Attitude and rate estimation without rate sensors: A Lie group approach

Nikolas Koumpis

Master of Science Thesis

TU Delft Delft University of Technology

Delft Center for Systems and Control

Attitude and rate estimation without rate sensors: A Lie group approach

Master of Science Thesis

For the degree of Master of Science in Systems and Control at Delft University of Technology

Nikolas Koumpis

April 9, 2020

Faculty of Mechanical, Maritime and Materials Engineering (3mE) · Delft University of Technology

This MSc thesis is performed at Airbus Defense and Space GmbH in Friedrichshafen (Germany) in the department AOCS/GNC and Flight Dynamics.

Delft University of Technology
Department of
Delft Center for Systems and Control (dcsc)

The undersigned hereby certify that they have read and recommend to the Faculty of Mechanical, Maritime and Materials Engineering (3mE) for acceptance a thesis entitled

Attitude and rate estimation without rate sensors: A Lie group approach

by

Nikolas Koumpis

in partial fulfillment of the requirements for the degree of

Master of Science Systems and Control

Dated: April 9, 2020

Supervisor(s): _____

Reader(s): _____

Abstract

Attitude estimation is one of the basic processes of the attitude and orbit control system (AOCS) hosted by every satellite. The spacecraft should be orientated properly during the orbit, according to the mission objectives. The estimation is based on a plethora of sensors, like rate gyroscopes, star trackers, sun and earth sensors, magnetometers (for low earth orbits). However, during the orbit the spacecraft operates according to some protocol, which is characterized from its different modes. During safe mode conditions, gyroscopes and/or star trackers might be turned off for energy saving reasons. Then the satellite senses only the geomagnetic field's direction and directions of other celestial bodies. In addition, the motion of a rigid body is described from well-known dynamic equations up to model uncertainties.

The goal is to estimate how the satellite faces the Earth at each instant of time, by utilizing the two aforementioned sources of information. Limited research has been undertaken for this problem; on the contrary, the vast majority of published works provide solutions to the problem at the presence of rate sensors. Within a general non-linear framework, the problem of orientation and rate estimation can be set under the Bayesian formulation of estimation. The available tools provided by this approach, are the widely-known Kalman-Based filters. An alternative compelling approach considers the dual optimal control problem thus being more compatible with the geometric nature of the problem. This approach is also more promising, since it overcomes some critical difficulties of the former one and it is applicable in a coordinate free fashion.

In this thesis, the aforementioned observation is emphasized and looked into. Therefore, two specific filters are studied: the second-order-optimal minimum energy filter on the special orthogonal group $\mathbb{TSO}(3)$; and the predictive filter on $\mathbb{TSO}(3)$, which is derived on this work. The two filters perform very well with the predictive filter to be less tolerant in measurement noise. Both filters are analyzed and simulated in various scenarios and a comparison is demonstrated, given particular emphasis to their structure. Finally, the predictive filter on $\mathbb{TSO}(3)$ is extended and used in an attempt to provide an alternative approach for the problem of the orbital position estimation by utilizing a kinematic model for the Earth's geomagnetic field. Towards this direction it is proved that under some realistic assumptions this latter problem is equivalent to two separate problems and therefore the resulting filter can be decomposed into two distinct filters interconnected in series. The optimal correction

term which is determined through the estimation process provides information regarding the orbital rotation rate.

Table of Contents

Chapter 1

Introduction

1-1 Overview and motivation

Attitude estimation is an important aspect of guidance, navigation and control of vehicles, such as cars, drones and satellites [1]. In satellites -where accurate attitude estimation is of paramount importance- it is performed via data fusion of several attitude sensors like star trackers, coarse earth sun sensors and rate gyros. In order to achieve the desired accuracy, the relative orientation of all sensors with respect to each other and all of them with respect to the spacecraft's mounting base should be known precisely; an apt calibration is therefore necessary.

Furthermore, these systems are required to be cheap and have a robust operation. The sensors that are employed need to be cheap with minimum hardware implementation. During failure, the sensor should enter a safe mode operation, providing the necessary information in order to maintain the desired orientation dictated by the mission objective. In conjunction to this, the energy consumption during this mode should be minimized until the failure is restored. For the above reasons, estimation methods that require only reference sensors under safe mode conditions are utilized.

By setting the problem of attitude and rate estimation in a general nonlinear framework the goal is to estimate the state of the system which evolves according to a nonlinear differential equation from observations that are acquired from the state in a nonlinear fashion. For nonlinear state space models under the Bayesian formulation of estimation, the mostly known filtering techniques are the extended Kalman filter, the unscented Kalman filter and the particle filters. However, the interaction between the various coordinate systems and the architecture of the Bayesian filters indicates the following three difficulties: For some coordinate systems the motion cannot be represented uniquely by coordinates. Highly nonlinear state space models emerge which is difficult to be handled by the Kalman-based filters. In case where a bilinear model is achieved, the Kalman-based filter architecture violates the algebraic constraint which expresses the geometry of the space in which the state evolves.

Considering the special structure of the underlying space of orientation, the problem of orientation and rate estimation without rate sensors can be formulated as a deterministic filtering

problem and be subjected to the theory of optimal control. Within the optimal control context the geometry of the space of orientations can be maintained. The model error of the dynamic equation, considered as unknown and deterministic signal, is utilized as a control port. The observer can be driven then towards the physical system by applying optimal corrections which result from well defined optimization problems. The kinematic equation then remains unaffected and it is led by the dynamics. In addition, the space of orientations admits the structure of a Lie group. Therefore, the aforementioned optimization problems can be formulated and solved in a coordinate-free fashion. Furthermore, the Hamiltonian formulation of optimal control combined with the theory of Lie groups, are the cornerstone for the second-order-optimal minimum energy filter on Lie groups [2]. The above observations and the inherent disadvantages of the Bayesian context constitute the main motivation for the presented thesis.

1-2 Contributions

From the framework that was mentioned in the previous section the filtering problem is set under the optimal control theory. In particular:

- The second-order-optimal minimum filter on the tangent bundle of the special orthogonal group $T\mathbb{SO}(3)$ [2] is implemented and basic points of its structure are commented.
- Formulating the control problem as in [3] a predictive filter is derived by modeling the system directly on $\mathbb{TSO}(3)$.
- The error dynamics are derived for both the second-order-optimal minimum energy filter and the predictive observer.
- The Riccati equation proposed but not provided by [4] is finally derived in this thesis in case where the state lies in $\mathbb{R}^n$.
- The aforementioned filters are compared and their performance is assessed.
- The predictive filter on $\mathbb{TSO}(3)$ is extended and used in order to provide alternative approach for the problem of orbital position estimation. Also, it has been proved that under some realistic assumptions, the whole problem can be decomposed into two different problems and therefore the resulting filter can be seen as a series interconnection of two distinct filters.

1-3 Organization

The presented thesis is organized as follows: Chapter 2 contains the literature survey. The Bayessian formulation of estimation and the deterministic filtering approach are discussed. The Kalman based filters and the optimal controllers are analyzed in detail. Subsequently, a brief discussion is made of how analytical mechanics are connected with the problem of attitude estimation. The next section is dedicated to the main conclusions of this survey while the last one presents the necessary numerical integrators regarding the rigid body motion.

Chapter 3 deals with the implementation of the second-order-optimal minimum energy filter on $T\mathbb{SO}(3)$ proposed by [2]. Experiments for various initial conditions and filter's parameters indicate its performance. The error dynamics of the filter are also derived.
Furthermore, in Chapter 4, the predictive filter on $\mathbb{TSO}(3)$ is derived. Simulations for various initial conditions and different filter's parameters reveal its performance. The error dynamics are also derived in this chapter.
In Chapter 5, the predictive filter on $\mathbb{TSO}(3)$ gives an alternative direction towards the problem of orbital position estimation. It has also been proved that the resulting filter can be decomposed in two distinct filters connected in series.
The last chapter contains the main conclusions of this work and some proposals for future research on this specific topic.
Finally, the appendix contains the proof regarding the Riccati equation of the second-order-optimal energy filter in case where the state lies in $\mathbb{R}^n$.

Chapter 2

Literature survey

2-1 Introduction

The main goal of this chapter is to present the two approaches that have been utilized by researchers in the available literature for the problem of orientation estimation from vector measurements. Initially, the Bayesian formulation of the estimation problem is studied, which culminates in the Kalman-based filtering framework and where the disadvantages of this approach are highlighted. In the sequel, the estimation problem is considered as a dual optimal control problem, which solves some of the main issues of the Bayesian approach. Before the two approaches be presented, some necessary mathematical preliminaries should be provided.

2-2 Mathematical preliminaries

A Lie group is a topological group that is also a smooth manifold, with some other properties [5]. Associated with every Lie group is a Lie algebra, which is a vector space discussed below. More importantly, a Lie group and its Lie algebra are closely related, allowing calculations in one to be mapped usefully into the other.

Definition 2-2.1. *An (n-dimensional) manifold is a pair (M, X) where M is a set and X is a maximal atlas into $\mathbb{R}^n$, such that the topology induced by X is Hausdorff and second-countable.*

Definition 2-2.2. *A **Lie group** is a manifold that has a group structure consistent with its manifold structure in the sense that group multiplication*

$$\mu : G \times G \to G \ : \ (g, h) \mapsto gh$$

*is a smooth map. The maps $L_g : G \to G \ : h \mapsto gh$ and $R_h : G \to G \ : g \mapsto gh$ are called the **the left and right translation maps** respectively.*

As an example, the set defined as

$$\mathbb{SO}(3) = \{R \in \mathbb{R}^{3\times 3} | R^T R = I, \det(R) = +1\} \tag{2-1}$$

is a Lie group called the **special orthogonal group** of dimension 3.

Definition 2-2.3. *A Lie algebra* $\mathfrak{g}$ *is a vector space that is closed under the Lie commutator* $[X, Y] = XY - YX$*, for* $X, Y \in \mathbb{R}^{n\times n}$*.*

Every Lie group has an associated Lie algebra, which is the tangent space at the identity of the group. The Lie algebra is a vector space generated by differentiating the group elements along chosen directions in the space at the identity. A very important feature (not true for any differentiable manifold) is that the tangent space has the same structure at all group elements, though tangent vectors undergo a coordinate transformation when moved from one tangent space to another.

The tangent space associated with a Lie group provides a space in which to represent differential quantities related to the group. The latter is because the tangent space is a vector space with the same dimension as the number of degrees of freedom of the group. Also, the exponential map converts any element of the tangent space exactly into an element of the group.

Finally, the adjoint action transforms linearly tangent vectors from one tangent space to another. The adjoint property is what ensures that the tangent space has the same structure at all points on the manifold, because a tangent vector can always be transformed back to the tangent space around the identity [6].

Considering the $\mathbb{SO}(3)$ the associated Lie algebra is defined as

$$\mathfrak{so}(3) = \{X \in \mathbb{R}^{3\times 3} | X^T = -X\} \tag{2-2}$$

For the rest of this thesis every definition and proposition will be given w.r.t. the $\mathbb{SO}(3)$.

Proposition 1. *The matrix Lie algebra* $\mathfrak{so}(3)$ *is isomorphic to* $(\mathbb{R}^3, \times)$ [1].

A 3×3 skew-symmetrtic matrix has dimension three. It is straightforward to transfer the matrix addition. The matrix commutator corresponds to the vector product.
For

$$X = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \in \mathbb{R}^3 \xrightarrow{(\,)^\times} X^\times = \begin{bmatrix} 0 & -X_3 & X_2 \\ X_3 & 0 & -X_1 \\ X_2 & X_1 & 0 \end{bmatrix} \in \mathfrak{so}(3)$$

Then for $X, Y \in \mathbb{R}^3$

$$X^\times Y = X \times Y = -Y \times X = -Y^\times X$$

and

$$[X^\times, Y^\times] = X^\times Y^\times - Y^\times X^\times = (X \times Y)^\times$$

[1] Where "$\times$" is referred to the vector product.

Theorem 1. *[7] The exponential map*

$$exp : \mathfrak{so}(3) \to \mathbb{SO}(3) \tag{2-3}$$

is well defined and surjective.

In Lie groups, a tangent vector from a tangent space can be mapped linearly to another tangent space. The adjoint performs this transformation. For an element R of a Lie group, the adjoint is written Adj_R. In the case of $\mathbb{SO}(3)$, the adjoint is [7]

$$Adj_R = R$$

In many cases we are interested to differentiate functions w.r.t. group elements. Consider for example $R \in \mathbb{SO}(3)$ and $X \in \mathbb{R}^3$. The representation X changes by R into the representation $Y \in \mathbb{R}^3$ as

$$Y = RX$$

We are interested to see how Y changes w.r.t. a small variation δR from R, with $\delta R \in T_R\mathbb{SO}(3)$, the tangent space in the group element R. Essentially, we seek the tangent map$df_R : T_R\mathbb{SO}(3) \to T_Y\mathbb{R}^3$. From the adjoint property it follows that $T_R\mathbb{SO}(3) = R\mathfrak{so}(3)$. The tangent space $T_Y\mathbb{R}^3$ can be identified with $\mathbb{R}^3$ by using the isomorphism induced by the parallel transport in $\mathbb{R}^3$.

For any tangent vector $\delta R \in T_R\mathbb{SO}(3)$, let $\phi(t) = Re^{t\Omega^\times}$ be a curve with $\phi(0) = R$ and $\frac{d}{dt}\phi(0) = R\Omega^\times$.

Then

$$df_R(R\Omega^\times) = \frac{d}{dt}(Re^{t\Omega^\times}X)(0) = R\Omega^\times X$$

For more details in the theory of manifolds and Lie groups the reader is referred to [8, 7, 5]

2-3 Bayesian sequential estimation

Consider the nonlinear dynamic system that can be described by a state-space model of the following form:

$$\begin{aligned} \dot{X}_1(t) &= f_1(X_1(t), X_2(t)) \\ \dot{X}_2(t) &= f_2(X_2(t), t) + BU(t) \\ Y(t) &= h(X_1(t); V(t), t) \end{aligned} \tag{2-4}$$

where $U \in \mathbb{R}^m$ and $V \in \mathbb{R}^q$ express the model uncertainty and the measurement noise respectively and $X{=}(X_1, X_2) \in \mathbb{R}^n$ is the state vector[2]. We are interested to sequentially estimate the state of the above system.

The dynamic state space model (DSSM), together with the known statistics of the noise random vectors as well as the prior distributions of the system states, defines a probabilistic

[2]At this point it is important to mention that the state is not necessarily a vector belonging to some vector space. However, this expression is commonly used in the literature.

generative model of how the system evolves over time and of how we observe the hidden state evolution. By following [9, 10] it is assumed that a discrete time system is available since it is easier for some important points to be revealed.

The problem statement of sequential estimation in the DSSM framework can be framed as follows: How do we optimally estimate the hidden system variables in a recursive way as noisy observations become available [11]. Given that in a Bayesian framework, the posterior density allows to calculate the optimal estimate for *any* cost functional, the problem can be recast as follows. How to recursively compute the posterior density as new observations arrive online?

By combining the Bayes chain rule and the DSSM we arrive in the following result:

$$p\left(X_k|Y_{1:k}\right) = \frac{p\left(Y_k|X_k\right)p\left(X_k|Y_{1:k-1}\right)}{p\left(Y_k|Y_{1:k-1}\right)} \tag{2-5}$$

Although in principle (2-5) propagates the posterior distribution, implementing the above relation in a closed form is in general not feasible. In the particular case where the DSSM is linear (not necessarily time invariant) and the noisy sequences Gaussian, the posterior distribution can be parametrized by its first two moments and be implemented by the Kalman filter [12].

However, in the MSE sense the above recursion can be optimally implemented by a linear estimator [12]. Therefore, even with a non-Gaussian, non-linear system, the optimal MSE estimator; the conditional expectation, is a linear function of the observations:

$$\begin{aligned} \hat{X}_k &= \hat{X}_k^- + K_k\left(Y_k - \hat{Y}_k^-\right) \\ P_{X_k} &= P_{X_k}^- - K_k P_{\overline{Y}_k} K_k^T \end{aligned} \tag{2-6}$$

where the optimal terms in this recursion are given by:

$$\begin{aligned} \hat{X}_k^- &= E\left[f\left(X_{k-1}, U_k\right)\right] \\ \hat{Y}_k^- &= E\left[h\left(X_k^-, V_k\right)\right] \\ K_k &= E\left[\left(X_k - \hat{X}_k^-\right)\left(Y_k - \hat{Y}_k^-\right)^T\right] E\left[\left(Y_k - \hat{Y}_k^-\right)\left(Y_k - \hat{Y}_k^-\right)^T\right]^{-1} \\ &= P_{X_k\overline{Y}_k} P_{\overline{Y}_k}^{-1} \end{aligned} \tag{2-7}$$

Nevertheless, the calculation of the above expected values requires all the moments of the random state vectors and all the derivatives of the nonlinear functions involved [10].

2-3-1 Extended Kalman filter (EKF)

The Extended Kalman Filter makes the following assumption

$$\begin{aligned} \hat{X}_k^- &\approx f\left(\hat{X}_{k-1}, \overline{V}, U_k\right) \\ \hat{Y}_k^- &\approx h\left(\hat{X}_k^-, \overline{V}\right) \\ K_k &\approx \hat{P}_{X_k Y_k}\left(\hat{P}_{Y_k}\right)^{-1} \end{aligned} \tag{2-8}$$

Clearly these approximations cannot be measured[3] [4]. They are valuable only when the nonlinearity of the system is small compared with the uncertainty of the state. However, this is a condition that cannot be inferred in a strict mathematical way. This also reveals how sensitive the EKF is to initial estimation errors. The EKF, arrives in these approximations by using a first-order truncation of the multidimensional Taylor series expansion for the nonlinear functions [11]. The first derivatives for both the nonlinear functions have to be calculated and stored.

The expected values of (2-7) require knowledge of all the derivatives of f and h and all the moments of the random states. By linearizing around the current state, implicitly, the expected values are determined by using only the first moment of the random variables involved and up to the first derivative of the functions in (2-7). In other words not the whole information that describes the uncertainty of the random variables is taken into account. There is no meaning of linearizing around a *random* vector. It has a meaning though to linearize around its first moment. The latter approach however, ignores its remaining inherent uncertainty.

Although the stochastic nature of the state is ignored when linearizing, the estimation error covariance matrix takes place in the EKF algorithm and therefore attributing, stochastic meaning to the estimated state. Therefore there exists an incompatibility regarding the way EKF operates.

Since the linearization step considers only the first moment, the filter "trusts" its own estimates more than is warranted by the true underlying state space evolution and observation sequence. This has large implications for the accuracy and consistency of the resulting EKF algorithm [11]. These approximations often introduce large errors in the EKF's calculated posterior mean and covariance of the (nonlinear) random variable, which may lead to suboptimal performance and sometimes divergence of the filter [9, 10, 13].

2-3-2 Unscented Kalman filter (UKF)

From the above discussion there is one major shortcoming of the EKF that need to be addressed. The disregard for the "uncertainty" of the underlying system state and noise random vectors during the linearization of the system equations, or in other words the limited first-order accuracy of propagated means and covariances resulting from a first-order truncated Taylor-series.

Due to the above flaws, derivativeless, deterministic sampling based Kalman filters have been developed that consistently outperform the EKF not only in terms of estimation accuracy, but also in filter robustness and ease of implementation, for no added computational cost [14, 15, 16, 17, 18]. These filters use a deterministic sampling approach to approximate the optimal gain and prediction terms in the Kalman filter framework and belong to a more general class of filters called sigma-point Kalman filters. The Uncended Kalman Filters is one of these and its core will be briefly analyzed here.

The UKF uses another approach in order to capture the statistics of the random state vector up to the second moment, without using the Taylor expansion. Instead, it employes an unscented transform (UT).

[3] It is not possible to measure how close the filter performs to the true stochastic mean.

Essentially, the UT is the answer to the following question: Assuming that we can calculate a set of *weighted* samples that accurately capture the relevant prior statistics of a random variable, what should the value of these weights and sample locations be, such that we can accurately approximate the posterior mean and covariance using only weighted sample mean and covariance estimators operating on functional (nonlinear) evaluations of these samples [19]?

Thus, the UT is a method for calculating the statistics of a random variable which undergoes a nonlinear function and is based on the principle that it is easier to approximate a probability distribution instead of an arbitrary nonlinear function [14, 20]. This is the main advantage over the EKF method. The UKF is a straightforward application of the scaled unscented transformation to the recursive Kalman filter framework. For more detail, see [20]. Another advantage of the UKF (since it does not use the Taylor expansion) is that it does not require the calculation of Jacobian matrices. In many applications, Jacobian matrix evaluation can be nontrivial and lead to implementation difficulties [15].

2-3-3 Application of Kalman based filters in the problem of attitude estimation

The work of [21] can be considered as the first one where the problem of attitude and rate estimation is studied in the absence of rate sensors. In this work, the Euler angles are used for the motion to be expressed with coordinates. The well known problem of Euler angles singularities is avoided by using a temporary shifted reference frame- the orientation is estimated w.r.t. the previous orientation estimate. By doing so, the orientation is estimated at each time instant w.r.t. a moving reference frame and therefore, the coordinated orientation remains away from singular points[4]. Although the state space model is highly nonlinear an EKF scheme is deployed. The flaws of the EKF mentioned in Sections 2-3-1 to 2-3-2 appear in this work. The nonlinearity of the system is increased w.r.t. to the uncertainty of the estimated state which leads into the filter's poor performance.

Later works [22, 23] try to address the problem differently by employing different attitude representations for the model to become quasi-linear. The only attitude representation that achieves quasi-linear kinematics-dynamics is quaternions. However, only the (unit) quaternion algebra is valuable for representing orientations and it does so by the quaternion multiplication. The unit constraint is maintained only by the quaternion multiplication. On the other hand, in the EKF algorithm, the previous state vector[5] is *added* to the scaled estimated error, thus violating the unit sphere constraint. The resulting estimates then, outside of the three-unit sphere, do not represent orientation. Generally, two approaches are used to deal with this issue: a Euclidean, or brute force normalization after the update stage [22], and the more complex multiplicative approach of [23].

In the work of [22] the orientation is represented by using the directional cosine matrix (DCM). The matrix is parametrized by using the algebra of quaternions but in its vectorial nature. That parametrization allows the problem to be treated in a vectorial way. The true orientation then is expressed as the sum of an estimate quaternion and a quaternion error. As state vector is considered the quaternion error that is connected with the output error vector through a

[4] Of course the time interval of two successive measurements should be reasonably small

[5] The term vector here should not be considered as an element of a vector space but rather as an element of $\mathbb{R}^n$

linear output equation. The algorithm for estimating the state is a simple recursive least squares scheme in the underlying quaternion algebra. In order for the parametrization of the DCM w.r.t. the quaternion to remain valid after the algebraic manipulations of the recursive least squares algorithm, the quaternion has to be normalized on the unit three-sphere. As was shown in [24] an optimal normalization in the Euclidean sense is achieved by simply dividing the quaternion estimate by its Euclidean norm.

The second approach [23] tries to address this issue by performing the quaternion update based on the Euler's theorem. The kinematic equation is presented in the matrix quaternion form. The underlying assumption is that, since instantaneously the motion is a rotation, the updated quaternion is the rotation of the previous quaternion. The rotation is performed by the constructed (according to the Rodrigues formula) skew symmetric matrix. In this way the update is performed in a multiplicative fashion and the unit constraint is maintained.

It is worth to be noticed that the work of [22] is closely related and in some sense evolved to the so called Multiplicative Extended Kalman filter (MEKF). In the MEKF [25] the attitude is represented by a unit quaternion. The key idea is that the true orientation is represented as a quaternion product of an error quaternion and a reference quaternion, with the error quaternion parametrized by a Euclidean (axis) vector. This non-zero vector is a way to indicate the deviation between an existing, imaginary reference frame and the estimated frame. The estimation procedure is applied in a vector space in order for the well known addition issues (with quaternions and EKF) to be avoided (this can be considered as the central idea of [25], i.e to apply the EKF without problems with the underlying algebra). The vectorial error is estimated, which implies an estimate for the quaternion w.r.t. the reference quaternion. Then the last estimate becomes the new estimate. The later articles indicate a way in order algebraic quaternion normalization issues to be avoided, i.e by parameterizing the orientation with vectorial elements, in order for the EKF to be applicable in a meaningful way.

