\cos(\phi),0),\cos(\phi),\phi\}$, $\vartheta \in \{F(\theta,r),\theta\}$ a mapping $l_{i,j}^{-1}$ which transforms the considered states to its corresponding quaternion with $i \in \{1,2,3\}$ and $j \in \{1,2\}$. The mapping of the indices i,j to the considered states is the following, for $i=1$, the state is considered to be $\max(\cos(\phi),0)$, for $i=2$ it is $\cos(\phi)$ and for $i=3$ the mapping corresponds to the state ϕ. Analogously, for $j=1$ the mapping corresponds to the state $F(\theta,r)$ and for $j=2$ it shall be associated to θ. To allow a compact notation, we write for the cosine of a vector $\cos(\phi)=[\cos(\phi_1),\dots\cos(\phi_6)]$ and for the maximum of the cosine vector we write $\max(\cos(\phi),0)=[\max(\cos(\phi_1),0),\dots\max(\cos(\phi_6),0)]$ and for the vector valued form factor $F(\theta,r)=[F(\theta_1,r),\dots F(\theta_6,r)]$.

Starting with the first function $l_{1,1}^{-1}:\Phi_1^6\times\Theta_1^6\to\mathbb{S}_3$ which maps $\max(\cos(\theta),0)$ and the form factor $F(\theta,r)$ to its attitude q we use the Equations (6.29) and (6.43) developed in the previous section for the definition

$$l_{1,1}^{-1}(\varphi,\vartheta) = f^{\text{Wahba}}\left(\begin{bmatrix}1&1\end{bmatrix}^\top, \begin{bmatrix}\frac{s}{\|s\|} & \frac{r}{\|r\|}\end{bmatrix}^\top, \begin{bmatrix}\frac{(W_\phi N^\top)^+ W_\phi\varphi}{\|(W_\phi N^\top)^+ W_\phi\varphi\|} & \frac{-(W_\phi N^\top)^+ W_\phi\vartheta}{\|(W_\phi N^\top)^+ W_\phi\vartheta\|}\end{bmatrix}^\top\right) \tag{6.45}$$

where W_ϕ,W_θ and N are defined as in the previous section. This allows to define the functions for the other states $l_{i,j}^{-1}:\Phi_i^6\times\Theta_j^6\to\mathbb{S}_3$ where

$$l_{2,j}^{-1}(\varphi,\vartheta) = l_{1,j}^{-1}(\max(\varphi,0),\vartheta) \tag{6.46}$$

$$l_{3,j}^{-1}(\varphi,\vartheta) = l_{1,j}^{-1}(\max(\cos(\varphi),0),\vartheta) \tag{6.47}$$

$$l_{i,2}^{-1}(\varphi,\vartheta) = l_{i,1}^{-1}(\varphi,F(\vartheta,r)) \tag{6.48}$$

with $i \in \{1,2,3\}$ and $j \in \{1,2\}$. These are functions from the six dimensional (ϕ,θ)-space into the three dimensional attitude space. Generally, these kind of functions are not invertible, but we choose the domains to be

$$\Phi_1 = [0,1]\,, \qquad \Phi_2 = [-1,1]\,, \qquad \Phi_3 = [0,\pi]\,, \qquad \Theta_1 = [0,1]\,, \qquad \Theta_2 = [0,\pi]$$

such that an inverse function exists. These domains can easily be augmented if it is required at the cost of the loss of injectivity. In order to obtain the inverse functions, we consider $l_{3,2} : \mathbb{S}_3 \to [0,\pi]^6 \times [0,\pi]^6$, the function that maps the quaternion to the states associated directly to the angles ϕ and θ. It is defined by $l_{3,2} = [l_{3,2,1}, \ldots, l_{3,2,12}]$ and

$$l_{3,2,k}(q) = \begin{cases} \arccos(s^\top A(q)^\top n_k) & \text{if } k \in \mathbb{N}_{\leq 6} \\ \arccos(-r^\top A(q)^\top n_{k-6}) & \text{if } k \in \mathbb{N}_{\geq 7} \end{cases} \tag{6.49}$$

where $\mathbb{N}_{\leq 6}$ denotes all natural numbers less than or equal to six and $\mathbb{N}_{\geq 7}$ all natural numbers higher than six. This allows to define the other inverse functions $l_{i,j} : \mathbb{S}_3 \to \Phi_i^6 \times \Theta_j^6$ with $l_{i,j} = [l_{i,j,1}, \ldots, l_{i,j,12}]$. These individual functions are defined as $l_{i,j,k} : \mathbb{S}_3 \to Y$ with either $Y = \Phi_i$ if $k \in \mathbb{N}_{\leq 6}$ or $Y = \Theta_j$ if $k \in \mathbb{N}_{\geq 7}$ and their respective function value is

$$l_{2,j,k}(q) = \begin{cases} \cos(l_{3,2,k}(q)) & \text{if } k \in \mathbb{N}_{\leq 6} \\ l_{3,j,k}(q) & \text{if } k \in \mathbb{N}_{\geq 7} \end{cases} \tag{6.50a}$$

$$l_{1,j,k}(q) = \begin{cases} \max(\cos(l_{3,2,k}(q)), 0) & \text{if } k \in \mathbb{N}_{\leq 6} \\ l_{3,j,k}(q) & \text{if } k \in \mathbb{N}_{\geq 7} \end{cases} \tag{6.50b}$$

$$l_{i,1,k}(q) = \begin{cases} l_{i,2,k}(q) & \text{if } k \in \mathbb{N}_{\leq 6} \\ F(l_{i,2,k}(q), r) & \text{if } k \in \mathbb{N}_{\geq 7} \end{cases} \tag{6.50c}$$

with $i \in \{1,2,3\}$ and $j \in \{1,2\}$. The proof that these are the inverse functions is a simple conclusion of the analysis in the previous section. The fact that all these functions exist, allows to define an observer that estimates the states φ, ϑ and then use the function $l_{i,j}^{-1}$ to estimate the attitude. We define the corresponding observers to see the differences in the choice of the states.

For every state φ, ϑ we define the corresponding temperature T dynamics $f_T^{i,j}(T,\varphi,\vartheta,t)$, angle φ dynamics $f_\varphi^{i,j}(T,\varphi,\vartheta,\omega)$ and angle ϑ dynamics $f_\vartheta^{i,j}(T,\varphi,\vartheta,\omega,t)$ which are given explicitly in the next sections. This leads to the observer designs in the sense of Section 2.3 as

$$\begin{bmatrix} \dot{\hat{T}} \\ \dot{\hat{\varphi}} \\ \dot{\hat{\vartheta}} \\ \dot{\hat{\omega}} \\ \dot{\hat{t}} \end{bmatrix} = \begin{bmatrix} f_T^{i,j}(T,\varphi,\vartheta,t) \\ f_\varphi^{i,j}(T,\varphi,\vartheta,\omega) \\ f_\vartheta^{i,j}(T,\varphi,\vartheta,\omega,t) \\ f_\omega(\omega,u) \\ f_t \end{bmatrix} + \mathrm{d}\mathcal{O}_{i,j}(T,\varphi,\vartheta,\omega,t)^{-1} L \left(\begin{bmatrix} T \\ \omega \\ t \end{bmatrix} - \begin{bmatrix} \hat{T} \\ \hat{\omega} \\ \hat{t} \end{bmatrix} \right) \tag{6.51}$$

where $\mathrm{d}\mathcal{O}_{i,j}(T,\varphi,\vartheta,\omega,t)$ denotes the Jacobian of the observability mapping and L the tuning matrix given in detail in the following. The observability mapping $\mathcal{O}_{i,j} : \mathbb{R}^{28} \to \mathbb{R}^{28}$ is defined

in a manner such that the first entries are always a pair consisting of the temperature measurements and its corresponding derivative. Thus, for $k \in \{1, \dots 28\}$ the k-th entry of the observability mapping $\mathcal{O}_{i,j}$ has the form

$$\mathcal{O}_{i,j,k}(T,\varphi,\vartheta,\omega,t) = \begin{cases} T_{\bar{k}} & \text{if } k < 25 \text{ and } k \notin 2\mathbb{N} \text{ with } \bar{k} = \frac{k+1}{2} \\ f^{i,j}_{T,\bar{k}}(T_{\bar{k}}, \varphi_{\bar{k}}, \vartheta_{\bar{k}}, t) & \text{if } k < 25 \text{ and } k \in 2\mathbb{N}_{\leq 6} \text{ with } \bar{k} = \frac{k}{2} \\ f^{i,j}_{T,\bar{k}}(T_{\bar{k}}, \varphi_{\bar{k}-6}, \vartheta_{\bar{k}-6}, t) & \text{if } k < 25 \text{ and } k \in 2\mathbb{N}_{>6} \text{ with } \bar{k} = \frac{k}{2} \\ \omega_{k-24} & \text{with } k \in \{25, 26, 27\} \\ t & \text{with } k = 28 \end{cases} \tag{6.52}$$

where we denote by $f^{i,j}_{T,\bar{k}}(T,\varphi,\vartheta,t)$ the $\bar{k}$-th entry of $f^{i,j}_{T}(T,\varphi,\vartheta,t)$ and by $2\mathbb{N}_{\leq 6}$ the first six even numbers $\{2, 4, \dots, 12\}$. Further, we have identified $f^{i,j}_{T,\bar{k}}(T,\varphi,\vartheta,t)$ with $f^{i,j}_{T,\bar{k}}(T_{\bar{k}}, \varphi_{\bar{k}}, \vartheta_{\bar{k}}, t)$ to emphasise that the $\bar{k}$-th vector field depends only on $(T_{\bar{k}}, \varphi_{\bar{k}}, \vartheta_{\bar{k}}, t)$ and not on the remianing states. Note that the temperature measurements $T_7, \dots T_{12}$ depend on the same angles $\phi_1, \dots, \phi_6, \theta_1, \dots, \theta_6$ as the temperature measurements $T_1, \dots T_6$. The Jacobian of the observability matrix $\mathrm{d}\mathcal{O}_{i,j}(T,\varphi,\vartheta,\omega,t) \in \mathbb{R}^{28\times 28}$ is defined by

$$\mathrm{d}\mathcal{O}_{i,j}(T,\varphi,\vartheta,\omega,t) = \begin{bmatrix} \mathrm{d}\mathcal{O}^{\top}_{x_{1:24},1} & 0 & \mathrm{d}\mathcal{O}^{\top}_{t,1} \\ \vdots & \vdots & \vdots \\ \mathrm{d}\mathcal{O}^{\top}_{x_{1:24},24} & 0 & \mathrm{d}\mathcal{O}^{\top}_{t,24} \\ 0 & I & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{6.53}$$

where the first column is defined by the vector valued entries $\mathrm{d}\mathcal{O}_{x_{1:24},k}$ with

$$\mathrm{d}\mathcal{O}^{\top}_{x_{1:24},k} = \begin{cases} \frac{\partial T_{\bar{k}}}{\partial(T,\varphi,\vartheta)} & \text{if } k \notin 2\mathbb{N} \text{ with } \bar{k} = \frac{k+1}{2} \\ \frac{\partial f^{i,j}_{T,\bar{k}}(T,\varphi,\vartheta,t)}{\partial(T,\varphi,\vartheta)} & \text{if } k \in 2\mathbb{N} \text{ with } \bar{k} = \frac{k}{2} \end{cases} \tag{6.54a}$$

$$= \begin{cases} e^{\top}_{\bar{k}} & \text{if } k \notin 2\mathbb{N} \text{ with } \bar{k} = \frac{k+1}{2} \\ \begin{bmatrix} \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{1,\bar{k}})}{\partial T_{\bar{k}}} e^{\top}_{\bar{k}} & \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{1,\bar{k}})}{\partial \varphi_{\bar{k}}} e^{\top}_{\bar{k}} & \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{1,\bar{k}})}{\partial \vartheta_{\bar{k}}} e^{\top}_{\bar{k}} \end{bmatrix} & \text{if } k \in 2\mathbb{N}_{\leq 6} \text{ with } \bar{k} = \frac{k}{2} \\ \begin{bmatrix} \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{2,\bar{k}})}{\partial T_{\bar{k}}} e^{\top}_{\bar{k}} & \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{2,\bar{k}})}{\partial \varphi_{\bar{k}-6}} e^{\top}_{\bar{k}-6} & \frac{\partial f^{i,j}_{T,\bar{k}}(x^{\mathcal{O}}_{2,\bar{k}})}{\partial \vartheta_{\bar{k}-6}} e^{\top}_{\bar{k}-6} \end{bmatrix} & \text{if } k \in 2\mathbb{N}_{>6} \text{ with } \bar{k} = \frac{k}{2} \end{cases} \tag{6.54b}$$

where we used the notation $x^{\mathcal{O}}_{1,\bar{k}} = (T_{\bar{k}}, \varphi_{\bar{k}}, \vartheta_{\bar{k}}, t)$ and $x^{\mathcal{O}}_{2,\bar{k}} = (T_{\bar{k}}, \varphi_{\bar{k}-6}, \vartheta_{\bar{k}-6}, t)$. Note that $e^{\top}_{\bar{k}}$ denotes the unit vector of appropriate size, i.e. in Equation (6.54) $e_{\bar{k}}$ is of dimension 24,

12 or 6. The last column of the Jacobian consists of scalar entries of the form

$$\mathrm{d}\mathcal{O}_{t,k}^{\top} = \begin{cases} 0 & \text{if } k \notin 2\mathbb{N} \text{ with } \bar{k} = \frac{k+1}{2} \\ \frac{\partial f_{T,\bar{k}}^{i,j}(x_{1,\bar{k}}^{\mathcal{O}})}{\partial t} & \text{if } k \in 2\mathbb{N}_{\leq 6} \text{ with } \bar{k} = \frac{k}{2} \\ \frac{\partial f_{T,\bar{k}}^{i,j}(x_{2,\bar{k}}^{\mathcal{O}})}{\partial t} & \text{if } k \in 2\mathbb{N}_{>6} \text{ with } \bar{k} = \frac{k}{2} \end{cases}$$

for $k \in \{1, \dots 24\}$.

The tuning matrix L is defined by the vectors $L_{T,1}, \dots, L_{T,12} \in \mathbb{R}^{2\times 1}$ and $L_{w,1}, L_{w,2}, L_{w,3}, L_t \in \mathbb{R}$ which are the design parameters for the corresponding decoupled differentiators. This gives $L \in \mathbb{R}^{28\times 16}$ as

$$L = \mathrm{diag}(L_{T,1}, \dots, L_{T,12}, L_{w,1}, L_{w,2}, L_{w,3}, L_t) \tag{6.55}$$

which can be designed in the spirit of Section 2.2 for the system in transformed coordinates.

The rising question is now if the choice of the set of states matters and which one is the best to choose for the proposed observer. Clearly, dependent on the choice of the state φ and ϑ, i.e. the choice of i and j, the dynamics $f_{\varphi}^{i,j}$ and inverse observability matrix $\mathrm{d}\mathcal{O}_{i,j}$ are different. The inverse of the latter might also not exist. Therefore, the different states are not equivalent and we point out in the next section the differences in order to evaluate which state is most suitable for the observer design. We start by introducing the resulting temperature dynamics.

6.5.1.1 Temperature Dynamics

The temperature dynamics have the form $f_T^{i,j} = [f_{T,1}^{i,j}, \dots f_{T,12}^{i,j}]$ for $i \in \{1,2,3\}$ and $j \in \{1,2\}$ where every dynamic depends only on a single temperature T_k, a single state $\varphi_{\bar{k}}$ and a single state $\vartheta_{\bar{k}}$ for $k \in \{1, \dots, 12\}$ and $\bar{k} \in \{1, \dots, 6\}$ and the time t.

For the states $\max(\cos(\phi), 0)$ and the form factor $F(\theta, r)$ the dynamics are defined as

$$f_{T,k}^{1,1}(T, \varphi, \vartheta, t) = \begin{cases} \alpha_k(t)\varphi_k + \Big(\max\left(\beta_k(t), 0\right) + \gamma_k\Big)\vartheta_k - \delta_k T_k^4 + \frac{Q_k}{C_k} & k \in \mathbb{N}_{\leq 6} \\ \alpha_k(t)\varphi_{k-6} + \Big(\max\left(\beta_k(t), 0\right) + \gamma_k\Big)\vartheta_{k-6} - \delta_k T_{k-6}^4 + \frac{Q_k}{C_k} & k \in \mathbb{N}_{\geq 7} \end{cases} \tag{6.56}$$

for $k \in \{1, \dots 12\}$. The other definitions of $f_{T,i}^{i,j}(T, \varphi, \vartheta)$ can then be written as

$$\begin{aligned} f_{T,k}^{2,j}(T, \varphi, \vartheta, t) &= f_{T,k}^{1,j}(T, \max(\varphi, 0), \vartheta, t) \\ f_{T,k}^{3,j}(T, \varphi, \vartheta, t) &= f_{T,k}^{1,j}(T, \max(\cos(\varphi), 0), \vartheta, t) \\ f_{T,k}^{i,2}(T, \varphi, \vartheta, t) &= f_{T,k}^{i,1}(T, \varphi, F(\vartheta, r), t)\,. \end{aligned}$$

Any of these dynamics are Lipschitz continuous. The vector field $f_{T,k}^{1,1}$ is even linear in φ and ϑ. The choice of $i = 2$ introduces a nonlinearity in the form of a maximum function where $i = 3$ introduces also an additional cosine. However the main issue arises from the differentiability of the functions. The derivatives of the temperature dynamics with respect to their individual solar states φ_k have the form

$$\frac{f_{T,k}^{1,j}(T,\varphi,\vartheta,t)}{\partial\varphi_k} = \alpha_k(t) \tag{6.57a}$$

$$\frac{f_{T,k}^{2,j}(T,\varphi,\vartheta,t)}{\partial\varphi_k} = \begin{cases}\alpha_k(t) & \varphi_k > 0\\ 0 & \varphi_k < 0\end{cases} \tag{6.57b}$$

$$\frac{f_{T,k}^{3,j}(T,\varphi,\vartheta,t)}{\partial\varphi_k} = \begin{cases}-\alpha_k(t)\sin(\varphi_k) & \cos(\varphi_k) > 0\\ 0 & \cos(\varphi_k) < 0\end{cases} \tag{6.57c}$$

for $k \in \mathbb{N}_{\leq 6}$. It can be seen that the derivatives (6.57b)and (6.57c) for $i = 2, 3$ are not defined at $\varphi_k = 0$ and $\cos(\varphi_k) = 0$, respectively. The resulting Jacobian matrix from these states used in the observer (6.51) is therefore also not defined at this point. However, the derivative in Equation (6.57a) for $i = 1$ is constant and therefore defined for any φ_k.

This makes $i = 1$, i.e. using $\max(\cos(\phi), 0)$ as the state, the most suitable choice if the existence of the Jacobian matrix was the only criterion. However, this phenomenon is reversed for the angle dynamics.

6.5.1.2 Angle Dynamics

The dynamics of φ are calculated using the definitions (6.49),(6.50) and the quaternion dynamics such that

$$f_{\varphi,k}^{1,j}(T,\varphi,\vartheta,\omega) = \frac{\mathrm{d}}{\mathrm{d}t} l_{1,j,k}(q) = \begin{cases} s^\top \frac{\partial A(q) n_k}{\partial q}\dot{q} & \cos(\varphi_k) > 0\\ 0 & \cos(\varphi_k) < 0\end{cases}. \tag{6.58a}$$

$$f_{\varphi,k}^{2,j}(T,\varphi,\vartheta,\omega) = \frac{\mathrm{d}}{\mathrm{d}t} l_{2,j,k}(q) = s^\top \frac{\partial A(q) n_k}{\partial q}\dot{q} \tag{6.58b}$$

$$f_{\varphi,k}^{3,j}(T,\varphi,\vartheta,\omega) = \frac{\mathrm{d}}{\mathrm{d}t} l_{3,j,k}(q) = \frac{\partial \arccos(\cos(\phi_k))}{\partial x} s^\top \frac{\partial A(q)^\top n_k}{\partial q}\dot{q} \tag{6.58c}$$

with $\dot{q} = \frac{1}{2}\Xi(q)\omega$ for $q = l_{i,j}^{-1}(\varphi,\vartheta)$ and $k \in \{1,\ldots,6\}$. Here, the opposite phenomenon can be seen. The dynamics for $i = 1$ are not defined for $\cos(\varphi_k) = 0$ while no definition issues arises for $i = 2, 3$.

6.5.1.3 Discussion

As presented in the previous sections, problems with the definition of the Jacobi matrix and the dynamics arise dependent on the choice of the states. Clearly, there is no ideal solution

but the drawbacks have to be weighted against each other. In this section, we discuss the choice of the states.

