UNIVERSITY OF VERONA

DEPARTMENT OF
Computer Science

DOCTORAL PROGRAM IN
Computer Science
XXXV cycle (2019)

Geometric methods for designing optimal filters on Lie groups

S.S.D. ING - INF/05

Coordinator: **Prof. Ferdinando Cicalese**

Advisor: **Prof. Riccardo Muradore**
Co-Advisor: **Prof. Nicola Sansonetto**

Candidate: **Damiano Rigo**

Contents

Abstract

In control theory, the problem of having available good measurements is of primary importance in order to perform good tracking and control. Unfortunately, in real-life applications, sensing systems do not provide direct measurements about the pose (and its rate) of mechanical systems, while, in other situations, measurements are so noisy that require pre-processing to filter out disturbances and biases. These problems could be faced by using filters and observers.

In this thesis, we apply a second-order optimal minimum-energy filter constructed on Lie groups to several planar bodies. We start by studying the application of the filter to the matrix Lie group TSE(2), i.e. the tangent bundle of the Special Euclidean group SE(2); moreover, a comparison with the extended Kalman filter is presented. After that, we consider the Chaplygin sleigh case, that is a mechanical system with a nonholonomic constraint. Then, we move our attention to the case of an articulated convoy with hooking constraints. Finally, we apply the filter to a real case scenario consisting of a scaled model representing a parking truck semi-trailer system.

Particular attention is posed to the description of the geometric structure that underlies the dynamics and to the choice of the measurement equation, the affine connection, and the other parameters that define the filters. Simulations show the effectiveness of the proposed filters. The use of Lie groups theory for designing the filters is challenging, but the accuracy of the results, obtained considering the geometric structure and the symmetries of the system justifies the effort.

List of symbols

Notations used for the design of the filters in this thesis.

G	a connected Lie group;
n	dimension of the group G;
g, h	element of G;
T_gG	tangent space at g;
$\mathfrak{g}$	the Lie algebra associated with G;
X, Y	elements of the Lie algebra $\mathfrak{g}$;
$[\cdot, \cdot]$	the Lie bracket of $\mathfrak{g}$;
$\mathfrak{g}^*$	the dual of the Lie algebra $\mathfrak{g}$;
μ	an element of $\mathfrak{g}^*$;
$L_g : G \to G$	left translation $L_g h = gh$;
$T_h L_g$	the tangent map of L_g at $h \in G$;
gX	shorthand for $T_e L_g(X) \in T_gG$;
$\langle \cdot, \cdot \rangle$	duality paring $\langle \mu, X \rangle = \mu(X)$;
V	finite-dimensional vector space;
$f : G \to V$	differentiable map;
$\mathrm{d}f(g)$	differential of f at g, $\mathrm{d}f(g) : T_gG \to V$ identifying $T_{f(g)}V$ with V;
$\mathrm{d}_1, \mathrm{d}_2$	differentials with respect to individual arguments of a multiple-argument map;
$\nabla_X Y$	covariant derivative;
$\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$	connection function associated with ∇;
$\omega_X : \mathfrak{g} \to \mathfrak{g}$	$\omega_X(Y) = \omega(X, Y)$;
$\omega_Y^{\leftleftarrows} : \mathfrak{g} \to \mathfrak{g}$	$\omega_Y^{\leftleftarrows}(X) = \omega_X(Y)$;
$\omega_X^* : \mathfrak{g}^* \to \mathfrak{g}^*$	$\langle \omega_X^{*\leftleftarrows}(\mu), Y \rangle = \langle \mu, \omega_X Y \rangle$;
$\omega_\mu^{*\leftleftarrows} : \mathfrak{g} \to \mathfrak{g}^*$	$\langle \omega_\mu^{*\leftleftarrows}(X), Y \rangle = \langle \omega_X^*(\mu), Y \rangle = \langle \mu, \omega_X Y \rangle$;
$\omega_Y^{\leftleftarrows *} : \mathfrak{g}^* \to \mathfrak{g}^*$	$\langle \omega_Y^{\leftleftarrows *}(\mu), X \rangle = \langle \mu, \omega_Y^{\leftleftarrows}(X) \rangle = \langle \mu, \omega_X Y \rangle$;
$T(X, Y) \in \mathfrak{g}$	torsion function associated with ω;
$T_X : \mathfrak{g} \to \mathfrak{g}$	partial torsion function $T_X Y = T(X, Y)$;
$\mathrm{Hess} f(g) : T_gG \to \mathcal{L}(T_gG, V)$	Hessian operator of a twice differentiable function $f : G \to \mathbb{R}$ (or a map $f : G \to V$);
$(\phi)^W : \mathcal{L}(W, U) \to \mathcal{L}(W, V)$	exponential functor $(\cdot)^W$ applied to a linear map $\phi : U \to V$, exponential functor lifts ϕ to $\phi^W : \mathcal{L}(W, U) \to \mathcal{L}(W, V)$ defined by $\phi^W(\psi) = \phi \circ \psi$;
$\mathbb{I} : \mathfrak{g} \times \mathfrak{g} \to \mathbb{R}$	inner product on $\mathfrak{g}$;
$\mathbb{I}^\sharp : \mathfrak{g}^* \to \mathfrak{g}$	$\sharp$-map associated to the inner product $\mathbb{I}$;