So until now, the bad behavior of the EKF coupled with the Euler angles representation, has been replaced with the quasi-linear dynamics provided by the quaternion representation. However the issue that is shown up in that case has to do with the fact that the EKF does not maintain the unit sphere constraint of the unit quaternions and therefore in order to produce meaningful quantities the results have to be re-projected onto the unit sphere.

As indicated in Section 2-3-1, the UKF has the advantage of handling nonlinearities, through the UT, better that the EKF. This makes reasonable the choice for using it in conjunction with the Euler angles. An attempt towards this direction can be found in [26]. On the other hand, when the quaternion representation is used, the UKF in a standard format cannot be implemented straightforwardly. The reason is again the quaternion's unit constraint. If the kinematics are used in the filter directly, there is no guarantee that the predicted quaternion mean of the UKF will satisfy the unit norm constraint and therefore express an orientation. In [27] however, the authors overcome this problem by using an unconstrained vector to represent an attitude-error quaternion instead of using all its four components. They represent the local error-quaternion using the so called general Rodrigues parameters (GRP) coordinate system [28]. A comparison between the EKF and the UKF under the quaternion representation can be found in [29]. The conclusion is that the UKF performs better compared with the EKF, when the kurtosis and the higher order moments in the state error distributions are significant. The sampling rate is another cause why the UKF does

not perform better when estimating quaternion motion. In general, if the sampling rate is sufficiently high, the quaternion dynamics behave in a quasilinear fashion since, with small time steps, the integration steps propagate the quaternions only small deviations away from the unit sphere, making the error in linearization minimal.

Until now, the problem of attitude and rate estimation without rate sensors has been set under the Bayesian formulation of estimation. In the previous paragraphs it is revealed how the Kalman-based filters interact with the different coordinate systems and the quaternion algebra. The next section presents an alternative approach for the estimation problem, outside the Bayesian framework, using control theory.

2-4 Estimation problem as dual of the optimal control problem

The method of [30] determines the corrections added to the assumed model such that the model and corrections yield an accurate representation of the system behavior. The model uncertainty is considered as an unknown but deterministic signal within a Hilbert space. The goal is to estimate the states so that the resulting measurements approximate the measured observations while keeping the considered model as valid as possible. This is done by minimizing the total norm of the augmented error-uncertainty vector. However, since there is not a limit on how much the cost function can be decreased, the optimization problem incorporates a covariance constraint. This is done in order the state estimates to have statistical meaning. The problem as formulated here is related to the classical optimal control problem.

The covariance constraint is satisfied by tuning a positive definite matrix that reflects the amount of influence of the uncertain signal in the cost. By not allowing the uncertainty to influence significantly the total cost (equivalently, by keeping strong faith in the model) we might not be able to achieve good state estimates. On the other hand, by abandoning the assumed model more, we might achieve estimates that are without meaning, i.e. estimates that do not satisfy the covariance constraint. Essentially, the goal is to estimate the state by keeping faith in the model as much as possible, or in other words, by allowing only for a minimum model error. The necessary conditions for the minimization of the cost are called Pontryagin's necessary conditions. These conditions introduce an additional co-state vector and lead to a two-point boundary value problem (TPBVP).

The above method is applied in the work of [31]. The state equations consist of the attitude kinematics and dynamics. The output equations relate the known and the measured pointing directions with the unknown orientation matrix. The orientation is represented by using quaternions. One important fact that has to be mentioned here is that with this approach, the kinematic equation remains isolated. This is due to the fact that the uncertainty is considered as a deterministic signal that is determined separately during the estimation process and is fed as a correction term only through the rigid body dynamics. In this way, problems regarding the algebra of orientations can be avoided by updating the kinematics separately. However, the above filter has a very important drawback. A Large sequence of observations and many realizations have to be available since essentially we are talking about a batch (offline) estimation process. Another difficulty that is posed is that the TPBVP requires a tuning procedure in order for the boundary conditions to be satisfied. Despite the above mentioned difficulties, the main reason this filter is introduced at this point is because it reinforces the intuition for this non-Bayesian approach and reveals some initial benefits.

In [3] a predictive controller for nonlinear systems is developed. The optimal controller aims to minimize a prediction error based on a minimum current control expenditure. This formulation results into an online filter. Another important characteristic of the above method is that the optimization problem is not subject to the system's vector field. Instead, the information about the vector field is fed implicitly into the cost through the prediction that is expressed via a Lie-Taylor expansion and therefore the boundary-value problem induced from the Pontryagin's minimum principle is avoided.

Based on the previously described controller and the duality principle, a filter can be derived. The model uncertainty has to be considered as a deterministic and unknown signal of a function space. Since the divergence of the next state depends on the error of the current sate, the model error signal is considered as a function of the state. This state feedback function is determined from an optimization problem. The minimization targets to achieve the best output prediction while keeping the current model uncertainty low.

The prediction horizon step and the influence of the model uncertainty are conflict. By allowing for a large prediction step a small prediction error can be achieved by remaining close to the assumed model. On the other hand a small prediction error might not be achievable for a small prediction step without abandoning the model. Although a big step ahead might reduce the considered uncertainty set for a given prediction error, the results might not be correct in a statistical sense. In order to do so, the optimization problem for the filter derivation is subject to the covariance constraint referred to [30]. So, the goal is to choose for a proper time step that achieves a statistically meaningful output prediction, while the considered model remains as valid as possible.

2-4-1 Hamiltonian formulation of the optimal control problem

In [4] the modal trajectory estimator is studied. This approach is based entirely on the Hamiltonian formulation of optimal control. The initial cost[6] regulates the initial state and this has to be done by minimizing the total error between the measured and estimated outputs. The minimization is done w.r.t. both the deterministic model uncertainty signal and the initial state that produce these measurements. This optimization problem has to be solved for every time instant and is essentially the tracking problem of optimal control theory. The necessary conditions of optimality lead to a boundary value problem which has to be solved for each time instant.

This issue is solved by applying the Hamilton-Jacobi-Bellman theory and as a result the desired observer's differential equation is derived. The optimal estimate is then the final value of the optimal state trajectory. The very important point that is implicitly mentioned from [4] is the significance of the boundary conditions when they are used in combination with the Hamilton-Jacobi-Bellman equation. This is the key fact for deriving the differential equation for the filter. Another important fact is that a differential Ricatti-type equation analogous to the one of [12] can be derived. However the second-order Ricatti-type equation results by assuming that available information only up to second order is available for the estimate error uncertainty. Essentially, maximum-likelihood nonlinear filtering suffers from the same problem of moments[7] as the minimum variance nonlinear filters mentioned in Section 2-3.

[6] The corresponding terminal cost of the control problem
[7] This terminology is given by [4]

This is reasonable since the uncertainty is inherent and independent of how the problem is formulated. However, remarkably, under this approach, the problem of moments is reflected only in the gain operator. In particular in the fact that its evolution can be described up to some finite order. Therefore, the problem of moments is appeared to affect the filter in a less harmful way compared with the Bayesian formulation.

2-4-2 Second-order-optimal minimum energy filter for attitude estimation

The space of orientations admits the structure of a Lie group, the special orthogonal group $\mathbb{SO}(3)$. The manifold structure on a set can be defined without an algebraic structure. However, the ability of defining an algebraic structure provides some extra advantages for handling problems defined on the space. In the case of the matrix Lie groups, there is no need of working with coordinate maps and transition functions since the manifold structure itself is also equipped with an algebraic structure. As in the case of a vector space, algebraic operations are defined without using coordinates. It is completely valid to add and subtract vectors, in a geometric sense, without using coordinates. After all this is the distinction between geometry and linear algebra. The same is true if we consider the points that lie on the unit circle represented by the complex numbers of unit length. In this later case, every point of the circle operates on it and performs one of its symmetries.

However, coordinate system maps and transition functions are necessary in order to establish the differential structure and thus being able to perform differential calculus in a coordinate free fashion later on. The very important fact to be mentioned is that although it is feasible (in principle) to embed the original space within a Euclidean space and treat it as a submanifold, in some cases such information is not available and an intrinsic approach is needed. This is where the theory of differentiable manifolds is truly necessary. For more details on the theory of manifolds and Lie groups see [7, 8].

The next cornerstone needed in order to form a full picture about the work of [2] is analytical mechanics. This is another way of modeling mechanics. The variational principle was an idea based on the law of the lever proposed by Archimedes. Johann Bernoulli originally proposed the variational method and Lagrange, by giving an explanation on Bernoulli's method ended up with the famous method of multipliers. One of the main results of analytical mechanics is Hamilton's principle of least action. Every physical system evolves in such a way that minimizes a well defined length integral in an abstract differentiable manifold.

The Euler Lagrange equations (the mathematical expression of Hamilton's principle or a generalization of Newton's second law) give the evolution of the system, which appears to be a particle within the manifold space, and express equations of straight lines (geodesics). The optimal control problem then is the extension of Hamilton's principle. This can be seen by considering the constrained problem formed by incorporating the system's dynamics (vector field). By applying the variational method to the constrained (optimal control) problem, Pontryagin's principle gives the necessary conditions for optimality [32]. Instead of asking for the trajectory that minimizes the cost functional, the optimal control problem incorporates an additional port, the control commands, and asks for these control orders that produce a trajectory which minimizes the cost functional.

The aforementioned points are first utilized by [4], and extended in the work of [2] where a second-order-optimal minimum energy filter on Lie groups has been designed. The really

remarkable fact, which is also the reason for its performance, is its structure. It is the Lie group structure combined with basic principles of analytical mechanics and the Hamiltonian formulation of optimal control that produce this particular structure.

Based on the discussion of the previous section, Kalman-based filters perform a correction to both the kinematic and dynamic equation. However, working with bilinear equations, in order to maintain the geometry an undesirable re-projection step is needed. Working with the dual optimal control problem formulated as in [3], the attitude correction is performed through the dynamic equation and the kinematic equation is isolated to be updated separately. The success of observer proposed in [2] lies on the fact that it maintains the geometry and performs corrections to both the kinematic and the dynamic part. The update of the attitude equation is performed, in a geometric fashion, by utilizing the Lie group structure. This is the advantage of working without coordinates. By incorporating coordinate systems or the quaternion representation the natural algebraic structure of the space is not preserved.

2-5 Summary

In a nutshell, within the Bayesian context it can be observed how the Kalman based filters developed for solving the estimation problem interact with the geometry of the attitude space through the different coordinate systems the attitude can be expressed with. Either we end up with a highly nonlinear state space model (difficult to be handled by EKF,UKF scheme) and a filter without geometric problems[8] or, with a bilinear coordinate system and a filter that has problems to maintain its evolution within the space of orientations. By using the quaternion representation we end up with a bilinear filter that can easily be treated by the EKF or the UKF. However the next optimal estimate for the orientation is based on the previous optimal re-projected estimate. The projection step induces a projection error that is accumulated with time and therefore it might be considered as harmful for long time simulations. In the case where the kinematics and dynamics are expressed with Euler angles, a singular coordinate system, the orientation should be considered w.r.t. the previous estimate as in [21]. This is how the coordinates remain away from the singular point.

By realizing that the Kalman filter framework is applicable in nonlinear dynamics it is perfectly valid to work within it. The resulting theory reveals two different tools, i.e. the EKF and the UKF. One of the main features of the above strategies is that they require knowledge of the first and second order statistics and in particular the estimate error covariance. The estimate error covariance however requires knowledge of both the model and the measurement error. While measurement errors may be determined in numerous ways, model errors are generally unknown by definition [30]. The aforementioned filter strategies deal with the model error knowledge requirement by assuming that the model error is a symmetrically distributed white noise sequence of known covariance, normally called "process noise". However, this assumption often has no theoretical basis, and in fact, for physical systems, model errors are more often smooth functions resulting from typical model simplification assumptions [30].

In [33], the author questions the appropriateness of a probabilistic treatment of the estimation problem. It is sufficient to think about the reason of the statistical treatment. One of the

[8] The term "geometric problems" refers to problems of maintain the geometry of the space in the orientation's update step.

main ingredients of the Bayesian formulation is the availability of probability distributions. A probability distribution is required in order to localize the uncertainty and makes possible to talk about intervals. This is of course a general approach, independent on any particular state space. However, for the present problem, of the attitude and rate estimation, the state can be considered to evolve in a compact state space. As an example, someone can think of the two dimensional case, where the largest distance between any two orientations is 180^o. The present problem maintains this characteristic. There is a compact bounded set where the orientation lies in. There is no need to localize this information by using a probability distribution. The state space is already compact [7]. Of course, an assumption for the range where the angular rate vary should be made as well. It is reasonable to consider the angular rate of the spacecraft being bounded in its measure and free to have any direction[9].

It is also a fact that none of the above mentioned Kalman based methods shows up as the result of a well defined optimization criterion, rather they use an optimal result of the Bayesian context of the recursive estimation problem. Also, although it is possible to perform close to the optimum, there is no way to measure this performance in a rigorous mathematical way.

Trying to maintain the orientation's geometry while the filter operates, the problem of attitude and rate estimation is attacked within a deterministic framework as a dual optimal control problem. By doing so, the kinematic equation does not suffer from the above mentioned geometric issues. However the question of how the kinematic equation is properly integrated in time remains valuable.

All the approaches lie in the stochastic framework utilize some coordinate map and subsequently derive the filter 's differential equations. This is a commonly used strategy. By using coordinate system (bijective) maps, and smooth transition functions, the abstract space of interest is modeled (under some additional assumptions) as a differentiable manifold. Then the problem is transferred on the manifold where the well known tools from calculus in $\mathbb{R}^n$ can be utilized. The results are mapped back in the original space and conclusions are inferred. Nevertheless, the problem of attitude estimation can be studied under the theory of Lie groups. Lie groups are constructions with both a differential structure and an algebraic structure. There is no need for establishing coordinates and the problem can be defined directly on the space it self. The second-order-optimal minimum energy filter on Lie groups essentially transfers the derivation of [4] in the context of Lie groups. In some sense, the problem is transformed to the Lie algebra which is a vector space resulting the trivialized dynamics and as a result ends up with a filter which has familiar structure.

2-6 Geometric integrators for the rigid body motion

This section introduces the adapted numerical methods for Lie groups that are generated by left invariant vector fields. General numerical integration algorithms approximate the evolution of a dynamical system by utilizing only the information from the vector field. Nevertheless, geometric integrators are designed in order to preserve inherent characteristics of the continuous system, like the symplectic structure, the geometry of the configuration space. The main advantage of a geometric integrator is that it can be used for long time simulations.

[9]The directions are in one-to-one correspondence with the points on the unit sphere embedded in the three dimensional Euclidean space

Modeling the rigid body motion within the context of Lie group theory provides us with a coordinate-free and therefore a singularity-free approach. The continuous time equations of a controlled rigid body are given by

$$\begin{aligned} \dot{R} &= R\Omega^{\times} \\ \dot{\Omega} &= \mathbb{I}^{-1}((\mathbb{I}\Omega)^{\times}\Omega + T) \end{aligned} \tag{2-9}$$

where $T \in \mathbb{R}^3$ is the control torque, $R \in \mathbb{SO}(3)$ the rigid body orientation, $\Omega \in \mathbb{R}^3$ the rigid body angular rate and $\mathbb{I} \in \mathbb{R}^{3\times3}$ is the inertia tensor. As mentioned above, although it is possible to express the motion with coordinates we prefer a numerical integration on the configuration space directly.

2-6-1 Implicit second-order integrators

The challenge of designing an integrator for updating the rigid body dynamics equations is that the differential equations are defined on a Lie group which as we saw in Section 2-2 is a smooth manifold with curvature which is equipped with a group structure. For all Hamiltonian systems the continuous time flow is symplectic and preserves the Hamiltonian. Utilization of the Legendre transform implies that the total energy of the system is also preserved [34].

The evolution of the system in a configuration manifold can be comprehended as a geodesic of the deformed due to the potential energy function space. The state space equations are the equations which describe the geodesics. Contrary, these differential equations, the so called Lagrange equations, show as necessary conditions for optimality of a cost functional which is expressed through an integral of an, most of the times, artificial Lagrangian function defined on the deformed space. The differential equations are derived by applying the variational principle, which basically results into a generalization of the differential for a general functional.

There are two approaches to create an update numerical scheme. The first one discretizes the geodesics and the second discretizes the cost functional directly and derives the integration algorithm as a result of applying the variational principle. The advantage of the second approach is that it produces a numerical algorithm through a machinery, which naturally leads to integrators that preserve essential characteristics of the continuous time system, as for example the symplectic form or the momentum. Note that the Euler-Lagrange equations are always second order differential equations. A numerically transcendent integrator can be set as follows [34] :

$$\begin{aligned} R_{k+1} &= R_k\tau(h\Omega_k) \\ C_{\tau}^{\mathrm{T}}(h\Omega_{k+1})\mathbb{I}\Omega_{k+1} - C_{\tau}^{\mathrm{T}}(-h\Omega_k)\mathbb{I}\Omega_k &= hT(\Omega_k) \end{aligned} \tag{2-10}$$

Where for the map τ can be used either the exponential map or the Cayley map :

$$\begin{aligned} \exp(\Omega) &= I_{3\times3} + \frac{\sin(||\Omega||)}{||\Omega||}\Omega^{\times} + \frac{1-\cos(||\Omega||)}{||\Omega||^2}(\Omega^{\times})^2 \\ \mathrm{cay}(\Omega) &= I_{3\times3} + \frac{4}{4+||\Omega||^2}\Big(\Omega^{\times} + \frac{(\Omega^{\times})^2}{2}\Big) \end{aligned} \tag{2-11}$$

Finally the map $C_\tau : \mathbb{R}^3 \to \mathbb{L}(\mathbb{R}^3, \mathbb{R}^3)$ is, depending on the choice of the group map τ :

$$\begin{aligned} C_{\exp}(\Omega) &= \mathbb{I}_{3\times 3} - \frac{1}{2}\Omega^\times + \frac{1}{12}(\Omega^\times)^2 \\ C_{\text{cay}}(\Omega) &= \mathbb{I}_{3\times 3} - \frac{1}{2}\Omega^\times + \frac{1}{4}\Omega\Omega^T \end{aligned} \tag{2-12}$$

By ignoring the quadratic terms in these later matrices, thus resulting in the map

$$C(\Omega) \approx I_{3\times 3} - \frac{1}{2}\Omega^\times \tag{2-13}$$

the integration scheme ends up to be the trapezoidal Lie-Newmark method [35]. This map can be substituted for either $C_{\exp}$ or C_{cay} without losing the second-order accuracy of the discrete Euler Poincare equations. For the purpose of this thesis we use the exponential map as a difference map which results in the following implicit integration scheme :

$$R_{k+1} = R_k e^{h\Omega_k^\times} \tag{2-14}$$

$$C_{\exp}(-h\Omega_{k+1})(\mathbb{I}\Omega_{k+1}) = C_{\exp}(h\Omega_k)(\mathbb{I}\Omega_k) + hT(\Omega_k) \tag{2-15}$$

where h is the integration step. The integrator update step is summarized in the following pseudococe [36] :

Given (R_k, Ω_k)

- pick $\tau = \exp$ or $\tau = \text{cay}$
- Compute C_τ
- Solve (2-15) for Ω_{k+1} using a Newton solver
- Update (R_{k+1}, Ω_{k+1})

However, using the untruncated expressions (2-12) the resulting scheme is a variational symplectic integrator which preserves the spatial momentum in the absence of applied torques.

2-6-2 Time step bounds

Both the trapezoidal and the symplectic methods are implicit, i.e they require the solution of a nonlinear algebraic equation. Usually this nonlinear equation is solved by employing a Newton method. In this case, a maximum allowed time-step rule can be derived which guarantees solvability of the equation. For the trapezoidal algorithm particularly, the nonlinear equation can be formulated as [34]:

$$\mathbb{E}(\Omega) = \Big(\mathbb{I}_{3\times 3} - \frac{1}{2}\Omega^\times + \frac{1}{12}(\Omega^\times)^2\Big)\mathbb{I}\Omega - a_k = 0 \tag{2-16}$$

where

$$a_k = C_{\exp}(-h\Omega_k)(\mathbb{I}\Omega_k) + hT(\Omega_k)$$

Newton solvers extract a solution for the above nonlinear algebraic equation by using the initial guess $\Omega_{n_0} = \Omega_k$, $\forall k \in \mathbb{N}^*$ which is then iterated according to

$$\Omega_{n+1} = \Omega_n - \left(\mathbf{d}_\Omega \mathbb{E}(\Omega_n)\right)^{-1} \mathbb{E}(\Omega_n)$$

where

$$\mathbf{d}_\Omega \mathbb{E} = \mathbb{I} - \frac{h}{2}(\mathbb{I}\Omega)^\times + \frac{h}{2}\Omega^\times \mathbb{I} + \frac{h}{4}\Omega^T \mathbb{I}\Omega I_{3\times 3} + \frac{h}{2}\Omega\Omega^T\mathbb{I}$$

is the Jacobian of $\mathbb{E}$ which of course has to be invertible. From the above expressions can easily be observed that for a very small time-step $h \approx 0$ the Jacobian matrix is always invertible (since the inertia tensor is always considered to be positive definite w.r.t. the body fixed frame). However, for larger time steps we can write [34]

$$\mathbf{d}_\Omega \mathbb{E} \geq \Big(\mathbb{I}_{min} - \frac{h}{4}||\Omega||(\mathbb{I}_{max} - \mathbb{I}_{min}) + h^2||\Omega||^2\big(\mathbb{I}_{min} - \frac{1}{4}\mathbb{I}_{max}\big)\Big) I_{3\times 3}$$

From the above inequality we can conclude that if the time step is chosen such that $0 < h < \tilde{h}(\Omega)$ where the upper bound is defined by

$$\tilde{h}(\Omega) = \begin{cases} \infty & \text{if } \kappa \geq 4\sqrt{7} - 7 \text{ or } ||\Omega|| = 0 \\ \frac{\kappa - 1 - \sqrt{\kappa^2 + 14\kappa - 63}}{(8 - 2\kappa)||\Omega||} & \text{otherwise} \end{cases} \tag{2-17}$$

with $\kappa = \frac{\mathbb{I}_{max}}{\mathbb{I}_{min}}$, the Jacobian matrix remains positive definite.

2-6-3 Comparison between explicit implicit and symplectic method

In what follows, a rigid body its inertia tensor w.r.t. to the principal axes has the diagonal representation $\mathbb{I} = \text{diag}(2, 5, 3)$ Kg m^2 w.r.t the body frame. We assume that initially there is no deviation of orientation, i.e that $R(0) = I_{3\times 3}$ and the initial angular velocity is $\Omega_0 = \begin{bmatrix} 0.5 & 0.6 & 0.4 \end{bmatrix}^T$. The input torque is $T(t) = \begin{bmatrix} \sin(\frac{2\pi}{3}t) & \cos(\frac{2\pi}{1}t) & \sin(\frac{2\pi}{5}t) \end{bmatrix}^T$. The output equations relate the inertial directions $\alpha_{i=1,2}$, with the directions expressed w.r.t. the body frame $y_{i=1,2}$.

$$\dot{\Omega} = \mathbb{I}^{-1}((\Omega\mathbb{I})^\times \Omega + T)$$
$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} R^T a_1 \\ R^T a_2 \end{bmatrix} \quad ,$$

where $R \in \mathbb{SO}(3)$, $\Omega \in \mathbb{R}^3$, $(.)^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$, $a_i \in \mathbb{R}^3$. The Jacobian remains invertible even for $h = 0.1$. Figure 2-1 shows the angular velocities and unit directions respectively.