First, it became evident in Section 6.5.1.1 that the choice of $i = 3$ and $i = 2$ has the same definition issues for the Jacobi matrix. More importantly however, Section 6.5.1.2 shows that both angle dynamics have no singularities. In this case it is advisable to choose $i = 2$ over $i = 3$, because the latter introduces and additional nonlinearity in form of a cosine that is then compensated in $l_{3,j}$. This is not a real issue but it impedes the calculations. Thus $i = 2$ is deemed a better choice than $i = 3$.

In order to judge whether $i = 1$ or $i = 2$ is more suitable, we need to have a closer look at the definition issues. For the Jacobi matrix as well as the solar states dynamics the issue arises when the angle to the Sun passes 90 degrees. If this is only an event at a single time t this does not pose a problem. In the next section, the Jacobi matrix will be trimmed such that always three out of the six Equations are used. This makes this matrix undefined if and only if the opposing angles are both identical to 90 degrees. This, however, happens only twice every orbit.

The same cannot be said for the solar angle dynamics. Simply because integrating the dynamics (6.58a) for an initial value of $\phi_k(0) < 0$ leads to the trajectory $\phi_k(t) = 0$ independent of the other entries of ϕ and θ. This is an issue that can and needs to be fixed by using the estimated angle resulting from all entries of ϕ and θ instead of the integrated ϕ_k. Then the dynamics read

$$\frac{\mathrm{d}}{\mathrm{d}t}\max(\cos(\phi_k), 0) = \begin{cases} s^\top \frac{\partial A(q) n_k}{\partial q}\dot{q} & \cos(\phi_k) > 0 \\ 0 & \cos(e_k^\top l_{1,1}(l_{1,1}^{-1}(\phi, \theta))) < 0 \end{cases}. \tag{6.59}$$

Then the definition problem again only arises twice every orbit and the issue is similar to the Jacobi matrix.

The numerical simulation of these dynamics is illustrated in Figure 6.14. The quaternion dynamics $\dot{q} = \frac{1}{2}\Xi(q)\omega$ are simulated for two orbits with q_0 and ω as in the previous section. Additionally, the dynamics (6.58b) and (6.59) with $\phi_k(0) = s^\top A(q)e_k$ and $\theta_k(0) = r^\top A(q)e_k$ are simulated to obtain the angle ϕ_1. The error e of the variable $\cos(\phi_1)$ to a simulation with quaternions q as state is displayed in Figure 6.14, i.e.

$$e = r^\top A(q)e_1 - \cos(\phi_1)\,.$$

Dashed lines show the error for a step size of 10s while solid lines show the error for a step size of 1s. For $\cos(\phi)$ as the state, the error is continuous and always comparatively small. It can be seen that for $\max(\cos(\phi), 0)$ as states, the error rises discontinuously at the points where the ϕ angle crosses 90 degrees (at 0.13 and 0.63 orbits). The rise depends on the step size and it can be seen at 0.13 orbits that a smaller step size leads in general to a smaller rise. The main issue can be seen at 0.63 orbits for a step size of $h = 1$s. The error becomes extremely large at this point and does not converge back to zero. This problem arises only for some step sizes and at different points during the orbit. While these errors may of course be, at least theoretically, compensated by the correction term in the observer, the choice of the states as $\cos(\phi)$ appears to give overall better results. Therefore, we use $i = 2$, i.e.

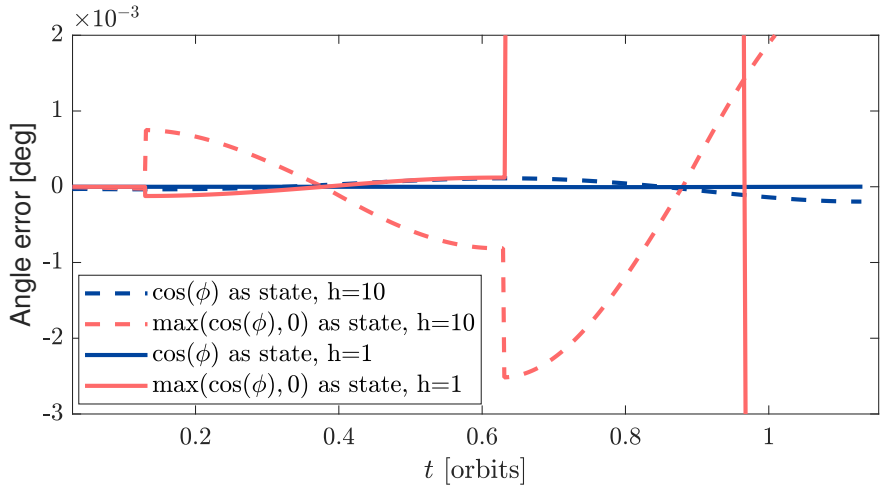


Figure 6.14: Error for the angle obtained from the integrated dynamics with $\cos(\phi)/\max(\cos(\phi), 0)$ as state.

$\cos(\phi)$ as the states for the remainder of this section. The analogue considerations can be made for j leading to $j = 1$, i.e. the state $F(\theta, r)$.

6.5.2 Observer Design

Finishing the choice of the states allows to define the observer dynamics. The task remains to shape the dynamics such that only the angles are reconstructed that directly influence the temperature signals. As can be seen in (6.56), the dynamics $f_{T,i}$ do not depend on ϕ_i if $\cos(\phi_i) < 0$ for $i \in \{1, \dots, 12\}$. Therefore, as shown in (6.57b), the i-th column of the Jacobi matrix $\mathrm{d}\mathcal{O}$ is identical to zero. As discussed in Section 6.4.2 always three out of the six ϕ_i are identical or smaller than zero. Analogously, at most one Form Factor F_i may correspond to a zero column of the Jacobi matrix. Consequently, we remove these states that correspond to columns identical to zero from the observability matrix and define an observer that estimates the remaining states.

We define the vector $k = [k_1 \ \dots \ k_m] \in \mathbb{N}^m_{\leq 28}$ as the vector containing the m column indices of the observability matrix that are identical to zero. This give rises to the definition of the matrix $V(k)$ defined according to Equation (A.49). A simple multiplication with this matrix $V(k)^\top$ allows to remove the entries k of a vector as discussed in Proposition A.1. We denote the state vector by $x = [T^\top \ \cos(\phi)^\top \ F^\top \ \omega^\top \ t]^\top$ and the reduced state vector by $\bar{x} = V^\top(k)x$. The latter is defined to contain only states that really influence the temperature dynamics, i.e. the corresponding Jacobian matrix has full rank. We denote the corresponding dynamics by $\bar{f}$ with $\bar{f}(\bar{x}) = V^\top(k) f(V(k)\bar{x})$ and $f = f_{2,1}$ as in (6.51).

The observability mapping $\mathcal{O} : \mathbb{R}^{28} \to \mathbb{R}^{28}$ for x containing T and its first derivatives is defined as $\mathcal{O} := \mathcal{O}_{2,1}$ in (6.52). The reduced observability mapping $\bar{\mathcal{O}} : \mathbb{R}^{28-m} \to \mathbb{R}^{28}$ with

$m \in \{3, 4\}$ and

$$\bar{\mathcal{O}}(\bar{x}) = \mathcal{O}(V(k)\bar{x}) \tag{6.60}$$

has the same codomain but its domain does not require the states specified in k as they do not act on the dynamics. By the choice of $V(k)$ according to Section 6.4.2, the observability mapping (and therefore the temperature derivatives) are only dependent on the state $\bar{x}$, i.e. it is

$$\mathcal{O}(x) = \mathcal{O}(V(k)\bar{x}) = \bar{O}(\bar{x})\,. \tag{6.61}$$

As a simple consequence, it can be seen that the dynamics of the transformed reduced state $\bar{z} = \bar{\mathcal{O}}(\bar{x})$ are identical to the dynamics of the original transformed state $z = \mathcal{O}(x)$, i.e. $\dot{\bar{z}} = \dot{z}$.

The Jacobian of the reduced observability mapping can be written as

$$\frac{\partial \bar{\mathcal{O}}(\bar{x})}{\partial \bar{x}} = \frac{\partial \mathcal{O}(V(k)\bar{x})}{\partial x} V(k)\,. \tag{6.62}$$

Note that Condition (6.61) ensures that the Jacobian of the original and reduced observability mapping differs only in columns which are identical to 0, i.e.

$$\frac{\partial \mathcal{O}(x)}{\partial x} = \frac{\partial \bar{\mathcal{O}}(\bar{x})}{\partial \bar{x}} V(k)^\top\,. \tag{6.63}$$

Further note, that the Jacobian of the reduced observability mapping is not quadratic $\frac{\partial \bar{\mathcal{O}}(\bar{x})}{\partial \bar{x}} \in \mathbb{R}^{28\times 28-m}$. We continue by proposing two options how to cope with the additional outputs that are not necessary for the estimation.

6.5.2.1 Minimal Output Observer

The first option is to simply use the minimum amount of outputs required for the estimation. This results in an observation problem in the conventional sense and allows an analytical proof of its convergence. Define the vector $\tilde{k}$ which contains the indices of the derivatives of the temperature T in the state $\bar{z}$ that are not required for the estimation. For arbitrary systems, these can be found for example by searching for the rows which removal does not change the rank of the observability matrix. In our case, the index vector can chosen to be $\tilde{k} = 2(k - 12)$. This is due to the structure of the natural state $x = [T_1, \ldots, T_{12}, \phi_1, \ldots, \phi_6, F^\top, \omega^\top, t]^\top$ and the structure of the transformed state $z = [T_1, \dot{T}_1, T_2, \dot{T}_2, \ldots, T_{12}, \dot{T}_{12}, \omega^\top, t]^\top$. Thus, the vector k contains the indices of the states that are not required for the estimation and the definition of the dynamics in (6.9) ensures that $\tilde{k}$ then contains the indices corresponding to the derivatives of the temperatures that can be neglected. Note that there are of course other choices of $\tilde{k}$ possible for the observer design. Then the further reduced observability mapping $\tilde{\mathcal{O}} : \mathbb{R}^{28-m} \mapsto \mathbb{R}^{28-m}$ with

$$\tilde{\mathcal{O}}(\bar{x}) = V(\tilde{k})^\top \mathcal{O}(V(k)\bar{x}) \tag{6.64}$$

is invertible with the Jacobian of the observability mapping

$$\frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}} = V(\tilde{k})^\top \frac{\partial\mathcal{O}(V(k)\bar{x})}{\partial x} V(k)\,. \tag{6.65}$$

We define $\tilde{L} \in \mathbb{R}^{28-m\times 16}$ using the differentiator matrix L in (6.55) for the complete system as

$$\tilde{L} = V(\tilde{k})^\top L \tag{6.66}$$

to define a differentiator for the system in z coordinates. This makes the observer in real coordinates

$$\dot{\hat{\bar{x}}} = \bar{f}(\hat{\bar{x}}) + \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}}^{-1} \tilde{L}(y-\hat{y}) \tag{6.67}$$

$$\hat{y} = \begin{bmatrix}\hat{T}^\top & \hat{\omega}^\top & \hat{t}\end{bmatrix}^\top \tag{6.68}$$

stable by design as shown in the following proposition.

Proposition 6.8. *Define the system*

$$\dot{\bar{x}} = \bar{f}(\bar{x}) + Bu \tag{6.69}$$

$$y = \begin{bmatrix}T^\top & \omega^\top & t\end{bmatrix}^\top \tag{6.70}$$

and the observer candidate

$$\dot{\hat{\bar{x}}} = \bar{f}(\hat{\bar{x}}) + Bu + \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}}^{-1} \tilde{L}(y-\hat{y}) \tag{6.71a}$$

$$\hat{y} = \begin{bmatrix}\hat{T}^\top & \hat{\omega}^\top & \hat{t}\end{bmatrix}^\top \tag{6.71b}$$

with the gain matrix $\tilde{L}$ as in (6.66) *and the Jacobian of the observability matrix $\tilde{\mathcal{O}}$ as in* (6.65)*. Let the transformed coordinates defined by $\tilde{\mathcal{O}}$ fulfil Proposition 2.12. Then it is*

$$\lim_{t\to\infty} \|\bar{x}(t) - \hat{\bar{x}}(t)\| = 0\,. \tag{6.72}$$

Proof. We show the claim by transforming the estimated coordinates into z coordinates, the coordinates describing T and its derivatives, and show that these converge. Consider the transformed coordinates with $\tilde{z} = \tilde{\mathcal{O}}(\bar{x})$ and $\hat{\tilde{z}} = \tilde{\mathcal{O}}(\hat{\bar{x}})$. The transformed dynamics are

$$\begin{aligned}\dot{\hat{\tilde{z}}} &= \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}} \left(\bar{f}(\hat{\bar{x}}) + Bu + \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}}^{-1} \tilde{L}(y-\hat{y})\right) \\ &= \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}} \bar{f}(\hat{\bar{x}}) + \frac{\partial\tilde{\mathcal{O}}(\hat{\bar{x}})}{\partial\bar{x}} Bu + \tilde{L}(y-\hat{y})\end{aligned}$$

which corresponds to a system in integrator form as in Section 2.2. By the design of L it is

$$\lim_{t\to\infty} \|\bar{z}(t) - \hat{\bar{z}}(t)\| = 0 \tag{6.73}$$

and because of Proposition 2.12 it follows

$$\lim_{t\to\infty} \|\bar{x}(t) - \hat{\bar{x}}(t)\| = 0\,. \tag{6.74}$$

□

The proposed observer uses the minimum amount of derivatives of T necessary for the estimation. However, the neglected outputs do contain information that may be used to improve the estimation. One possible way to achieve this is given in the upcoming section.

6.5.2.2 Optimal Observer Design

In this section, we propose to use the pseudo inverse for the Jacobian matrix to incorporate all derivatives that contain information about the attitude. Then the observer has the form

$$\dot{\hat{\bar{x}}} = \bar{f}(\hat{\bar{x}}) + Bu + \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}}^{+} L(y - \hat{y}) \tag{6.75a}$$

$$\hat{y} = \left[\hat{T}^\top \; \hat{\omega}^\top \; \hat{t}\right]^\top \tag{6.75b}$$

with the gain matrix L as in (6.55) and the Jacobian of the observability matrix $\bar{\mathcal{O}}$ as in (6.62). This observer is optimal in a sense that is illustrated in the following. Any observer of the form

$$\dot{\hat{\bar{x}}} = \bar{f}(\hat{\bar{x}}) + Bu + P(\hat{\bar{x}})L(y - \hat{y}) \tag{6.76a}$$

$$\hat{y} = \left[\hat{T}^\top \; \hat{\omega}^\top \; \hat{t}\right]^\top \tag{6.76b}$$

with a matrix $P(\hat{\bar{x}}) \in \mathbb{R}^{28-m\times 28}$ has the transformed dynamics

$$\dot{\hat{\bar{z}}} = \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} \bar{f}(\hat{\bar{x}}) + \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} Bu + \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} P(\hat{\bar{x}}) L(y - \hat{y})\,.$$

The idea is to drive the difference between $L(y - \hat{y})$ and $\frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} P(\hat{\bar{x}}) L(y - \hat{y})$ to zero as this would allow the same proof as in the previous Section 6.5.2.1 to show stability of the observer. However, as $\frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}}$ is not quadratic this is not possible and we relax the condition in order to minimise at least the norm of the difference

$$\min_P \|L(y - \hat{y}) - \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} P(\hat{\bar{x}}) L(y - \hat{y})\| \tag{6.77}$$

for all L and $y, \hat{y}$. This linear optimisation problem is solved for $P = \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}}^{+}$ which motivates the choice of our observer. Note however, that this solution is in general not equivalent to the solution of the actual optimisation problem in mind

$$\min_{P} \lim_{t\to\infty} \|\bar{z}(t) - \hat{\bar{z}}(t)\| .$$

Furthermore, the stability proof of Proposition 6.8 cannot be carried out for the observer (6.75) due to $\frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}} \frac{\partial \bar{\mathcal{O}}(\hat{\bar{x}})}{\partial \bar{x}}^{+} \neq I$. Consequently, its stability needs to be proven in another way. In this work we settle for the numerical validation in the upcoming section. It remains to show, how these two proposed observers can be cast back into the observer for the complete states x.

6.5.2.3 Augmentation on the Complete State Space

As established in the previous section, there are states that we have not incorporated into the observer design for $\bar{x}$. These are

$$x^{\mathrm{no}} = V^{\perp\top}(k)x \tag{6.78}$$

with $V^{\perp}(k)$ defined as in Equation (A.51) separating all states that are indexed in the vector k. These states induce a rank deficit into the observability matrix $\mathcal{O}$ by the definition of k and are not included in the estimate $\hat{\bar{x}}$. The most natural way to estimate complete x is to add these states to the observers (6.71) and (6.75) for $\bar{x}$ by incorporating its dynamics without a correction term.

Proposition 6.9. *The extended observer*

$$\dot{\hat{x}} = f(\hat{x}) + V(k)M(V^{\top}(k)\hat{x})(y - \hat{y}) \tag{6.79}$$

with $M(\bar{x}) \in \mathbb{R}^{28-m\times 16}$ *has the reduced states following its designed dynamics*

$$\dot{\hat{\bar{x}}} = \bar{f}(\bar{x}) + M(\hat{\bar{x}})(y - \hat{y}) \tag{6.80}$$

while the remaining states obey their nominal dynamics without correction

$$\dot{\hat{x}}^{\mathrm{no}} = V^{\perp\top}(k)f(\hat{x}) . \tag{6.81}$$

Proof. Multiplying Equation (6.79) with $V^{\top}(k)$ from the left side yields with $V(k)V^{\top}(k) = I$ to the dynamics

$$\begin{aligned}\dot{\hat{\bar{x}}} &= V^{\top}(k)f(V(k)V^{\top}(k)\hat{x}) + V^{\top}(k)V(k)M(V^{\top}(k)\hat{x})(y - \hat{y}) \\ &= \bar{f}(\bar{x}) + M(\hat{\bar{x}})(y - \hat{y}) .\end{aligned}$$

Multiplying with ${V^{\perp}}^{\top}(k)$ from the left side yields with ${V^{\perp}}^{\top}(k)V(k) = 0$ that

$$\begin{aligned}\dot{\hat{x}}^{\mathrm{no}} &= {V^{\perp}}^{\top}(k)f(\hat{x}) + {V^{\perp}}^{\top}(k)V(k)M(V^{\top}(k)\hat{x})(y-\hat{y}) \\ &= {V^{\perp}}^{\top}(k)f(\hat{x})\,.\end{aligned}$$

□

This proposition shows that the observers

$$\dot{\hat{x}} = f(\hat{x}) + Bu + V(k)\frac{\partial\tilde{\mathcal{O}}(V^{\top}(k)\hat{x})}{\partial\bar{x}}^{-1}\tilde{L}(y-\hat{y}) \tag{6.82a}$$

$$\hat{y} = \left[\hat{T}^{\top}\ \hat{\omega}^{\top}\ \hat{t}\right]^{\top}. \tag{6.82b}$$

and

$$\dot{\hat{x}} = f(\hat{x}) + Bu + V(k)\frac{\partial\bar{\mathcal{O}}(V^{\top}(k)\hat{x})}{\partial\bar{x}}^{+}L(y-\hat{y}) \tag{6.83a}$$

$$\hat{y} = \left[\hat{T}^{\top}\ \hat{\omega}^{\top}\ \hat{t}\right]^{\top} \tag{6.83b}$$

obey the reduced dynamics as established in the Sections 6.5.2.1 and 6.5.2.2. The desired attitude is then obtained simply by applying $\hat{q} = l_{2,1}^{-1}(\cos(\phi), F)$. However, for both observers there are two problems remaining. First, the states $k_1 \ldots, k_m$ are not influenced by any correction term. As the unobservable states change over the course of time with the motion of the satellite, this might lead to high errors at the switching instance. Secondly, the observers suffer from the problem addressed in the definition of $l_{2,1}^{-1} : [-1,1]^6 \times [0,1]^6 \to \mathbb{S}_3$ in (6.45). While the evolution of $\hat{x}_{13:24}$ may be in complete $\mathbb{R}^{12}$, only a subset defined by the image of $l_{2,1} : \mathbb{S}_3 \to [-1,1]^6 \times [0,\pi]^6$ in (6.50) contains the variables $(\cos(\phi), F)$ corresponding to an attitude. The function $l_{2,1}^{-1}$ maps points $(\cos(\phi), F) \notin \mathrm{im}(l_{2,1})$ to the optimal attitude in the sense of Wahba. The observer starts with $(\cos(\phi), F) \in \mathrm{im}(l_{2,1})$ and should ideally not leave $\mathrm{im}(l_{2,1})$ which is not ensured in the current form. Both problems can be tackled simultaneously by introducing a correction term of the current states $(\cos(\phi), F)$ and the states $l_{2,1}(l_{2,1}^{-1}(\cos(\phi), F))$ corresponding to the estimated attitude. The resulting observers have the form

$$\dot{\hat{x}} = f(\hat{x}) + Bu + V(k)\frac{\partial\tilde{\mathcal{O}}(V^{\top}(k)\hat{x})}{\partial\bar{x}}^{-1}\tilde{L}(y-\hat{y}) + K\left(l_{2,1}(l_{2,1}^{-1}(\cos(\hat{\phi}), \hat{F})) - \left[\cos(\hat{\phi})\ \hat{F}\right]^{\top}\right) \tag{6.84a}$$

$$\hat{q} = l_{2,1}(\cos(\hat{\phi}), \hat{F})) \tag{6.84b}$$

$$\hat{y} = \left[\hat{T}^{\top}\ \hat{\omega}^{\top}\ \hat{t}\right]^{\top} \tag{6.84c}$$

and

$$\dot{\hat{x}} = f(\hat{x}) + Bu + V(k)\frac{\partial \bar{\mathcal{O}}(V^\top(k)\hat{x})^+}{\partial \bar{x}} L(y-\hat{y}) + K\left(l_{2,1}(l_{2,1}^{-1}(\cos(\hat{\phi}), \hat{F})) - \left[\cos(\hat{\phi})\ \hat{F}\right]^\top\right) \tag{6.85a}$$

$$\hat{q} = l_{2,1}(\cos(\hat{\phi}), \hat{F}) \tag{6.85b}$$

$$\hat{y} = \left[\hat{T}^\top\ \hat{\omega}^\top\ \hat{t}\right]^\top \tag{6.85c}$$

with the matrix

$$K = \left[0_{12\times 12}\ \mathrm{diag}(k_1^{\mathrm{dp}}, \ldots, k_{12}^{\mathrm{dp}})\ 0_{12\times 4}\right]^\top \in \mathbb{R}^{28\times 12}$$

and the design parameters $k_1^{\mathrm{dp}}, \ldots, k_{12}^{\mathrm{dp}} \in \mathbb{R}_{\geq 0}$.