$\mathbb{I}^\flat : \mathfrak{g} \to \mathfrak{g}^*$	$\flat$-map associated to the inner product $\mathbb{I}$;
$\text{ad} : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$	adjoint map on $\mathfrak{g}$;
$\text{ad}^* : \mathfrak{g} \times \mathfrak{g}^* \to \mathfrak{g}^*$	dual adjoint map;
$0_{n\times n}$	null matrix of dimension $n \times n$;
$I_{n\times n}$	identity matrix of dimension $n \times n$.
T_s	sample time
$\boldsymbol{\delta}$	Dirac delta
$\mathbb{E}\{v\} = \mu$	mean value of a causal variable v
$\text{Var}\{v\}$	variance of a causal variable v
$\sigma\{v\}$	standard deviation of a causal variable v
$\text{rms}\{v\}$	root mean square of a causal variable v
∇_x	gradient operator
tr(A)	trace of a matrix A
$\|\|A\|\|_F = \sqrt{\text{tr}(AA^T)}$	Frobenius norm
SE(2)	special Euclidean group in $\mathbb{R}^2$
$\mathfrak{se}(2)$	Lie algebra of SE(2)
$\Sigma_I = \{e_x, e_y\}$	inertial frame of reference
$\Sigma_b = \{e_1^b, e_2^b\}$, $\Sigma_b = \left\{e_\parallel^b, e_\perp^b\right\}$	bodies frames of reference
$\vee : \mathfrak{se}(2) \to \mathbb{R}^3$	"vee" isomorphism
$\wedge : \mathbb{R}^3 \to \mathfrak{se}(2)$	"wedge" isomorphism
$\lambda = g^{-1}\dot{g}$	left-trivialized dynamics
δ	model error
ε	measurement error
$\widehat{(\cdot)}$	estimated quantity
J	inertia
m	mass

In this thesis, we will use the notation $\overline{G}$ when we will refer to a Lie group that extends the Lie group SE(2). We will use the same notation for Lie algebras and their elements.

Chapter 1

Introduction

In control theory, the problem of having good measurements is of primary importance in order to perform good tracking and control. Unfortunately, in real-life applications, the sensing system does not often provide direct measurements about the pose (and its rate) of mechanical systems, while, in other situations, measurements are noisy and require pre-processing to filter out disturbances and biases. The design of pose estimators for robotic systems is then of paramount importance to produce an effective regulation on a desired stationary position and tracking of a reference trajectory, [24]. Moreover, the pose of the robot is crucial for unmanned aerial and ground vehicles (UAV, UGV) that exploit SLAM (Simultaneous Localization and Mapping) algorithms for computing their position and for planning their trajectory [42]. These problems are solved by designing filters and observers. In the last decades, many linear and nonlinear, deterministic and stochastic filters/observers have been proposed in the literature. The most famous approach is based on Kalman filtering [19] and its many extensions.

The Kalman filter is a recursive algorithm that, through measurements over time, produces accurate estimates of unknown variables by computing at run-time the joint probability distribution. The Kalman filter achieved great success for its simplicity and its wide range of applicability, but it is optimal only for linear and Gaussian systems: such assumption is usually too strong in UAV and UGV, where the model is nonlinear and where nonholonomic constraints may arise and must be taken into account.

After the formulation of its standard form, generalizations aiming at extending it to nonlinear systems, such as the Extended Kalman Filter (EKF), have been proposed [1]. The EKF computes at each step the linear approximation of the dynamics and measurement equations and applies the same algorithm as its linear form.

An attempt to overcome the Gaussian assumption is by using the Unscented Kalman Filter [18] and the Particle Filter [2] that approximate the probability density function with a certain number of sampling points and updates their values according to the past state and current measurements. Nonlinear maps can then be considered to get more accurate estimations of the mean and covariance of the state vector.