The simplest Euler-method on $\mathbb{SO}(3)$ has the form

$$\begin{aligned} R_{k+1} &= R_k e^{h\Omega_k^\times} \\ \Omega_{k+1} &= \Omega_k + h\mathbb{I}^{-1}\big((\mathbb{I}\Omega_k)^\times \Omega_k + T_k\big) \end{aligned} \tag{2-18}$$

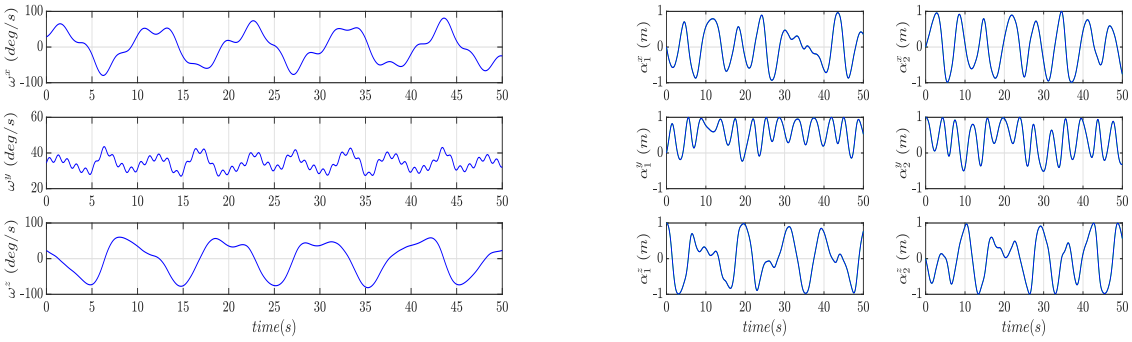


Figure 2-1: Angular velocity components (left) and unit directions (right) w.r.t. the body fixed frame for symplectic integration

Since for small h the matrix $R_k^T R_{k+1}$ is close to the identity where $\exp : \mathfrak{so}(3) \to \mathbb{SO}(3)$ is a diffeomorphism, the above integration scheme explicitly updates the next state given the current state.

Similarly to any other Euler method, the explicit scheme is only accurate up to the first order and has numerical stability issues especially for large time-steps [34]. In order to see the essential difference between an explicit Euler scheme (2-18) and the symplectic scheme, we can look at the rigid body's kinetic energy $\mathcal{K}_k = \frac{1}{2}\Omega_k^T \mathbb{I}\Omega_k$ transferred from the input torque.

From Figure 2-2 one can observe the main difference between the two integration schemes. The symplectic method conserves the energy while the explicit Euler scheme seems to add, artificially, energy into the system. The latter can also be seen from the fact that the angular rate norm diverges.

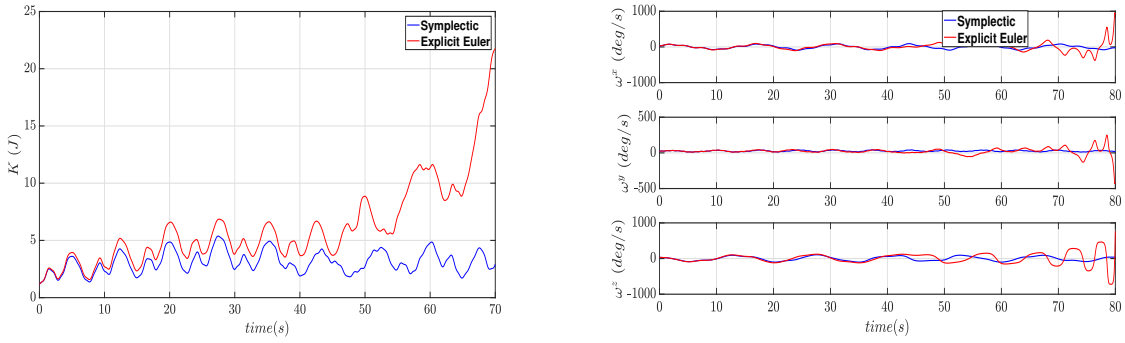


Figure 2-2: Kinetic energy and angular rate for explicit and symplectic scheme

Figure 2-2 - although correct - does not convince for our claim since indeed, with a proper torque input, the body could accelerate without limit. Another experiment that provides a better intuition can be formed by eliminating the external input torque $T(t) = \begin{bmatrix} 0 & 0 & 0 \end{bmatrix}^{\mathrm{T}}$ $\forall t \in \mathbb{R}^{+}$. Thereafter, the transferred work is zero and therefore the kinetic energy

$$\mathcal{K}_k = \frac{1}{2}\Omega_k^T \mathbb{I}\Omega_k = \mathcal{K}_0 = \frac{1}{2}\Omega_0^T \mathbb{I}\Omega_0 = 1.19 \ \forall k \in \mathbb{N}^*$$

However, in Figure 2-3, the integrator seems to artificially add an amount of energy into the autonomous system, which also means that the norm of the angular rate increases without limit as indicated the second plot in Figure 2-3. In the implicit Euler method the update Ω_{k+1} appears on both sides of the equation and thus the method needs to solve an algebraic equation for the unknown Ω_{k+1}. The implicit integration rule reads

$$\Omega_{k+1} = \Omega_k + h\Big((\mathbb{I}\Omega_{k+1})^{\times}\Omega_{k+1} + T_k\Big) \tag{2-19}$$

and by defining

$$\mathbb{E}_{imp}(\Omega) = \Omega - \Omega_k - h\Big((\mathbb{I}\Omega)^{\times}\Omega + T_k\Big) \tag{2-20}$$

the nonlinear algebraic equation $\mathbb{E}_{imp}(\Omega) = 0$ needs to be solved by means for example of a Newton solver, with starting point $\Omega_{n=0} = \Omega_k$. We make the comparison for all the three integrators regarding their ability to conserve the kinetic energy as shown in Figure 2-4

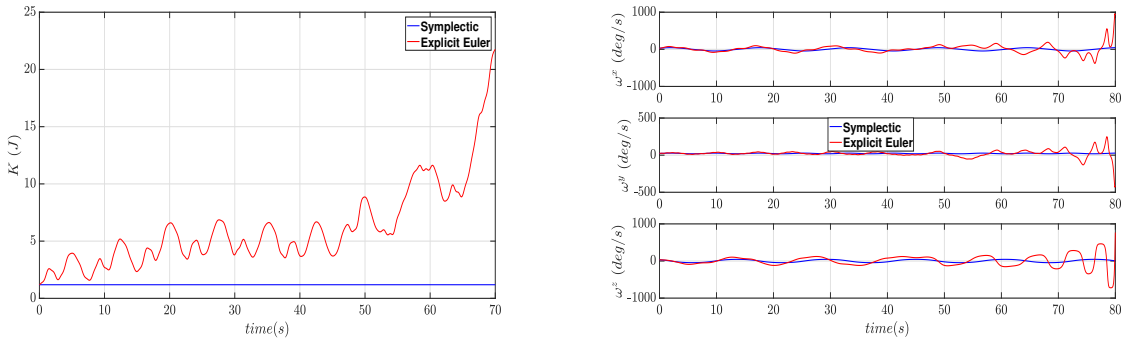


Figure 2-3: Kinetic energy and angular rate for explicit and symplectic scheme (zero input torque)

From Figure 2-4, it is concluded that the symplectic scheme conserves the kinetic energy. The explicit Euler method adds energy into the system, while the implicit integrator works as an artificial damper, i.e. it removes energy from the system. Since the initial assumption for the continuous time model demands the kinetic energy to remain constant, both the explicit and the implicit methods are not valid for simulating the rigid body motion. Throughout the presented thesis and its results, the symplectic method is therefore applied.

Since the implicit numerical scheme is characterized by a poor performance, the symplectic's scheme property of conserving the kinetic energy is not implied from its implicit character.

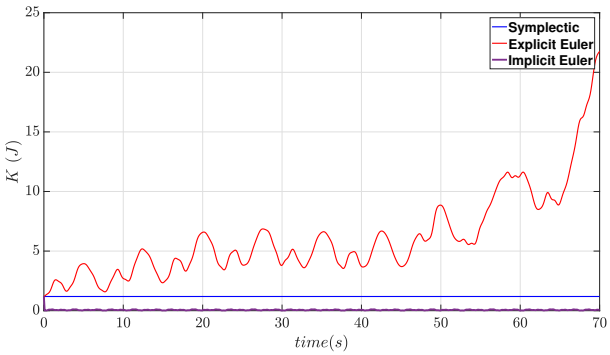


Figure 2-4: Kinetic energy for all three integration methods

Symplecticity should be understood as a distinct property. The symplectic integrator is produced under the machinery of the variational principle of mechanics by discretizing directly the cost functional. This is the map which transferes all the desired properties to the numerical scheme.

2-7 Problem statement

The problem is stated into two parts. The first part aims to estimate the orientation and rate of the spacecraft w.r.t. the local-vertical-local-horizontal frame (L.V.L.H.)[10] frame. The second part aims to estimate the orientation w.r.t. the Earth-centered inertial (E.C.I)[11] frame.

- Part I

  Given the Sun's and the Geomagnetic field's direction w.r.t the L.V.L.H. frame and measuring the Sun's and the Geomagnetic field's direction, the goal is to estimate the orientation and the angular rate of the spacecraft w.r.t. the L.V.L.H. frame.

  Afterwards, the estimated rate is the relative angular rate of the body frame w.r.t. the L.V.L.H. frame. Let us declare with $R(t) \in \mathbb{SO}(3)$ the orientation of the body frame w.r.t. the L.V.L.H. frame at the time instant $t \in \mathbb{R}$ and with $\Omega(t) \in \mathbb{R}^3$ the corresponding angular rate w.r.t. the same frame. Let us declare with $(R, \Omega), R \in \mathbb{SO}(3), \Omega \in \mathbb{R}^3$ the state of the system. Euler's theorem states that the motion of the rigid body w.r.t. its center of mass is instantaneously a rotation, or in other words:

  $$R^T(t)\dot{R}(t) = \Omega^{\times}(t) \tag{2-21}$$

  The time variation of R, when it is accounted from the body frame, is seen for a small time as a rotation. This equation is of the form $\dot{X}_1 = f_1(X_1, X_2)$. A second equation, Euler's equation, describes how the axis of rotation Ω changes with time while is affected by the external torques T_{ext} and the angular momentum $\mathbb{H} = \mathbb{I}\Omega$. The external torques are considered to include both environmental and control torques $T_{\text{ext}} = T_{\text{e}} + T_{\text{c}}$. The angular rate expressed in body fixed frame evolves according to

  $$\dot{\Omega} = \mathbb{I}^{-1}((\Omega\mathbb{I})^{\times}\Omega + T_{\text{ext}}) \quad , \tag{2-22}$$

  where $\mathbb{I} \in \mathbb{R}^{3\times3}$ is the inertia tensor and $T \in \mathbb{R}^3$ the environmental torques represented w.r.t. the body frame. This equation is of the form $\dot{X}_2 = f_2(X_2, U)$. In addition, the environmental torques are not (precisely) known and the same holds for the inertia tensor. This lack of knowledge makes the true angular rate Ω at time t, to deviate from what the given model indicates by an unknown and deterministic term $\delta(t)$. Therefore, the model can be recast to

  $$\dot{\Omega}(t) = \mathbb{I}^{-1}((\Omega(t)\mathbb{I})^{\times}\Omega(t) + T(t)) + \delta(t) \tag{2-23}$$

  where T refers to the control torques and both unknown factors are accumulated within the term $\delta(t)$.

  On the other hand, the measured directions $y_{i=1,2}(t)$ are represented w.r.t. the body frame and related to the corresponding, known, inertial directions $a_{i=1,2}(t)$ w.r.t. the L.V.L.H. frame by the body's orientation, as a change of coordinates,

  $$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} R^T a_1 \\ R^T a_2 \end{bmatrix} \tag{2-24}$$

[10] The L.V.L.H. frame is analogous to the Frenet-Serret frame which appears in differential geometry.
[11] An E.C.I frame has its origin in the Earth's center of mass and does not rotate with respect to the stars.

However, measurements are subject to additive sensor noise ϵ and (2-24) should be modified to

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} R^T a_1 \\ R^T a_2 \end{bmatrix} + \epsilon \tag{2-25}$$

or in general to $Y = h(X_1, \epsilon, t)$.

Based on the above, given the control torques T and the available on-board measurements $y_{i=1,2}$ the goal is to estimate the system's state (R, Ω).

- Part II

  However, as the problem stated above, a unique R might correspond to many orientations of the spacecraft w.r.t. the E.C.I. frame. For the configuration to be estimated uniquely, knowledge of how the L.V.L.H. frame is oriented w.r.t. the E.C.I. frame is necessary. Since the L.V.L.H. is essentially the Frenet frame on the orbit, the previous requirement is equivalent with estimating the orbital position θ. Given that no position information is available on board, the goal is to find a way to estimate the orbital position only from vector measurements. Thereupon, the pair (θ, R) determines uniquely the orientation of the spacecraft w.r.t. the E.C.I. frame.

  Intuitively, the motion can be thought of as the concept of a robot manipulator's motion. The first degree of freedom corresponds to the orbital position θ and the second degree of freedom to the rigid body's orientation R w.r.t. the L.V.L.H. frame. Although kinematically the two motions result the orientation of the body fixed frame w.r.t. the E.C.I. frame the motion dynamics of the center of mass are uncoupled with the attitude dynamics. However, a proper state space model has to contain output equations which, ideally, involve both R and θ. In other words, ideally we want a state space model of the form

$$\begin{aligned} R^T \dot{R} &= \Omega^{\times} \\ \dot{\Omega} &= \mathbb{I}^{-1}((\Omega\mathbb{I})^{\times}\Omega + T) + B\delta \\ \ddot{\theta}(t) &= f(\dot{\theta}, \theta, \xi) \\ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} R^T R(\theta)^T a_1 \\ R^T R(\theta)^T a_2 \end{bmatrix} + D\epsilon \end{aligned} \tag{2-26}$$

  where $a_{i=1,2}$ refer to representations w.r.t. the E.C.I. frame, $\xi \in \mathbb{R}$ to some torque input which affects the orbital motion of the spacecraft and $f : \mathbb{R} \to \mathbb{R}$ a nonlinear function.

Chapter 3

Second-order-optimal minimum energy filter on $T\mathbb{SO}(3)$

3-1 Filter's structure

This section deals with the implementation of the second order-optimal minimum energy filter. The cornerstone of the minimum energy filtering is analytical mechanics and the Hamiltonian formulation of optimal control. The basic problem of analytical mechanics is to derive geodesics by minimizing a well defined cost functional. The minimization is performed, by utilizing the variational principle. The geodesics are expressed from the Euler-Lagrange equations and essentially emerge as the necessary conditions for optimality.

Subsequently, the optimal control problem can be comprehended as a constrained optimization problem. The search is performed only among these trajectories of the manifold space which can be produced from a parameterized vector field in some space of admissible controls. The constrained optimization problem can be recast into an unconstrained problem by introducing the Lagrange multiplier. Since the whole formulation is based on physical quantities, the multiplier is introduced by a work function which is the fundamental quantity in analytical mechanics. The Lagrange multiplier acts from the co-tangent space. It represents how the control commands are introduced into the problem represented by a physical quantity. The latter observation can be checked also from Pontryagin's maximum principle which provides the necessary conditions for optimality. These conditions express a boundary value problem. Based on the above, the estimation problem can be formulated as a dual optimal control problem.

The main contribution of Mortensen in [4] is that he managed to overcome the difficulty from the boundary value problem and to extract a differential equation for the filter. The filter contains a gain operator the evolution of which can be described up to a countable ordered differential equation. In the second-order-optimal minimum energy filter a gain operator is evolved according to a second-order Riccati-type equation. A second-order-optimal minimum energy filter is derived from [2] in the case where the considered state space is modeled as a Lie group. The main contribution of [2] lies in the derivation of the disturbed Riccati equation.

Given the system

$$\begin{aligned} R^T\dot{R} &= \Omega^\times \\ \dot{\Omega} &= \mathbb{I}^{-1}((\Omega\mathbb{I})^\times\Omega + T) + B\delta \\ \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} &= \begin{bmatrix} R^T a_1 \\ R^T a_2 \end{bmatrix} + D\epsilon \end{aligned} \tag{3-1}$$

where $R \in \mathbb{SO}(3)$, $\Omega \in \mathbb{R}^3$, $(.)^\times : \mathbb{R}^3 \to \mathfrak{so}(3)$, $D \in \mathbb{R}^{6\times 6}$, $a_i \in \mathbb{R}^3$, the second order minimum energy filter is expressed by the following differential equations:

$$\begin{aligned} \hat{R}^T\dot{\hat{R}} &= \left(\hat{\Omega} + K_{11}r^R + K_{12}r^\Omega\right)^\times \\ \dot{\hat{\Omega}} &= \mathbb{I}^{-1}\left((\hat{\mathbb{I}}\hat{\Omega})^\times\hat{\Omega} + T\right) + K_{21}r^R + K_{22}r^\Omega \end{aligned} \tag{3-2}$$

where the residual $r_t = [r_R\ r_\Omega]^T$ and the second order optimal gain

$$\mathbb{K} = \begin{bmatrix} \mathbb{K}_{11} & \mathbb{K}_{12} \\ \mathbb{K}_{21} & \mathbb{K}_{22} \end{bmatrix}$$

are given bellow. Let $\hat{y}_1 = \widehat{R}^T a_1$ and $\hat{y}_2 = \widehat{R}^T a_2$ then the residual r_t is given by

$$r_t = \begin{bmatrix} r^R \\ r^\Omega \end{bmatrix} = \begin{bmatrix} -\left(q_1/d_1^2\right)\left(\hat{y}_1 \times y_1\right) - \left(q_2/d_2^2\right)\left(\hat{y}_2 \times y_2\right) \\ 0 \end{bmatrix}$$

while the gain $\mathbb{K}$ is the solution of the perturbed matrix Riccati differential equation:

$$\begin{aligned} \dot{\mathbb{K}} = &-\alpha\mathbb{K} + A\mathbb{K} + \mathbb{K}A^\top - \mathbb{K}E\mathbb{K} + BR^{-1}B^\top \\ &- W\left(\mathbb{K}, r_t\right)\mathbb{K} - \mathbb{K}W\left(\mathbb{K}, r_t\right)^\top \end{aligned} \tag{3-3}$$

where

$$\begin{aligned} A &= \begin{bmatrix} -\widehat{\Omega}^\times & I \\ 0 & \mathbb{I}^{-1}\left[(\mathbb{I}\widehat{\Omega})^\times - \widehat{\Omega}^\times\mathbb{I}\right] \end{bmatrix} \\ E &= \begin{bmatrix} \sum_{i=1}^{2} -\left(q_i/d_i^2\right)\left(\hat{a}_i^\times a_i^\times + a_i^\times \hat{a}_i^\times\right)/2 & 0 \\ 0 & 0_{3\times 3} \end{bmatrix} \\ BR^{-1}B^\top &= \begin{bmatrix} 0_{3\times 3} & 0 \\ 0 & B_2R^{-1}B_2^\top \end{bmatrix}, \quad \text{and} \\ W\left(\mathbb{K}, r_t\right) &= \begin{bmatrix} 1/2\left(\mathbb{K}_{11}r^R + \mathbb{K}_{12}r^\Omega\right)^\times & 0 \\ 0 & 0_{3\times 3} \end{bmatrix} \end{aligned}$$

for the (0)-connection function and

$$
\begin{aligned}
A &= \begin{bmatrix} 0 & I \\ 0 & \mathbb{I}^{-1}\left[(\widehat{\mathbb{I}\Omega})^{\times} - \widehat{\Omega}^{\times}\mathbb{I}\right] \end{bmatrix} \\
E &= \begin{bmatrix} \sum_{i=1}^{2} - \left(q_i/d_i^2\right) a_i^{\times}\hat{a}_i^{\times} & 0 \\ 0 & 0 \end{bmatrix} \\
BR^{-1}B^{\top} &= \begin{bmatrix} 0_{3\times 3} & 0 \\ 0 & B_2 R^{-1} B_2^{\top} \end{bmatrix}, \quad \text{and} \\
W\left(\mathbb{K}, r_t\right) &= 0
\end{aligned}
$$

for the (-)-connection function. The remarkable feature of the above filter is its structure. Since only the rigid body dynamic equation is considered uncertain we would expect a correction term acting only on the second equation. However, the filter incorporates an additional term which affects also the kinematic equation. By introducing the map $(\ )^{-\times} : \mathfrak{so}(3) \to \mathbb{R}^3$ we can write :

$$
\begin{bmatrix} (\hat{R}^T\dot{\hat{R}})^{-\times} \\ \dot{\hat{\Omega}} \end{bmatrix} = \begin{bmatrix} \hat{\Omega} \\ \mathbb{I}^{-1}\left((\mathbb{I}\hat{\Omega})^{\times}\hat{\Omega} + T\right) \end{bmatrix} + \begin{bmatrix} \mathbb{K}_{11} & \mathbb{K}_{12} \\ \mathbb{K}_{21} & \mathbb{K}_{22} \end{bmatrix} \begin{bmatrix} r_R \\ r_{\Omega} \end{bmatrix} \tag{3-4}
$$

or

$$
\dot{X} = f(X) + \mathbb{K}r_t \tag{3-5}
$$

which is exactly the structure of an unbiased estimator like for example the Kalman filter. So the above filter achieves the structure of a Kalman filter without violating the geometry of the space. The Kalman filter was originally developed by utilizing basic tools from the theory of Hilbert spaces, an approach which is not applicable in general since it is based on the assumption that the underlying space is a vector space. The alternative, Hamiltonian formulation of optimal control provides more freedom for deriving the aforementioned structure in more general cases.

This is the advantage of formulating the estimation problem as an optimal control problem instead of utilizing the Bayesian formulation. The former, more conservative approach, applies successfully in this specific problem due to the fact that the orientations are mapped into a compact manifold.

3-2 Simulations

The filter is implemented by employing the symplectic integration method from Section 2-6. Without loss of generality it is assumed that the inertia tensor is expressed w.r.t the principal axes and therefore represented with the diagonal matrix $\mathbb{I} = \text{diag}(2\ 5\ 3)$ Kgm^2. Initially the rigid body deviates $R(0) = (60^o, \begin{bmatrix} 0 & 0 & 1 \end{bmatrix})$ from the identity and rotates with $\Omega_0 = \begin{bmatrix} 0.5 & 0.6 & 0.4 \end{bmatrix}^{\mathrm{T}}$. The tracking directions w.r.t. the inertial frame are considered to be $a_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ and $a_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$. As in [2], the measured output is obtained by adding a Gaussian white noise with standard deviation $\sigma_n = 0.5/3m$, corresponding to assume that with probability 99.7% the output samples have a distance of less than $0.5m$ from the nominal output samples. A sample error of $0.5m$ corresponds to an error of about 30 deg in the measurement of the reference directions. Figure 3-1 represents, for the (0)- and (-)-connection functions, the attitude error $e^R \in [0,\pi]$ and the angular velocity error e^Ω. These errors are defined as

$$\begin{aligned} e^R(t) &:= \text{acos}\left(1 - \frac{\text{tr}\left(I - R^T(t)\hat{R}(t)\right)}{2}\right) \\ e^\Omega(t) &:= \widehat{\Omega}(t) - \Omega(t) \end{aligned} \tag{3-6}$$

Both filters show similar performances, with the (0)-connection function performing best.

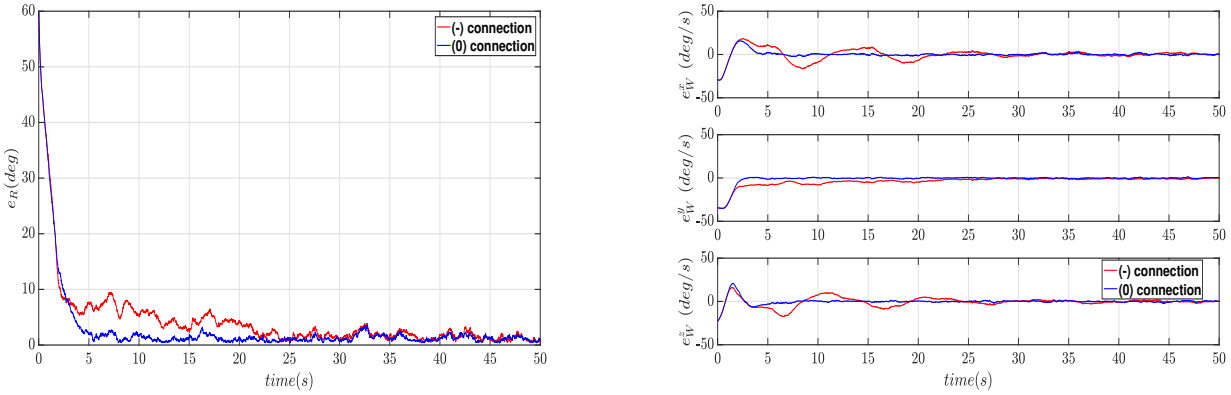


Figure 3-1: Plots for attitude error e_R and angular rate error e_Ω for both connection functions

By varying the inertia tensor to $\mathbb{I} = \text{diag}(1, 8, 3)$ Kg m^2 the role of the connection function becomes crucial as shown in Figure 3-2.