We can write the system in a compact form with $M(x) \in \left\{\frac{\partial \tilde{\mathcal{O}}(V^\top(k)\hat{x})}{\partial \bar{x}}^{-1}\tilde{L}, \frac{\partial \bar{\mathcal{O}}(V^\top(k)\hat{x})}{\partial \bar{x}}^{+}L\right\}$, $\hat{x}_{13:24} = \left[\cos(\hat{\phi})\ \hat{F}\right]^\top$ and $\lambda(\hat{x}_{13:24}) = l_{2,1}(l_{2,1}^{-1}(\cos(\hat{\phi}), \hat{F})$

$$\dot{\hat{x}} = f(\hat{x}) + Bu + V(k)M(\hat{x})(y-\hat{y}) + K\left(\lambda(\hat{x}_{13:24}) - \hat{x}_{13:24}\right) \tag{6.86a}$$

$$\hat{y} = \left[\hat{T}^\top\ \hat{\omega}\ \hat{t}\right]^\top . \tag{6.86b}$$

Stability of the resulting observer is discussed numerically in the next section. In this section we analyse the analytic error dynamics to identify stabilising and destabilising terms. We first consider the error dynamics between the estimated states and the corresponding estimated attitude. Define $e^{\mathrm{att}} = \lambda(\hat{x}_{13:24}) - \hat{x}_{13:24}$, $k^{\mathrm{att}} = [13, \ldots, 24]$. As discussed in Appendix A.7 using $V(k^{\mathrm{att}})^\perp \in \mathbb{R}^{28\times 12}$ allows to write $V(k^{\mathrm{att}})^{\perp\top}\hat{x} = \hat{x}_{13:24}$. Then the dynamics have the form

$$\begin{aligned}
\dot{e}^{\mathrm{att}} &= \frac{\mathrm{d}}{\mathrm{d}t}\left(\lambda(V(k^{\mathrm{att}})^{\perp\top}\hat{x}) - V(k^{\mathrm{att}})^{\perp\top}\hat{x}\right) \\
&= \left(\frac{\partial\lambda(\hat{x}_{13:24})}{\partial x_{13:24}} - I\right) V(k^{\mathrm{att}})^{\perp\top}\left(f(\hat{x}) + Bu + V(k)M(\hat{x})(y-\hat{y}) + K\left(\lambda(\hat{x}_{13:24}) - \hat{x}_{13:24}\right)\right) \\
&= \left(\frac{\partial\lambda(\hat{x}_{13:24})}{\partial x_{13:24}} - I\right) \mathrm{diag}(k_1^{\mathrm{dp}}, \ldots, k_{12}^{\mathrm{dp}}) e^{\mathrm{att}} \\
&\quad + \left(\frac{\partial\lambda(\hat{x}_{13:24})}{\partial x_{13:24}} - I\right) V^{\perp\top}(k^{\mathrm{att}})\left(f(\hat{x}) + Bu + V(k)M(\hat{x})(y-\hat{y})\right) .
\end{aligned}$$

It can be seen that the first term is stabilising with some assumptions for $\frac{\partial\lambda(\hat{x}_{13:24})}{\partial x_{13:24}}$ as the matrix $\mathrm{diag}(k_1^{\mathrm{dp}}, \ldots, k_{12}^{\mathrm{dp}})$ is positive definite. The second term impairs the convergence as it consists of a part driven by the eigendynamics f and the correction term driven by the error $y - \hat{y}$. Because the parameters k_i^{dp} can be chosen freely, the second term can be dominated and practical stability can be ensured.

Next we determine the dynamics of the unobservable states. It can be readily verified that $V^{\perp\top}(k)V^{\perp}(k^{\mathrm{att}})\hat{x}_{13:24} = x^{\mathrm{no}}$ because $\{k_1, \ldots, k_m\} \subset \{k_1^{\mathrm{att}}, \ldots, k_{12}^{\mathrm{att}}\}$. This leads to the dy-

namics of the unobservable states as

$$\begin{aligned}\dot{\hat{x}}^{\mathrm{no}} &= V^{\perp\top}(k)f(\hat{x}) + V^{\perp\top}(k)K\left(\lambda(\hat{x}_{13:24}) - \hat{x}_{13:24}\right)\\ &= V^{\perp\top}(k)f(\hat{x}) + V^{\perp\top}(k)KV^{\perp\top}(k^{\mathrm{att}})V^{\perp}(k)V^{\perp\top}(k)V^{\perp}(k^{\mathrm{att}})\left(\lambda(\hat{x}_{13:24}) - \hat{x}_{13:24}\right)\\ &= V^{\perp\top}(k)f(\hat{x}) + V^{\perp\top}(k)KV^{\perp\top}(k^{\mathrm{att}})V^{\perp}(k)\left(V^{\perp\top}(k)V^{\perp}(k^{\mathrm{att}})\lambda(\hat{x}_{13:24}) - \hat{x}^{\mathrm{no}}\right)\\ &= f_{k^{\mathrm{no}}}(\hat{x}) + \mathrm{diag}(k^{\mathrm{dp}}_{k_1},\ldots,k^{\mathrm{dp}}_{k_m})\left(\lambda_{k^{\mathrm{no}}}(\hat{x}_{13:24}) - \hat{x}^{\mathrm{no}}\right)\,.\end{aligned}$$

with $k^{\mathrm{no}} = [k_1,\ldots,k_m]$ and the notation $f_k(x) = [f_{k_1}(x),\ldots,f_{k_m}(x)]$ for a vector valued function f with $f(x) = [f_1(x),\ldots,f_n(x)]$. Note that the second equality holds as shown in Lemma A.2. This makes clear that the unobservable states converge to the states resulting from the reconstructed attitude $\lambda_{k^{\mathrm{no}}}(\hat{x}_{13:24})$ for sufficiently high k^{no}.

In the next section we simulate the two proposed observers and show the performance in comparison to the observer in transformed coordinates.

6.5.3 Simulation

We implement the observers (6.84) and (6.85) and compare their results to the observer in transformed coordinates proposed in Section 6.4 with the same settings as before. Note that the implemented observer in transformed coordinates does not incorporate the filter from Section 3.3.2 which makes it independent from angular velocity measurements. This is done to underline the added value if angular velocity measurements are used.

Figure 6.15 shows the angle error in the transient phase. It can be seen that the observer in transformed coordinates has a faster convergence to its stationary behaviour than the observers in natural coordinates. The transient behaviour of the observers in natural coordinates can be divided into three different phases that have approximately exponential convergence to a steady state error. In the final phase the two observers converge to a value close to zero. The observer with the optimal estimate has the slowest convergence time with approximately 10 % of an orbit. This is sufficient for most applications.

Figure 6.16 shows the error of the three observers for one orbit in the science as well as the slew mode from the data set discussed in Section 6.1. It can be seen that all three observers have qualitatively the same behaviour. This suggests that the error is dominated by the impurity of the model and not by the choice of the observer. A significant difference can be seen in the beginning of the slew mode. Here the observer in natural coordinates which incorporate the angular velocity show smaller errors. Note that the angular velocity can also be incorporated into the design of the observer in transformed coordinates as in Section 6.4.4. Than the error of this observer is also smaller as can be seen in Figure 6.13. These results are underlined by the mean and standard deviation shown in Table 6.8. The optimal observer has errors slightly smaller than the one with minimal coordinates. However, all errors are in a similar range.

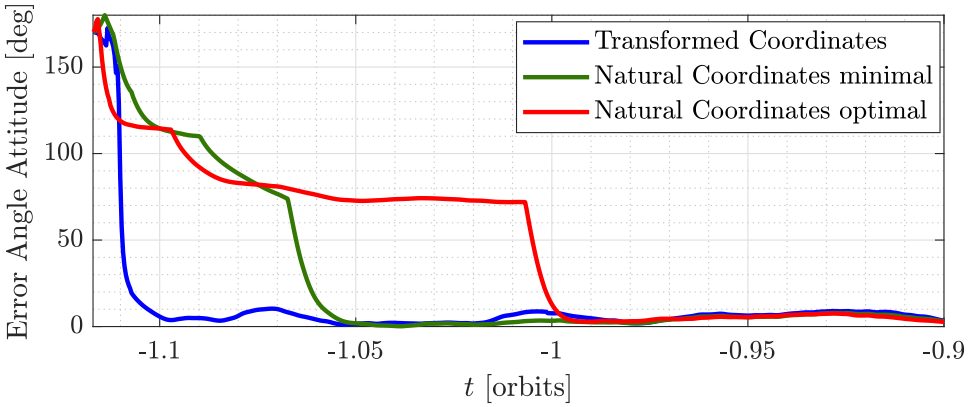


Figure 6.15: Evolution of the attitude angle error in degree in the transient phase from the observers in transformed and in natural coordinates.

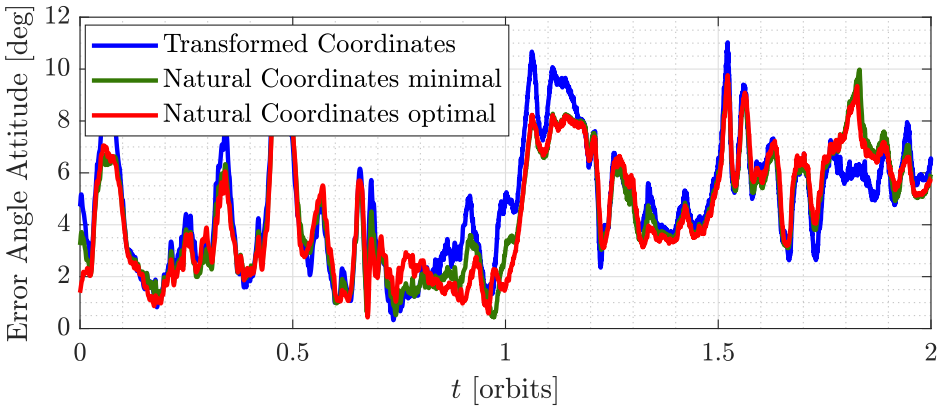


Figure 6.16: Evolution of the attitude angle error in degree in science and slew mode.

Table 6.8: Mean and standard deviation of the angle errors.

	Mean			Standard Deviation		
Error Angle in degree	SM	Slew	Total	SM	Slew	Total
Transformed coordinates	3.51	5.99	4.75	2.11	1.75	2.30
Natural Coordinates minimal	3.14	5.87	4.51	1.89	1.64	2.23
Natural Coordinates optimal	3.10	5.83	4.47	1.88	1.64	2.23

Table 6.9: Main sensors and their accuracies [44].

Sensor	Accuracy [deg]
Sun Sensor	0.01 - 11
Earth Horizon Sensor	0.06 - 1.2
Magnetometer	0.5 - 2.8
Star Tracker	0.0002 - 0.02

6.6 Application and Fusion

In this section we put the proposed method using temperature measurements into context with other means to estimate the attitude. We discuss what the proposed method can offer and in which scenario it can be gainfully applied.

We start with the most common sensors for attitude control and their accuracies. An overview is given in Table 6.9 of the main sensors that are most commonly used to estimate the attitude. These sensors provide vector measurements from which at least two must be used in order to estimate the complete attitude. Additionally, angular velocity measurements from gyroscopes provided by the IMU are used in almost every mission for attitude estimation as they have a very high sampling frequency of multiple kHz which complements the slow sampling frequency of the sensors listed in Table 6.9 of about 1 Hz.

As already discussed in Section 4.2, the underlying temperature dynamics as well as the small sampling time suggest that our proposed method of attitude estimation is not suitable for highly accurate or dynamic manoeuvres. Nevertheless, with the preceding investigations there are three scenarios which come immediately to mind that allow a fruitful application of the proposed method, namely *safe mode operation*, *low cost satellites* and *sensor failure*.

In most conventional operation modes, the attitude of a spacecraft is obtained using star trackers as they give the highest accuracy. It is of tens of magnitudes higher than any other sensor producing vector measurements as can be seen in Table 6.9. Adversely, its computational effort and costs are high. Furthermore, for every mission there exists a safe mode for the spacecraft. This mode is usually designed in order to guarantee thermal and power survival of the spacecraft and their electrical components. Consequently, this mode focuses on detumbling, ensuring that the solar panels see enough sunlight and not letting any of the electrical components overheat. Therefore, non-essential and power consuming electrical components such as star trackers are usually turned off and the desired pointing accuracies are often not more than 10deg [133]. With the temperature sensors being cheap and having a low power consumption, the method proposed in this work is designated to be applied in this scenario. Hereby, one option is to use the temperature measurements as the only source to determine the attitude. However, this means that the attitude during the eclipse phase is estimated by using only an integration of the measured angular velocity, as it can not be fully estimated by the algorithm proposed in this section during the absence of solar irradiation. Instead, the temperature measurements and their resulting estimates can also be combined with other low cost measurements such as magnetometer or acceleration measurements. For the same reasons, it is clear, that the proposed method can also be

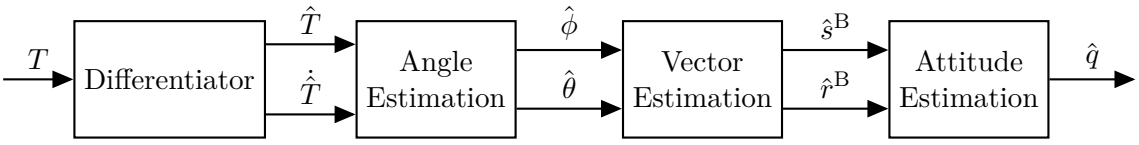


Figure 6.17: Block diagram of the attitude estimation based on temperature.

applied for low cost satellites. Additionally, spacecraft need fall back options in case their nominal attitude sensors fail. Temperature sensors are extremely cheap such that a multitude of them can be used in a hot redundant manner. This makes this method a reliable fall back option if for example the usually used star trackers fail.

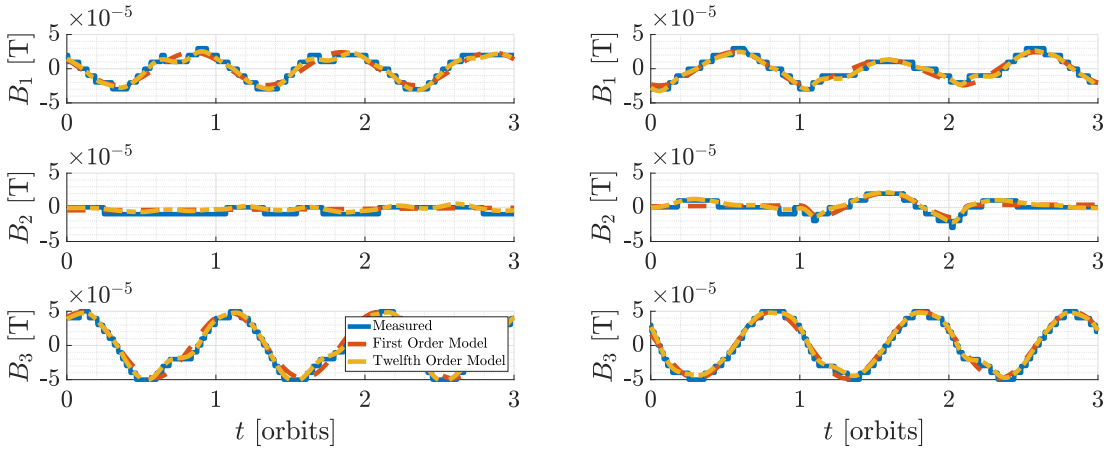
As already mentioned before, it is desirable for each of these three use cases (safe mode, low cost attitude estimation and fallback method) to fuse the temperature measurements with other measurements in order to obtain a more accurate and reliable estimation. We use the illustration of the algorithm in Figure 6.17 in order to discuss at which levels a fusion of the temperature measurements and their resulting estimation is possible with other measurements and estimates. At the highest level, a fusion using the final estimated attitude from the temperature data $\hat{q}^1$ with all other attitude estimates $\hat{q}^2, \dots \hat{q}^{\mathrm{n}}$ is possible. This can be achieved by using for example an Kalman Filter approach as suggested in Chapter 6.2 and [41]. This fusion has the advantage that it allows all the algorithms to work completely decoupled from each other and only use their final estimated attitude. Another option is to use the vector estimates of $\hat{s}^{\mathrm{B}}$ and $\hat{r}^{\mathrm{B}}$ and fuse them with other vector measurements from magnetometers, for instance. This can be achieved as well using the Kalman Filter approach. In the proposed framework the additional measurements can also simply be introduced to the attitude estimation, i.e. Equation (6.43) is replaced by

$$\hat{q} = f^{\mathrm{Wahba}}\left(\alpha^\top, \left[\tfrac{s}{\|s\|} \ \tfrac{r}{\|r\|}, v_1, \dots v_n\right]^\top, \left[\tfrac{\hat{s}^{\mathrm{B}}}{\|\hat{s}^{\mathrm{B}}\|} \ \tfrac{\hat{r}^{\mathrm{B}}}{\|\hat{r}^{\mathrm{B}}\|}, \hat{v}_1^{\mathrm{B}}, \dots \hat{v}_n^{\mathrm{B}}\right]^\top\right) \tag{6.87}$$

where α denotes the gain vector, $v_1, \dots, v_n$ the additional vector measurements in ECI and $\hat{v}_1^{\mathrm{B}}, \dots, \hat{v}_n^{\mathrm{B}}$ in body frame. Finally, if there are vector measurements from Sun and horizon sensors available, they can be fused with the vector estimates $\hat{s}^{\mathrm{B}}$ and $\hat{r}^{\mathrm{B}}$ obtained from the temperature measurements.

We illustrate the potential of the fusion on the vector measurement level, considering additional vector measurements in the form of magnetic field and accelerometer measurements.

The magnetic field measurements are obtained by the on board magnetometer and are displayed in Figure 6.18a and 6.18b. Additionally, the magnetic field values using a first and a twelfth order model for the science mode attitude are shown. These models use spherical harmonics in the spirit of the International Geomagnetic Reference Field [134]. For the twelfth order model we use the implementation of [135], based on [136]. For the first order



(a) Magnetic Field in the Science Mode Data Set. (b) Magnetic Field in the Slew Mode Data Set.