Minimum energy filters

Another way to design a filter is by formulating it as an optimization problem. In [26] the author considers nonlinear dynamics with nonlinear measurement equations, an initial cost, and a cost functional (energy) in which the incremental cost weights the contribution of the model and measurement errors. The dynamic programming principle ensures the optimality for the estimated quantity along the trajectory.

This algorithm presents an innovative way of tackling the problem, but it does not take into consideration the symmetries and the geometric structure of the systems that can be described with the theory of Lie groups. A Lie group is a differentiable manifold that has a continuous operation which gives it a group structure. From the manifold structure, the Lie group inherits all the notions related to differentiation, while from the group structure, it inherits symmetry notions. Thanks to these properties, Lie groups are the natural mathematical tool to describe mechanical and robotic systems.

The first results that exploit Lie groups to design minimum energy filters come from the works [33, 34, 35]. Unlike the stochastic approach, both uncertainty and noise related to model and measurements are considered unknown deterministic signals and the optimal filter is obtained by minimizing the square of the estimation error (energy). The solution is obtained by differentiating the boundary conditions of the associated optimal control problem. It is called *second-order* optimal in the sense that it is a truncation of the exact solution that would be an infinite dimensional system. The filter takes the form of a gradient observer coupled with a kind of Riccati differential equation that updates its output-injection gain (similarly to the standard Kalman filter).

Outline of the thesis

In Chapter 2 we recall the main results on minimum energy optimal filters and on second-order optimal filters designed for Lie groups. After that, we design observers for planar rigid bodies exploiting the second-order filter presented in [35]. Chapters 3 to 6 are devoted to the application of the second-order filter to the following cases:

- free rigid body;
- Chaplygin sleigh;
- n-trailer vehicle;
- truck and semi-trailer system.

In all of these cases, we study the geometric structures and the needed operators exploiting their matrix representations. Then, we design explicitly the filters with the proofs of the related theorems. The accuracy of the filters is shown with numerical simulations. They are elaborated using external inputs that generate the real (nominal) trajectories, while Gaussian white processes

model plant uncertainty and measurement noises. The measurements are taken as inputs of the filter to find the estimated trajectories. The general scheme is reported in Figure 1.1.

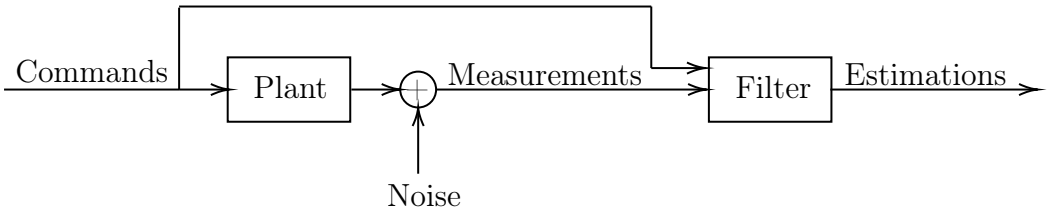


Figure 1.1: Scheme of the simulations setup. The external inputs (i.e. commands) generate the actual trajectories; Gaussian white noises corrupt the measurements.

Planar rigid body

The literature dealing with Lie groups and dynamic systems focuses mainly on simple manifolds without examining the case of their tangent bundle. Therefore filters designed on tangent bundles are less common.

In Chapter 3 we apply the second-order filter to the case of a planar rigid body ([7]). The dynamics evolves on the tangent bundle TSE(2) of the base Lie group SE(2) that can be trivialized as the Cartesian product $\mathrm{TSE}(2) \simeq \mathrm{SE}(2) \times \mathfrak{se}(2)$. The study of this type of manifold is particularly relevant in real applications. For example, it can represent many marine vehicle systems such as hovercraft. The dynamics is well described by the Euler-Poincaré equation for the rigid body that fits well with the formulation via Lie groups. Particular attention is devoted to the choice of the measurement equations and we compare three of them. In the last part of this chapter, a comparison with the extended Kalman filter is presented ([1], [28]), since the second-order filter could be seen as an improvement of the extended Kalman filter.

Chaplygin sleigh

In real cases, most vehicles are forced to follow only certain trajectories due to constraints originated by the use of wheels (such as bicycles or cars) or blades (such as sleighs). When the set of constraints cannot be integrated as constraints of the position we talk about nonholonomic constraints, [3]. The presence of a nonholonomic constraint changes the configuration space and a deeper analysis is needed. In particular, the state space is no more described by the tangent space of a base manifold, but by a distribution of it.