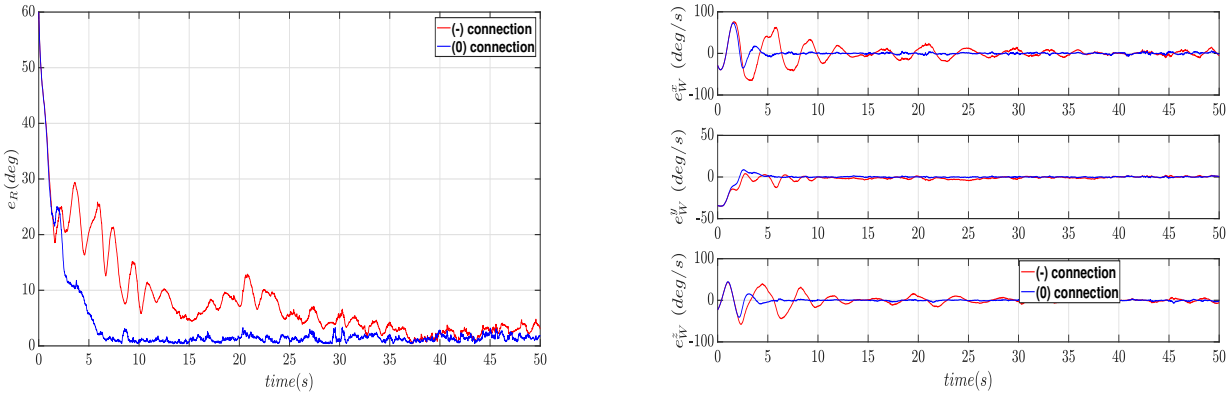


Figure 3-2: Plots for attitude error e_R and angular rate error e_Ω for both connection functions

3-3 Time varying reference directions

Constant reference directions $a_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^{\mathrm{T}}$ and $a_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^{\mathrm{T}}$ represented w.r.t. the L.V.L.H. frame declare that the direction of the sun is always perpendicular to the orbital plane and also the geomagnetic field is tangent on the orbit which is not the case in practice. This also implies that the position information is known from a separate source. However, in general both unit vectors vary with the orbital position $\theta(t)$. We can write the reference directions as vector functions of the orbital position. This also implies that in this case, the L.V.L.H. frame is considered not as reference but as a frame which is rotated w.r.t. the E.C.I. frame. So, both $a_{i=1,2}$ are considered time varying representations w.r.t. the E.C.I. frame encoding also the position information. The radius of the orbit is considered to be known and constant. The model error is simply simulated as Gaussian white noise, in order to declare a model error which we are not interested to determine.

Both the reference directions are vector functions of the position

$$a_1(\theta(t)) = \alpha_1(t) = \begin{bmatrix} \cos(\theta(t)) & \sin(\theta(t)) & 0 \end{bmatrix}$$

$$a_1(\theta(t)) = \alpha_2(t) = \begin{bmatrix} \sin(\theta(t)) & -\cos(\theta(t)) & 0 \end{bmatrix}$$

with $\theta(t) = \frac{2\pi}{5}t$ and are chosen such that

$$\langle \alpha_1(t), \alpha_2(t) \rangle = 0, \quad \forall t > 0 \tag{3-7}$$

Figure 3-3 shows the performance of the filter when both directions vary with time. Decreasing the measurement noise the filter presents a more smooth orientation estimation error which however is affected from the model uncertainty.

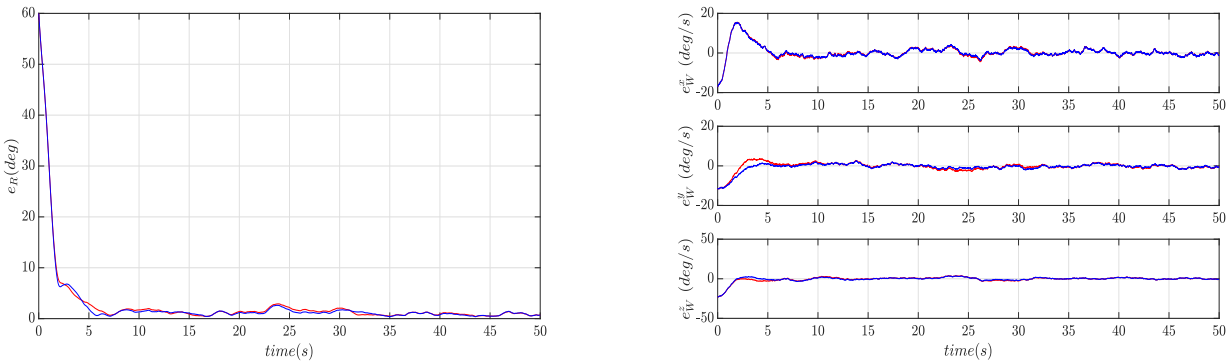


Figure 3-3: Attitude error e_R and angular rate error e_Ω for (0) and (-)-connection and time varying directions. Initial deviations: $(R_0, \Omega_0) = ((60, [1\ 4\ 1]^T), [.3\ .2\ .4]^T)$. Model error standard deviation $\sigma_m = 0.1$, measurement noise standard deviation $\sigma_\epsilon = 0.01$

3-4 Error dynamics for the second order optimal minimum energy filter

For deriving the error dynamics we consider the left invariant group error for the orientation

$$R_0 = \hat{R}^T R \in \mathbb{SO}(3) \tag{3-8}$$

The above error function replaces the commonly used error function $\hat{x} - x$ in observer theory when $\hat{x}, x$ are elements of a vector space. Notice that

$$R_0 = I_{3\times 3} \Leftrightarrow \hat{R} = R \tag{3-9}$$

By differentiating (3-8) we obtain

$$\begin{aligned} \dot{R}_0 &= \dot{\hat{R}}^T R + \hat{R}^T \dot{R} \\ &= -(\hat{\Omega} + \mathbb{K}_{11} r_R)^\times \hat{R}^T R + \hat{R}^T R (\Omega)^\times \\ &= -(\hat{\Omega} + \mathbb{K}_{11} r_R)^\times R_0 + R_0 \Omega^\times \end{aligned} \tag{3-10}$$

or

$$\begin{aligned} R_0^T \dot{R}_0 &= -R_0^T (\hat{\Omega} + \mathbb{K}_{11} r_R)^\times R_0 + \Omega^\times \\ &\overset{(!)}{=} (\Omega - R_0^T (\hat{\Omega} + \mathbb{K}_{11} r_R))^\times \\ &= (\Omega - R_0^T \hat{\Omega} + R_0^T \mathbb{K}_{11} r_R)^\times \end{aligned} \tag{3-11}$$

where (!) holds because $X^T Y^\times X = (X^T Y)^\times$ with $X \in \mathbb{SO}(3)$ and $Y^\times \in \mathfrak{so}(3)$.

Setting

$$\Omega_0 = \Omega - R_0^T \hat{\Omega} \quad , \tag{3-12}$$

the orientation's estimation error dynamics read

$$R_0^T \dot{R}_0 = (\Omega_0 + R_0^T \mathbb{K}_{11} r_R)^\times \tag{3-13}$$

Please notice that (2-23) for the angular rate of the rigid body, has this particular form because the angular rate is represented w.r.t. the body fixed frame. At this point it might be helpful to think of the observer as a distinct frame of reference. Both R and $\hat{R}$ declare the deviation of the body's and the filter's frame w.r.t. L.V.L.H frame. However, Ω and $\hat{\Omega}$ are representations of the body's and filter's angular rate w.r.t. the body and filter's frame respectively. Therefore, both Ω and $\hat{\Omega}$ refer to different frames. On the other hand, R_0 is the deviation of the filter's frame w.r.t. the body frame. Thereupon, it can be understood that the expression $\Omega_0 = \Omega - R_0^T \hat{\Omega}$ transfers the filter's angular rate from the filter's frame to the body frame and then computes the error in the same frame. There is no meaning of adding representations referred to different frames. Nonetheless, the difference $e_\Omega = \Omega - \hat{\Omega}$ has a meaning since both Ω and $\hat{\Omega}$ are axes of rotation and thus valid representations if considered w.r.t. the inertial frame.

For deriving the angular rate error, we first multiply by the inertia tensor and then apply the time derivative operator in both sides of (3-12) which then reads

$$\begin{aligned}
\mathbb{I}\dot{\Omega}_0 &= \mathbb{I}\dot{\Omega} - \mathbb{I}\frac{d}{dt}(R_0^T \hat{\Omega}) \\
&= (\mathbb{I}\Omega)^\times \Omega + T + \mathbb{I}B\delta - \mathbb{I}(-(\Omega_0 + R_0^T \mathbb{K}_{11} r_R)^\times R_0^T \hat{\Omega} + R_0^T \dot{\hat{\Omega}}) \\
&= (\mathbb{I}\Omega)^\times \Omega + T + \mathbb{I}B\delta - \mathbb{I}(R_0^T \hat{\Omega})(\Omega_0 + R_0^T \mathbb{K}_{11} r_R) - \mathbb{I}R_0^T \dot{\hat{\Omega}} \\
&= ((\mathbb{I}\Omega)^\times - (R_0^T \hat{\Omega})^\times \mathbb{I} - \mathbb{I}(R_0^T \hat{\Omega})^\times)\Omega_0 + T + \mathbb{I}B\delta - (R_0^T \hat{\Omega})^\times \mathbb{I}(R_0^T \hat{\Omega}) - \mathbb{I}(R_0 \hat{\Omega})^\times R_0^T \mathbb{K}_{11} r_R - \mathbb{I}R_0^T \mathbb{I}^{-1} \mathbb{I}\dot{\hat{\Omega}}
\end{aligned}$$

By setting

$$\mathbb{H}^\times = (\mathbb{I}\Omega)^\times - (R_0^T \hat{\Omega})^\times \mathbb{I} - \mathbb{I}(R_0^T \hat{\Omega})^\times$$

it can be verified that $\mathbb{H}^\times \in \mathfrak{so}(3)$. Furthermore,

$$\mathbb{I}\dot{\hat{\Omega}} = (\mathbb{I}\hat{\Omega})^\times \hat{\Omega} + T + \mathbb{I}\mathbb{K}_{21} r_R$$

and therefore

$$\begin{aligned}
\mathbb{I}\dot{\Omega}_0 &= \mathbb{H}^\times \Omega_0 + T + \mathbb{I}B\delta - (R_0^T \hat{\Omega})^\times \mathbb{I}(R_0^T \hat{\Omega}) - \mathbb{I}(R_0 \hat{\Omega})^\times R_0^T \mathbb{K}_{11} r_R - \mathbb{I}R_0^T \mathbb{I}^{-1}((\mathbb{I}\hat{\Omega})^\times \hat{\Omega} + T + \mathbb{I}\mathbb{K}_{21} r_R) \\
&= \mathbb{H}^\times \Omega_0 + \underbrace{(I_{3\times 3} - \mathbb{I}R_0^T \mathbb{I}^{-1})T + \mathbb{I}B\delta - \mathbb{I}((R_0 \hat{\Omega})^\times R_0^T \mathbb{K}_{11} + R_0^T \mathbb{K}_{21}) r_R + R_0^T (XY + YX)\mathbb{I}\hat{\Omega}}_{\gamma(R_0, \hat{\Omega}, y)}
\end{aligned} \tag{3-14}$$

where, $X = \hat{\Omega}^\times$ and $Y = R_0 \mathbb{I} R_0^T \mathbb{I}^{-1}$.

Finally, the angular rate estimation error can be written as

$$\dot{\Omega}_0 = \mathbb{I}^{-1}(\mathbb{H}^\times \Omega_0 + \gamma(R_0, \hat{\Omega}, y)) \tag{3-15}$$

where $\gamma : \mathbb{SO}(3) \times \mathfrak{so}(3) \times \mathbb{R}^3 \to \mathbb{R}^3$ is independent on Ω_0.

Overall the error dynamics are given by the following structure.

$$\begin{aligned} R_0^T \dot{R}_0 &= (\Omega_0 + R_0^T \mathbb{K}_{11} r_R)^\times \\ \dot{\Omega}_0 &= \mathbb{I}^{-1}\big(\mathbb{H}^\times \Omega_0 + \gamma(R_0, \hat{\Omega}, y)\big) \end{aligned} \tag{3-16}$$

3-5 Summary

In this chapter the second-order-optimal minimum energy filter on $\mathbb{TSO}(3)$ [2] is implemented in order to stress the significance of the symplectic integration method mentioned in Section 2-6. The filter presents an excellent performance and it is extremely tolerant to model errors. Although the kinematics essentially express Euler's theorem and therefore cannot be uncertain, the filter acts directly through the kinematics as well. Finally, the error dynamics are derived and can be used towards a stability analysis of the filter.

Chapter 4

Predictive observer for attitude estimation

In this chapter the dual optimal control problem is considered and a predictive filter is derived based on the nonlinear predictive controller of [3]. The filter predicts the necessary model error which drives the attitude estimate towards the true orientation of the spacecraft under the presence of significant model errors. As stated in Section 2-4 a benefit of this approach lies on the fact that the correction is performed through the dynamics while the kinematic equation remains isolated and thus can be integrated geometrically. This was the problem faced when a Kalman-based filter is deployed in conjunction with the algebra of quaternions. The observer is derived in a coordinate free fashion directly on the special orthogonal group.

4-1 Observer's structure

Given the following state space model:

$$\begin{aligned} R^T\dot{R} &= \Omega^\times \\ \dot{\Omega} &= \mathbb{I}^{-1}\left((\Omega\mathbb{I})^\times\Omega + T\right) + B\delta \\ y &= \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} R^T a_1 \\ R^T a_2 \end{bmatrix} + D\epsilon \quad , \end{aligned} \tag{4-1}$$

the goal is to find the value of δ at time t which minimizes the cost

$$J\{\delta(t)\} = \frac{1}{2}(y(t+h) - \hat{y}(t+h))^T Q(y(t+h) - \hat{y}(t+h)) + \frac{1}{2}\delta^T(t)\Sigma\delta(t) \tag{4-2}$$

subject to

$$\begin{aligned} \hat{R}(t)^T\dot{\hat{R}}(t) &= \hat{\Omega}^\times(t) \\ \dot{\hat{\Omega}}(t) &= \mathbb{I}^{-1}\left((\hat{\Omega}(t)\mathbb{I})^\times\hat{\Omega}(t) + T(t)\right) + B\delta(t) \\ \hat{y}(t) &= \begin{bmatrix} \hat{y}_1(t) \\ \hat{y}_2(t) \end{bmatrix} = \begin{bmatrix} \hat{R}^T(t)a_1(t) \\ \hat{R}^T(t)a_2(t) \end{bmatrix} \end{aligned} \tag{4-3}$$

The value of the uncertainty term δ at time t influences the state (R, Ω) at a posterior instant of time $t+h$ and subsequently the same is true for the output since the state-output relation is expressed via a memoryless system. Therefore, it has a meaning to ask for the value of δ at time t that minimizes the weighted norm of the prediction error $y - \hat{y}$ at time $t+h$ for each time instant t. The matrices penalize the prediction error and the correction term are $Q = qI_{3\times3}$ and $\Sigma = \sigma I_{3\times3}$ respectively, with $q, \sigma > 0$.

The estimated output $\hat{y}$ at time $t+h$ can be expanded in time by using Taylor series up the second order[1]derivative as

$$\hat{y}_i(t+h) = \hat{y}_i(t) + \dot{\hat{y}}_i(t)h + \ddot{\hat{y}}_i(t)\frac{h^2}{2} \tag{4-4}$$

However, this approach requires expressions for the time derivatives of $a_{1,2}$ and thus, both directions should have been modeled as known functions of time. This latter requirement though cannot always be sutisfied.

Instead of using time derivatives, expansions from [37, 38] can be utilized. These expansions are based on the concept of relative derivative. Since the second order derivative of the output depends on the control input, each output estimate $\hat{y}_{i=1,2}$ can be written as

$$\hat{y}_i(t+h) \approx \hat{y}_i(t) + z(\hat{R}, \hat{\Omega}, h) + \Lambda(h)\mathbb{W}(\hat{R}, \hat{\Omega})\delta(t) \tag{4-5}$$

where

$$z_i\left(\hat{R}, \hat{\Omega}^{\times}, h\right) = h\mathcal{L}_f^1(\hat{y}_i) + \frac{h^2}{2!}\mathcal{L}_f^2(\hat{y}_i) \tag{4-6}$$

and

$$\Lambda(h) = \frac{h^2}{2}\mathbb{I}_{3\times3}$$

whith h being the selected time step. The $\mathcal{L}_{\phi}^{k=1,2}(\hat{y}_i)$ denotes the kth order Lie derivative of $\hat{y}_i$ w.r.t. the system.

Equation (4-5) declares that the correction term at time t influences the state times a factor of $O(h^2)$. The larger the time step the larger the influence the control term has. This is because the output depends on the orientation only and the correction term δ at time t affects the orientation less compared to the angular velocity since they distant two derivatives.

Please notice how the time-delay dependency between the input and the output has been transformed from a derivative expression into a polynomial expression. Essentially, this expansion expresses the time dependency from the input to the output with polynomial instead of differential terms. In this way, the optimization problem becomes unconstraint which is another benefit of this approach.

After substituting (4-5) in the cost function, the necessary condition for optimality

$$d_{\delta(t)}(J\{\delta(t)\}) = 0$$

reads

[1]The reason why we choose to expand until the 2nd order derivative will become clear later on.

$$\begin{aligned}
\mathbf{d}_{\delta(t)}(J\{\delta(t)\}) &= 0 \\
\mathbf{d}_{\delta(t)}((y(t+h)-\hat{y}(t+h))^T Q(y(t+h)-\hat{y}(t+h))+\delta^T(t)\Sigma\delta(t)) &= 0 \\
\mathbf{d}_{\delta(t)}(-2y^T(t+h)Q\hat{y}(t+h)+\hat{y}^T(t+h)Q\hat{y}(t+h)+\delta^T(t)\Sigma\delta(t)) &= 0 \\
\mathbf{d}_{\delta(t)}(-2y^T(t+h)Q\hat{y}(t+h))+\mathbf{d}_{\delta(t)}(\hat{y}^T(t+h)Q\hat{y}(t+h))+\mathbf{d}_{\delta(t)}(\delta^T(t)\Sigma\delta(t)) &= 0
\end{aligned} \tag{4-7}$$

The last equality of (4-7) has three terms. Since only these at time instant $(t+h)$ depend on the correction $\delta(t)$, for the first one we obtain

$$\begin{aligned}
\mathbf{d}_{\delta(t)}(-2y^T(t+h)Q\hat{y}(t+h)) = \\
\mathbf{d}_{\delta(t)}(-2y^T(t+h)Q(\hat{y}(t)+z\left(\hat{R},\hat{\Omega}^{\times},h\right)+\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})\delta(t))) = \\
\mathbf{d}_{\delta(t)}(-2y^T(t+h)Q(\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})\delta(t))) = \\
-2y^T(t+h)Q(\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega}))
\end{aligned} \tag{4-8}$$

The second term of (4-7) can be written as

$$\begin{aligned}
\mathbf{d}_{\delta(t)}(\hat{y}^T(t+h)Q\hat{y}(t+h)) = \\
\mathbf{d}_{\delta(t)}\{(\hat{y}_i(t)+z_i(\hat{R},\hat{\Omega},h)+\Lambda(h)W_i(\hat{R},\hat{\Omega})\delta)^T Q(\hat{y}_i(t)+z_i(\hat{R},\hat{\Omega},h)+\Lambda(h)W_i(\hat{R},\hat{\Omega})\delta)\} = \\
\mathbf{d}_{\delta(t)}\{\hat{y}^T(t)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})\delta(t)+z^T\left(\hat{R},\hat{\Omega}^{\times},h\right)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})\delta(t)+ \\
\delta^T(t)\mathbb{W}^T(\hat{R},\hat{\Omega})(\Lambda(h))^2Q\mathbb{W}(\hat{R},\hat{\Omega})\delta(t)\} = \\
\hat{y}^T(t)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})+z^T\left(\hat{R},\hat{\Omega}^{\times},h\right)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})+2\delta^T(t)\mathbb{W}^T(\hat{R},\hat{\Omega})(\Lambda(h))^2Q\mathbb{W}(\hat{R},\hat{\Omega})
\end{aligned} \tag{4-9}$$

and the quadratic term w.r.t. $\delta(t)$ is simply given by

$$\mathbf{d}_{\delta(t)}(\delta^T(t)\Sigma\delta(t)) = 2\delta^T(t)\Sigma \tag{4-10}$$

By combining (4-8), (4-9), (4-10) we obtain

$$\begin{aligned}
-2y^T(t+h)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})+ \\
2\delta^T(t)\Sigma+\hat{y}^T(t)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})+ \\
z^T\left(\hat{R},\hat{\Omega}^{\times},h\right)Q\Lambda(h)\mathbb{W}(\hat{R},\hat{\Omega})+ \\
2\delta^T(t)\mathbb{W}^T(\hat{R},\hat{\Omega})(\Lambda(h))^2Q\mathbb{W}(\hat{R},\hat{\Omega}) = 0
\end{aligned} \tag{4-11}$$

Transposition in both parts of the above equation and inversion yield

$$\begin{aligned}
\delta(t) = -\frac{1}{2}\{(\Lambda(h))^2\mathbb{W}^T(\hat{R},\hat{\Omega})Q\mathbb{W}(\hat{R},\hat{\Omega})+\Sigma^T\}^{-1}\left(\mathbb{W}^{\mathbb{T}}(\hat{R},\hat{\Omega})Q\Lambda(h)\right)\cdot \\
\{-y(t+h)+\hat{y}(t)+z\left(\hat{R},\hat{\Omega}^{\times},h\right)\}
\end{aligned} \tag{4-12}$$

Finally, (4-3) and (4-12) define the structure of the predictive observer. The terms $z(\hat{R},\hat{\Omega},h)$ and $\mathbb{W}(\hat{R},\hat{\Omega})$ remain to be determined. In the case where $a_{i=1,2}$ are constant, (4-4) can be used with

$$\begin{aligned}\dot{\hat{y}}_i(t) &= \mathbf{d}_t(\hat{R}^T)a_i \\ &= -\hat{\Omega}^\times \hat{R}^T a_i\end{aligned} \tag{4-13}$$

and

$$\begin{aligned}\ddot{\hat{y}}_i(t) &= -\mathbf{d}_t(\hat{\Omega}^\times \hat{R}^T a_i) \\ &= -\mathbf{d}_t(\hat{\Omega}^\times)\hat{R}^T a_i - \hat{\Omega}^\times \mathbf{d}_t(\hat{R}^T)a_i \\ &= -(\dot{\hat{\Omega}})^\times \hat{R}^T a_i - \hat{\Omega}^\times \dot{\hat{R}}^T a_i \\ &= -(\dot{\hat{\Omega}})^\times \hat{R}^T a_i + (\hat{\Omega}^\times)^2 \hat{R}^T a_i \\ &\overset{(!)}{=} (\hat{R}^T a_i)^\times \dot{\hat{\Omega}} + (\hat{\Omega}^\times)^2 \hat{R}^T a_i \\ &= (\hat{R}^T a_i)^\times \dot{\hat{\Omega}} + (\hat{\Omega}^\times)^2 \hat{R}^T a_i \\ &= (\hat{R}^T a_i)^\times \mathbb{I}^{-1}\left((\hat{\Omega}\mathbb{I})^\times \hat{\Omega} + T + \mathbb{I}B\delta\right) + (\hat{\Omega}^\times)^2 \hat{R}^T a_i\end{aligned} \tag{4-14}$$

where (!) is because in general for two vectors A, $B \in \mathbb{R}^3$

$$(A)^\times B = A \times B = -B \times A = -(B)^\times A$$

and therefore

$$\begin{aligned}\hat{y}_i(t+h) &= \hat{y}_i(t) + (-\hat{\Omega}^\times \hat{R}^T a_i)h + ((\hat{R}^T a_i)^\times \mathbb{I}^{-1}\left((\hat{\Omega}\mathbb{I})^\times \hat{\Omega} + T + \mathbb{I}B\delta\right) + (\hat{\Omega}^\times)^2 \hat{R}^T a_i)\frac{h^2}{2} \\ &= \hat{y}_i(t) + \underbrace{(-\hat{\Omega}^\times \hat{R}^T a_i)h + \left\{(\hat{R}^T a_i)^\times \mathbb{I}^{-1}\left((\hat{\Omega}\mathbb{I})^\times \hat{\Omega} + T\right) + (\hat{\Omega}^\times)^2 \hat{R}^T a_i\right\}\frac{h^2}{2}}_{z_i(\hat{R},\hat{\Omega},h)} + \frac{h^2}{2}\underbrace{(\hat{R}^T a_i)^\times B}_{\mathbb{W}_i(\hat{R},\hat{\Omega})}\delta \\ &= \hat{y}_i(t) + z_i(\hat{R},\hat{\Omega},h) + \Lambda(h)\mathbb{W}_i(\hat{R},\hat{\Omega})\delta\end{aligned}$$

For the more general case where both inertial directions vary with time, the expansion from (4-6) should be employed.