Figure 6.18: Measured and modelled magnetic field vector.

dipole model, as described in [41], the model is simple and has the form

$$B^{\text{Dipole}}(r) = \frac{3m^\top rr - \|r\|^2 m}{\|r\|^3} \text{ where } m = a^3 \begin{bmatrix} g_1^1 \\ h_1^1 \\ g_1^0 \end{bmatrix}$$

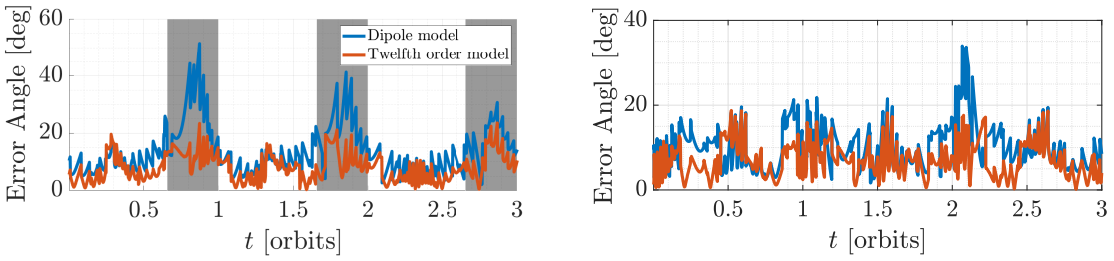
with the position in ECEF r and the parameters $g_1^1 = -1.5011\,\mu\text{T}$, $h_1^1 = 4.7962\,\mu\text{T}$, $g_0^1 = -29.4385\,\mu\text{T}$. The magnetic field is then transformed into body coordinates using the science mode attitude. An in depth analysis of preprocessing and calibration of the magnetic data can be found in [137]. The data obtained for this work had only an accuracy of up to $10\,\mu\text{T}$ as can be seen in Figure 6.18. In comparison to this accuracy and the dynamics exhibited over the orbit the difference between the models of different order are not significant. The slew manoeuvre in Figure 6.18b can be observed mainly in the second vector field component B_2. While this vector field is almost constant during science mode, it exhibits a significant change during the manoeuvre.

Figure 6.19 shows the angle error between the measured magnetic field vector and the magnetic field vector obtained from the first and twelfth order model provided with the actual attitude, time and position. For both data sets, the twelfth order model is around 5 degree more accurate than the dipole model, admitting a mean accuracy of $6 - 8$ degree. For the science mode data set it can be seen that the errors are largest during the eclipse phases, especially for the first order model.

The real accelerometer readings were unfortunately not accessible. Therefore virtual measurements are generated by using

$$a^{B,\text{meas}} = a^B + \eta_a$$

where a^B denotes the real acceleration, $a^{B,\text{meas}}$ the measured acceleration and η_a the measurement noise with $\sigma_\text{a} = 1 \cdot 10^{-3}\,\frac{\text{m}}{\text{s}^2}$. Note that we have deliberately chosen a model that



(a) Error in the Science Mode Data Set.

(b) Error in the Slew Mode Data Set.

Figure 6.19: Angle error between the measured and modelled magnetic field vectors. Black background corresponds to phases when the spacecraft is in solar eclipse.

simulates all the uncertainties as Gaussian white noise. This is sufficient for this kind of analysis as it covers roughly the accuracy that can be obtained from these kind of measurements while still exhibiting the main problem that this vector measurement is parallel to the vector pointing towards Earth r^{B}.

The angular velocity measurements of the IMU are not available as it was malfunctioning and no back-up was available [133]. Therefore they are generated using the measurement model given in [41]

$$\omega^{\mathrm{meas}} = (I + S)\omega + \beta + \eta_\omega \tag{6.88}$$

with

$$S = \begin{pmatrix} s_1 & k_{\mathrm{U},1} & k_{\mathrm{U},2} \\ k_{\mathrm{L},1} & s_2 & k_{\mathrm{U},3} \\ k_{\mathrm{L},2} & k_{\mathrm{L},3} & s_3 \end{pmatrix}$$

and the parameters $\beta_i = 0.49 \cdot 10^{-6} \frac{\mathrm{rad}}{\mathrm{s}}$, $s = [1500\,1000\,1500] \cdot 10^{-6}$, $k_{\mathrm{U}} = [1000\,1500\,2000] \cdot 10^{-6}$ and $k_{\mathrm{L}} = [500\,1000\,1500] \cdot 10^{-6}$ and the noise parameter $\sigma_\omega = \sqrt{10} \cdot 10^{-7} \frac{\mathrm{rad}}{\mathrm{s}}$. Note that these virtual measurements were not corrected by any algorithm that estimates the bias β or the matrix S, thus some deterministic errors are introduced.

We use the algorithm proposed in Section 6.4 to estimate the position of the Sun s^{B} and Earth r^{B} based on temperature data. Then we combine these estimates with the measurements from the magnetic field vector B^{B} and accelerometer a^B, and use a combination of them to obtain an estimate $\hat{q}^{\mathrm{vm}}$ of the attitude, i.e. for $m \in \{2, 3, 4\}$ it is

$$\hat{q}^{\mathrm{vm}} = f^{\mathrm{Wahba}}\left(\alpha^\top, \left[v_1, \ldots, v_m\right]^\top, \left[v_1^{\mathrm{B}}, \ldots, v_m^{\mathrm{B}}\right]^\top\right) \tag{6.89}$$

where $v_i \in \left\{\frac{s}{\|s\|}, \frac{r}{\|r\|}, \frac{B}{\|B\|}, \frac{a}{\|a\|}\right\}$ is the vector in ECI frame obtained by the associated model and v_i^{B} the corresponding measurement in body frame for $i \in \{1, \ldots m\}$. We investigate the following four combinations:

1. Only Magnetometer and Accelerometer. The gain vector is chosen to be $\alpha = [1\ 1]$

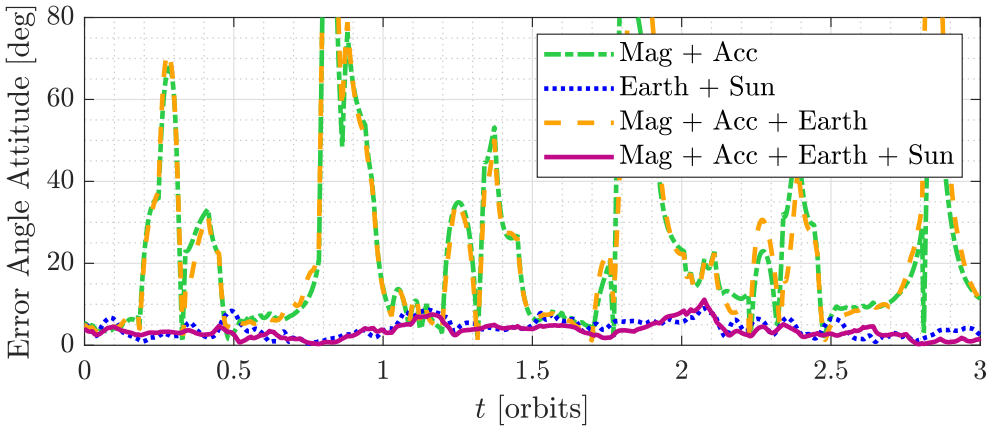


Figure 6.20: Attitude Estimation Error in Slew Mode for different combinations of the vector measurements.

2. Only the estimate of the Earth and Sun vector based on the temperature measurements. The gain vector is chosen to be $\alpha = [1\ 1]$.
3. A combination of the magnetometer, accelerometer measurements with the estimate of the Earth vector. The gain vector is chosen to be $\alpha = [1\ 1\ 1]$
4. The combination of all four measurements and estimates. The gain vector is chosen to be $\alpha = [1\ 1\ 0.5\ 1]$.

Finally, the resulting estimated attitude $\hat{q}^{\mathrm{vm}}$ is used in the nonlinear observer (3.24) with the angular velocity to obtain a better estimation. The gain $k = 0.017$ is chosen such that the nonlinear observer relies more on the information of the vector measurements than on the angular velocity.

We apply the proposed algorithm on the Slew data set. The initial estimated attitude is the attitude rotated by 180 degrees around the x-axis. This shall underline the global stability of the proposed algorithm. Indeed, the algorithm converges in a similar manner as in Figure 6.10 within 0.1 orbits and is therefore not explicitly displayed here. The steady state phase is shown in Figure 6.20 giving the error between the estimated and the real attitude when using a combination of the proposed vector measurements. We have used the simple first order model for the magnetic field vector B to show that it is already good enough to provide sufficiently accurate estimates. It can be seen that any combination of magnetic field vector with an Earth pointing vector measurement such as a^{B} or r^{B} leads to very high errors close to the poles. This is due to the fact that the magnetic field becomes almost parallel to the gravity vector near the poles. Thus, these three measurements are not sufficient to estimate the complete attitude over an entire orbit. This problem can be solved by either relying more on the angular velocity measurements close to the poles as discussed in the next analysis or by incorporating the estimated position of the Sun. The combination of Earth and Sun vector shows the same results as in the previous section. It can be seen that refining these measurements with the additional magnetic field and acceleration measurements improves the estimation. The mean error of the estimation for the fusion of

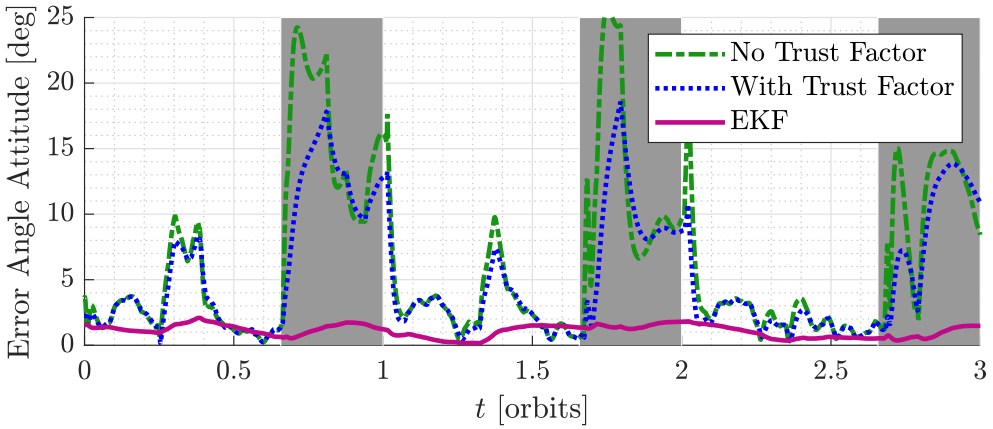


Figure 6.21: Attitude Estimation Error in Science Mode for different observer designs.

all four measurements is 3.49 deg over the course of the three orbits. This is about 0.7 deg better than the estimation using only the temperature measurements. The exact accuracy of the estimation in general depends on the obtained measurements and theory accuracy. Here, the fusion gains were chosen to weigh all measurements equally except the magnetic measurements due to their poor accuracy. In order to obtain the best results, these gains can be optimised dependent on the errors of the individual measurements and estimates. This however, is out of the scope of this chapter as the choice of the gain does not impair the estimation significantly in comparison to the actual measurement accuracies and the underlying models.

In order to improve the accuracy in steady state, it is also advisable to rely more on the angular velocity measurements if the measurements are sufficiently accurate. We underline this point by considering the Science Mode data set. In comparison to the Slew data set, the satellite experiences an eclipse phase for approximately one third of the orbit. During this phase, the Sun position cannot be estimated as discussed in Section 6.4, i.e. we need to choose the gain $\alpha_1 = 0$. Thus, the algorithm needs to rely on the remaining measurements and the angular velocity to estimate the attitude. As was shown in the previous simulation, the measurements without Sun position are not sufficient if the latitude is close to ± 90 deg.

We suggest two possible approaches to deal with this issue. First, we introduce a trust factor $k_{\text{trust}} \in [0, 1]$ for the nonlinear observer (3.24) that is multiplied to the gain such that $k = k_{\text{trust}} k_{\text{const}}$. This trust term uses the available vector measurements to determine how much the model dynamics shall be influenced by the correction term to determine the final attitude. For a trust factor of 0, the nonlinear observer does not use the vector measurements but relies solely on the accuracy of the current estimated position as well as the estimated angular velocity. A trust factor of 1 on the other hand means that the previously designed observer with $k = k_{\text{const}}$ is used. In order to determine this trust factor we calculate the four singular values $\sigma_1, \ldots, \sigma_4$ of the vector measurement matrix

$$R^{\text{vm}} = \left[\alpha_1 \frac{s}{\|s\|}, \alpha_2 \frac{r}{\|r\|}, \alpha_3 \frac{g}{\|g\|}, \alpha_4 \frac{B}{\|B\|} \right] .$$

When $\alpha_i \neq 0$ and $\alpha_i \leq 1$ for all $i \in \{1, 2, 3, 4\}$ the first and second singular value is always bigger than zero and smaller 1 because of the normalisation of the column vectors and the Sun vector and position vector never being parallel. The second singular value becomes zero if $\alpha_1 = 0$ and the remaining three vectors become parallel. This happens during the eclipse phase. Therefore we choose the second singular value as our trust factor $k_{\text{trust}} = \sigma_2(R^{\text{vm}})$.

The second option is to use an observer with a variable gain as in the case of the Extended Kalman Filter (EKF) proposed in Section 3.3.2.2. In this case, the gain is high in the beginning when the covariance of the estimation is high but becomes smaller when the observer transient phase is over. Here, the algorithm deals internally with the problem of the parallel vectors and relies mostly on the angular velocity when the observer has converged. As covariance matrices we have used $Q = 10^{-16} I_3$ and $R = 10^{-9} I_{12}$.

The original algorithm and the two proposed adapted ones are illustrated in Figure 6.21. It can be seen that despite using all four available measurements, the nonlinear observer designs admit high errors during the eclipse phase. On the one hand, these errors are due to the measurements being almost parallel close to the poles. This can be seen by the fact that the trust factor slows down the divergence at the beginning of the eclipse phase. On the other hand, the magnetic field measurements are of comparatively very low accuracy during the eclipse phase as discussed in Figure (6.19a). This makes the estimations during the eclipse phase worse as it relies heavily on these measurements during this period. For the EKF on the other side, it can be seen during the steady state phase the algorithm relies mostly on the angular velocity measurements as the quality of the estimation is barely influenced by the worse magnetic field measurements during the eclipse phase. This leads to a mean error of about 1 degree through the course of these three complete orbits.

7 Conclusions

7.1 Conclusions

This work investigates the topic of attitude estimation for spacecraft using temperature measurements. In order to achieve this task, a mathematical model is developed that describes the spacecraft dynamics including the underlying dynamics of a temperature sensor. The influence of the attitude on the temperature is incorporated as the resulting angle between the heat sources (Earth and Sun) and the normal vector of the temperature sensor. The dynamics of the system are modelled to be time variant as the position is modelled as a function of time and not the state. The validity of the model is shown by fitting its dynamics to the temperature data obtained from the GRACE mission. It can be seen that the model is well suited for this data with R^2 values varying between 78.8% and 99.9%. Deviations mainly occur due to the absence of knowledge about the dissipated heat from the spacecraft as well as the coarse model of albedo and infra-red irradiation.

This model is then used to investigate the two configurations considered in this work. The first configuration uses only a single temperature sensor in order to estimate the attitude. In order to achieve this, derivatives up to at least third order have to be incorporated to enable the attitude estimation. Additionally, a number of necessary conditions must be fulfilled to make the estimation possible. This comprises amongst others that the angular velocity must not be identical to zero and that the aforementioned time variance is not neglected. The latter rules out the application for this method for spacecraft on orbits with high altitude. Most importantly however, there exists a connected manifold with unobservable points that makes the estimation only locally possible. This issue can also be seen in the observer designs. The case in which only infra-red irradiation is acting on the spacecraft is considered for the observer design. This case has quantitatively the worst observability properties while admitting the simplest description of the dynamics which makes it a natural first candidate for the observer design.

One observer in natural and one in transformed coordinates is designed using the system in observability normal form. These observers use a new transformation that allows to incorporate the angle between two vectors directly into the attitude description. This allows a system order reduction of the underlying inversion problem. In the case of the transformed observer, this inversion problem has to be solved numerically to obtain the estimated attitude. The optimisation is circumvented with the design of the observer in natural coordinates. However, this design suffers severely under the issue of the non-observable points as this might lead to instability in the transient phase. An additional function that maps every point into a previously specified subset of possible attitudes allows to solve this problem. This comes at the price of reducing the basin of attraction for the real attitude even more.

To conclude, it was shown that the complete attitude can be estimated using only a single temperature sensor. However, the main difficulties lie in the required accuracies and update rate of the temperature measurements as well as in the absence of a possible global reconstruction.

The second configuration uses twelve temperature sensors. Two in the direction of every single unit vector with different physical properties. This allows using either an observer in natural or in transformed coordinates to estimate the angles between the individual unit vectors and the heat sources Sun/Earth. Subsequent calculations allow then, to obtain an estimate of the Earth and solar vector in body coordinates which is then finally used to obtain the estimated attitude. The observer in transformed coordinates requires only a first order differentiator and some basic calculations to obtain the angles. No numerical solution of a non-linear algebraic system is required as it is usually common for such an observer design. Additionally, the angular velocity is not required to estimate the attitude for the observer in transformed coordinates. However, the estimates become more accurate if a subsequent standard observer is used to refine the estimation. The observer in natural coordinates uses directly the angular velocity measurements but requires some additional add-ons in order to converge. First, the choice of the observer states is not trivial. Additionally, how to deal with the measurements that include redundant information to estimate the angles is vital. Finally, the angles that are not corrected with the measurements need to be corrected using the complete estimated attitude to obtain a smoother estimate.

The estimations for the second configuration are incorporated into other common means to estimate the attitude. The method is capable to provide coarse attitude estimation in case of safe mode, sensor failures or for low cost satellites. It can be used either on its own to estimate the attitude with an accuracy of a few degrees or the method may be fused with other measurements to obtain more accurate results. In the considered case in which the method is fused with magnetometer and accelerometer measurements, a valid estimate can be obtained even near the poles where the accelerometer and magnetometer measurement are almost parallel. Despite this, it can be seen that fusing all of the considered measurements allows a more accurate estimate than using only a part of them. These estimates can even further be improved using an extended Kalman filter which incorporates an adaptive gain and allows to rely more on angular velocity measurements if the estimate is close to the stationary phase.

Concluding, it can be said that the measurements of temperature data are very well suited to obtain or refine the attitude estimate of low Earth orbit spacecraft. It is shown that in theory, already a single temperature sensor measurement is sufficient to estimate the complete attitude by itself when only infra-red irradiation is acting. This however requires an accuracy of the measurements that is not available yet from the current hardware and further needs very accurate and time consuming numerical computations and a good estimate of the initial attitude. These drawbacks disappear when the twelve sensor configuration is used. In this case an attitude estimate within the accuracy of some few degrees can be obtained at any point during the sunlit phase.

7.2 Outlook

In this thesis, the attitude estimations are carried out using exactly one or exactly twelve temperature sensors. A next step is to develop a framework which estimates the attitude for any number of temperature sensors. This requires the combination of higher order derivatives and weight matrices to use only the sensors that carry relevant information.

Then, the observer and control design should be combined and used simultaneously. This means that the attitude control is supplied with estimates from the proposed observer. A fine tuning of the observer and control gain matrices for different missions is required for future application.

Finally, improvements for the accuracy of the proposed methods can be achieved refining the underlying thermal model and using more detailed descriptions of the emitted albedo and infrared irradiation from the Earth.

A Appendix

A.1 Notations and Definitions

This section contains some notations and definitions well established in the context of control theory.

We denote the natural numbers excluding zero by $\mathbb{N}$. Consequently, we denote by $\mathbb{N}_{>n}$ for $n \in \mathbb{N}$ all natural numbers higher than n. The corresponding notations are also used for $\geq$, $>$, $<$, $\leq$ and the real numbers $\mathbb{R}$.

We denote by I_n and 0_n the identity and zero matrix of dimension n for $n \in \mathbb{N}$. We use $0_{m,n}$ to denote the zero matrix of dimension $m \times n$ and we write $I_{m,n} = [I_m \, 0_{m,n-m}]$ for $m, n \in \mathbb{N}$ and $n > m$. We neglect subscripts when the dimension of the matrix is clear. Further, we denote by e_i the i-th unity vector of appropriate size. For any vector $x = [x_1, \ldots, x_n] \in \mathbb{R}^n$ we denote by $x_{i:j}$ with $i, j \in \{1, \ldots n\}$ and $i \leq j$ the entries i to j of x, i.e. $x_{i:j} = [x_i, \ldots, x_j]$. With $A^+ := (A^\top A)^{-1} A$ we denote the Moore-Penrose inverse of the matrix $A \in \mathbb{R}^{n \times n}$. For matrices $A_i \in \mathbb{R}^{m_i \times p_i}$ with $i \in \{1, \ldots, n\}$ and $m_i, p_i \in \mathbb{N}$, we define by $\operatorname{diag}(A_1, \ldots, A_n)$ the diagonal matrix resulting from these matrices, i.e.

$$\operatorname{diag}(A_1, \ldots, A_n) := \begin{bmatrix} A_1 & 0 & 0 & 0 \\ 0 & A_2 & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & A_n \end{bmatrix} \tag{A.1}$$

where 0 denotes the zero matrices of appropriate size.