4-2 Lie derivatives in $\mathbb{SO}(3)$

In this Section it is presented another way to derive the correction term and it is used while assuming reference directions that vary with time and cannot be modeled by specific known functions. This derivation does not process the inertial directions.

The first term of (4-6) reads

$$\begin{aligned}\mathcal{L}^1_\phi(\hat{y}_i) &= (\mathbf{d}_{(\hat{R},\hat{\Omega})}\hat{y}_i)f \\ &= (\mathbf{d}_{(\hat{R},\hat{\Omega})}\hat{R}^T a_i)f \\ &= \begin{bmatrix} \mathbf{d}_{\hat{R}}\hat{R}^T a_i & \mathbf{d}_{\hat{\Omega}}\hat{R}^T a_i \end{bmatrix}\begin{bmatrix} \hat{\Omega} \\ \mathbb{I}^{-1}\left((\hat{\Omega}\mathbb{I})^\times \hat{\Omega} + T\right) + B\delta \end{bmatrix}\end{aligned} \tag{4-15}$$

where $\hat{\Omega} \in \mathbb{R}^3$. The first term can be written as

$$\begin{aligned} \mathbf{d}_{\hat{R}} \hat{R}^T a_i &= \mathbf{d}_{\hat{R}} (\hat{R}^T a_i)(\delta \hat{R}) = \mathbf{d}_{\hat{R}} (\hat{R} \delta \hat{R})^T a_i \\ &= \frac{\partial}{\partial \hat{\Omega}} (e^{-\hat{\Omega}^\times} \hat{R}^T a_i)|_{\hat{\Omega}=0} = (\hat{R}^T a_i)^\times = \hat{y}_i^\times \end{aligned} \tag{4-16}$$

The latter results by expanding the exponential matrix, and from the fact that with $X = x^\times \in \mathfrak{so}(3)$

$$\begin{aligned} \mathbf{d}_X (X^T a) &= \mathbf{d}_x ((x^\times)^T a) \\ &= \mathbf{d}_x (-x^\times a) \\ &= \mathbf{d}_x (a \times x) \\ &= \mathbf{d}_x (a^\times x) \\ &= a^\times \in \mathfrak{so}(3) \end{aligned} \tag{4-17}$$

Finally, since $\mathbf{d}_{\hat{\Omega}} \hat{R}^T a_i = 0$ we obtain

$$\mathcal{L}_\phi^1 (\hat{y}_i) = (\hat{R}^T a_i)^\times \hat{\Omega} = -\hat{\Omega}^\times (\hat{R}^T a_i) + 0 \tag{4-18}$$

Furthermore, the second-order Lie derivative of (4-6) reads

$$\begin{aligned} \mathcal{L}_\phi^2 (\hat{y}_i) &= \mathcal{L}^1 (-\hat{\Omega}^\times (\hat{R}^T a_i)) \\ &= - \begin{bmatrix} \mathbf{d}_{\hat{R}} (\hat{\Omega}^\times (\hat{R}^T a_i)) & \mathbf{d}_{\hat{\Omega}} (\hat{\Omega}^\times (\hat{R}^T a_i)) \end{bmatrix} \begin{bmatrix} \hat{\Omega} \\ \mathbb{I}^{-1} \left((\hat{\Omega} \mathbb{I})^\times \hat{\Omega} + T \right) + B\delta \end{bmatrix} \\ &= - \begin{bmatrix} \hat{\Omega}^\times (\hat{R}^T a_i)^\times & -(\hat{R}^T a_i)^\times \end{bmatrix} \begin{bmatrix} \hat{\Omega} \\ \mathbb{I}^{-1} \left((\hat{\Omega} \mathbb{I})^\times \hat{\Omega} + T \right) + B\delta \end{bmatrix} \\ &= -\hat{\Omega}^\times (\hat{R}^T a_i)^\times \hat{\Omega} + (\hat{R}^T a_i)^\times \left(\mathbb{I}^{-1} \left((\hat{\Omega} \mathbb{I})^\times \hat{\Omega} + T \right) + B\delta \right) \\ &= (\hat{\Omega}^\times)^2 (\hat{R}^T a_i) + (\hat{R}^T a_i)^\times \left(\mathbb{I}^{-1} \left((\hat{\Omega} \mathbb{I})^\times \hat{\Omega} + T \right) + B\delta \right) \end{aligned} \tag{4-19}$$

where the term $\mathbf{d}_{\hat{R}} (\hat{\Omega}^\times (\hat{R}^T a_i))$ is computed using the product rule:

$$\begin{aligned} \mathbf{d}_{\hat{R}} (\hat{\Omega}^\times (\hat{R}^T a_i)) &= \hat{\Omega}^\times \frac{\partial}{\partial \hat{\Omega}} (\hat{R}^T a_i)|_{\hat{\Omega}=0} \\ &= \hat{\Omega}^\times \Big(\frac{\partial}{\partial \hat{\Omega}} (e^{-\hat{\Omega}^\times}) \hat{R}^T a_i \Big)|_{\hat{\Omega}=0} \\ &= \hat{\Omega}^\times (\hat{R}^T a_i)^\times \end{aligned} \tag{4-20}$$

Notice that the optimal correction term makes the observer a non-causal system since δ at time t uses the value of the measured output at the next time instant which of course is not known a priori. By reducing the integration step h we can manage to have $y(t+h) \to y(t)$. However, as (4-5) indicates, by decreasing h, the influence of the correction term weakens for constant penalty matrices Q and Σ.

4-3 The covariant constraint

In the present method the optimal state trajectory estimate is determined on the basis of the assumption that consistent estimates of the state trajectories must match the available measurements with a residual error covariance which is approximately equal to the known measurement error covariance [30]. In other words the uncertainty of the estimation error between the estimated outputs and the measured observations cannot be less than the covariance of the measurement noise as shown in (4-21).

$$\mathbb{E}\left\{(\hat{y}(t)-y(t))(\hat{y}(t)-y(t))^T\right\} \geq \mathbb{E}\left\{\epsilon(t)\epsilon(t)^T\right\} \tag{4-21}$$

This constraint has to be imposed next by the dual optimal control problem in order the estimates to be statistically consistent. Furthermore, in order to achieve optimal and statistically consistent estimates the penalty matrix Σ has to be chosen such that the following equality approximately holds:

$$\mathbb{E}\left\{(\hat{y}(t)-y(t))(\hat{y}(t)-y(t))^T\right\} \approx \sigma_\epsilon I_{6\times 6} \tag{4-22}$$

The estimated outputs should fit the measurements with approximately the same uncertainty as the actual measurements fit the truth (without noise) output. While $\delta(t)$ may in fact be large, it should be as small as possible. The presence of $||\delta||_\Sigma$ in the cost functional J produces somewhat ambivalent results. If the original state model equations contain significant errors, then the addition of large model correction terms should enable the estimate to better fit the measurements. Thus, the first term in the cost J is decreased. However, the addition of the weighted norm for $\delta(t)$ overall increases cost J. The proper balance between the two competing effects depends on the choice of Σ.

The matrix Σ is determined in order the covariance constraint to be satisfied. If Σ is decreased below this value, the estimate matches the measurements too closely and becomes too affected by the measurement noise. If Σ is too large, the estimate is too far from the measurements. If Σ is too small, the estimate is based more on the measurements. From (4-22) it can be seen that the proper value for Σ is decided only after the output estimates have been obtained.

4-4 Simulations

Various simulation scenarios are set for testing the observer's performance. Both constant and time varying inertial directions $a_{i=1,2}$ are considered with different initial states. The attitude estimation error is given from the following function

$$e_R(t) = \cos^{-1}\left(1 - \frac{\mathrm{tr}(I - R^T(t)\hat{R}(t))}{2}\right) \tag{4-23}$$

and the angular rate estimation error is calculated as $e_\Omega = \Omega(t) - \hat{\Omega}(t)$. For representing the initial orientation the principal rotation vector notation (θ, K) is used. From the principal rotation vector representation we can derive the orientation matrix $R \in \mathbb{SO}(3)$ by using the Rodrigues formula [39]:

$$R = I + (\sin(\theta))K^\times + \frac{1-\cos^2(\theta)}{2}(K^\times)^2 \tag{4-24}$$

The control torque is considered to be $T(t) = \begin{bmatrix} \sin\left(\frac{2\pi}{10}\right) & \cos\left(\frac{2\pi}{2}\right) & \sin\left(\frac{2\pi}{50}\right) \end{bmatrix}^T$ and exerts in a rigid body with inertia tensor $\mathbb{I} = \text{diag}(2, 5, 3)$ Kg m^2 w.r.t. the principal axes.The inertial directions $a_1 = \begin{bmatrix} 0 & 0 & 1 \end{bmatrix}^T$ and $a_2 = \begin{bmatrix} 0 & 1 & 0 \end{bmatrix}^T$ are used.

Initially we perform simulations for various initial conditions (R_0, Ω_0) using a fixed time step $h = 0.1$. The prediction and uncertainty error are penalized with $Q = 150 I_{6\times 6}$ and $\Sigma = 0.001 I_{3\times 3}$ respectively. The measured output is obtained by adding zero mean Gaussian white noise with standard deviation $\sigma_\epsilon = 0.1$ m. The model error is considerred to be zero. The results are shown in Figure 4-1

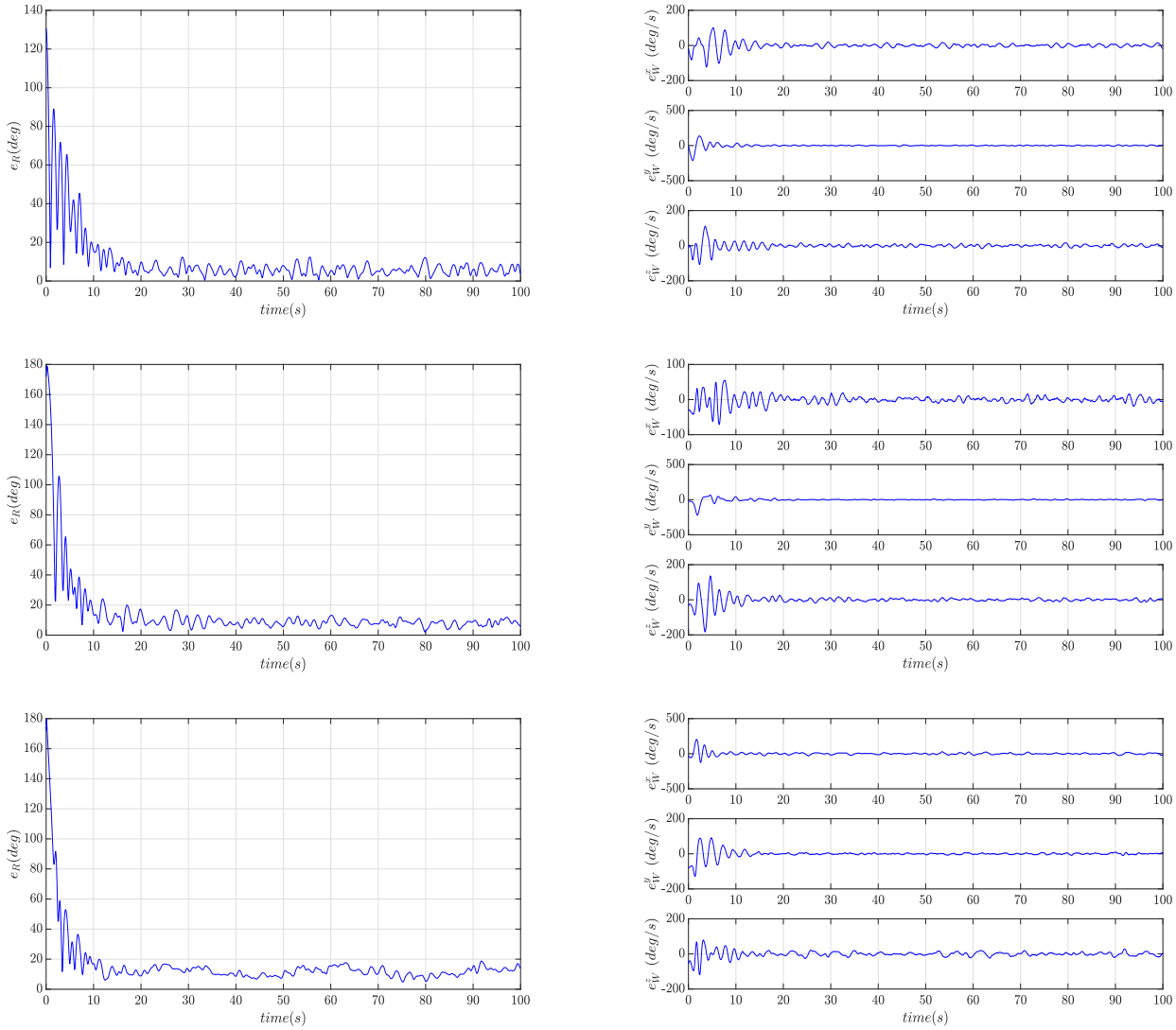


Figure 4-1: Performance of predictive observer for the three different initial conditions (R_0, Ω_0). The top, middle, and bottom figures correspond to $((4, [1\ \ 4\ \ 1]^T), [.3\ \ .1\ \ .2]^T)$, $((3, [.1\ 4\ 10]^T), [.3\ \ .1\ \ .2]^T)$ and $((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$ respectively.

The observer reduces the estimation error starting from large initial deviations both for the angular rate and the orientation. However, does not behave well in steady state since the even the rate estimation error is not even close to zero. Essentially, this means that measuring the orientation error in degrees is misleading since the axes of rotations of the system and the estimator are not aligned. One idea is to increase the penalty $Q = qI_{6\times 6}$ while abandoning the considered model more by also decreasing $\Sigma = \sigma I_{3\times 3}$. The results for $Q = 50I_{6\times 6}$, $\Sigma = 0.0001I_{3\times 3}$ and $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$are shown in Figure 4-2.

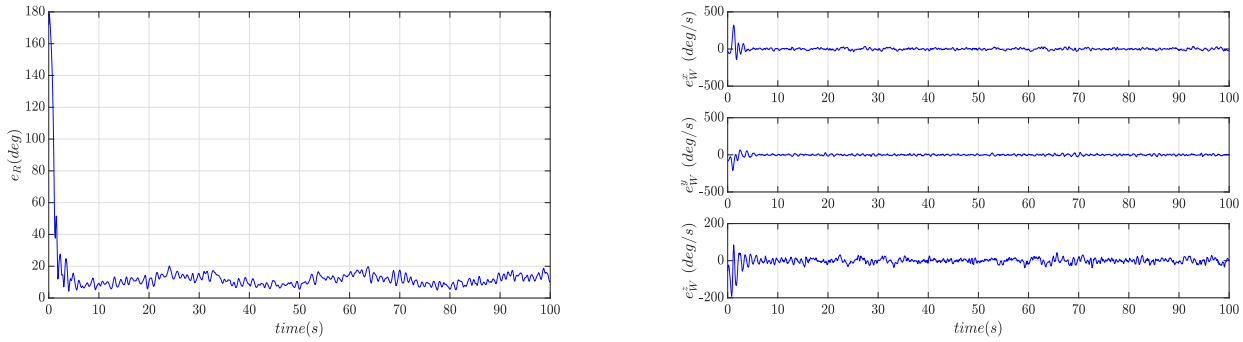


Figure 4-2: Performance of predictive observer for $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$ and penalties $Q = 50I_{6\times 6}$, $\Sigma = 0.0001I_{3\times 3}$.

We observe that the estimation error converges faster but the steady state error oscillates heavily around 10 deg. The results become even worse by increasing Q since then the estimated outputs are forced to follow in more detail the noisy measurements as discussed in Section 4-3.

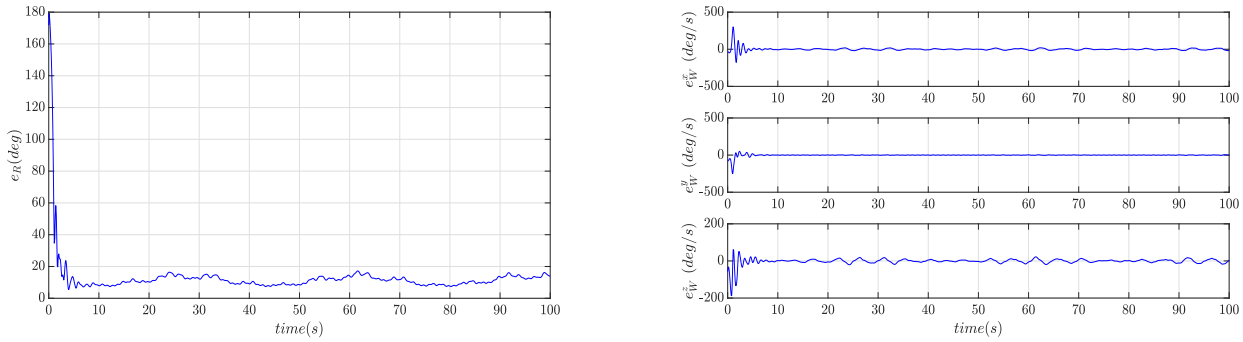


Figure 4-3: Performance of predictive observer for $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$ and penalties $Q = 50I_{6\times 6}$, $\Sigma = 0.0001I_{3\times 3}$.

At this point it is worth to be mentioned that by decreasing the horizon step h, the influence of the correction term δ is decreased for constant penalty matrices. However, by decreasing the time horizon and also tuning the penalty matrices we can achieve the result that is shown in Figure 4-4. Simulations reveal that for any given estimation error accuracy we can find time step h and penalty matrices Q and Σ that achieve the given error accuracy in an also statistical valid way.

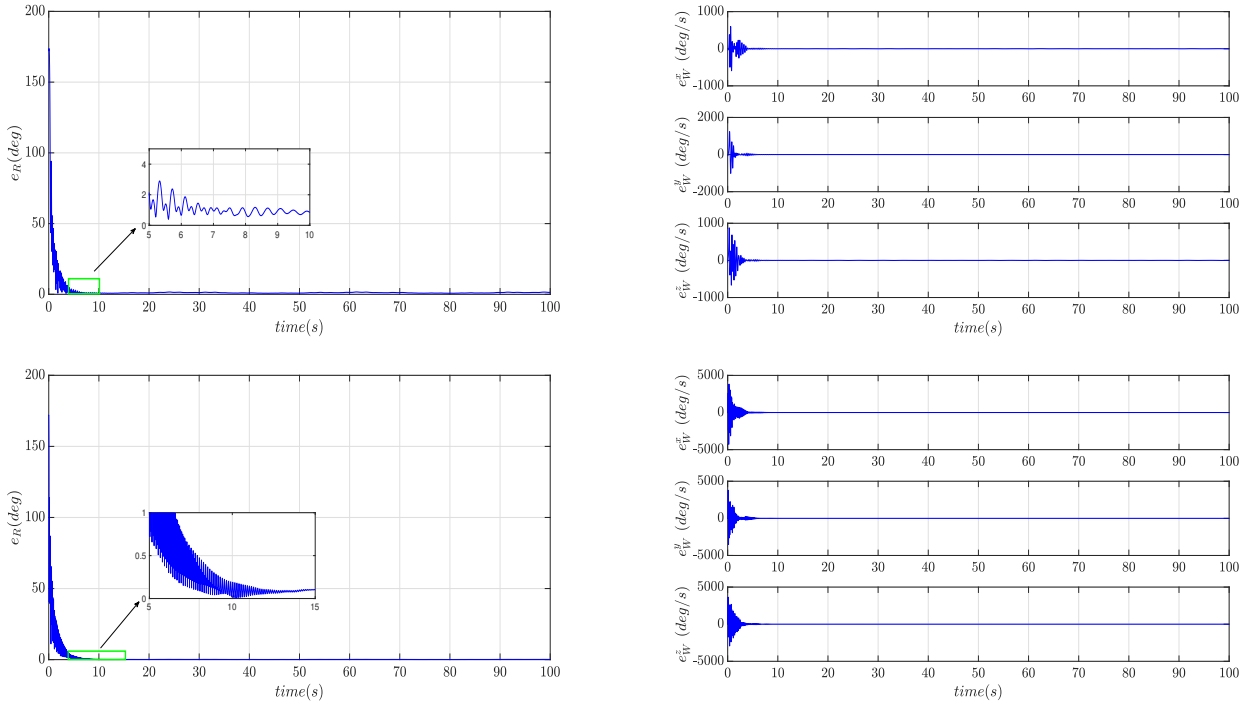


Figure 4-4: Performance of predictive observer for $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$, $h = 0.01$ and penalties $Q = 10^3 I_{6\times6}$, $\Sigma = 10^{-5} I_{3\times3}$ (top), and for $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$, $h = 0.001$ and penalties $Q = 10^5 I_{6\times6}$, $\Sigma = 10^{-6} I_{3\times3}$ (bottom).

Finally, Figure 4-5 shows the result when the observer is tested in the presence of model error and measurement noise with the time step decreased. In this case, the model error is simply simulated as the deterministic signal $\delta(t) = \begin{bmatrix} 0.9\cos\frac{2\pi}{5} & 0 & 0 \end{bmatrix}^T$

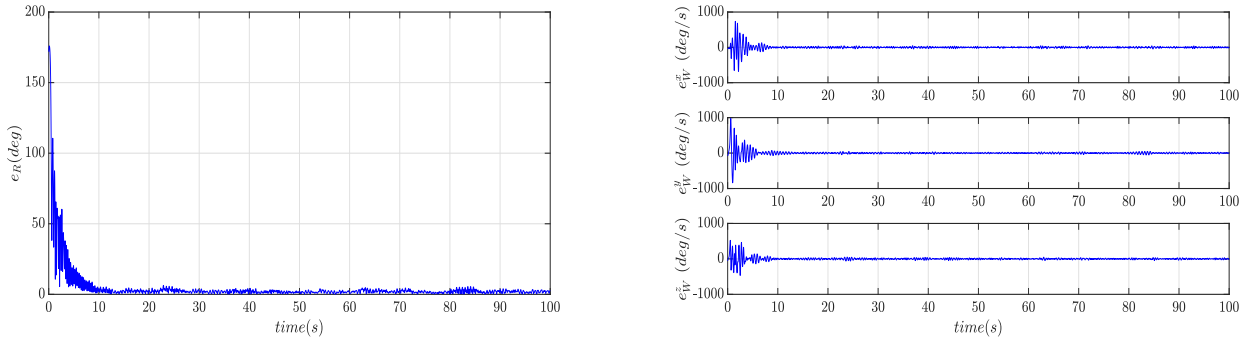


Figure 4-5: Performance of predictive observer for $(R_0, \Omega_0) = ((3, [.1\ 4\ 10]^T), [.5\ 1.5\ 1]^T)$, $h = 0.01$ and penalties $Q = 10^3 I_{6\times6}$, $\Sigma = 10^{-4} I_{3\times3}$. The measurement noise is set to have standard deviation $\sigma_\epsilon = 0.1$ and the model error is equal to δ.

From Figures 4-4 and 4-5 we can conclude the following: By decreasing the horizon step h the observer seems to perform better. As it was mentioned earlier, the influence of the correction term δ is affected by the size of the horizon step. By choosing a large step, the observer is

influenced more from the correction term δ. However, a smaller step allows to reduce the penalty Σ more or equivalently to abandon the considered model more by allowing for larger model errors.