We define the k-th Lie derivative of h with respect to f in a recursive fashion for $k \geq 1$ as $\mathcal{L}_f^k h(x) = \frac{\partial \mathcal{L}_f^{k-1} h(x)}{\partial x} f(x)$ with the zero order Lie derivative $\mathcal{L}_f^0 h(x) = h(x)$. We denote $\mathcal{L}_f h(x) = \mathcal{L}_f^1 h(x)$. These notations are used to define the relative degree of the system $\dot{x} = f(x) + g(x)u$. It has relative degree r if

$$\mathcal{L}_g \mathcal{L}_f^k h(x) = 0 \qquad \forall \, k \in \{0, \ldots, r-2\} \tag{A.2}$$

$$\mathcal{L}_g \mathcal{L}_f^{r-1} h(x) \neq 0 \, . \tag{A.3}$$

A.2 Attitude Representations

The topic of attitude description and control is vast and the notation used varies dependent on the author. In this work we have used [41] as an orientation for the descriptions and

functions used throughout this work. We use in this work three attitude representations, namely rotation matrices, quaternions and rotation vector and angle representation. For the most parts, quaternions are used as a basis to define all the other representations. This section gives the required transformations between the individual representations as well as some helpful metrics and identities used in this work. For detailed information about the individual representations, and their flaws and benefits the interested reader is referred to the literature, e.g. [41, 44, 138].

The first most common attitude representation are rotation matrices. These are defined on the special orthogonal group

$$\mathcal{SO}_3 = \{A \in \mathbb{R}^{3\times 3} \mid AA^\top = I, \det(A) = 1\}\,.$$

The second representation we consider are quaternions which are defined to be an element of the third unit sphere $\mathbb{S}_3$, where the i-th unit sphere is defined by

$$\mathbb{S}_i := \{x \in \mathbb{R}^{i+1} \mid \|x\| = 1\}\,,$$

with the Euclidean norm $\|\cdot\|$. The final attitude representation considered in this work is the rotation vector and angle representation which consists of a normalised vector $n \in \mathbb{S}_2$ and an angle $\phi \in [0, 2\pi)$.

We denote with $\operatorname{atan2}(\cdot,\cdot)$ as defined in [139] the extension of the function atan which uses two arguments and allows the mapping of a two dimensional vector to an angle between $-\pi$ and π. With $\arcsin(\cdot)$, $\arccos(\cdot)$ we denote the inverse of the sine, cosine function as in [139]. We use this definition to define the angle $\angle(v_1, v_2) \in [0, \pi]$ between two vectors $v_1, v_2 \in \mathbb{R}^n$ as

$$\angle(v_1, v_2) := \arccos\left(\frac{v_1^\top v_2}{\|v_1\|\|v_2\|}\right)\,.$$

Note that in order to avoid an overflow of variables we denote by a slight abuse of notation an element of an attitude representation with the same variable as their corresponding transformation. For example, may the variable q denote a quaternion $q \in \mathbb{S}_3$ or the transformation from attitude matrices to quaternions $q : \mathcal{SO}_3 \to \mathbb{S}_3$ with $q(A) \in \mathbb{S}_3$. Additionally, any definition for a normalised vector $r \in \mathbb{S}_2$ is augmented if necessary for $\bar{r} \in \mathbb{R}^3$ by simply using $r = \frac{\bar{r}}{\|\bar{r}\|}$. In the following we define the transformations required to switch between the individual attitude representations.

Transformation between Quaternion and Rotation Vector and Angle Representation

For every quaternion q the corresponding rotation axis n and angle ϕ as

$$n : \mathbb{S}_3 \to \mathbb{S}_2 \qquad \phi : \mathbb{S}_3 \to [0, 2\pi] \tag{A.4a}$$

$$q \mapsto n(q) = \frac{q_{1:3}}{\sqrt{1 - q_4}} \qquad q \mapsto \phi(q) = 2\arccos(q_4)\,. \tag{A.4b}$$

Conversely, a quaternion can be considered as a function of a rotation axis n and a rotation angle ϕ where

$$\begin{aligned} q : \mathbb{S}_2 \times [0, 2\pi] &\to \mathbb{S}_3 \\ (n, \phi) &\mapsto q(n, \phi) = \left[n_1 \sin(\tfrac{\phi}{2}) \; n_2 \sin(\tfrac{\phi}{2}) \; n_3 \sin(\tfrac{\phi}{2}) \; \cos(\tfrac{\phi}{2}) \right]^\top . \end{aligned} \tag{A.5}$$

Transformations from and to Rotation Matrix Representation

We denote by A the function which maps any attitude representation onto its rotation matrix. In particular, we use the symbol A as a function of q and $r \in \mathbb{S}_2$ and $\phi \in [0, 2\pi]$ with

$$A(q) = (q_4^2 - \|q_{1:3}\|^2) I_3 - 2q_4 [q_{1:3}\times] + 2q_{1:3} q_{1:3}^\top \tag{A.6}$$

$$A(r, \phi) = \cos(\phi) I_3 - \sin(\phi)[r\times] + (1 - \cos(\phi)) r r^\top . \tag{A.7}$$

Conversely, the transformation from rotation matrices to quaternions $q : \mathcal{SO}_3 \to \mathbb{S}_3$ is defined as in [140] with the four quantities

$$\begin{aligned} M_1 &= 1 + a_{11} + a_{22} + a_{33} & M_2 &= 1 + a_{11} - a_{22} - a_{33} \\ M_3 &= 1 - a_{11} + a_{22} - a_{33} & M_4 &= 1 - a_{11} - a_{22} + a_{33} . \end{aligned}$$

These define the four formally identical but numerically differently conditioned quaternions

$$u_1 = \frac{1}{2} \begin{bmatrix} \sqrt{M_1} \\ \frac{a_{32} - a_{23}}{\sqrt{M_1}} \\ \frac{a_{13} - a_{31}}{\sqrt{M_1}} \\ \frac{a_{21} - a_{12}}{\sqrt{M_1}} \end{bmatrix} \quad u_2 = \frac{1}{2} \begin{bmatrix} \frac{a_{32} - a_{23}}{\sqrt{M_2}} \\ \sqrt{M_2} \\ \frac{a_{12} + a_{21}}{\sqrt{M_2}} \\ \frac{a_{31} - a_{13}}{\sqrt{M_2}} \end{bmatrix} \quad u_3 = \frac{1}{2} \begin{bmatrix} \frac{a_{21} - a_{12}}{\sqrt{M_3}} \\ \frac{a_{12} + a_{21}}{\sqrt{M_3}} \\ \sqrt{M_3} \\ \frac{a_{23} - a_{32}}{\sqrt{M_3}} \end{bmatrix} \quad u_4 = \frac{1}{2} \begin{bmatrix} \frac{a_{21} - a_{12}}{\sqrt{M_4}} \\ \frac{a_{31} + a_{13}}{\sqrt{M_4}} \\ \frac{a_{32} + a_{23}}{\sqrt{M_3}} \\ \sqrt{M_4} \end{bmatrix} . \tag{A.8}$$

The desired quaternion is obtained using the best conditioned one according to [140] as

$$q(A) = u_i \qquad \text{with } i \in \{ i \in \{1, 2, 3, 4\} \mid M_i = \max(M_1, M_2, M_3, M_4) \} . \tag{A.9}$$

Consequently, the transformation from rotation matrices to rotation vector and angle can be obtained using the concatenation of the transformation from rotation matrices to quaternions (A.9) and from quaternions to rotation vector and angle (A.4).

Transformation identities

Some useful identities that use the previously introduced transformations are

$$q(n,\phi) = -q(n,\phi - 2\pi) \tag{A.10}$$
$$A(q) = A(-q) \tag{A.11}$$
$$A(r,\theta) = A(r,\theta + 2\pi)\,. \tag{A.12}$$

Additionally, for two vectors $r, n \in \mathbb{S}_3$ with $e_i \times r \neq 0$ and an angle $\theta \in [0,\pi]$ the following equalities hold

$$r^\top A(e_i \times r,\theta)\, r = \cos\theta \tag{A.13}$$
$$A(n,\theta)\, n = n\,. \tag{A.14}$$

A.3 Quaternion Identities

This section gives a brief overview of identities incorporating quaternions that are used throughout this work. Most of them are borrowed from [41].

Cross Product and Quaternion Matrices

Fundamental for the quaternion algebra is the definition of the cross product matrix for a vector $u \in \mathbb{R}^3$ as

$$[u\times] := \begin{pmatrix} 0 & -u_3 & u_2 \\ u_3 & 0 & -u_1 \\ -u_2 & u_1 & 0 \end{pmatrix}.$$

This can be used to define the three matrices

$$\Xi(q) = \begin{pmatrix} q_4 I_3 + [q_{1:3}\times] \\ -q_{1:3}^\top \end{pmatrix}, \quad \Psi(q) = \begin{pmatrix} q_4 I_3 - [q_{1:3}\times] \\ -q_{1:3}^\top \end{pmatrix}, \quad \Omega(\omega) = \begin{pmatrix} -[\omega\times] & \omega \\ -\omega^\top & 0 \end{pmatrix}$$

Some useful identities incorporating these matrices have the form

$$\Xi(q)\omega = \Omega(\omega)q \tag{A.15}$$
$$\Xi^\top(q)\Xi(q) = \|q\|^2 I_3 \tag{A.16}$$
$$\Xi(q)\Xi^\top(q) = \|q\|^2 I_4 - qq^\top \tag{A.17}$$
$$\Xi^\top(q)q = 0\,. \tag{A.18}$$

Note that the norm of the quaternion is identical to one for the considered normalised quaternion and can therefore be omitted in these formulae.

Quaternion Multiplication

The operator for the quaternion multiplication is denoted by $\otimes$ and defined as

$$q \otimes \bar{q} = \begin{bmatrix} \Psi(q)\ q \end{bmatrix} \bar{q} = \begin{bmatrix} \Xi(\bar{q})\ \bar{q} \end{bmatrix} q\,. \tag{A.19}$$

Its definition confers directly to the corresponding attitude matrix

$$A(q \otimes \bar{q}) = A(q)^\top A(\bar{q})^\top\,. \tag{A.20}$$

In order to define an inverse with respect to this multiplication we define the conjugate quaternion denoted by q^* as

$$q^* = \begin{bmatrix} -q_{1:3}\ q_4 \end{bmatrix}^\top = I_{\mathrm{mp}} q \tag{A.21}$$

with $I_{\mathrm{mp}} = \mathrm{diag}(-I_3, 1)$. We also denote the conjugate quaternion by q^{-1} as it is also the inverse quaternion with respect to the multiplication

$$q^* \otimes q = q \otimes q^* = \begin{bmatrix} 0\ 0\ 0\ 1 \end{bmatrix}^\top\,.$$

This property inherits directly to the attitude matrix

$$A(q)^{-1} = A(q)^\top = A(q^*)\,. \tag{A.22}$$

Quaternion Error

The quaternion error between two quaternions q, $\bar{q}$ is defined as

$$q^e(q, \bar{q}) = q \otimes \bar{q}^{-1} = \begin{bmatrix} \Xi^\top(\bar{q}) \\ \bar{q}^\top \end{bmatrix} q = \begin{bmatrix} -\Xi^\top(q) \\ q^\top \end{bmatrix} \bar{q}\,. \tag{A.23}$$

The matrix that allows the definition of this quaternion error also admits the following helpful equalities

$$\begin{bmatrix} \Xi^\top(q) \\ q^\top \end{bmatrix} \Xi(q)\omega = \begin{bmatrix} \omega \\ 0 \end{bmatrix} \tag{A.24}$$

$$\begin{bmatrix} \Xi^\top(q) \\ q^\top \end{bmatrix} \Omega(\omega) \begin{bmatrix} \Xi^\top(q) \\ q^\top \end{bmatrix}^\top = \Omega(\omega)\,. \tag{A.25}$$

This error quaternion can moreover be used to define the error angle between two attitudes as

$$\begin{aligned} \angle^{\mathrm{e}} : \mathbb{S}_3 \times \mathbb{S}_3 &\to [0, \pi] \\ (q_1, q_2) &\mapsto \phi(q^{\mathrm{e}}(q_1, q_2))\,. \end{aligned}$$

and the error vector as

$$\begin{aligned} n^{\mathrm{e}} : \mathbb{S}_3 \times \mathbb{S}_3 &\to \mathbb{S}_2 \\ (q_1, q_2) &\mapsto n(q^{\mathrm{e}}(q_1, q_2)) \end{aligned}$$

with the transformation defined in (A.4).

Quaternion Dynamic Identities

The quaternion dynamics are governed by the Equation

$$\dot{q} = \frac{1}{2}\Xi(q)\omega = \frac{1}{2}\Omega(\omega)q \tag{A.26}$$

which leads to the useful identities

$$\Xi^{\top}(q)\dot{q} = -\Xi^{\top}(\dot{q})q \tag{A.27}$$

$$\Xi^{\top}(\dot{q})\Xi(q) = -\frac{1}{2}q^{\top}q[\omega\times] \tag{A.28}$$

$$\Xi^{\top}(\dot{q})q = -\frac{1}{2}q^{\top}q\omega\,. \tag{A.29}$$

A.4 Parameters and Orbit Data

As a reference orbit we use a Sun-synchronous orbit as described in Section 3.2.3. However, in order to allow a more compact description it is approximated using the function

$$r(t) = \begin{bmatrix} a_1 \sin(\omega t + b_1) \\ a_2 \sin(\omega t + b_2) \\ a_3 \sin(\omega t + b_3) \end{bmatrix}. \tag{A.30}$$

We use the parameters and constants given in Table A.1. The initial states of the analysis are given in (A.2). This leads to the sets

$$\begin{aligned} \mathcal{T}_{\mathrm{ecl}} &= [0, 1168] \cup [5793, 5801] \\ \mathcal{T}_{\mathrm{alb}} &= [1168, 2017] \cup [4944, 5793] \\ \mathcal{T}_{\mathrm{sun}} &= [2017, 4944]\,. \end{aligned}$$

Table A.1: Parameters and Constants

Parameter	Value	Parameter	Value
α_s	0.17	ε_e	0.88
A	0.16 m	n	e_3
C	500 $\frac{\mathrm{J}}{\mathrm{s}}$	s	$[0.1\ 1.4\ 6.0]^\top \cdot 10^{11}$ m
I_IR	239 $\frac{\mathrm{W}}{\mathrm{m}^2}$	$\rho_\mathrm{alb,av}$	0.38
I	$\mathrm{diag}[5.4\ 5.4\ 0.9]\,\mathrm{kg\,m}^2$	a_1	$6.90 \cdot 10^6$ m
a_2	$1.05 \cdot 10^6$ m	a_3	$6.89 \cdot 10^6$ m
ω_o	0.001 $\frac{\mathrm{rad}}{\mathrm{s}}$	b_1	2.56 rad
b_2	-1.64 rad	b_3	-2.17 rad
G_s0	1361 W	au	149 597 870 700 m
$r_\oplus$	6 371 000 m	σ_s	$5.670\,367 \cdot 10^{-8}\,\mathrm{W/m^2K^4}$
$R_\oplus$	6 378 137 m	$GM_\oplus$	$3.986\,004\,418 \cdot 10^{14}$ $\frac{\mathrm{m}^3}{\mathrm{s}^2}$
J_2	$1.086 \cdot 10^{-3}$		

Table A.2: Initial states in Chapter 5

Parameter	Value
T	292 K
θ	$\frac{\pi}{2}$ rad
w	$0.01[0.5774\ 0.5774\ 0.5774]$ $\frac{\mathrm{rad}}{\mathrm{s}}$
t	100 s

A.5 Analytical calculations and proofs of Section 5.2

Derivatives

For the partial derivatives of f_T at (T^*, q^*, t^*) with $r^* = r(t^*)$ and $\theta^* = \theta(q^*, t^*)$ we obtain

$$\frac{\partial f_T}{\partial T} = -4\delta T^{*3} \tag{A.31a}$$

$$\begin{aligned}\frac{\partial f_T}{\partial q} = {} & \frac{\partial \max(\alpha(t^*)\cos(\phi^*), 0)}{\partial x_1}\alpha(t^*)\frac{s^\top}{\|s\|}\frac{\partial A(q^*)^\top n}{\partial q} \\ & + \left(\frac{\partial \max(\beta(t)F(\theta^*), 0)}{\partial x_1}\beta(t) + \gamma\right)\frac{\mathrm{d}F(\theta^*)}{\mathrm{d}\theta}\frac{\partial \theta(t^*, q^*)}{\partial q}\end{aligned} \tag{A.31b}$$

$$\begin{aligned}\frac{\partial f_T}{\partial t} = {} & \frac{\partial \max(\beta F(\theta^*), 0)}{\partial x_1}\dot{\beta}(t^*)F(\theta(t^*, q^*)) \\ & + \frac{\partial \max(\beta F(\theta^*), 0)}{\partial x_1}\beta(t)\frac{dF(\theta^*)}{d\theta}\frac{\partial \theta(t^*, q^*)}{\partial t}\end{aligned} \tag{A.31c}$$

$$+ \gamma\frac{dF(\theta^*)}{d\theta}\frac{\partial \theta(t^*, q^*)}{\partial t} + (\beta(t) + \gamma)\frac{\partial F}{\partial t} \tag{A.31d}$$

with $H = \frac{\|r\|}{r_\oplus}$ and the derivative of the maximum function

$$\frac{\partial \max(x_1, x_2)}{\partial x_1} = \begin{cases} 1 & x_1 > x_2 \\ 0 & x_2 < x_1 \\ \text{undefined} & x_1 = x_2 \end{cases}.$$

The derivative of the form factor is

$$\frac{\partial F(\theta^*)}{\partial \theta} = \begin{cases} -\frac{\sin(\theta^*)}{H^2} & \theta \leq \frac{\pi}{2} - \arcsin\left(\frac{1}{H}\right) \\ \frac{\partial F_2(\theta^*)}{\partial \theta} & \frac{\pi}{2} - \arcsin\left(\frac{1}{H}\right) < \theta < \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \\ 0 & \theta > \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \\ \text{undefined} & \theta = \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \end{cases}$$

$$\frac{\partial F(\theta^*)}{\partial t} = \begin{cases} 2\frac{r_\oplus^2 r^\top}{\|r(t)\|^4}\dot{r}(t)\cos(\theta^*) & \theta \leq \frac{\pi}{2} - \arcsin\left(\frac{1}{H}\right) \text{ and } r \neq 0 \\ \frac{\partial F_2(\theta^*)}{\partial t} & \frac{\pi}{2} - \arcsin\left(\frac{1}{H}\right) < \theta < \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \text{ and } r \neq 0 \\ 0 & \theta > \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \\ \text{undefined} & \theta = \frac{\pi}{2} + \arcsin\left(\frac{1}{H}\right) \text{ and } r = 0 \end{cases}$$

where the analytical expression of $\frac{\partial F_2(\theta^*)}{\partial \theta}$ is omitted here due to space requirements. The partial derivatives of the angle θ are described by

$$\begin{aligned} \frac{\partial \theta(q^*, t^*)}{\partial t} &= \frac{d \arccos(-\frac{r^{*\top}}{\|r^*\|} A(q^*)^\top n)}{dx} \frac{\partial \cos(\theta^*)}{\partial t} \\ \frac{\partial \cos(\theta^*)}{\partial t} &= -n^\top A(q) \frac{\partial}{\partial r}\left(\frac{r}{\|r\|}\right)\dot{r} \\ \frac{\partial}{\partial r}\left(\frac{r}{\|r\|}\right) &= \frac{1}{\|r\|^3}\begin{bmatrix} r_2^2 + r_3^2 & -r_1 r_2 & -r_1 r_3 \\ -r_1 r_2 & r_1^2 + r_3^2 & -r_2 r_3 \\ -r_1 r_3 & -r_2 r_3 & r_1^2 + r_2^2 \end{bmatrix} \\ \frac{\partial \theta(q^*, t^*)}{\partial q} &= \frac{d \arccos(-\frac{r^{*\top}}{\|r^*\|} A(q^*) n)^\top}{dx} \frac{-r^{*\top}}{\|r^*\|} \frac{\partial A(q^*)^\top n}{\partial q} \end{aligned}$$

and the derivative of the rotation matrix is given in [41] as

$$\frac{\partial A(q^*)n}{\partial q} = 2\|q^*\|^{-2}[A(q^*)n\times]\Xi^\top(q^*)\,. \tag{A.32}$$

and thus with (A.21) and (A.22)

$$\frac{\partial A(q^*)^\top n}{\partial q} = 2\|q^*\|^{-2}[A(I_{\text{mp}}q^*)n\times]\Xi^\top(I_{\text{mp}}q^*)I_{\text{mp}}\,.$$

Finally, the derivative of the β variable with respect to time takes the form

$$\dot{\beta}(t) = \frac{1}{C}\rho_{\text{alb}}\alpha_{\text{s}}G_s A\frac{s^\top}{\|s\|}\frac{\partial}{\partial r}\left(\frac{r(t)^\top}{\|r(t)\|}\right)\dot{r}(t)\,.$$

Lie derivatives

It is easy to see that

$$\frac{\partial}{\partial T}\frac{\partial f_T(T,q,t)}{\partial q} = 0 \tag{A.33}$$

$$\frac{\partial}{\partial T}\frac{\partial f_T(T,q,t)}{\partial t} = 0 \tag{A.34}$$

$$\frac{\partial}{\partial q}\left(\frac{\partial f_T(T,q)}{\partial T}f_q(q,\omega)\right) = \frac{\partial f_T(T,q)}{\partial T}\frac{\partial f_q(q,\omega)}{\partial q}\,. \tag{A.35}$$

We use this to obtain the Lie derivatives

$$h(x) = T \tag{A.36a}$$

$$\mathcal{L}_f h(x) = f_T(T,q,t) \tag{A.36b}$$

$$\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q,t)}{\partial T}f_T(T,q,t) + \frac{\partial f_T(T,q,t)}{\partial q}f_q(q,\omega) + \frac{\partial f_T(T,q,t)}{\partial t} \tag{A.36c}$$

$$\begin{aligned}\mathcal{L}_f^3 h(x) = {} & \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T}f_T(T,q,t)\right)f_T(T,q,t) \\ & + \left(\frac{\partial f_T(T,q,t)}{\partial T}\frac{\partial f_T(T,q,t)}{\partial q} + f_q(q,\omega)^\top\frac{\partial^2 f_T(T,q,t)}{\partial q^2}\right. \\ & \quad \left. + \frac{\partial f_T(T,q,t)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial q} + \frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t}\right)f_q(q,\omega) \\ & + \frac{\partial f_T(T,q,t)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial \omega}f_\omega(\omega) \\ & + \frac{\partial f_T(T,q,t)}{\partial T}\frac{f_T(T,q,t)}{\partial t} + \frac{\partial}{\partial t}\frac{\partial f_T(T,q,t)}{\partial q}f_q(q,\omega) + \frac{\partial^2 f_T(T,q,t)}{\partial t^2}\end{aligned} \tag{A.36d}$$

$$\mathcal{L}_g\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q,t)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial \omega}g_\omega \tag{A.36e}$$

and

$$\begin{aligned}
\frac{\partial}{\partial T}\mathcal{L}_f^2 h(x) &= \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q)}{\partial T} f_T(T,q)\right) \\
\frac{\partial}{\partial q}\mathcal{L}_f^2 h(x) &= \frac{\partial f_T(T,q)}{\partial T}\frac{\partial f_T(T,q)}{\partial q} + f_q(q,\omega)^\top \frac{\partial^2 f_T(T,q)}{\partial q^2} \\
&\quad + \frac{\partial f_T(T,q)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial q} + \frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t} \\
\frac{\partial}{\partial \omega}\mathcal{L}_f^2 h(x) &= \frac{\partial f_T(T,q)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial \omega} \\
\frac{\partial}{\partial t}\mathcal{L}_f^2 h(x) &= \frac{\partial f_T(T,q,t)}{\partial T}\frac{f_T(T,q,t)}{\partial t} \\
&\quad + \frac{\partial}{\partial t}\frac{\partial f_T(T,q,t)}{\partial q} f_q(q,\omega) \\
&\quad + \frac{\partial^2 f_T(T,q,t)}{\partial t^2}
\end{aligned}$$

Still Satellite

For $\omega = 0$ the Lie derivatives simplify to

$$h(x) = T \tag{A.37a}$$

$$\mathcal{L}_f h(x) = f_T(T,q,t) \tag{A.37b}$$

$$\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t) + \frac{\partial f_T(T,q,t)}{\partial t} \tag{A.37c}$$

$$\begin{aligned}
\mathcal{L}_f^3 h(x) &= \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t)\right) f_T(T,q,t) \\
&\quad + \frac{\partial f_T(T,q,t)}{\partial T}\frac{f_T(T,q,t)}{\partial t} + \frac{\partial^2 f_T(T,q,t)}{\partial t^2}
\end{aligned} \tag{A.37d}$$

$$\mathcal{L}_g \mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q,t)}{\partial q}\frac{\partial f_q(q,\omega)}{\partial \omega} g_\omega \tag{A.37e}$$

and its derivatives

$$\frac{\partial}{\partial T}\mathcal{L}_f^2 h(x) = \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q)}{\partial T} f_T(T,q)\right) \tag{A.38a}$$

$$\frac{\partial}{\partial q}\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q)}{\partial T}\frac{\partial f_T(T,q)}{\partial q} + \frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t} \tag{A.38b}$$

$$\frac{\partial}{\partial \omega}\mathcal{L}_f^2 h(x) = 0 \tag{A.38c}$$

$$\frac{\partial}{\partial t}\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q,t)}{\partial T}\frac{f_T(T,q,t)}{\partial t} + \frac{\partial^2 f_T(T,q,t)}{\partial t^2} \tag{A.38d}$$

and

$$\begin{aligned}\frac{\partial}{\partial q}\mathcal{L}_f^3 h(x) = &\frac{\partial}{\partial q}\frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t)\right) f_T(T,q,t)\\ &+ \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t)\right)\frac{\partial f_T(T,q,t)}{\partial q}\\ &+ \frac{\partial f_T(T,q,t)}{\partial T}\frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t} + \frac{\partial}{\partial q}\frac{\partial^2 f_T(T,q,t)}{\partial t^2}\end{aligned}$$

Further it is

$$\begin{aligned}&\frac{\partial}{\partial q}\frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t)\right) f_T(T,q,t)\\ &= \frac{\partial}{\partial q}\left(\frac{\partial^2 f_T(T,q,t)}{\partial T^2} f_T(T,q,t) + \left(\frac{\partial f_T(T,q,t)}{\partial T}\right)^2\right) f_T(T,q,t)\\ &= f_T(T,q,t)\frac{\partial^2 f_T(T,q,t)}{\partial T^2}\frac{f_T(T,q,t)}{\partial q}\end{aligned}$$

and

$$\mathrm{d}\mathcal{O}_q = \begin{bmatrix} 0 \\ \frac{\partial}{\partial q}\mathcal{L}_f^1 h \\ \frac{\partial}{\partial q}\mathcal{L}_f^2 h \\ \frac{\partial}{\partial q}\mathcal{L}_f^3 h \\ 2q^\top \end{bmatrix} \tag{A.39}$$

with

$$\begin{aligned}\frac{\partial}{\partial q}\mathcal{L}_f^3 h = &f_T(T,q,t)\frac{\partial^2 f_T(T,q,t)}{\partial T^2}\frac{f_T(T,q,t)}{\partial q} + \frac{\partial}{\partial T}\left(\frac{\partial f_T(T,q,t)}{\partial T} f_T(T,q,t)\right)\frac{\partial f_T(T,q,t)}{\partial q}\\ &+ \frac{\partial f_T(T,q,t)}{\partial T}\frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t} + \frac{\partial}{\partial q}\frac{\partial^2 f_T(T,q,t)}{\partial t^2}\end{aligned} \tag{A.40}$$

A.5.1 Still Spacecraft

Proposition. For a still spacecraft, i.e. $\omega = 0$, the rank of the Jacobian is eight or less.

Proof. We show the loss of rank analytically for Case 4 and 5 with $F = \frac{\cos(\theta)}{H^2}$. Case 3 is a special case shown in the next proposition and the remaining cases are validated numerically. For Case 4 and 5 we show that every row of $\mathrm{d}\mathcal{O}_q$ is in the linear subspace $U := \mathrm{span}\{r^\top \frac{\partial A(q)n}{\partial q}, \dot{r}^\top \frac{\partial A(q)n}{\partial q}, q^\top\}$. For Case 4 and $F = \frac{\cos(\theta)}{H^2}$ we obtain

$$\frac{\partial f_T(T,q,t)}{\partial q} = (\beta+\gamma)\frac{r_\oplus^2}{\|r\|^3} r^\top \frac{\partial A(q)n}{\partial q} \tag{A.41}$$

and

$$\frac{\partial f_T(T,q,t)}{\partial t} = \dot{\beta}\frac{r_\oplus^2}{\|r\|^3} r^\top A(q) n + (\beta+\gamma)\frac{r_\oplus^2}{\|r\|^3}\dot{r}^\top A(q) n - 3(\beta+\gamma)\dot{r}^\top r \frac{r_\oplus^2}{\|r\|^5} r^\top A(q) n\,. \tag{A.42}$$

Further we have

$$\frac{\partial}{\partial q}\mathcal{L}_f^2 h(x) = \frac{\partial f_T(T,q)}{\partial T}\frac{\partial f_T(T,q)}{\partial q} + \frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t} \tag{A.43a}$$

$$\frac{\partial}{\partial q}\mathcal{L}_f^3 h(x) = f_T(x)\frac{\partial^2 f_T(x)}{\partial T^2}\frac{f_T(x)}{\partial q} + \frac{\partial}{\partial T}\left(\frac{\partial f_T(x)}{\partial T} f_T(x)\right)\frac{\partial f_T(x)}{\partial q} + \frac{\partial f_T(x)}{\partial T}\frac{\partial}{\partial q}\frac{\partial f_T(x)}{\partial t} + \frac{\partial}{\partial q}\frac{\partial^2 f_T(x)}{\partial t^2}\,. \tag{A.43b}$$

Inspecting (A.41) it is clear that $\frac{\partial f_T(T,q,t)}{\partial q} \in U$. Further, it is straightforward to see with (A.42) that $\frac{\partial}{\partial q}\frac{\partial f_T(T,q,t)}{\partial t}$ is in U. Thus, it follows from (A.43a) that $\frac{\partial}{\partial q}\mathcal{L}_f^2 h(x)$ is in U as well. By the periodicity of r we have $\ddot{r} \in \operatorname{span}\{r\}$. Now with (A.42) it is evident that $\frac{\partial^2 f_T(T,q,t)}{\partial t^2}$ is in U and we see with (A.43b) that $\frac{\partial}{\partial q}\mathcal{L}_f^3 h(x)$ is also an element of U. Since all rows of $\mathrm{d}\mathcal{O}_q$ are elements of the linear subspace U which is of dimension 3, the rank of $\mathrm{d}\mathcal{O}_q$ is 3 at most and the claim is shown for Case 4. Case 5 is a special case of Case 4. For the remaining cases the analytical proof is more difficult due to the complicated structure of the form factor and the additional solar irradiation. Hence, the remaining cases are verified in a numeric manner by calculating the singular values of $\mathrm{d}\mathcal{O}_q$. Carrying out the procedure for a number of randomly distributed values yields one singular value identical to zero. Consequently, the matrix does not have full rank. □

A.5.2 Still Time-Invariant Spacecraft

Proposition. For a spacecraft described by (5.2) without time variance and $\omega = 0$ the rank of the Jacobian is 7.

Proof. First, we show that for this case the Lie derivatives take a very simple form and state the entries in the Jacobian. Then we show that every row of the Jacobian $\mathrm{d}\mathcal{O}_q$ is in the span of $\frac{\partial f_T(x)}{\partial q}$.

Lemma. For (5.2) without time variance and $\omega = 0$ holds

$$\mathcal{L}_f^i h_1(x) = \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x) f_T(x) \tag{A.44}$$

$$\frac{\partial}{\partial q}\mathcal{L}_f^i h_1(x) = \mathcal{L}_{f_T}\frac{\partial}{\partial q}\mathcal{L}_{f_T}^{i-1} h_1(x) + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q} \tag{A.45}$$

$$\frac{\partial}{\partial q}\mathcal{L}_f^i h_1(x) \in \operatorname{span}\left\{\frac{\partial f_T(x)}{\partial q}\right\}. \tag{A.46}$$

Proof. Equation (A.44) follows from the definition of Lie derivative and the fact that $f_\omega(\omega) = f_q(q,\omega) = 0$ and that $\frac{\partial f_T(x)}{\partial t} = 0$. In order to show (A.45) we use Schwarz's theorem [119] and obtain

$$\begin{aligned}\frac{\partial}{\partial q}\mathcal{L}_{f_T}^i h_1(x) &= \frac{\partial}{\partial q}\left(\frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x) f_T(x)\right)\\ &= \mathcal{L}_{f_T}\frac{\partial}{\partial q}\mathcal{L}_{f_T}^{i-1} h_1(x) + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q}.\end{aligned}$$

We then proof (A.45) by induction. For $i = 0$ and $i = 1$ the equation holds. Assume the equation holds for $i-1$. Then there exists $a(x) \in \mathbb{R}$ such that $\frac{\partial}{\partial q}\mathcal{L}_f^{i-1} h_1(x) = a(x)\frac{\partial f_T(x)}{\partial q}$. With (A.45) and $\frac{\partial}{\partial T}\frac{\partial f_T(T,q,t)}{\partial q} = 0$ we obtain

$$\begin{aligned}\frac{\partial}{\partial q}\mathcal{L}_f^i h_1(x) &= \mathcal{L}_{f_T}\frac{\partial}{\partial q}\mathcal{L}_{f_T}^{i-1} h_1(x) + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q}\\ &= \mathcal{L}_{f_T}\left(a(x)\frac{\partial f_T(x)}{\partial q}\right) + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q}\\ &= \mathcal{L}_{f_T} a(x)\frac{\partial f_T(x)}{\partial q} + a(x)\mathcal{L}_{f_T}\frac{\partial f_T(x)}{\partial q}\\ &\quad + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q}\\ &= \mathcal{L}_{f_T} a(x)\frac{\partial f_T(x)}{\partial q} + \frac{\partial}{\partial T}\mathcal{L}_{f_T}^{i-1} h_1(x)\frac{\partial f_T(x)}{\partial q}.\end{aligned}$$

□

With (A.46) from the lemma and that $\operatorname{span}\left\{\frac{\partial f_T(x)}{\partial q}\right\}$ is a linear subspace of dimension 1, we obtain that the Jacobian (5.5) has rank seven. □

A.5.3 Earth and Sun Pointing

Some spacecraft modes require the pointing into a specific direction to obtain valuable information, e.g. the Earth and the Sun pointing mode. In our case, Earth pointing means

that q is of the form such that

$$A(q)^\top n = -\frac{r}{\|r\|}\,.$$

In this mode only Case 1 and Cases 4–6 can occur.

Proposition. In Case 4–6 for Earth pointing mode, i.e. $A(q)^\top n = -\frac{r}{\|r\|}$, the rank of the Jacobian is at most 8.

Proof. By definition of the form factor the albedo and infra-red irradiations are maximal. Thus, their derivative with respect to q is identical to zero. Then, because of no solar irradiation, we have $\frac{\partial}{\partial q} f_T(t, q, \omega) = 0$ and the Jacobian cannot have full rank. □

For a Sun pointing mode $A(q)^\top n = \frac{s}{\|s\|}$ the observability is affected in the same way in Case 3.

A.5.4 Rotation Around the Surface Normal

Spin stabilised spacecraft show a better robustness in comparison to three-axis stabilised satellites. However, if the rotation axis aligns with the surface normal, i.e. $\omega = a\,n$ for $a \in \mathbb{R}$, observability cannot be ensured.

Proposition. For a spacecraft rotating around the surface normal, i.e. $\omega = a\,n$, the rank of the Jacobian is at most 8.

Proof. We show that the Lie derivatives $\mathcal{L}_f^2 h_1(x)$ and $\mathcal{L}_f^3 h_1(x)$ simplify to the same Lie derivatives as for $\omega = 0$. Then the proof is identical to Proof A.5.1. First, we show that $\frac{\partial F(t,q)}{\partial q} f_q(q, \omega) = 0$. For $F = \frac{\cos(\theta)}{H^2}$ we obtain

$$\frac{\partial A(q)^\top n}{\partial q} f_q(q, \omega) = -2a[A(I_{\mathrm{mp}}q)n\times]A(q)^\top n = 0\,.$$

This directly yields $\frac{\partial F(t,q)}{\partial q} f_q(q, \omega) = 0$. For $F = F_{\mathrm{f},2}$ the expression is validated symbolically. Thus, it follows that $\frac{\partial f_T(T,q,t)}{\partial q} f_q(q, \omega) = 0$. Then the Lie derivative is identical to the second Lie derivative with $\omega = 0$. For the third Lie derivative we use that

$$\begin{aligned}
f_q(q, a\,n)^\top \frac{\partial^2 f_T(T, q, t)}{\partial q^2} f_q(q, a\,n) &= 0 \\
\frac{\partial f_T(T, q, t)}{\partial q} \frac{\partial f_q(q, a\,n)}{\partial q} f_q(q, a\,n) &= 0 \\
\frac{\partial}{\partial q} \frac{\partial f_T(T, q, t)}{\partial t} f_q(q, a\,n) &= 0\,.
\end{aligned}$$

This is shown symbolically for $q = q(n, \varsigma)$ and numerically for other q. Since $f_\omega(a\,n) = 0$, we have

$$\frac{\partial f_T(T,q,t)}{\partial q}\frac{\partial f_q(q,a\,n)}{\partial \omega} f_\omega(\omega) = 0\,. \tag{A.47}$$

Consequently, the Lie derivative $\mathcal{L}_f^3 h_1(x)$ simplifies to the same as for $\omega = 0$ and the resulting Jacobian is identical. This makes the rank identical to eight at most. □

A.5.5 Spacecraft Only Under Solar Irradiation

Proposition. For a spacecraft only under solar irradiation, i.e. Case 3, the Jacobian has rank at most 8.

Proof. The proof follows from calculating matrix $\mathrm{d}\mathcal{O}_q$ symbolically and determining its rank numerically. □

The missing time variance in Case 3 leads to drop of rank. Thus, neglecting albedo and infra-red irradiation in the thermal model also changes the observability properties.

A.5.6 No Irradiation

Proposition. For a spacecraft surface that is not under the influence of any irradiation, i.e. Case 6, the Jacobian has rank 6.

Proof. In this configuration the temperature dynamics $f_T(T, q, \omega)$ are not dependent on the attitude any more. As a simple consequence the first three rows of $\mathrm{d}\mathcal{O}_q$ are identical to zero and the rank is equal to six. □

A.6 Realisation of a Transfer Function

Let $G(s) = \frac{Z(s)}{N(s)} U(s)$ be a transfer function with

$$\begin{aligned} Z(s) &= Z_p s^p + Z_{p-1} s^{p-1} + \cdots + Z_1 s + Z_0 \\ N(s) &= s^n + N_{n-1} s^{n-1} + \cdots + N_1 s + N_0 \end{aligned}$$

for $p, n \in \mathbb{N}$ with $p \leq n$ and $Z_i, N_j \in \mathbb{R}$ for $i \in \{1, \ldots, p\}$ and $j \in \{1, \ldots, n-1\}$. A state space realisation of $G(s)$ has the form

$$\dot{x} = Ax + Bu \tag{A.48a}$$
$$y = Cx + Du \tag{A.48b}$$

such that $G(s) = C(sI - A)^{-1}B + D$ [141]. A realisation in canonical form is given by the system (A.48) and the dynamic and input matrix

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & 0 \\ 0 & \cdots & \cdots & 0 & 1 \\ -N_0 & -N_1 & \cdots & \cdots & -N_{n-1} \end{bmatrix} \qquad B = \begin{bmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

The output and feedthrough matrix depend on the relative degree of the transfer function. For $p < n$ the they have the form

$$C = \begin{bmatrix} Z_0 \, Z_1 \, \dots \, Z_p \end{bmatrix}$$
$$D = 0$$

while for $p = n$ they are defined by

$$C = \begin{bmatrix} Z_0 - Z_n N_0 & Z_1 - Z_n N_1 & \dots & Z_{n-1} - Z_n N_{n-1} \end{bmatrix}$$
$$D = Z_n \, .$$

A.7 Matrix Row and Column Manipulation

For any vector or matrix it is possible to add zeros or remove entries by a simple matrix multiplication. This can be achieved defining vector $k = [k_1 \, \dots, k_m] \in \mathbb{N}^m_{\leq n}$ for $m \in \mathbb{N}_{\leq n}$ with $k_i < k_{i+1}$ and the set $K_i = \{x \in \{k_1, \dots, k_m\} \mid x < i\}$ which contains all indices of k that are smaller than i for $i \leq n$ for $n \in \mathbb{N}$. We denote by $|K_i|$ the cardinality of K_i and define the matrix $V^\top(k) = [v_1(k), \dots, v_n(k)] \in \mathbb{R}^{n-m \times n}$ with the columns

$$v_i(k) = \begin{cases} e_{i-j} & i \notin \{k_1, \dots, k_m\} \text{ with } j = |K_i| \\ 0_{n-m \times 1} & i \in \{k_1, \dots, k_m\} \end{cases}. \tag{A.49}$$

A multiplication with the matrix $V^\top(k)$ to a matrix $A \in \mathbb{R}^{n \times p}$ from the left *removes* the rows $k_1, \dots, k_m$. Multiplying by the transposed matrix $V(k)$ from the left to a matrix $A \in \mathbb{R}^{n-m \times p}$ adds zeros at the $k_1, \dots, k_m$ rows.