4-5 Time varying reference directions

By representing both directions in the L.V.L.H. frame we also implicitly assume knowledge about the orbital position $\theta(t)$. In an attempt to incorporate the information about the orbital position, both directions should vary with $\theta(t)$. This also implies that in this case, the L.V.L.H. frame is considered not as reference but as a frame which is rotated w.r.t. the E.C.I. frame. The later frame is then the inertial one. The radius of the orbit is considered to be known and almost constant.

Both reference directions vary with time

$$a_1(t) = \begin{bmatrix} \cos\left(\frac{2\pi}{5}t\right) & \sin\left(\frac{2\pi}{5}t\right) & 0 \end{bmatrix}$$

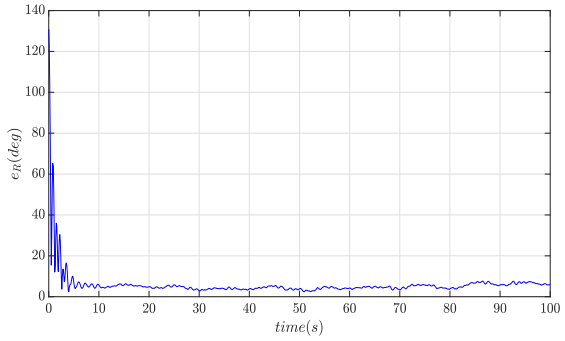
$$a_2(t) = \begin{bmatrix} \sin\left(\frac{2\pi}{5}t\right) & -\cos\left(\frac{2\pi}{5}t\right) & 0 \end{bmatrix}$$

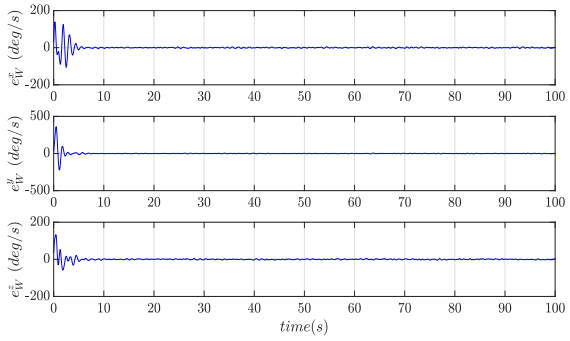
and are chosen such that

$$\langle a_1(t), a_2(t) \rangle = 0, \quad \forall t > 0 \tag{4-25}$$

In this case we assume that the sensors used, introduce measurement noise with small (positive definite) covariance matrix, $\sigma_\epsilon = 0.05$ m. In this section the predictive nonlinear filter is tested w.r.t. to its tolerance in model errors and with the reference directions to vary with time as mentioned earlier.

A model uncertainty $\delta(t) = \begin{bmatrix} \cos\frac{2\pi}{5} & 0 & 0 \end{bmatrix}^T$ drives the actual system. We expect that the filter will show a good performance by determining, through the estimation process, the optimal $\delta^*(t)$ that drives the estimates. Furthermore, in order to stress the significancy of the time step, we simulate first for $h = 0.1$ and $Q = 1000 I_{6\times6}$, $\Sigma = 0.001 I_{3\times3}$ and subsequently for $h = 0.01$, $Q = 10^5 I_{6\times6}$ and $\Sigma = 10^{-5} I_{3\times3}$. The simulations involve different initial conditions (R_0, Ω_0). The results for $h = 0.1$ are given in Figure 4-6.





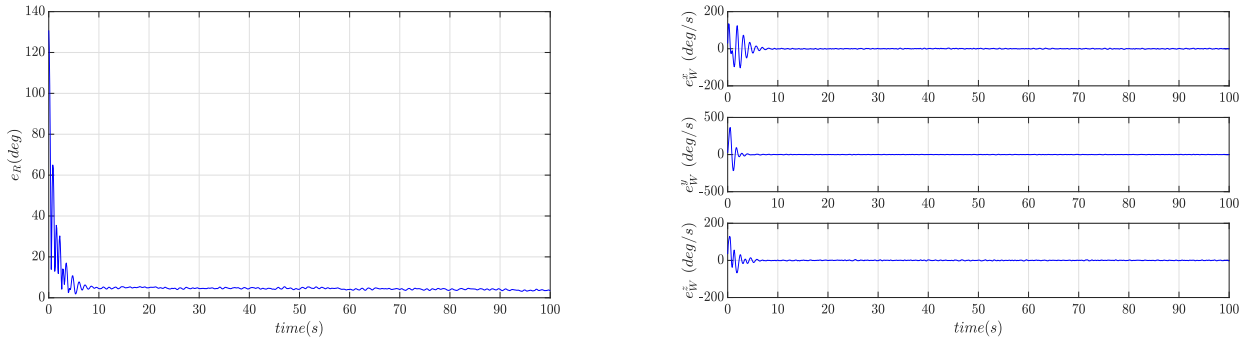


Figure 4-6: Attitude and angular rate error of the predictive observer for different model uncertainties and time varying reference directions. Initial condition $(R_0, \Omega_0) = ((4, [1\ 4\ 1]^T), [.4\ .5\ .6]^T)$, and horizon step $h = 0.1$ sec. Inertia tensor $\mathbb{I} = \text{diag}(11\ 18\ 15)$ (top), and $\mathbb{I} = \text{diag}(60\ 80\ 50)$ (bottom) and model uncertainty a model uncertainty $\delta(t)$ rad/sec.

The model uncertainties tested correspond to large model deviations which indicates that the filter presents tolerance in model errors. This is a very important fact since the environmental torques are not known and can deviate notably. The filter's tuning parameters do not seem to depend on the model uncertainty and the initial conditions which is of course desirable. Do they depend however on the inertia tensor. This is reasonable since the inertia tensor determines the amount of nonlinearity which characterizes the motion. A larger penalty q, forces the estimated output to follow in a more detailed way the "more" nonlinear output. This also requires a smaller penalty in the control action δ. Of course a larger inertia tensor results in a more "slow", "not so nonlinear" motion given the fact that the control torque is "small" as the one used in these simulations.

Finally the observer is examined for smaller time steps and for adapted penalties $Q = 10^5 I_{6\times 6}$ and $\Sigma = 10^{-5} I_{3\times 3}$ under the same model errors and for measurement noise with standard deviation $\sigma_\epsilon = 0.01$ m. The results for $h_1 = 0.01$ and $h_2 = 0.001$ are shown in Figure 4-7 and Figure 4-8 respectively.

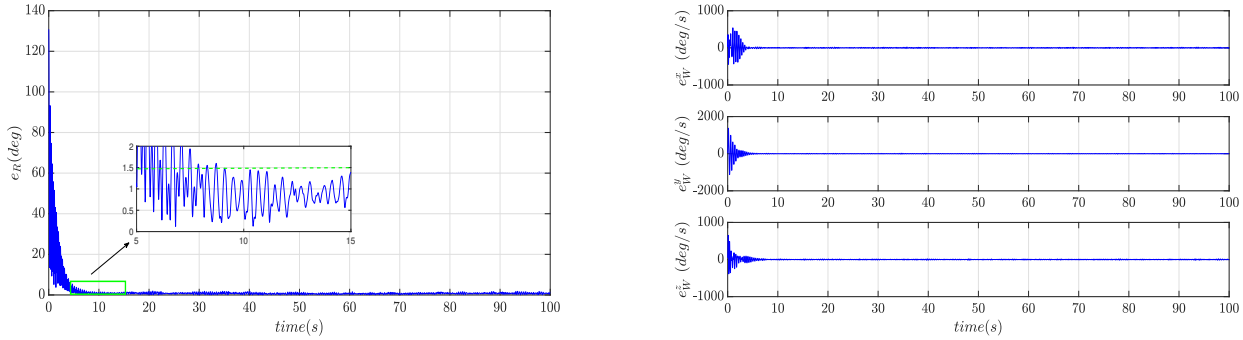


Figure 4-7: Attitude and angular rate error of the predictive observer. Initial condition $(R_0, \Omega_0) = ((4, [1\ 4\ 1]^T), [.4\ .5\ .8]^T)$, and horizon step $h = 0.01$ sec. Model error $\delta(t)$ rad/sec and inertia tensor $\mathbb{I} = \text{diag}(20\ 50\ 30)$.

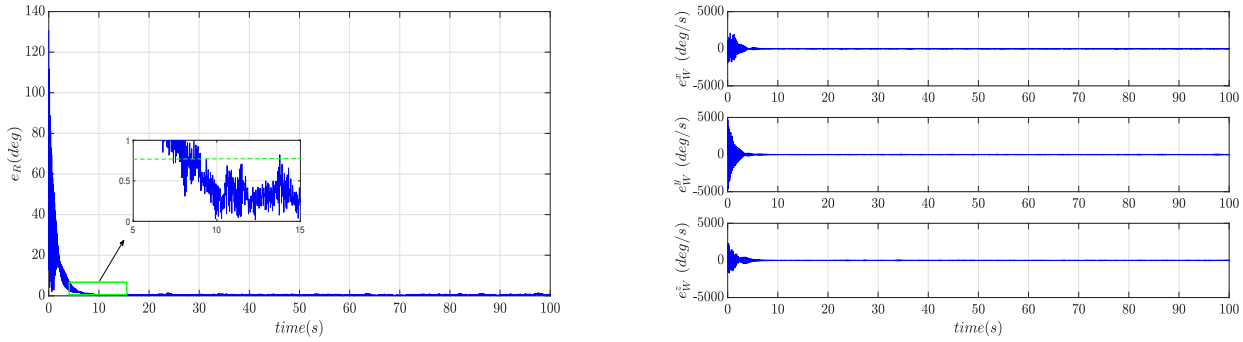


Figure 4-8: Attitude and angular rate error of the predictive observer. Initial condition $(R_0, \Omega_0) = ((4, [1\ 4\ 1]^T), [.4\ .5\ .8]^T)$, and horizon step $h = 0.001$ sec. Model error $\delta(t)$ rad/sec and inertia tensor $\mathbb{I} = \mathrm{diag}(20\ 50\ 30)$.

4-6 Error dynamics for the predictive observer

The error dynamics for the predictive observer are based on the same method followed in Section 3-4. Again, the left invariant group error for the orientation is used

$$R_0 = \hat{R}^T R \in \mathbb{SO}(3) \tag{4-26}$$

Applying the time derivative we obtain

$$R_0^T \dot{R}_0 = (\Omega_0)^\times \tag{4-27}$$

where

$$\Omega_0 = \Omega - R_0^T \hat{\Omega} \tag{4-28}$$

By time differentiation of the latter equation we get

$$\begin{aligned} \mathbb{I}\dot{\Omega}_0 &= \mathbb{H}^\times \Omega_0 + T + \mathbb{I}B\delta - (R_0^T\hat{\Omega})^\times \mathbb{I}(R_0^T\hat{\Omega}) - \mathbb{I}R_0^T\mathbb{I}^{-1}((\mathbb{I}\hat{\Omega})^\times\hat{\Omega} + T + \mathbb{I}\delta^*(\hat{R},\hat{\Omega})) \\ &= \mathbb{H}^\times\Omega_0 + \underbrace{(I_{3\times3} - \mathbb{I}R_0^T\mathbb{I}^{-1})T + \mathbb{I}B\delta - \mathbb{I}R_0^T\delta^*(\hat{R},\hat{\Omega})}_{\gamma(R_0,\hat{\Omega},y)} \end{aligned} \tag{4-29}$$

where $\delta^*(\hat{R},\hat{\Omega})$ is defined in (4-12). The error dynamics for the predictive observer maintain the well known form of the rigid body dynamics. A skew symmetric matrix $\mathbb{H}^\times$- the angular momentum of the error frame- affects the angular rate error Ω_0. In addition, a correction term γ independent on Ω_0 acts as the input torque.

4-7 Comparison between the predictive observer and the second-order-optimal minimum energy filter on $\mathbb{TSO}(3)$

In this section a comparison between the two filters is performed and justifications are made regarding their performance. Setting $\left(\hat{R}^T\dot{\hat{R}}\right)^{-\times} = \dot{\hat{X}}_1 \in \mathbb{R}^3$, and focusing in their structure, both filters can be written as follows:

$$\begin{bmatrix} \dot{\hat{X}}_1 \\ \dot{\hat{X}}_2 \end{bmatrix} = \begin{bmatrix} \hat{X}_2 \\ \phi(\hat{X}_2, \delta) \end{bmatrix} + \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix} \begin{bmatrix} r_1(\hat{X}_1, \hat{X}_2, y) \\ 0 \end{bmatrix} \tag{4-30}$$
$$\dot{\mathbb{K}} = E(\mathbb{K})$$

$$\begin{bmatrix} \dot{\hat{X}}_1 \\ \dot{\hat{X}}_2 \end{bmatrix} = \begin{bmatrix} \hat{X}_2 \\ \phi(\hat{X}_2, \delta) \end{bmatrix} + \begin{bmatrix} 0_{3\times 3} \\ B \end{bmatrix} \delta(\hat{X}_1, \hat{X}_2, y) \tag{4-31}$$

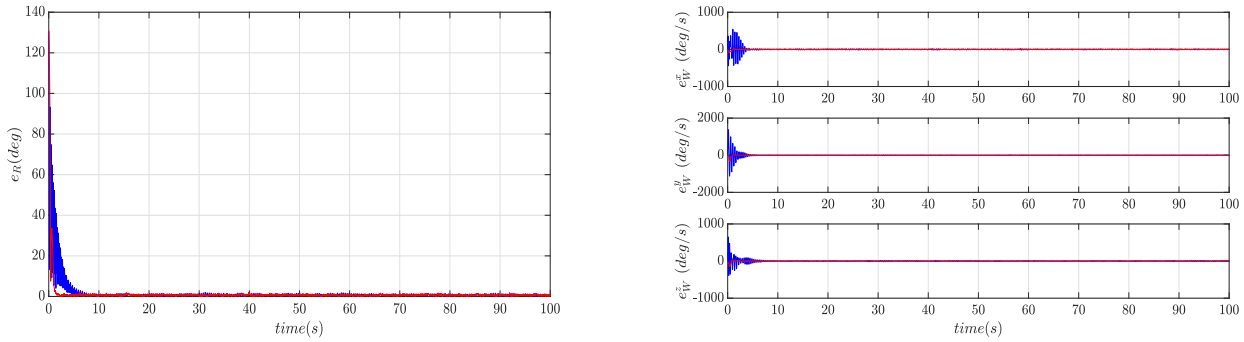


Figure 4-9: Comparison between the second order minimum energy filter (red) and the predictive observer (blue). Tuning parameters: $Q = 1$, $R = 10^{-2}$ (second-order-optimal min. energy filter) and $Q = 10^4$, $\Sigma = 10^{-4}$ predictive filter. Model error $\delta(t) = \left[0.1\cos(\frac{2\pi}{3})\ 0.1\sin(\frac{2\pi}{5})\ 0\right]^T$.

For the comparison, to both filters exerts an input torque $T(t) = \left[\ \sin\left(\frac{2\pi}{3}\right) \quad \cos\left(\frac{2\pi}{1}\right) \quad \sin\left(\frac{2\pi}{5}\right)\ \right]^{\mathrm{T}}$, the reference directions are chosen as

$$a_1(t) = \left[\ \cos\left(\tfrac{2\pi}{5}t\right) \quad \sin\left(\tfrac{2\pi}{5}t\right) \quad 0\ \right]$$
$$a_2(t) = \left[\ \sin\left(\tfrac{2\pi}{5}t\right) \quad -\cos\left(\tfrac{2\pi}{5}t\right) \quad 0\ \right]$$

and the inertia tensor w.r.t. the principal axes is $\mathbb{I} = \mathrm{diag}(20,\ 50,\ 30)$. Both filters are integrated with the same time step $h = 0.01$. The measurement noise is simulated as gaussian white noise sequence with variance $\sigma_\epsilon = 0.05\ m$. The results shown in Figure 4-9 indicate that the two estimators perform equally with the second-order-optimal minimum energy filter to converge faster than the predictive filter.

However, by increasing the measurement noise $\sigma_\epsilon = 0.1$ the minimum energy filter is expected to be superior against the predictive observer. Indeed from Figure 4-10 can be seen that the predictive observer is not tolerant to measurement noise. One way to solve this issue, is to decrease the penalty matrix Σ. The estimated outputs are then not overfitted to the noisy

measurements. However, the value of Σ should be chosen carefully according to the discussion of Section 4-3 in order for a reasonable level of accuracy to be maintained and also for the filter to provide statistically consistent estimates.

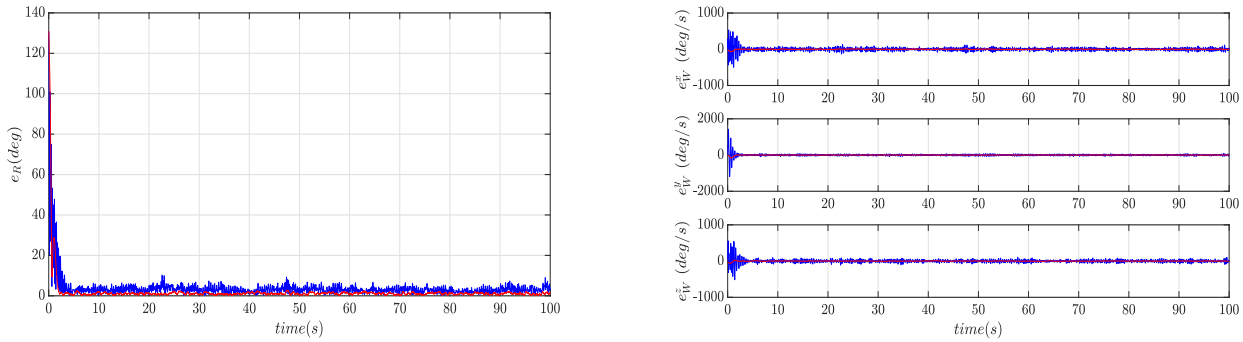


Figure 4-10: Comparison between the second order minimum energy filter (red) and the predictive filter (blue) with increased measurement noise. Tuning parameters: $Q = 1$, $R = 10^{-2}$ (second-order-optimal min. energy filter) and $Q = 10^4$, $\Sigma = 10^{-4}$ predictive filter. Model error $\delta(t) = \left[0.1\cos(\frac{2\pi}{3})\ 0.1\sin(\frac{2\pi}{5})\ 0\right]^T$.

Although in the second order minimum energy filter the measured observations influence linearly the estimates too, the noise effect is reduced by the matrix $\mathbb{K}$ which corresponds to the covariance matrix of the Kalman filter [4]. The matrix $\mathbb{K}$ combines linearly, optimally, at each instant of time, the prior knowledge and the incoming measurements and evolves according to the perturbed Riccati equation. This is the reason why, with increased noisy measurements, the second-order-optimal minimum energy filter performs better than the predictive filter.

The predictive filter performs a correction through the dynamics only. Instead, the minimum energy filter introduces two correction terms, one in the dynamics and a second one in the kinematics. Overall, we can conclude that the filter's performances emerge from their structure as a result of formulating the optimization problem differently.

Under the assumption that the sensors are of good quality and precision, both the predictive filter on $\mathbb{TSO}(3)$ and the minimum energy filter can be considered that perform almost equally as shown in Figure 4-9. The main advantage of both filters is that they are extremely tolerant to model errors as a result of the deterministic filtering approach.

4-8 Summary

In this chapter the predictive filter on $\mathbb{TSO}(3)$ is derived. The filter's performance was assessed both for constant and time varying reference inertial directions. Finally the predictive filter on $\mathbb{TSO}(3)$ is compared with the second-order-optimal minimum energy filter on $\mathbb{TSO}(3)$ [2]. The comparison study shown that both filters are extremely tolerant to model errors. The predictive filter is however, disturbed heavily from the measurement noise compered with the minimum energy filter. Another significant observation is that the predictive filter comes with the covariant constraint which requires a tuning procedure for the model error penalty

matrix. Finally, the error dynamics are also derived and can be used for analyzing the filter w.r.t. its stability properties.

Chapter 5

Geomagnetic field estimation for orbital position estimation

In this chapter the second part of the problem of attitude and rate estimation discussed in Section 2-7 is studied. The method proposed takes advantage of the deterministic filtering approach. The goal is to estimate the orientation of the spacecraft w.r.t. the E.C.I. reference frame. This can be done by employing a kinematic model for the geomagnetic field up to some model error. This model error is considered to be deterministic and unknown within a Hilbert space. The deterministic filtering approach presented in Chapter 4 is appropriate for this task. In order to extract the kinematic model for the geomagnetic field the transport theorem [40] is utilized. In order to estimate the orientation, angular rate and geomagnetic field of the spacecraft, the predictive filter in $\mathbb{TSO}(3)$ derived in Chapter 4 is extended and used. The important characteristic of the deterministic filtering approach is that the model error is determined during the estimation procedure. The main idea is that after estimating the Earth's geomagnetic field we can impose a specific structure on the model uncertainty and extract information about the orbital angular velocity of the spacecraft which when integrated gives the (angular) orbital position. This can be stated after assuming a dipole model for the geomagnetic field. The chapter starts with the problem statement. Subsequently the optimization problem is solved and the extended filter emerges. Subsequently it is proved that under realistic assumptions, the problem can be decomposed into two separate problems where the resulting filter is equivalent with the series interconnection of two predictive filters. As in Chapter 4 the problem is set and solved without using coordinates.

Before stating the problem it is important to present the transport theorem of analytical mechanics.

Theorem 2. *[28] Let A and B be two frames with relative angular velocity vector $\omega_{\mathbf{A/B}}$, and let $\mathbf{r}$ be a vector. The time derivative of $\mathbf{r}$ in the A frame can be related to the time derivative of $\mathbf{r}$ in the B frame as*

$$\frac{d^A}{dt}\mathbf{r} = \frac{d^B}{dt}\mathbf{r} + \omega_{\mathbf{A/B}} \times \mathbf{r} \tag{5-1}$$

Please notice that the theorem does not specify some particular representation for the vector $\mathbf{r}$ and therefore it can be represented arbitrarily either w.r.t. A or w.r.t. B. This note is important in order to form the output equation for the geomagnetic field. The kinematic equation refers to the time derivative of the geomagnetic field as shown from the body frame while it is represented in the inertial frame. Furthermore, since $\omega_{\mathbf{A/B}}$ indicates the instantaneous angular rotation vector of A w.r.t. B its representation is the same for both frames.

Consider a nonlinear dynamic system that can be described by a state space model of the following form:

$$\begin{aligned} R^T(t)\dot{R}(t) &= \Omega^{\times}(t) \\ \dot{\Omega}(t) &= \mathbb{I}^{-1}((\Omega(t)\mathbb{I})^{\times}\Omega(t) + T) + B_1\delta(t) \\ \dot{g}(t) &= \Omega^{\times}(t)g(t) + B_2\sigma(t) \\ y(t) = \begin{bmatrix} y_1(t) \\ y_2(t) \\ y_3(t) \end{bmatrix} &= \begin{bmatrix} R^T(t)a_1(t) \\ R^T(t)a_2(t) \\ R^T(t)g(t) \end{bmatrix} + D\epsilon(t) \end{aligned} \tag{5-2}$$

where $R(t) \in \mathbb{SO}(3)$, $g(t) \in \mathbb{R}^3$ is the geomagnetic field and, $\Omega(t), \delta(t), \sigma(t), \epsilon(t) \in \mathbb{R}^3$. For the output noise, it is considered that there is no inference between the various output channels and also that is white and Gaussian. Therefore $D = \text{diag}(d_1, d_2, d_3)$, $d_i \in \mathbb{R}^{3\times3}$. The transport theorem is used for extracting the kinematic equation for the geomagnetic field g. The orbital radius is considered constant. It is worth to be noticed that both inertial directions $a_{i=1,2}$ are considered known w.r.t. the inertial frame and correspond to the Sun's direction and the direction of (possibly) another celestial body respectively.

As can be seen from Theorem 2 the time variation of the geomagnetic field as seen from the body frame and represented w.r.t. the inertial has two terms. The kinematic equation for the geomagnetic field in (5-2) considers only the second term of (5-1). This variation expresses how the field changes w.r.t. body frame due to the fact that the body frame is instantaneously rotating. The second part refers to how the geomagnetic field changes w.r.t. the inertial frame. This component is unknown and it is modeled as an unknown and deterministic model error. As a first step, we hope that the (deterministic) predictive filtering approach will reveal the (minimum) model error that produces the available noisy measurements.