To allow a compact notation we define the complement vector of k to be $k^\mathrm{c} = [k^\mathrm{c}_1, k^\mathrm{c}_2, \dots, k^\mathrm{c}_{n-m}]$ with

$$k^\mathrm{c}_i \in \{j \in \{1, \dots, n\} \mid j \notin \{k_1, \dots, k_m\}\} \tag{A.50}$$

for $n \in \mathbb{N}$ and $k^\mathrm{c}_i < k^\mathrm{c}_{i+1}$. This allows to write the matrix V as

$$V(k) = [e_{k^\mathrm{c}_1}, e_{k^\mathrm{c}_2}, \dots, e_{k^\mathrm{c}_{n-m}}] \in \mathbb{R}^{n \times n-m} \, .$$

Thus, for the matrices $A = [a_1, \ldots, a_n]^\top$, $B = [b_1, \ldots, b_{n-m}]^\top$, $C = [c_1, \ldots, c_{n+m}]^\top$ with the rows $a_i^\top, b_i^\top, c_i^\top \in \mathbb{R}^{1\times p}$ and $B = V^\top(k)A \in \mathbb{R}^{n-m\times n}$, $C = V(k)A \in \mathbb{R}^{n+m\times n}$ holds

$$b_i = a_{k_i^c}\,,$$
$$c_i = \begin{cases} a_{i-j} & i \notin \{k_1, \ldots, k_m\} \text{ with } j = |K_i| \\ 0_{n\times 1} & i \in \{k_1, \ldots, k_m\} \end{cases}.$$

A multiplication with the matrix $V(k)$ from the right has the same effect but for the columns and not the rows, i.e. for the matrices $A = [a_1, \ldots, a_n]$, $B = [b_1, \ldots, b_{n-m}]$, $C = [c_1, \ldots, c_{n+m}]$ with the columns $a_i^\top, b_i^\top, c_i^\top \in \mathbb{R}^{1\times p}$ and $B = AV(k)$, $C = AV^\top(k)$ holds

$$b_i = a_{k_i^c}\,,$$
$$c_i = \begin{cases} a_{i-j} & i \notin \{k_1, \ldots, k_m\} \text{ with } j = |K_i| \\ 0_{n\times 1} & i \in \{k_1, \ldots, k_m\} \end{cases}.$$

We define by $V^{\perp\top}(k) \in \mathbb{R}^{m\times n}$ the matrix that if multiplied from the left *keeps* only the rows $k_1, \ldots, k_m$. It has the form

$$V^\perp(k) = [e_{k_1}, e_{k_2}, \ldots, e_{k_m}, \ldots, e_{\bar{m}}] \in \mathbb{R}^{n\times m+\bar{m}-k_m} \tag{A.51}$$

where $\bar{m} = k_m$ or $\bar{m} = n$ such that $V^\perp(k)$ is of the required minimal size for the matrix multiplication. The matrix $V(k^c)$ is either identical to $V^\perp(k)$ or contains additional unit vectors, confer e.g. Example (A.55h) and (A.55i). Note that the definition of $V(k)$ and $V^\perp(k)$ is not unique, if only k is given. However, both definitions are unique if the dimensions of the matrix which it shall be multiplied to is known. The following properties hold for all matrices $V(k)$, $V^\perp(k)$:

$$\begin{aligned} V^{\perp\top}(k)V(k) &= 0_{m\times n-m} & V^\top(k)V^\perp(k) &= 0_{n-m\times m} \\ V^\top(k)V(k) &= I_{n-m} & V^{\perp\top}(k)V^\perp(k) &= I_m \\ V(k)V^\top(k) &= I_n - V^\perp(k)V^{\perp\top}(k) & V^\perp(k)V^{\perp\top}(k) &= I_n - V(k)V^\top(k) \\ V(k)V^\top(k)e_i &= \begin{cases} 0 & \text{if } i \in \{1, \ldots, k_m\} \\ e_i & \text{if } i \notin \{1, \ldots, k_m\} \end{cases} & V^\perp(k)V^{\perp\top}(k)e_i &= \begin{cases} 0 & \text{if } i \notin \{1, \ldots, k_m\} \\ e_i & \text{if } i \in \{1, \ldots, k_m\} \end{cases}. \end{aligned}$$

Furthermore, the matrix V allows characterisations of functions which are independent of certain variables.

Proposition A.1. *The following equivalences for a function $f \in \mathcal{C}^1(\mathbb{R}^n, \mathbb{R}^n)$ hold:*

i) The function f is independent of the component x_i for all $i \in \{k_1, \ldots, k_m\}$, i.e. for all $\tilde{x}_i, x_j \in \mathbb{R}$ with $j \in \{1, \ldots n\}$ holds $f(x_1, \ldots, x_i, \ldots, x_n) = f(x_1, \ldots, \tilde{x}_i, \ldots, x_n)$.

ii) Setting the entries x_i identical to zero does not change the function value of f for all $i \in \{k_1, \ldots, k_m\}$, i.e. for all $x \in \mathbb{R}^n$ holds $f(V(k)V^\top(k)x) = f(x)$.

iii) *The Jacobian of f with respect to x_i is identical to zero for all $i \in \{k_1, \dots, k_m\}$, i.e.* $\frac{\partial f(x)}{\partial x_i} = 0$.

iv) *The i-th column of the Jacobian matrix is identical to zero for all $i \in \{k_1, \dots, k_m\}$, i.e.* $\frac{\partial f(x)}{\partial x} = \frac{\partial f(x)}{\partial x} V(k) V^\top(k)$.

v) *The i-th column of the Jacobian matrix is identical to zero for all $i \in \{k_1, \dots, k_m\}$, i.e.* $\frac{\partial f(x)}{\partial x} V^\perp(k) = 0$.

Proof. *ii) $\Leftrightarrow$ i):* By the definition of $V(k)$ the equality $f(V(k)V^\top(k)x) = f(x)$ has the form $f(\tilde{x}) = f(x)$ where

$$\tilde{x}_i = \begin{cases} 0 & \text{if } i \in \{k_1, \dots k_m\} \\ x_i & \text{if } i \notin \{k_1, \dots k_m\} \end{cases}.$$

This is equivalent to the proposition that for all $i \in \{k_1, \dots k_m\}$ and $x_j \in \mathbb{R}^n$ with $j \in \{1, \dots n\}$ holds $f(x_1, \dots, 0, \dots, x_n) = f(x_1, \dots, \tilde{x}_i, \dots, x_n)$. This is equivalent to *i)*.

i) $\Leftrightarrow$ iii): The equivalence follows from simple derivation and integration.

iii) $\Leftrightarrow$ iv): Can be shown as 'ii) $\Leftrightarrow$ i)' using the fact that $\frac{\partial f(x)}{\partial x_i} = \frac{\partial f(x)}{\partial x} e_i$.

iv) $\Leftrightarrow$ v): Can be seen using $V^\perp(k){V^\perp}^\top(k) = I_n - V(k)V^\top(k)$ and ${V^\perp}^\top(k)V^\perp(k) = I_m$. □

A useful identity employed in Section 6.5.2.3 is given in the following Lemma.

Lemma A.2. *For the vectors $k \in \{13, \dots, 24\}^m$ with $k_i < k_j$ for $i < j$ and $k^{\text{att}} = \{13, \dots, 24\}$ and the matrix $K = \left[0_{12\times 12}\ \text{diag}(k_{13}^{\text{dp}}, \dots, k_{24}^{\text{dp}})\ 0_{12\times 4}\right]^\top \in \mathbb{R}^{28\times 12}$ holds*

$$V^{\perp\top}(k)K = V^{\perp\top}(k)K{V^{\perp}}^\top(k^{\text{att}})V^\perp(k){V^{\perp}}^\top(k)V^\perp(k^{\text{att}})\,. \tag{A.52}$$

Proof. It is

$$V^{\perp\top}(k)K{V^{\perp}}^\top(k^{\text{att}})V^\perp(k){V^{\perp}}^\top(k)V^\perp(k^{\text{att}}) = \text{diag}(k_{k_1}^{\text{dp}}, \dots, k_{k_m}^{\text{dp}}){V^{\perp}}^\top(k)V^\perp(k^{\text{att}})$$

By the definition of $V^\perp$ it is

$$\text{diag}(k_{k_1}^{\text{dp}}, \dots, k_{k_m}^{\text{dp}}){V^{\perp}}^\top(k)V^\perp(k^{\text{att}})e_i = \begin{cases} k_i^{\text{dp}} e_j & \text{if } i \in k \text{ with } j = |K_i| + 1 \\ 0 & \text{if } i \notin k \end{cases}.$$

Consequently, it is

$$e_l^\top \operatorname{diag}(k^{\mathrm{dp}}_{k_1}, \ldots, k^{\mathrm{dp}}_{k_m}) {V^\perp}^\top(k) V^\perp(k^{\mathrm{att}}) e_i = \begin{cases} k^{\mathrm{dp}}_{k_l} & \text{if } i \in k \text{ and } l = j \text{ with } j = |K_i| + 1 \\ 0 & \text{otherwise} \end{cases} \tag{A.53a}$$

$$= \begin{cases} k^{\mathrm{dp}}_{k_l} & \text{if } i = k_l \\ 0 & \text{otherwise} \end{cases} . \tag{A.53b}$$

Further, it is $e_l^\top {V^\perp}^\top(k) K = k^{\mathrm{dp}}_{k_l} e_{k_l}^\top$ and thus

$$e_l^\top {V^\perp}^\top(k) K e_i = \begin{cases} k^{\mathrm{dp}}_{k_l} & \text{if } i = k_l \\ 0 & \text{otherwise} \end{cases} \tag{A.54}$$

The equality of (A.53) and (A.54) shows the claim. □

Some example calculations to observe the properties and form of the matrices $V(k)$ and $V^\perp(k)$ are shown in the following.

Example 1. For $k = [1, 3]$ holds

$$V(k)^\top A = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 5 & 6 & 7 & 8 \end{pmatrix} \quad \text{with } n = 3 \tag{A.55a}$$

$$V(k)A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \quad \text{with } n = 5 \tag{A.55b}$$

$$AV(k)^\top = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 & 2 & 3 & 4 \\ 0 & 5 & 0 & 6 & 7 & 8 \\ 0 & 9 & 0 & 10 & 11 & 12 \end{pmatrix} \quad \text{with } n = 6 \tag{A.55c}$$

$$AV(k) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 4 \\ 6 & 8 \\ 10 & 12 \end{pmatrix} \quad \text{with } n = 4 \tag{A.55d}$$

$$V(k)^{\perp\top} A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 9 & 10 & 11 & 12 \end{pmatrix} \quad \text{with } n = 3 \tag{A.55e}$$

$$V(k)^\perp A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 0 & 0 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \quad \text{with } n = 4 \tag{A.55f}$$

$$AV(k)^{\perp\top} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 2 & 3 & 4 \\ 5 & 0 & 6 & 7 & 8 \\ 9 & 0 & 10 & 11 & 12 \end{pmatrix} \quad \text{with } n = 5 \tag{A.55g}$$

$$AV(k)^\perp = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 3 \\ 5 & 7 \\ 9 & 11 \end{pmatrix} \quad \text{with } n = 4 \tag{A.55h}$$

$$AV(k^\mathrm{c}) = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 9 & 11 & 12 \end{pmatrix} \quad \text{with } n = 4 \tag{A.55i}$$

A.8 Homogeneity

Homogeneity is a concept widely used in the design of sliding mode controllers and observers. In this section, the basic definitions are introduced and fundamental properties of

homogeneous systems are stated following [82].

Definition A.3. *For every weight m_i for a coordinate x_i we define the* homogeneity degree *to be*

$$\deg(x_i) = m_i$$

and the corresponding homogeneity dilatation d_κ *with parameter κ to be*

$$\mathrm{d}_\kappa : \mathbb{R}^n \to \mathbb{R}^n \tag{A.56a}$$

$$(x_1, \ldots, x_n) \mapsto (\kappa^{m_1} x_1, \ldots, \kappa^{m_n} x_n)\,. \tag{A.56b}$$

Consequently, we define the homogeneous degree of a function.

Definition A.4. *A function $f : \mathbb{R}^n \to \mathbb{R}$ is called* homogeneous of degree q *with dilatation* (A.56) *if the identity*

$$f(\mathrm{d}_\kappa x) = \kappa^q f(x)$$

for any $\kappa > 0$. We write

$$\deg(f) = q\,.$$

Some properties of the homogeneous degree can be found in [82]. The homogeneous degree of a vector field is defined by

Definition A.5. *A vector field $f : \mathbb{R}^n \to \mathbb{R}^n$ is called* homogeneous of degree q *with dilatation* (A.56) *if the identity*

$$f(x) = \kappa^{-q} \mathrm{d}_\kappa^{-1} f(\mathrm{d}_\kappa x)$$

for any $\kappa > 0$ holds. We write

$$\deg(f) = q\,.$$

Note the ambiguity occurring for scalar vector fields. Any linear time invariant differential equation is homogeneous with homogeneity degree 0.

Example 2. Consider the system

$$\dot{x}_1 = x_2 \tag{A.57a}$$

$$\dot{x}_2 = -x_1^{\frac{1}{3}} - |x_2|^{\frac{1}{2}} \operatorname{sgn}(x_2) \tag{A.57b}$$

We denote the left hand side of example (A.57) by $f(x)$ and use the weights 3 and 2 for x_1 and x_2 to obtain for any $\kappa > 0$ that

$$\begin{aligned}\kappa^{-q}\mathrm{d}_\kappa^{-1} f(\mathrm{d}_\kappa x) &= \kappa \mathrm{d}_\kappa^{-1} \begin{pmatrix} \kappa^2 x_2 \\ -\kappa x_1^{\frac{1}{3}} - |\kappa||x_2|^{\frac{1}{2}} \operatorname{sgn}(\kappa^2 x_2) \end{pmatrix} \\ &= \kappa \begin{pmatrix} \kappa^{-1} x_2 \\ -\kappa^{-1} x_1^{\frac{1}{3}} - \kappa^{-1}|x_2|^{\frac{1}{2}} \operatorname{sgn}(x_2) \end{pmatrix} \\ &= \begin{pmatrix} x_2 \\ -x_1^{\frac{1}{3}} - |x_2|^{\frac{1}{2}} \operatorname{sgn}(x_2) \end{pmatrix} = f(x)\,.\end{aligned}$$

Thus, the system (A.57) is homogeneous of the degree -1.

Homogeneous systems have useful properties. A homogeneous system which is locally stable is automatically globally stable [142] [82]. Further, if the homogeneous degree of a system is negative, asymptotic stability implies finite-time stability [142] [82].

Bibliography

[1] IEEE Spectrum, "Space station incident demands independent investigation." `https://spectrum.ieee.org/space-station-accident-needs-independant-investigation`, 2021. Accessed: 2021-10-19.

[2] Gizmodo, "The ISS backflipped out of control after russian module misfired, new details reveal." `https://gizmodo.com/the-iss-backflipped-out-of-control-after-russian-module-1847415359`, 2021. Accessed: 2021-10-19.

[3] C. Schwarz, J. Brembeck, and B. Heckmann, "Dynamics observer for the longitudinal behavior of a wheelset on a roller rig," *Proceedings of the Institution of Mechanical Engineers*, vol. 233, no. 10, pp. 1112–1119, 2019.

[4] J. Brembeck, "Nonlinear constrained moving horizon estimation applied to vehicle position estimation," *Sensors*, vol. 19, no. 10, p. 2276, 2019.

[5] J. Brembeck, A. Pfeiffer, M. Fleps-Dezasse, M. Otter, K. Wernersson, and H. Elmqvist, "Nonlinear state estimation with an extended FMI 2.0 co-simulation interface," in *International Modelica Conference*, vol. 96, pp. 53–62, March 2014.

[6] C. Schwarz and A. Keck, "Simultaneous estimation of wheel-rail adhesion and brake friction behaviour," *IFAC-PapersOnLine*, vol. 53, pp. 8470–8475, 2020.

[7] A. Keck, C. Schwarz, T. Meurer, A. Heckmann, and G. Grether, "Estimating the wheel lateral position of a mechatronic railway running gear with nonlinear wheel–rail geometry," *Mechatronics*, vol. 73, p. 102457, 2021.

[8] B. Xu, "Disturbance observer-based dynamic surface control of transport aircraft with continuous heavy cargo airdrop," *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, vol. 47, no. 1, pp. 161–170, 2016.

[9] D. Wang and K.-Y. Lum, "Adaptive unknown input observer approach for aircraft actuator fault detection and isolation," *International Journal of Adaptive Control and Signal Processing*, vol. 21, no. 1, pp. 31–48, 2007.

[10] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of Basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[11] D. Luenberger, "An introduction to observers," *IEEE Transactions on Automatic Control*, vol. 16, no. 6, pp. 596–602, 1971.

[12] J.-j. E. Slotine, J. K. Hedrick, and E. A. Misawa, "Nonlinear state estimation using sliding observers," in *IEEE Conference on Decision and Control*, pp. 332–339, IEEE, 1986.

[13] J. Gauthier and G. Bornard, "Observability for any u (t) of a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 26, no. 4, pp. 922–926, 1981.

[14] G. Bornard and H. Hammouri, "A high gain observer for a class of uniformly observable systems," in *IEEE Conference on Decision and Control*, pp. 1494–1496, IEEE, 1991.

[15] K. Busawon, M. Farza, and H. Hammouri, "A simple observer for a class of nonlinear systems," *Applied Mathematics Letters*, vol. 11, no. 3, pp. 27–31, 1998.

[16] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Systems & Control Letters*, vol. 3, no. 1, pp. 47–52, 1983.

[17] A. J. Krener and W. Respondek, "Nonlinear observers with linearizable error dynamics," *SIAM Journal on Control and Optimization*, vol. 23, no. 2, pp. 197–216, 1985.

[18] J.-P. Gauthier, H. Hammouri, and S. Othman, "A simple observer for nonlinear systems applications to bioreactors," *IEEE Transactions on Automatic Control*, vol. 37, no. 6, pp. 875–880, 1992.

[19] P. Bernard, *Observer design for nonlinear systems*. Springer, 2019.

[20] J.-P. Gauthier and I. Kupka, *Deterministic observation theory and applications*. Cambridge University Press, 2001.

[21] G. Besançon, *Nonlinear observers and applications*, vol. 363. Springer, 2007.

[22] D. Luenberger, "Observers for multivariable systems," *IEEE Transactions on Automatic Control*, vol. 11, no. 2, pp. 190–197, 1966.

[23] R. E. Kálmán and R. S. Bucy, "New results in linear filtering and prediction theory," *Journal of Basic Engineering*, vol. 83, pp. 95–108, 1961.

[24] G. L. Smith, S. F. Schmidt, and L. A. McGee, "Application of statistical filter theory to the optimal estimation of position and velocity on board a circumlunar vehicle," Tech. Rep. TR R-135, NASA, 1962.

[25] D. Simon, "Kalman filtering with state constraints: a survey of linear and nonlinear algorithms," *IET Control Theory & Applications*, vol. 4, no. 8, pp. 1303–1318, 2010.

[26] S. J. Julier and J. K. Uhlmann, "New extension of the kalman filter to nonlinear systems," in *Signal processing, sensor fusion, and target recognition VI*, vol. 3068, pp. 182–193, International Society for Optics and Photonics, 1997.

[27] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[28] H. Michalska and D. Q. Mayne, "Moving horizon observers and observer-based control," *IEEE Transactions on Automatic Control*, vol. 40, no. 6, pp. 995–1006, 1995.

[29] C. V. Rao, J. B. Rawlings, and J. H. Lee, "Constrained linear state estimation—a moving horizon approach," *Automatica*, vol. 37, no. 10, pp. 1619–1628, 2001.

[30] N. J. Gordon, D. J. Salmond, and A. F. Smith, "Novel approach to nonlinear/non-gaussian bayesian state estimation," in *IEE Proceedings F-radar and signal processing*, vol. 140, pp. 107–113, IET, 1993.

[31] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 174–188, 2002.

[32] D. Simon, *Optimal state estimation: Kalman, H infinity, and nonlinear approaches*. John Wiley & Sons, 2006.

[33] J.-P. Barbot, T. Boukhobza, and M. Djemai, "Sliding mode observer for triangular input form," in *Conference on Decision and Control*, vol. 2, pp. 1489–1490, IEEE, 1996.

[34] J. Davila, L. Fridman, and A. Levant, "Second-order sliding-mode observer for mechanical systems," *IEEE Transactions on Automatic Control*, vol. 50, no. 11, pp. 1785–1789, 2005.

[35] M. Arcak and P. Kokotović, "Nonlinear observers: a circle criterion design and robustness analysis," *Automatica*, vol. 37, no. 12, pp. 1923–1930, 2001.

[36] J. A. Moreno, "Observer design for nonlinear systems: A dissipative approach," *IFAC Proceedings Volumes*, vol. 37, no. 21, pp. 681–686, 2004.

[37] J.-P. Barbot, M. Fliess, and T. Floquet, "An algebraic framework for the design of nonlinear observers with unknown inputs," in *Conference on Decision and Control*, pp. 384–389, IEEE, 2007.

[38] M. Fliess, C. Join, and H. Sira-Ramírez, "Non-linear estimation is easy," *International Journal of Modelling, Identification and Control*, vol. 4, no. 1, pp. 12–27, 2008.

[39] D. Astolfi, R. Postoyan, and D. Nešić, "Uniting observers," *IEEE Transactions on Automatic Control*, vol. 65, no. 7, pp. 2867–2882, 2019.

[40] M. J. Sidi, *Spacecraft dynamics and control: a practical engineering approach*, vol. 7. Cambridge University Press, 1997.

[41] F. L. Markley and J. L. Crassidis, *Fundamentals of Spacecraft Attitude Determination and Control*. Springer, 2014.

[42] M. Hobsch, F. Cossavella, S. Löw, and J. Herman, "Attitude control on TET-1-experiences of the first year of operations," in *International Conference on Space Operations*, 2014.

[43] J. Herman, D. Presti, A. Codazzi, and C. Belle, "Attitude control for GRACE: The first low-flying satellite formation," in *International Symposium on Space Flight Dynamics*, vol. 548, pp. 27–33, 2004.

[44] E. Canuto, C. Novara, D. Carlucci, C. P. Montenegro, and L. Massotti, *Spacecraft Dynamics and Control: The Embedded Model Control Approach*. Butterworth-Heinemann, 2018.

[45] J. L. Crassidis, F. L. Markley, and Y. Cheng, "Survey of nonlinear attitude estimation methods," *Journal of Guidance, Control, and Dynamics*, vol. 30, no. 1, pp. 12–28, 2007.

[46] H. Ma, Z. Lu, X. Zhang, W. Liao, and K. Briess, "High-accuracy and low-cost attitude measurement unit of the cubesat," *International Journal of Aerospace Engineering*, vol. 2020, pp. 1–13, 2020.

[47] X. Mao, X. Du, and H. Fang, "Precise attitude determination strategy for spacecraft based on information fusion of attitude sensors: Gyros/GPS/star-sensor," *International Journal of Aeronautical and Space Sciences*, vol. 14, no. 1, pp. 91–98, 2013.

[48] J. Ding, J. Wu, M. Deng, and M. Liu, "Reference-free adaptive attitude determination method using low-cost MARG sensors," in *International Conference on Computer Vision Systems*, pp. 35–48, Springer, 2019.

[49] A. Tayebi, S. McGilvray, A. Roberts, and M. Moallem, "Attitude estimation and stabilization of a rigid body using low-cost sensors," in *Conference on Decision and Control*, pp. 6424–6429, 2007.

[50] H. B. Unhelkar, Vaibhav V.and Hablani, "Spacecraft attitude determination with sun sensors, horizon sensors and gyros: Comparison of steady-state kalman filter and extended kalman filter," in *Advances in Estimation, Navigation, and Spacecraft Control*, pp. 413–437, Springer Berlin Heidelberg, 2015.

[51] S. Theil, P. Appel, and A. Schleicher, "Low cost, good accuracy-attitude determination using magnetometer and simple sun sensor," *AIAA/USU Small Satellite Conference*, 2003.

[52] F. N. Gourabi, M. Kiani, and S. H. Pourtakdoust, "Autonomous temperature-based orbit estimation," *Aerospace Science and Technology*, vol. 86, pp. 671–682, 2019.

[53] J. Levenhagen, N. Duske, and R. Wolters, "Earth oriented safe mode design based on the EADS Astrium CESS," *IFAC Proceedings Volumes*, vol. 40, no. 7, pp. 301–304, 2007.

[54] H. J. Kramer, *Observation of the Earth and its Environment: Survey of Missions and Sensors*. Springer Science & Business Media, 2002.

[55] H. B. Khaniki and S. M. H. Karimian, "Determining the heat flux absorbed by satellite surfaces with temperature data," *Journal of Mechanical Science and Technology*, vol. 28, no. 6, pp. 2393–2398, 2014.

[56] H. B. Khaniki and S. M. H. Karimian, "Satellite attitude determination using absorbed heat fluxes," *Journal of Aerospace Engineering*, vol. 29, no. 6, p. 04016053, 2016.

[57] A. Labibian, S. H. Pourtakdoust, M. Kiani, A. A. Sheikhi, and A. Alikhani, "Experimental validation of a novel radiation based model for spacecraft attitude estimation," *Sensors and Actuators A: Physical*, vol. 250, pp. 114–122, 2016.

[58] A. Labibian, A. Alikhani, and S. H. Pourtakdoust, "Performance of a novel heat based model for spacecraft attitude estimation," *Aerospace Science and Technology*, vol. 70, pp. 317–327, 2017.

[59] A. Labibian, S. Pourtakdoust, A. Alikhani, and H. Fourati, "Development of a radiation based heat model for satellite attitude determination," *Aerospace Science and Technology*, vol. 82, pp. 479–486, 2018.

[60] T. Posielek, "A modelica library for spacecraft thermal analysis," in *The American Modelica Conference*, vol. 154, p. 46–55, Linköping University Electronic Press, 2018.

[61] T. Posielek and J. Reger, "A novel attitude representation in view of spacecraft attitude reconstruction using temperature data," in *Conference on Modelling, Identification and Control of Nonlinear Systems*, pp. 526–531, 2021.

[62] T. Posielek and J. Reger, "Observability analysis for spacecraft attitude determination using a single temperature sensor," in *European Control Conference*, pp. 1432–1439, 2021.

[63] T. Posielek and J. Reger, "Attitude reconstruction of a spacecraft from temperature measurements - Analysis and observer design for a not globally observable non-linear system," *IEEE Transactions on Control Systems Technology*, 2022. submitted.

[64] D. Gerbet and K. Röbenack, "On global and local observability of nonlinear polynomial systems: A decidable criterion," *at-Automatisierungstechnik*, vol. 68, no. 6, pp. 395–409, 2020.

[65] R. Hermann and A. Krener, "Nonlinear controllability and observability," *IEEE Transactions on Automatic Control*, vol. 22, no. 5, pp. 728–740, 1977.

[66] H. L. Trentelman, A. A. Stoorvogel, and M. Hautus, *Control theory for linear systems*. Springer Science & Business Media, 2012.

[67] H. K. Khalil, *Nonlinear Systems*. Prentice Hall, third ed., 2002.

[68] M. Fliess, C. Join, and H. Sira-Ramírez, "Robust residual generation for linear fault diagnosis: an algebraic setting with examples," *International Journal of Control*, vol. 77, no. 14, pp. 1223–1242, 2004.

[69] J. Reger and J. Jouffroy, "On algebraic time-derivative estimation and deadbeat state reconstruction," in *IEEE Conference on Decision and Control*, pp. 1740–1745, IEEE, 2009.

[70] D.-Y. Liu, O. Gibaru, and W. Perruquetti, "Synthesis on a class of algebraic differentiators and application to nonlinear observation," in *Chinese Control Conference*, pp. 2592–2599, IEEE, 2014.

[71] Q. Guo, W. Perruquetti, and M. Gautier, "On-line robot dynamic identification based on power model, modulating functions and causal Jacobi estimator," in *IEEE/ASME International Conference on Advanced Intelligent Mechatronics*, pp. 494–499, IEEE, 2014.

[72] M. Mboup, C. Join, and M. Fliess, "A revised look at numerical differentiation with an application to nonlinear feedback control," in *Mediterranean Conference on Control & Automation*, pp. 1–6, IEEE, 2007.

[73] Z. Gao, X. Dai, T. Breikin, and H. Wang, "Novel parameter identification by using a high-gain observer with application to a gas turbine engine," *IEEE Transactions on Industrial Informatics*, vol. 4, no. 4, pp. 271–279, 2008.

[74] Z. Gao, T. Breikin, and H. Wang, "High-gain estimator and fault-tolerant design with application to a gas turbine dynamic system," *IEEE Transactions on Control Systems Technology*, vol. 15, no. 4, pp. 740–753, 2007.

[75] R. E. B. Serhal and H. K. Khalil, "Application of the extended high gain observer to underactuated mechanical systems," in *American Control Conference*, pp. 4727–4732, IEEE, 2012.

[76] F. Lafont, E. Busvelle, and J.-P. Gauthier, "An adaptive high-gain observer for wastewater treatment systems," *Journal of Process Control*, vol. 21, no. 6, pp. 893–900, 2011.

[77] S. Banerjee and A. K. Jana, "High gain observer based extended generic model control with application to a reactive distillation column," *Journal of Process Control*, vol. 24, no. 4, pp. 235–248, 2014.

[78] A. Atassi and H. Khalil, "Separation results for the stabilization of nonlinear systems using different high-gain observer designs," *Systems & Control Letters*, vol. 39, no. 3, pp. 183–191, 2000.

[79] A. N. Atassi and H. K. Khalil, "A separation principle for the stabilization of a class of nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 44, no. 9, pp. 1672–1687, 1999.

[80] L. K. Vasiljevic and H. K. Khalil, "Differentiation with high-gain observers the presence of measurement noise," in *IEEE Conference on Decision and Control*, pp. 4717–4722, IEEE, 2006.

[81] L. K. Vasiljevic and H. K. Khalil, "Error bounds in differentiation of noisy signals by high-gain observers," *Systems & Control Letters*, vol. 57, no. 10, pp. 856–862, 2008.

[82] Y. Shtessel, C. Edwards, L. Fridman, and A. Levant, *Sliding mode control and observation*. Springer, 2014.

[83] A. Filippov, *Differential Equations with Discontinuous Right-hand Sides*. Springer, Dordrecht, 1988.

[84] L. Fridman, A. Levant, *et al.*, "Higher order sliding modes," *Sliding mode control in engineering*, vol. 11, pp. 53–102, 2002.

[85] A. Levant, "Higher-order sliding modes, differentiation and output-feedback control," *International Journal of Control*, vol. 76, no. 9-10, pp. 924–941, 2003.

[86] A. Levant, "Robust exact differentiation via sliding mode technique," *Automatica*, vol. 34, no. 3, pp. 379–384, 1998.

[87] A. Levant, "Quasi-continuous high-order sliding-mode controllers," in *IEEE International Conference on Decision and Control*, vol. 5, pp. 4605–4610, IEEE, 2003.

[88] A. Levant, M. Livne, and X. Yu, "Sliding-mode-based differentiation and its application," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1699–1704, 2017.

[89] A. Levant, "Filtering differentiators and observers," in *International Workshop on Variable Structure Systems*, pp. 174–179, IEEE, 2018.

[90] A. Levant and X. Yu, "Sliding-mode-based differentiation and filtering," *IEEE Transactions on Automatic Control*, vol. 63, no. 9, pp. 3061–3067, 2018.

[91] M. Reichhartinger, S. Spurgeon, M. Forstinger, and M. Wipfler, "A robust exact differentiator toolbox for Matlab®/Simulink®," *IFAC-PapersOnLine*, vol. 50, no. 1, pp. 1711–1716, 2017.

[92] J. A. Moreno, "Lyapunov-based design of homogeneous high-order sliding modes," in *Advances in Variable Structure Systems and Sliding Mode Control—Theory and Applications*, pp. 3–38, Springer, 2018.

[93] H. Amann and J. Escher, *Analysis II*, vol. 3. Birkhäuser, 2008.

[94] N. Capitaine, B. Guinot, and D. McCarthy, "Definition of the celestial ephemeris origin and of UT1 in the international celestial reference frame," *Astronomy and Astrophysics*, vol. 355, pp. 398–405, 2000.

[95] O. Montenbruck and E. Gill, *Satellite Orbits: Models, Methods and Applications*. Springer, 2011.

[96] H. Schaub and J. L. Junkins, *Analytical Mechanics of Space Systems*. American Institute of Aeronautics and Astronautics, 2014.

[97] J. R. Wertz, *Introduction to attitude dynamics and control*. Springer, 1978.

[98] B. Wie, *Space vehicle dynamics and control*. American Institute of Aeronautics and Astronautics, 2008.

[99] M. D. Shuster *et al.*, "A survey of attitude representations," *Navigation*, vol. 8, no. 9, pp. 439–517, 1993.

[100] N. A. Chaturvedi, A. K. Sanyal, and N. H. McClamroch, "Rigid-body attitude control," *IEEE Control Systems Magazine*, vol. 31, no. 3, pp. 30–51, 2011.

[101] J. Stuelpnagel, "On the parametrization of the three-dimensional rotation group," *SIAM Review*, vol. 6, no. 4, pp. 422–430, 1964.

[102] R. E. Mortensen, "A globally stable linear attitude regulator," *International Journal of Control*, vol. 8, no. 3, pp. 297–302, 1968.

[103] B. Wie and P. M. Barba, "Quaternion feedback for spacecraft large angle maneuvers," *Journal of Guidance, Control, and Dynamics*, vol. 8, no. 3, pp. 360–365, 1985.

[104] J.-Y. Wen and K. Kreutz-Delgado, "The attitude control problem," *IEEE Transactions on Automatic Control*, vol. 36, no. 10, pp. 1148–1162, 1991.

[105] A. Tayebi, "Unit quaternion observer based attitude stabilization of a rigid spacecraft without velocity measurement," in *Conference on Decision and Control*, pp. 1557–1561, 2006.

[106] A. Tayebi, "Unit quaternion-based output feedback for the attitude tracking problem," *IEEE Transactions on Automatic Control*, vol. 53, no. 6, pp. 1516–1520, 2008.

[107] G. Wahba, "A least squares estimate of satellite attitude," *SIAM review*, vol. 7, no. 3, pp. 409–409, 1965.

[108] R. Karam, *Satellite Thermal Control for Systems Engineers*. AIAA, 1998.

[109] F. P. Incropera, D. P. DeWitt, T. L. Bergman, and A. S. Lavine, *Fundamentals of Heat and Mass Transfer*. John Wiley & Sons, 2006.

[110] J. Meseguer, I. Pérez-Grande, and A. Sanz-Andrés, *Spacecraft Thermal Control*. Woodhead Publishing, 2012.

[111] N. H. Juul, "Diffuse radiation view factors from differential plane sources to spheres," *Journal of Heat Transfer*, vol. 101, no. 3, pp. 558–560, 1979.

[112] W. J. Larson and J. R. Wertz, *Space Mission Analysis and Design*. Springer, 1991.

[113] Texas Instruments, *TMP117 High-Accuracy, Low-Power, Digital Temperature Sensor*, 2021.

[114] Maxim Integrated, *MAX30208, Low-Power, High-Accuracy Digital Temperature Sensor*, 2020.

[115] AMS, *AS6221, High accurate digital temperature sensor*, 2020.

[116] P. Acquatella, "Fast slew maneuvers for the high-torque-wheels biros spacecraft," in *International Symposium on Space Flight Dynamics*, June 2017.

[117] B. Wie, D. Bailey, and C. Heiberg, "Rapid multitarget acquisition and pointing control of agile spacecraft," *Journal of Guidance, Control, and Dynamics*, vol. 25, no. 1, pp. 96–104, 2002.

[118] G. Teschl, *Ordinary differential equations and dynamical systems*, vol. 140. American Mathematical Soc., 2012.

[119] H. Amann, J. Escher, S. Levy, and M. Cargo, *Analysis*, vol. 3. Birkhäuser, 2005.

[120] B. Addis, M. Locatelli, and F. Schoen, "Local optima smoothing for global optimization," *Optimization Methods and Software*, vol. 20, no. 4-5, pp. 417–437, 2005.

[121] P. Martin, P. Rouchon, and J. Rudolph, "Invariant tracking," *ESAIM: Control, Optimisation and Calculus of Variations*, vol. 10, no. 1, pp. 1–13, 2004.

[122] S. Bonnabel and P. Rouchon, "On invariant observers," in *Control and observer design for nonlinear finite and infinite dimensional systems*, pp. 53–65, Springer, 2005.

[123] P. Martin and E. Salaün, "An invariant observer for earth-velocity-aided attitude heading reference systems," *IFAC Proceedings Volumes*, vol. 41, no. 2, pp. 9857–9864, 2008.

[124] M. Barczyk, "Invariant observer design of attitude and heading reference system," *Control Theory and Technology*, vol. 17, no. 3, pp. 228–240, 2019.

[125] P. Bernard, V. Andrieu, and L. Praly, "Expressing an observer in preferred coordinates by transforming an injective immersion into a surjective diffeomorphism," *SIAM Journal on Control and Optimization*, vol. 56, no. 3, pp. 2327–2352, 2018.

[126] D. D. Mazanek, *Aerothermal analysis and design of the gravity recovery and climate experiment (GRACE) spacecraft*. NASA, 2000.

[127] M. Kirschner, O. Montenbruck, and S. Bettadpur, "Flight dynamics aspects of the GRACE formation flying," in *International Workshop on Satellite Constellations and Formation Flying*, pp. 1–8, 2001.

[128] K. M. Larson, N. Ashby, C. Hackman, and W. Bertiger, "An assessment of relativistic effects for low earth orbiters: the GRACE satellites," *Metrologia*, vol. 44, no. 6, p. 484, 2007.

[129] W. Bertiger, Y. Bar-Sever, S. Desai, C. Dunn, B. Haines, G. Kruizinga, D. Kuang, S. Nandi, L. Romans, M. Watkins, *et al.*, "GRACE: millimeters and microns in orbit," in *International Technical Meeting of the Satellite Division of The Institute of Navigation*, pp. 2022–2029, 2002.

[130] R. Peyrou-Lauga, "Using real earth albedo and earth ir flux for spacecraft thermal analysis," in *International Conference on Environmental Systems*, p. 142, 2017.

[131] H.-D. Joos, "A multiobjective optimisation-based software environment for control systems design," in *IEEE International Symposium on Computer Aided Control System Design*, pp. 7–14, 2002.

[132] M. Donabedian and D. G. Gilmore, *Spacecraft Thermal Control Handbook*. Aerospace Press, 2003.

[133] J. Herman, D. Presti, A. Codazzi, and C. Belle, "Attitude control for GRACE: The first low-flying satellite formation," in *International Symposium on Space Flight Dynamics*, vol. 548, pp. 27–33, 2004.

[134] E. Thébault, C. C. Finlay, C. D. Beggan, P. Alken, J. Aubert, O. Barrois, F. Bertrand, T. Bondar, A. Boness, L. Brocco, *et al.*, "International geomagnetic reference field: the 12th generation," *Earth, Planets and Space*, vol. 67, no. 1, pp. 1–19, 2015.

[135] L. E. Briese, A. Klöckner, and M. Reiner, "The DLR environment library for multidisciplinary aerospace applications," in *International Modelica Conference*, pp. 929–938, 2017.

[136] S. Maus, S. Macmillan, S. McLean, B. Hamilton, A. Thomson, M. Nair, and C. Rollins, "The US/UK world magnetic model for 2010-2015," 2010. NOAA Technical Report NESDIS/NGDC.

[137] N. Olsen, "Magnetometer data from the grace satellite duo," *Earth, Planets and Space*, vol. 73, no. 1, pp. 1–20, 2021.

[138] B. Wie, *Space vehicle dynamics and control*. AIAA, 1998.

[139] I. Bronshtein, K. Semendyayev, G. Musiol, and H. Muehlig, *Handbook of Mathematics*. Springer, 2007.

[140] S. W. Shepperd, "Quaternion from rotation matrix," *Journal of Guidance and Control*, vol. 1, no. 3, pp. 223–224, 1978.

[141] B. Friedland, *Control System Design: An Introduction to State Space Methods*. McGraw-Hill, 1986.

[142] A. Bacciotti and L. Rosier, *Liapunov functions and stability in control theory*. Springer Science & Business Media, 2006.