5-1 Problem statement

The predictive filter in this case emerges from the following optimization problem:

$$\begin{aligned}
\min_{U(t)} \quad & \frac{1}{2}(y(t+h)-\hat{y}(t+h))^T Q(y(t+h)-\hat{y}(t+h)) + \frac{1}{2}U^T(t)\Pi U(t) \\
\text{s.t.} \quad & \hat{R}^T(t)\dot{\hat{R}}(t) = \hat{\Omega}^\times(t) \\
& \dot{\hat{\Omega}}(t) = \mathbb{I}^{-1}((\hat{\Omega}(t)\mathbb{I})^\times \hat{\Omega}(t) + T) + B_1\delta(t) \\
& \dot{\hat{g}}(t) = \hat{\Omega}^\times(t)\hat{g}(t) + B_2\sigma(t) \\
& \hat{y}(t) = \begin{bmatrix} \hat{y}_1(t) \\ \hat{y}_2(t) \\ \hat{y}_3(t) \end{bmatrix} = \begin{bmatrix} \hat{R}^T(t)a_1(t) \\ \hat{R}^T(t)a_2(t) \\ \hat{R}^T(t)\hat{g}(t) \end{bmatrix}
\end{aligned} \tag{5-3}$$

where $Q \in \mathbb{R}^{9\times 9} > 0$, $\Pi \in \mathbb{R}^{6\times 6} > 0$, and $U(t) = \begin{bmatrix} \delta^T(t) & \sigma^T(t) \end{bmatrix}^T$. It is also assumed that $B_{i=1,2} = I_{3\times 3}$.

Each estimated output can be expanded according to [37, 38] as shown in Chapter 4. For $\hat{y}_{i=1,2}$ we have already shown that

$$\begin{aligned}
& \hat{y}_{i=1,2}(t+h) = \hat{y}_{i=1,2}(t) - h\hat{\Omega}^\times \hat{R}(t)^T a_{i=1,2} \\
& + \frac{h^2}{2}\Big\{(\Omega^\times)^2 \hat{R}^T(t)a_{i=1,2} + (\hat{R}^T(t)a_{i=1,2})^\times \Big(\mathbb{I}^{-1}((\mathbb{I}\hat{\Omega})^\times\hat{\Omega} + T) + \delta(t)\Big)\Big\} \\
& = \hat{y}_{i=1,2}(t) - h\hat{\Omega}^\times \hat{R}(t)^T a_{i=1,2} + \\
& \frac{h^2}{2}\Big\{(\Omega^\times)^2 \hat{R}^T(t)a_{i=1,2} + (\hat{R}^T(t)a_{i=1,2})^\times \Big(\mathbb{I}^{-1}((\mathbb{I}\hat{\Omega})^\times\hat{\Omega} + T)\Big\} + \frac{h^2}{2}(\hat{R}^T(t)a_{i=1,2})^\times \delta(t) \\
& = \hat{y}_{i=1,2}(t) + z_{i=1,2}(\hat{R},\hat{\Omega},h) + \Lambda_{i=1,2}(h) w_{i=1,2}(\hat{R},\hat{\Omega})\delta
\end{aligned} \tag{5-4}$$

Deriving the corresponding expression for $\hat{y}_{i=3}(t+h)$ is much easier since the least number of times $y_3(t)$ has to be differentiated for any component of $u(t)$ to appear in its time derivative is equal to one. This is because the geomagnetic field g is involved in $y_3(t)$, and subsequently because it is directly affected by the model error $\sigma(t)$ in $U(t)$. Therefore we have

$$\hat{y}_3(t+h) = \hat{y}_3(t) + h\mathcal{L}(\hat{R}^T(t)\hat{g}(t)) \tag{5-5}$$

where

$$\mathcal{L}(\hat{R}^T(t)\hat{g}(t)) = \begin{bmatrix} \mathbf{d}_{\hat{R}}\hat{R}^T(t)\hat{g}(t) & \mathbf{d}_{\hat{\Omega}}\hat{R}^T(t)\hat{g}(t) & \mathbf{d}_{\hat{g}}\hat{R}^T(t)\hat{g}(t) \end{bmatrix} \begin{bmatrix} \hat{\Omega} \\ \mathbb{I}^{-1}((\hat{\Omega}(t)\mathbb{I})^\times\hat{\Omega}(t) + T) + \delta(t) \\ \hat{\Omega}^\times(t)\hat{g}(t) + \sigma(t) \end{bmatrix}$$

The term $\mathbf{d}_{\hat{R}}\hat{R}^T(t)\hat{g}(t)$ has already been calculated in Section 4-2 and $\mathbf{d}_{\hat{\Omega}}\hat{R}^T(t)\hat{g}(t) = 0$ since the angular velocity is not involved in $y_3(t)$. Finally, $\mathbf{d}_{\hat{g}}\hat{R}^T(t)\hat{g}(t)$ is simply the Jacobian of the vector $\hat{R}^T(t)\hat{g}(t)$ w.r.t. $\hat{g}(t)$ and hence equal to $\hat{R}^T(t)$. Therefore, we have

$$\begin{aligned}\hat{y}_3(t+h) &= \hat{y}_3(t) + h\Big((\hat{R}^T(t)\hat{g}(t))^\times \hat{\Omega}(t) + \hat{R}^T(t)\hat{\Omega}^\times(t)\hat{g}(t)\Big) + h\hat{R}^T(t)\sigma(t) \\ &= \hat{y}_3(t) + z_3(\hat{R},\hat{\Omega},h) + \Lambda_{i=3}(h)w_3(\hat{R},\hat{\Omega})\sigma(t)\end{aligned} \tag{5-6}$$

By denoting

$$\begin{aligned} Y &= \begin{bmatrix} \hat{y}_1^T(t) & \hat{y}_2^T(t) & \hat{y}_3^T(t) \end{bmatrix}^T \\ Z(\hat{R},\hat{\Omega},h) &= \begin{bmatrix} z_1^T(\hat{R},\hat{\Omega},h) & z_2^T(\hat{R},\hat{\Omega},h) & z_3^T(\hat{R},\hat{\Omega},h) \end{bmatrix}^T \\ W(\hat{R},\hat{\Omega},h) &= \begin{bmatrix} w_1(\hat{R},\hat{\Omega},h) & 0_{3\times3} \\ w_2(\hat{R},\hat{\Omega},h) & 0_{3\times3} \\ 0_{3\times3} & w_3(\hat{R},\hat{\Omega},h) \end{bmatrix} \end{aligned}$$

we can write

$$\hat{Y}(t+h) = \hat{Y}(t) + Z(\hat{R},\hat{\Omega},h) + W(\hat{R},\hat{\Omega},h)U(t) \tag{5-7}$$

and therefore, the solution $U^*(t)$ of the optimization problem (5-3) as shown in Section 4-1 can be written as

$$\begin{aligned} U^*(t) = -\frac{1}{2}\Big\{W^T(\hat{R},\hat{\Omega},\hat{g},h)QW(\hat{R},\hat{\Omega},\hat{g},h) + \Pi\Big\}^{-1} W^T(\hat{R},\hat{\Omega},\hat{g},h)Q\cdot \\ \Big(-Y(t) + \hat{Y}(t) + Z(\hat{R},\Omega,\hat{g},h)\Big) \end{aligned} \tag{5-8}$$

where $U^*_{[1:3]}(t) = \delta^*(t)$ and $U^*_{[4:6]}(t) = \sigma^*(t)$.

At this point it is possible to prove that the aforementioned predictive filter can be decomposed in two separate predictive filters connected in series as follows:

$$\begin{aligned} W^TQ &= \begin{bmatrix} w_1^{3\times3} & 0^{3\times3} \\ w_2^{3\times3} & 0^{3\times3} \\ 0^{3\times3} & w_3^{3\times3} \end{bmatrix}^T \begin{bmatrix} Q_{11}^{6\times6} & Q_{12}^{6\times6} \\ Q_{21}^{3\times6} & Q_{22}^{3\times3} \end{bmatrix} = \begin{bmatrix} w_{12}^{6\times3} & 0^{6\times3} \\ 0^{3\times3} & w_3^{3\times3} \end{bmatrix}^T \begin{bmatrix} Q_{11}^{6\times6} & Q_{12}^{6\times3} \\ Q_{21}^{3\times6} & Q_{22}^{3\times3} \end{bmatrix} \\ &= \begin{bmatrix} w_{12}^{T,\ 3\times6}Q_{11}^{6\times6} & w_{12}^{T,\ 3\times6}Q_{12}^{6\times3} \\ w_3^{T,\ 3\times3}Q_{21}^{3\times6} & w_3^{T,\ 3\times3}Q_{22}^{3\times3} \end{bmatrix} \end{aligned} \tag{5-9}$$

where $w_{12}^{6\times3} = \begin{bmatrix} w_1^{3\times3} \\ w_2^{3\times3} \end{bmatrix}$

Therefore, for W^TQW we have

$$\begin{aligned} W^TQW &= \begin{bmatrix} w_{12}^{T,3\times6}Q_{11}^{6\times6} & w_{12}^{T,3\times6}Q_{12}^{6\times3} \\ w_3^{T,3\times3}Q_{21}^{3\times6} & w_3^{T,3\times3}Q_{22}^{3\times3} \end{bmatrix} \begin{bmatrix} w_{12}^{6\times3} & 0^{6\times3} \\ 0^{3\times3} & w_3^{3\times3} \end{bmatrix} \\ &= \begin{bmatrix} w_{12}^{T,3\times6}Q_{11}^{6\times6}w_{12}^{6\times3} & w_{12}^{T,3\times6}Q_{12}^{6\times3}w_3^{3\times3} \\ w_3^{T,3\times3}Q_{21}^{3\times6}w_{12}^{6\times3} & w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3} \end{bmatrix} \end{aligned} \tag{5-10}$$

By using the inversion lemma [41] for $W^TQW+\Pi$ we have:

$$(W^TQW+\Pi)^{-1}W^TQ =$$

$$= \begin{bmatrix} w_{12}^{T,3\times6}Q_{11}^{6\times6}w_{12}^{6\times3}+\Pi_{11}^{3\times3} & w_{12}^{T,3\times6}Q_{12}^{6\times3}w_3^{3\times3}+\Pi_{12}^{3\times3} \\ w_3^{T,3\times3}Q_{21}^{3\times6}w_{12}^{6\times3}+\Pi_{21}^{3\times3} & w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3}+\Pi_{22}^{3\times3} \end{bmatrix}^{-1} \begin{bmatrix} w_{12}^{T,\,3\times6}Q_{11}^{6\times6} & w_{12}^{T,\,3\times6}Q_{12}^{6\times3} \\ w_3^{T,\,3\times3}Q_{21}^{3\times6} & w_3^{T,\,3\times3}Q_{22}^{3\times3} \end{bmatrix}$$

$$= \begin{bmatrix} I^{3\times3} & 0^{3\times3} \\ -(w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3}+\Pi_{22}^{3\times3})^{-1}(w_3^{T,3\times3}Q_{21}^{3\times6}w_{12}^{6\times3}+\Pi_{21}^{3\times3}) & I^{3\times3} \end{bmatrix}.$$

$$\begin{bmatrix} [(w_{12}^{T,3\times6}Q_{11}^{6\times6}w_{12}^{6\times3}+\Pi_{11}^{3\times3})-(w_{12}^{T,3\times6}Q_{12}^{6\times3}w_3^{3\times3}+\Pi_{12}^{3\times3})(w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3}+\Pi_{22}^{3\times3})^{-1}(w_3^{T,3\times3}Q_{21}^{3\times6}w_{12}^{6\times3}+\Pi_{21}^{3\times3})]^{-1} & ..... \\ 0_{3\times3} & ..... \end{bmatrix}$$

$$\begin{bmatrix} ..... & 0^{3\times3} \\ ..... & (w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3}+\Pi_{22}^{3\times3})^{-1} \end{bmatrix} \begin{bmatrix} I^{3\times3} & (w_{12}^{T,3\times6}Q_{12}^{6\times3}w_3^{3\times3}+\Pi_{12}^{3\times3})(w_3^{T,3\times3}Q_{22}^{3\times3}w_3^{3\times3}+\Pi_{22}^{3\times3})^{-1} \\ 0^{3\times3} & I^{3\times3} \end{bmatrix}.$$

$$\begin{bmatrix} w_{12}^{T,3\times6}Q_{11}^{6\times6} & w_{12}^{T,3\times6}Q_{12}^{6\times3} \\ w_3^{T,3\times3}Q_{21}^{3\times6} & w_3^{T,3\times3}Q_{22}^{3\times3} \end{bmatrix} = \mathcal{M}(\hat{R},\hat{\Omega},\hat{g},h) \tag{5-11}$$

Given the above discussion and by denoting $Y_{[1:6]} = \begin{bmatrix} y_1^T(t) & y_2^T(t) \end{bmatrix}^T$, $Y_{[7:9]} = \begin{bmatrix} y_3^T(t) \end{bmatrix}^T$ and $Z_{[1:6]} = \begin{bmatrix} z_1^T(t) & z_2^T(t) \end{bmatrix}^T$, $Z_{[7:9]} = \begin{bmatrix} \hat{z}_3^T(t) \end{bmatrix}^T$ we can write (5-8) as follows:

$$U^*(t) = -\frac{1}{2}\mathcal{M}(\hat{R}, \hat{\Omega}, \hat{g}, h) \begin{bmatrix} -Y_{[1:6]} + \hat{Y}_{[1:6]} + Z_{[1:6]} \\ -Y_{[7:9]} + \hat{Y}_{[7:9]} + Z_{[7:9]} \end{bmatrix} \tag{5-12}$$

By declaring $\hat{X} = (\hat{R}, \hat{\Omega}, \hat{g})$ the resulting filter has the form

$$\dot{\hat{X}} = f(\hat{X}) - \frac{\Gamma}{2}\mathcal{M}(\hat{X}, h)(-Y + \hat{Y} + Z) \tag{5-13}$$

where $\Gamma = \begin{bmatrix} 0^{3\times 6} \\ I^{6\times 6} \end{bmatrix}$. In particular, by writing $\mathcal{M} \in \mathbb{R}^{6\times 9}$ as a block matrix we obtain:

$$\begin{bmatrix} \delta^*(t) \\ \sigma^*(t) \end{bmatrix} = -\frac{1}{2} \begin{bmatrix} \mathcal{M}_{11} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{22} \end{bmatrix} \begin{bmatrix} -Y_{[1:6]} + \hat{Y}_{[1:6]} + Z_{[1:6]} \\ -Y_{[7:9]} + \hat{Y}_{[7:9]} + Z_{[7:9]} \end{bmatrix} \tag{5-14}$$

At this point, from (5-11) it is easy to verify that for $Q_{12} = 0^{6\times 3}$ and $Q_{21} = 0^{3\times 6}$ and Π diagonal, the matrix $\mathcal{M}$ becomes block diagonal with

$$\begin{aligned} \mathcal{M}_{11} &= (w_{12}^{T,3\times 6} Q_{11}^{6\times 6} w_{12}^{6\times 3} + \Pi_{11}^{3\times 3})^{-1} w_{12}^{T,3\times 6} Q_{11}^{6\times 6} \\ \mathcal{M}_{22} &= (w_3^{T,3\times 3} Q_{22}^{3\times 3} w_3^{3\times 3} + \Pi_{22}^{3\times 3})^{-1} w_3^{T,3\times 3} Q_{22}^{3\times 3} \end{aligned} \tag{5-15}$$

and therefore

$$\begin{aligned} \delta^*(t) &= -\frac{1}{2}(w_{12}^{T,3\times 6} Q_{11}^{6\times 6} w_{12}^{6\times 3} + \Pi_{11}^{3\times 3})^{-1} w_{12}^{T,3\times 6} Q_{11}^{6\times 6} (-Y_{[1:6]} + \hat{Y}_{[1:6]} + Z_{[1:6]}) \\ \sigma^*(t) &= -\frac{1}{2}(w_3^{T,3\times 3} Q_{22}^{3\times 3} w_3^{3\times 3} + \Pi_{22}^{3\times 3})^{-1} w_3^{T,3\times 3} Q_{22}^{3\times 3} (-Y_{[7:9]} + \hat{Y}_{[7:9]} + Z_{[7:9]}) \end{aligned} \tag{5-16}$$

By comparing the structure of the above equations with (5-8) and (4-12) we can conclude that the solution to the initial problem (5-3) can be given by taking the solutions of two separate problems with the resulting filters to be interconnected in series as shown in Figure 5-1.
The first problem is the one that was solved in Chapter 4 and the second one is the following:

$$\begin{aligned} \min_{\sigma(t)} \quad & \frac{1}{2}(y_3(t+h) - \hat{y}_3(t+h))^T Q (y(t+h) - \hat{y}_3(t+h)) + \frac{1}{2}\sigma^T(t)\Pi\sigma(t) \\ \text{s.t.} \quad & \dot{\hat{g}}(t) = \hat{\Omega}^{\times}(t)\hat{g}(t) + B_2\sigma(t) \\ & \hat{y}_3(t) = \hat{R}^T\hat{g}(t) \end{aligned} \tag{5-17}$$

with $Q = Q_{22}^{3\times 3}$, $\Pi = \Pi_{22}^{3\times 3}$ and $\hat{R}, \hat{\Omega}$ known.

By considering the penalty matrix Q such that $Q_{12} = 0^{6\times 3}$ and $Q_{21} = 0^{3\times 6}$ and the Π matrix to be diagonal, the prediction error of the first two measurements is independent of the prediction error of the third measurement and also the prediction error for the third measurement does not depend on the error of the first two measurements. Furthermore, it is worth to be noticed that the above decomposition makes the implementation of the above filter easier.

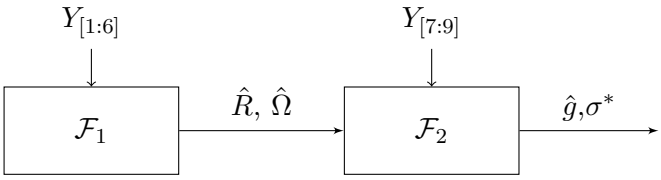
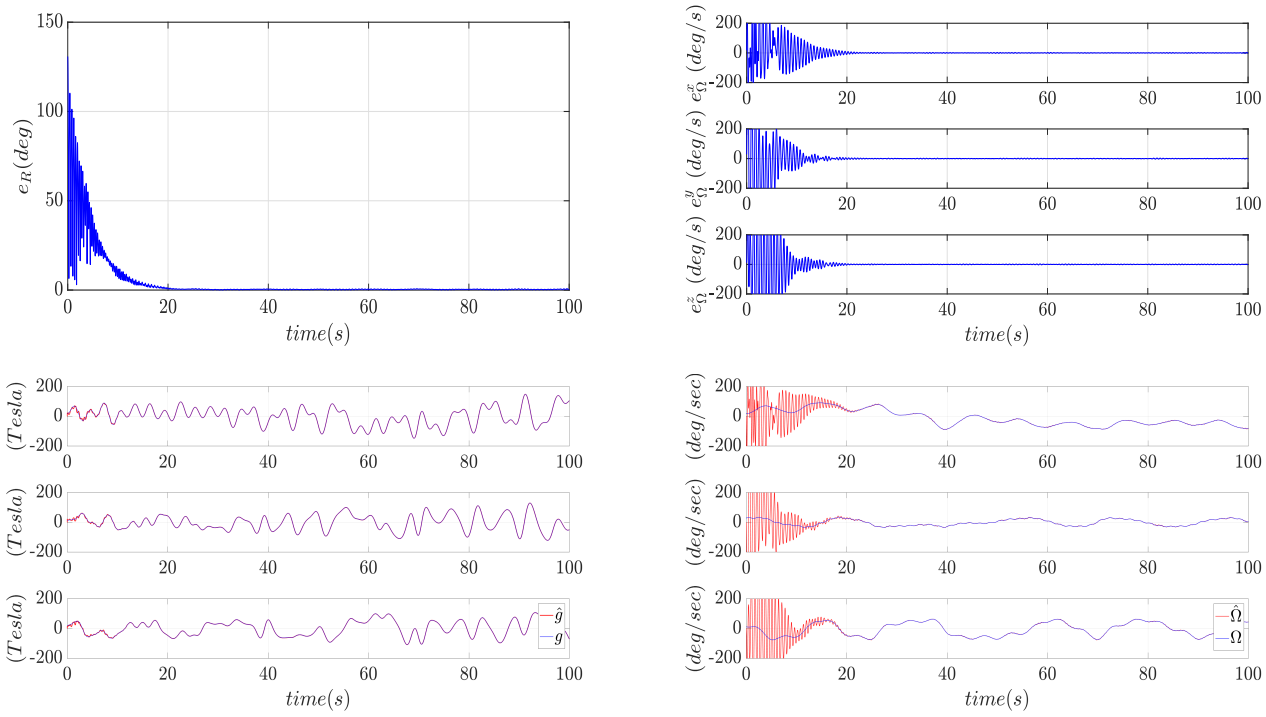


Figure 5-1: Series interconnection of the two filters. Filter $\mathcal{F}_1$ estimates the orientation $\hat{R}$ and angular velocity $\hat{\Omega}$ of the spacecraft using only two on-board measurements $Y_{[1:6]}$. Filter $\mathcal{F}_2$ takes as inputs the attitude and angular velocity estimates as well as the the third measurement $Y_{[7:9]}$ to produce the geomagnetic field estimate $\hat{g}$ and the optimal correction term σ^*.

5-1-1 Simulations

Here, we assume that the system is driven by the model error $\sigma(t) = \begin{bmatrix} 90\sin(\frac{2\pi}{3}t) & 0 & 0 \end{bmatrix}^T$. The orientation and angular rate estimation error in Figure 5-2 are not affected, as it was expected from the geomagnetic field observations. However, the performance of the second filter depends on $\hat{R}, \hat{\Omega}$ since good geomagnetic field estimates require both good orientation and angular velocity estimates. Finally, the right plot at the bottom shows that indeed the correction term determines the model error being disturbed however by the measurement noise.



We expect that in steady state, the optimal correction term σ^* will be equal to the model error. In other words, we expect that the filter will have determined the appropriate model error that drives the system (5-2). Please notice that the optimization problem (5-3) requests for the signal σ^* that minimizes the prediction error and which also has minimum energy.

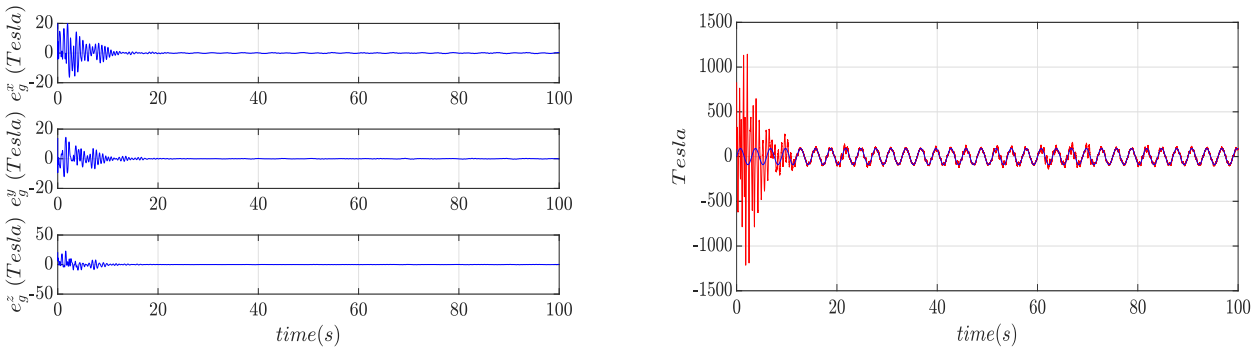


Figure 5-2: Attitude estimation error (top left) and angular velocity estimation error (top right). Geomagnetic field and geomagnetic field estimates (middle left), angular velocity and angular velocity estimates (middle right) geomagnetic field estimation error (bottom left) and model correction term (bottom right). Horizon step $h = 0.05$ sec., initial conditions $(R_0, \Omega_0, g_0) = ((4, [1\ 4\ 1]^T), [.4\ .5\ .8]^T, [20\ 14\ 21]^T)$, noise standard deviation $\sigma_\epsilon = 0.05$ m.

This of course does not mean that σ^* will be equal with the true model error signal that produces the actual measurements. The two signals however will have the same energy and therefore we can say that they differ up to a unitary transformation. In other words, by setting the true model error as given above, there is not guarantee that the σ^* resulting from the optimization problem will have its second and third component equal to zero for each time instant. The energy of σ^* can be distributed over all the three components.

However, in case where the subspace the model error lies in is known, we can achieve the desired structure by tuning properly the penalty matrix Π_{22}. By penalizing less some of the entries in matrix Π_{22} it would be possible to drive the energy flow of model error σ^* through the corresponding components of σ^* and therefore to achieve the desired structure. Equivalently, the aforementioned penalty matrix Π_{22} may result from the singular value decomposition (SVD) under the following reasoning:

$$\delta(t)^T \Pi_{22} \delta(t) = \delta(t)^T V \Sigma V^T \delta(t) = \delta(t)^T (\sigma_1 V_1 + \sigma_2 V_2 + \sigma_3 V_3) \delta(t)^T = \sum_{i=1}^{3} \sigma_i ||\delta(t)||_{V_i} \quad (5\text{-}18)$$

where $\text{rank}(V_i) = \text{rank}(V(:,i)V^T(:,i)) = 1$, $V(:,i)$ is the i-th column of V and $||\delta(t)||_{V_i}$ is the energy flow towards the direction $V(:,i)$. The columns $V(:,i)_i$ are chosen to constitute an orthonormal basis of the subspace the model error lies in, with the singular values σ_i to regulate the energy flow towards the directions $V(:,i)_{i=1,2,..}$.

5-2 Orbital position estimation

As stated at the beginning of this chapter, the goal is to use the geomagnetic field estimates in order to obtain (angular) orbital position estimates. Towards this direction, the next step is to apply the transport theorem for the geomagnetic field once again:

$$\frac{d^A}{dt} g = \frac{d^B}{dt} g + \Omega \times g \quad (5\text{-}19)$$

where the first term corresponds to the time variation of the geomagnetic field due to the fact that it is rotating by its nature. Therefore, the term $\frac{d^B}{dt}g$ can be written as

$$\frac{d^B}{dt}g = g \times \omega_o \tag{5-20}$$

where ω_o is the rotation rate of the geomagnetic field if the spacecraft had a fixed orientation. This first term of (5-20) was previously declared as a deterministic function corresponding to the model error. The model error is determined through the estimation process and therefore is a known function of time, depending linearly on the noisy measurements. Hence we can write:

$$\sigma^*(t) = (\hat{g} \times \hat{\omega}_o)(t) + n(t) \tag{5-21}$$

where $n(t) \in \mathbb{R}^3$ is a noise vector.

Given the fact that the frequency content of $n(t)$ does not overlap with the frequencies of the slower signal $g \times \omega_o$, the noise can be removed after applying a low pass filter on each component of $\sigma^*(t)$. By assuming a dipole model for the geomagnetic field, we can claim that the mean value of the instantaneous frequency of the filtered signal is half the frequency of the spacecraft orbiting the Earth. The reason is that the geomagnetic field is rotating two times during a full orbit. Figure 5-3 justifies this claim. Starting from the south pole and moving counterclockwise up to the north pole, the geomagnetic field makes a full turn. Subsequently, continue moving in a counterclockwise direction from the north to the south pole, the geomagnetic field performs another full turn.

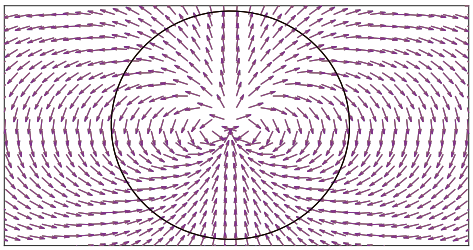


Figure 5-3: Dipole model for the Earth's geomagnetic field. For a fixed orientation, the geomagnetic field vector, as seen from the body frame is rotating.

5-2-1 Orbital rotational rate from geomagnetic field estimates

The faster the variation of $(\hat{g} \times \hat{\omega}_o)(t)$ is in time, the faster the spacecraft is orbiting the Earth. The rotation rate ω_o indicates the rate with which the geomagnetic field rotates if the spacecraft had a fixed orientation. Equivalently, ω_o indicates the geomagnetic field's rotation rate w.r.t. the L.V.L.H. frame. Therefore, ω_o belongs to the frequency content of $(\hat{g} \times \hat{\omega}_o)(t)$ and subsequently to the frequency content of σ^* as well.

The spectrograms shown in Figure 5-4 reveal the frequency content of the model error's and the signal's σ^* x-component. The window length is chosen to be large compared with the signal length (equal to 5000 samples) in order to achieve good frequency resolution. The

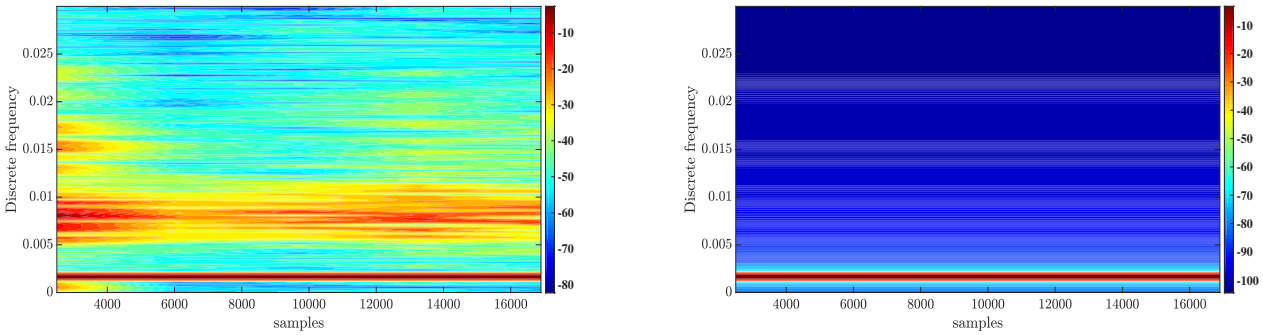


Figure 5-4: The magnitude of the short time Fourier transform (STFT) for the optimal correction term σ^* (left) and the model error signal (right).

vertical axes represent discrete frequencies. The x-component of $\sigma(t)$ is a sine wave with analog frequency $f_0 = 0.3$ Hz sampled with $f_s = 5$ KHz. Therefore, the corresponding discrete frequency is $\lambda_0 = 16 \cdot 10^{-3}$. Subsequently, the horizontal axis represent the samples of the discrete time signal.

The x-component of the model error $\sigma(t)$ is a pure sine wave with discrete frequency $\lambda_0 = 16 \cdot 10^{-3}$ and therefore is characterized by a constant instantaneous frequency law. The determined optimal correction term σ^* however contains additional higher frequencies due to the measurement noise. Nevertheless, the important fact is that the desirable frequency does appear in the frequency content of σ^* and can be filtered out using deterministic techniques in the frequency domain.

The reason why the spectrogram is used instead of the discrete Fourier transform aims to stress the fact that the actual model error might have a more complicated instantaneous frequency law that cannot be determined by classical frequency domain techniques. In this case, time-frequency techniques appear to be more suitable since the geomagnetic field rotates in a non-uniform way especially close to the geomagnetic poles as shown in Figure 5-3.

5-2-2 Summary

In this chapter, the predictive filter on $\mathbb{TSO}(3)$ has been extended and used to provide an alternative approach for the problem of the orbital position estimation by utilizing a kinematic model for the geomagnetic field. Towards this direction it has been proved that under some realistic assumptions the resulting filter can be seen as the series interconnection of two filters, which leads to the conclusion that the initial problem can be decomposed into two separate problems. We can clearly observe another benefit of the deterministic filtering approach as outcome of the dual optimal control theory: The fact that the correction term corresponding to the model error is actually determined through the estimation process allows us to estimate the orbital rotation rate and therefore of the orbital position of the spacecraft.

Chapter 6

Conclusions and future work

6-1 Conclusions

In this work, the problem of attitude and rate estimation has been studied for a low earth orbit satellite at the absence of rate sensors. The distinction between two main mathematical contexts for estimation is emphasized and investigated. On the one hand, in the stochastic context, which-is commonly used-the state space model is defined by stochastic differential equations and both model and measurement uncertainties with white noises. The objectives are then to produce an estimate of the mean state using the data at hand in the sense of the maximum likelihood estimator, or a mean square estimator, or a least square estimator. On the other hand, numerous contributions have been proposed in a purely deterministic context, where the model is described as a dynamical system. The errors are then unknown quantities defined in a Hilbert space with well defined function norms characterizing their weight.

The present work was focused on the deterministic description. One of the basic milestones of this problem is that the output equations are affected only by the orientation state, which lies in a curved space, the special orthogonal group $\mathbb{SO}(3)$. Otherwise, if a function of the angular rate could be measured, it would be possible through the model uncertainty to control the orientation via the angular velocity-an object which evolves in a vector space.

Within a general non-linear framework, the deterministic non-linear optimal observer can also be formulated and was first introduced by Mortensen [4]. This architecture was later employed by [2] to produce the second-order-optimal minimum energy filter on Lie groups. The controller derived in [3] allows to develop, a predictive observer directly on the tangent bundle of the special orthogonal group.

Both observers were analyzed and their performance was evaluated in different operating scenarios. In most simulated scenarios, the minimum energy filter was found to be superior compared to the predictive filter, where the latter was affected significantly from the measurement noise. Nonetheless, the derived predictive filter is extremely tolerant in model errors. It is evidently clear that this characteristic property is desired, since unknown environmental torques might vary significantly in practice. Furthermore, the error dynamics have been derived for both observers in an attempt for a Lyapunov stability evaluation.

Finally, the predictive filter on $\mathbb{TSO}(3)$ is extended and used in order to provide an alternative approach for the problem of the orbital position estimation by utilizing a kinematic model for the Earth's geomagnetic field. Towards this direction it is proved that under some realistic assumptions this later problem is equivalent to two separate problems and therefore the resulting filter can be seen as the series interconnection of two distinct filters. We can clearly observe another benefit of the deterministic filtering approach as an outcome of the dual optimal control theory. The fact that the correction term corresponding to the model error is actually determined through the estimation process allows us to estimate the orbital rotation rate and therefore, potentially, the orbital position of the spacecraft.

Overall, the main conclusion of this work could be summarized as follows: The dual optimal control approach with the resulting deterministic filters, outperforms the Bayesian techniques. The reason seems to be double. Through the deterministic approach the filtering problem can be set and solved in a coordinate free fashion under the theory of lie groups. In this way, nonlinearities induced from the coordinate system maps and the main issues of the Kalman-based architectures to preserve the space geometry are avoided. In addition, the fact that the model error is actually determined during the estimation process provides an alternative direction towards the problem of attitude and rate estimation.

6-2 Future work

Some directions for future research are the following:

- Improve the predictive filter on $\mathbb{TSO}(3)$ by finding an iterative way for updating the model error penalty matrix. This might lead into better, less noisy estimates.
- Formulate the problem by deploying D' Alembert principle might be another promising approach.
- Solve the problem stated in Chapter 5 with the minimum energy filter approach. By assuming two decomposed filters, the Riccati equation derived in Appendix A can be used for the second filter.

Appendix A

Proofs regarding the minimum energy filter

A-1 Derivation of the Riccati equation

It is

$$K(x^*,t) = \text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)$$

and we want to find the evolution of K w.r.t. time. In order to do so, we apply the time derivative operator in both sides of the above equation.

$$\frac{d}{dt}K(x^*,t) = \frac{d}{dt}\big(\text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)\big) \tag{A-1}$$

From the total derivative formula the right hand side of the above equation reads

$$\begin{aligned}\frac{d}{dt}\big(\text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)\big) &= \mathbf{d}_{x^*}\big(\text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)\big)\dot{x}^*(t) + \frac{\partial}{\partial t}\big(\text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)\big) \\ &= \mathbf{d}^3_{x^*}\big(\mathcal{V}(x^*,t;t_0,x_0)\big)\dot{x}^*(t) + \frac{\partial}{\partial t}\big(\text{Hess}_1 \mathcal{V}(x^*,t;t_0,x_0)\big) \\ &= \text{Hess}_1\big(\frac{\partial \mathcal{V}}{\partial t}(x^*,t;t_0,x_0)\big) + h.o.t\end{aligned} \tag{A-2}$$

but

$$\frac{\partial \mathcal{V}}{\partial t}(x^*,t;t_0,x_0) = \mathcal{H}(x^*, \mathbf{d}_1 \mathcal{V}(x^*,t;t_0,x_0), t) \tag{A-3}$$

therefore

$$\begin{aligned}
\mathrm{Hess}_1\Big(\frac{\partial \mathcal{V}}{\partial t}(x^*,t;t_0,x_0)\Big) &= -\mathrm{Hess}_1\Big(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0),t)\Big)\\
&= -\mathrm{Hess}_{x^*}\Big(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0),t)\Big)\\
&= \mathbf{d}_{x^*}\Big(\mathbf{d}_{x^*}\Big(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0),t)\Big)\Big)\\
&= \mathbf{d}_{x^*}\Big(\mathbf{d}_1\Big(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0),t)\Big)\Big)\\
&= \mathbf{d}_{x^*}\Big(\mathbf{d}_1(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))+\\
\mathbf{d}_1(\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)\Big)&\\
&= \mathbf{d}_{x^*}\Big(\mathbf{d}_1(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))+\\
\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0)\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)\Big)&
\end{aligned} \tag{A-4}$$

From the chain rule the first term reads

$$\begin{aligned}
\mathbf{d}_{x^*}\mathbf{d}_1(\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))) &=\\
&= \mathrm{Hess}_1\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))+\\
(\mathbf{d}_2\mathbf{d}_1\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))^T\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0)&
\end{aligned} \tag{A-5}$$

and the second term from the product rule reads

$$\begin{aligned}
\mathbf{d}_{x^*}(\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0)\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)) &=\\
\mathbf{d}_{x^*}(\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0))\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)+&\\
\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0)\mathbf{d}_{x^*}(\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0))&\\
= (\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0))\mathbf{d}_{x^*}(\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)) + h.o.t&
\end{aligned} \tag{A-6}$$

Again from chain rule the last equation reads

$$\begin{aligned}
(\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0))\mathbf{d}_{x^*}(\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)) + h.o.t&\\
= (\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0))\mathbf{d}_1\mathbf{d}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)+&\\
\mathrm{Hess}_2\mathcal{H}(x^*,\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0)\mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0) + h.o.t&
\end{aligned} \tag{A-7}$$

Given that $\mathbf{d}_1\mathcal{V}(x^*,t;t_0,x_0) = 0$ and $K(x^*,t) = \mathrm{Hess}_1\mathcal{V}(x^*,t;t_0,x_0)$, the evolution of the gain operator follows

$$\begin{aligned}
\dot{K}(x^*,t) &= \mathrm{Hess}_1\mathcal{H}(x^*,0,t)\\
&+\mathbf{d}_2\mathbf{d}_1\mathcal{H}(x^*,0,t)^TK(x^*,t)\\
+K(x^*,t)\mathbf{d}_1\mathbf{d}_2\mathcal{H}(x^*,0,t) &+ K(x^*,t)\mathrm{Hess}_2\mathcal{H}(x^*,0,t)K(x^*,t)
\end{aligned} \tag{A-8}$$

However we are interested in the evolution of the operator $\mathbb{K}(x^*,t) = (\text{Hess}_1\mathcal{H}(x^*,0,t))^{-1}$. Therefore we can write

$$\begin{aligned}\frac{d}{dt}\mathbb{K}(x^*,t) &= \frac{d}{dt}(\text{Hess}_1\mathcal{H}(x^*,0,t))^{-1} \\ &= \frac{d}{dt}(K^{-1}(x^*,t))\end{aligned} \tag{A-9}$$

but

$$\begin{aligned}0 = \frac{d}{dt}I &= \frac{d}{dt}\big(K(x^*,t)K^{-1}(x^*,t)\big) \\ &= \frac{d}{dt}\big(K(x^*,t)\big)K^{-1}(x^*,t) + K(x^*,t)\frac{d}{dt}\big(K^{-1}(x^*,t)\big)\end{aligned} \tag{A-10}$$

which means that

$$\frac{d}{dt}\big(K^{-1}(x^*,t)\big) = -K^{-1}(x^*,t)\frac{d}{dt}\big(K(x^*,t)\big)K^{-1}(x^*,t) \tag{A-11}$$

Finally by substituting (A-8) into (A-11) we have

$$\begin{aligned}\dot{\mathbb{K}}(x^*,t) = &-\mathbb{K}(x^*,t)\text{Hess}_1\mathcal{H}(x^*,0,t)\mathbb{K}(x^*,t) \\ &-\mathbb{K}(x^*,t)\mathbf{d}_2\mathbf{d}_1\mathcal{H}(x^*,0,t)^T - \mathbf{d}_1\mathbf{d}_2\mathcal{H}(x^*,0,t)\mathbb{K}(x^*,t) \\ &-\text{Hess}_2\mathcal{H}(x^*,0,t)\end{aligned} \tag{A-12}$$

or simply

$$\begin{aligned}\dot{\mathbb{K}}(t) = &-\mathbb{K}(t)\text{Hess}_1\mathcal{H}(x^*,0,t)\mathbb{K}(t) \\ &-\mathbb{K}(t)\mathbf{d}_2\mathbf{d}_1\mathcal{H}(x^*,0,t)^T - \mathbf{d}_1\mathbf{d}_2\mathcal{H}(x^*,0,t)\mathbb{K}(t) \\ &-\text{Hess}_2\mathcal{H}(x^*,0,t)\end{aligned} \tag{A-13}$$

Bibliography

[1] M. S. Grewal, L. R. Weill, and A. P. Andrews, *Global positioning systems, inertial navigation, and integration*. John Wiley & Sons, 2007.

[2] A. Saccon, J. Trumpf, R. Mahony, and A. P. Aguiar, "Second-order-optimal minimum-energy filters on lie groups," *IEEE Transactions on Automatic Control*, vol. 61, no. 10, pp. 2906–2919, 2015.

[3] P. Lu, "Nonlinear predictive controllers for continuous systems," *Journal of Guidance, Control, and Dynamics*, vol. 17, no. 3, pp. 553–560, 1994.

[4] R. Mortensen, "Maximum-likelihood recursive nonlinear filtering," *Journal of Optimization Theory and Applications*, vol. 2, no. 6, pp. 386–394, 1968.

[5] J. E. Marsden and T. S. Ratiu, "Introduction to mechanics and symmetry, volume 17 of texts in applied mathematics," 1999.

[6] E. Eade, *Monocular simultaneous localisation and mapping*. PhD thesis, University of Cambridge, 2009.

[7] J. Gallier, *Geometric methods and applications: for computer science and engineering*, vol. 38. Springer Science & Business Media, 2011.

[8] W. M. Boothby, *An introduction to differentiable manifolds and Riemannian geometry*, vol. 120. Academic press, 1986.

[9] E. Wan and R. van der Merwe, "Kalman filtering and neural networks, ser. adaptive and learning systems for signal processing, communications, and control," 2001.

[10] E. A. Wan, R. Van Der Merwe, and A. T. Nelson, "Dual estimation and the unscented transformation," pp. 666–672, 2000.

[11] R. Van Der Merwe *et al.*, *Sigma-point Kalman filters for probabilistic inference in dynamic state-space models*. PhD thesis, OGI School of Science & Engineering at OHSU, 2004.

[12] R. E. Kalman, "A new approach to linear filtering and prediction problems," *Journal of basic Engineering*, vol. 82, no. 1, pp. 35–45, 1960.

[13] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pp. 153–158, Ieee, 2000.

[14] S. J. Julier and J. K. Uhlmann, "A general method for approximating nonlinear transformations of probability distributions," tech. rep., Technical report, Robotics Research Group, Department of Engineering Science, 1996.

[15] 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.

[16] S. J. Julier, "Skewed approach to filtering," in *Signal and Data Processing of Small Targets 1998*, vol. 3373, pp. 271–282, International Society for Optics and Photonics, 1998.

[17] S. J. Julier, "The scaled unscented transformation," in *Proceedings of the 2002 American Control Conference (IEEE Cat. No. CH37301)*, vol. 6, pp. 4555–4559, IEEE, 2002.

[18] S. Julier, *Comprehensive process models for high-speed navigation*. PhD thesis, 1997.

[19] E. A. Wan and R. Van Der Merwe, "The unscented kalman filter for nonlinear estimation," in *Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No. 00EX373)*, pp. 153–158, Ieee, 2000.

[20] S. J. Julier and J. K. Uhlmann, "Unscented filtering and nonlinear estimation," *Proceedings of the IEEE*, vol. 92, no. 3, pp. 401–422, 2004.

[21] J. L. Farrell, "Attitude determination by kalman filtering," *Automatica*, vol. 6, no. 3, pp. 419–430, 1970.

[22] I. Bar-Itzhack and Y. Oshman, "Attitude determination from vector observations: Quaternion estimation," *IEEE Transactions on Aerospace and Electronic Systems*, no. 1, pp. 128–136, 1985.

[23] E. J. Lefferts, F. L. Markley, and M. D. Shuster, "Kalman filtering for spacecraft attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 5, no. 5, pp. 417–429, 1982.

[24] I. Y. Bar-Itzhack, "Optimum normalization of a computed quaternion of rotation," *IEEE transactions on aerospace and electronic systems*, no. 2, pp. 401–402, 1971.

[25] F. L. Markley, "Attitude error representations for kalman filtering," *Journal of guidance, control, and dynamics*, vol. 26, no. 2, pp. 311–317, 2003.

[26] D. Cilden, H. E. Soken, and C. Hajiyev, "Nanosatellite attitude estimation from vector measurements using svd-aided ukf algorithm," *Metrology and Measurement Systems*, vol. 24, no. 1, pp. 113–125, 2017.

[27] J. L. Crassidis and F. L. Markley, "Unscented filtering for spacecraft attitude estimation," *Journal of guidance, control, and dynamics*, vol. 26, no. 4, pp. 536–542, 2003.

[28] H. Schaub and J. L. Junkins, "Stereographic orientation parameters for attitude dynamics: A generalization of the rodrigues parameters," *Journal of the Astronautical Sciences*, vol. 44, no. 1, pp. 1–19, 1996.

[29] J. J. LaViola, "A comparison of unscented and extended kalman filtering for estimating quaternion motion," in *Proceedings of the 2003 American Control Conference, 2003.*, vol. 3, pp. 2435–2440, IEEE, 2003.

[30] D. J. Mook and J. L. Junkins, "Minimum model error estimation for poorly modeled dynamic systems," *Journal of Guidance, Control, and Dynamics*, vol. 11, no. 3, pp. 256–261, 1988.

[31] J. L. Crassidis and F. L. Markley, "Minimum model error approach for attitude estimation," *Journal of Guidance, Control, and Dynamics*, vol. 20, no. 6, pp. 1241–1247, 1997.

[32] M. Athans and P. L. Falb, *Optimal control: an introduction to the theory and its applications*. Courier Corporation, 2013.

[33] G. Likourezos, "The foundations of set-theoretic estimation," *Proceedings of the IEEE*, vol. 81, no. 2, pp. 181–181, 1993.

[34] C. L. Bottasso, *Multibody Dynamics: Computational Methods and Applications*, vol. 12. Springer Science & Business Media, 2008.

[35] N. Bou-Rabee and J. E. Marsden, "Hamilton–pontryagin integrators on lie groups part i: Introduction and structure-preserving properties," *Foundations of Computational Mathematics*, vol. 9, no. 2, pp. 197–219, 2009.

[36] M. Kobilarov, K. Crane, and M. Desbrun, "Lie group integrators for animation and control of vehicles," *ACM Trans. Graph.*, vol. 28, May 2009.

[37] R. W. Brockett, "Functional expansions and higher order necessary conditions in optimal control," in *Mathematical Systems Theory*, pp. 111–121, Springer, 1976.

[38] E. Gilbert, "Functional expansions for the response of nonlinear differential systems," *IEEE transactions on Automatic Control*, vol. 22, no. 6, pp. 909–921, 1977.

[39] L. Euler, "Formulae generales pro translatione quacunque corporum rigidorum," *Novi Commentarii academiae scientiarum Petropolitanae*, pp. 189–207, 1776.

[40] J. L. Junkins and H. Schaub, *Analytical mechanics of space systems*. American Institute of Aeronautics and Astronautics, 2009.

[41] M. Verhaegen and V. Verdult, *Filtering and system identification: a least squares approach*. Cambridge university press, 2007.