In Chapter 4 we extend the application of the filter to the case of nonholonomic dynamics. We consider the Chaplygin sleigh, which is a nonholonomic system that models a planar rigid body supported at three points, two of which slide freely while the third is a blade located ahead of the center of mass and that cannot move perpendicularly ([3], [29], [37]). To obtain the correct geometric structure, we exploit Hamel's approach to adapt the so-called Hamel's coordinates to the sleigh ([3], [4], [14], [46]). In this way, one coordinate vanishes. In this chapter, we analyze the choice of the affine connection and we

investigate the conditions that ensure the preservation of the nonholonomic constraint.

n-trailer vehicle

The study of articulated vehicles has grown in the last decades due to their use in real-life applications ([20], [23]). These systems model multi-bodies structures where the rigid bodies are linked by hooking constraints ([6]). They are very important in robotics and control theory since they can be exploited to model convoy systems where the first car pulls the trailers. Examples of these systems are luggage carriers in airports or cars with trailers in warehouses. A common strategy for deriving the equations of motion for nonholonomic systems consists in using Hamel's equations to write the system evolution in an arbitrary configuration-dependent frame.

In Chapter 5 we shift our attention from rigid bodies to articulated vehicles. We consider the case of a leading car pulling n trailers. In this new configuration, we consider a minimal set of measurement equations and we design different filters depending on whether the dynamic parameters are known or unknown.

Truck and semi-trailer

Special cases of n-trailer vehicles are truck trailer and truck semi-trailer systems ([10], [36], [43]). This type of system becomes unstable when the vehicles move in reverse, and this can give rise to the jackknifing problem. These problems become evident in parking maneuvers. The issue of autonomous or guided parking, facilitated through the measurements of the external environment or the knowledge of one's own state, has been developed in many areas (see e.g. [31]).

In Chapter 6 we consider the case of a truck semi-trailer system in a parking area. The experiments were done on a scaled model in a laboratory. All the measurements are simulated by adding noises to the actual state. In this case, we consider a likely set of sensors that can be applied in real life. Moreover, we consider different types of measurements in order to have better estimations when the system is moving in reverse, in which case the dynamics becomes unstable.

Chapter 2

Background on minimum energy filters

2.1 Maximum-likelihood recursive nonlinear filter

The construction of a minimum energy filter that we propose was derived and presented in [26]. The main idea consists in consider the estimation algorithm as an optimal control problem and exploit optimal control features to provide the estimation.

We consider a dynamics equation that describes the time evolution of a noisy system

$$\dot{\mathrm{x}}(t) = f(\mathrm{x}, t) + v(t)$$

where x is the vector of state dynamic variables, t is the time variable, f is a (nonlinear) function and v is a white noise, that is,

$$\mathbb{E}\{v(t)\} = 0, \quad \mathbb{E}\{v(t_1)v'(t_2)\} = R(t_1)\boldsymbol{\delta}(t_1 - t_2). \tag{2.1}$$

The system is able to take observations through the noisy measurement equation

$$\mathrm{y}(t) = h(\mathrm{x}(t), t) + w(t) \tag{2.2}$$

with $w(t)$ white noise independent of $v(t)$ and such that

$$\mathbb{E}\{w(t)\} = 0, \quad \mathbb{E}\{w(t_1)w'(t_2)\} = Q(t_1)\boldsymbol{\delta}(t_1 - t_2). \tag{2.3}$$

The initial state $\mathrm{x}(t_0)$ is a random vector with Gaussian distribution with covariance matrix Λ and mean value μ.

The idea that underlies a minimum energy filter is to find the dynamic vector $\mathrm{x}(t)$ that minimizes the "weighted" errors associated with the starting state, the model dynamics and the observations. This suggests considering the

likelihood functional

$$\begin{aligned} J_t = &+ \frac{1}{2}[\mathrm{x}(t_0) - \mu]\Lambda^{-1}[\mathrm{x}(t_0) - \mu] \\ &+ \frac{1}{2}\int_{t_0}^{t} \{[\dot{\mathrm{x}}(\tau) - f(\mathrm{x}(\tau), \tau)]' R^{-1}(\tau)[\dot{\mathrm{x}}(\tau) - f(\mathrm{x}(\tau), \tau)] \\ &+ [\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)]' Q^{-1}(\tau)[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)]\} \mathrm{d}\tau. \end{aligned} \tag{2.4}$$

Substituting the measurements $\mathrm{y}(\tau)$ into (2.4), J_t becomes a functional that depends only on $\mathrm{x}(\tau)$ for $0 \leq \tau \leq t$. Using Euler-Lagrange equations, it is possible to determine $\mathrm{x}^*(\tau)$ that minimizes the functional. This trajectory is defined on all the time intervals where the measurements are taken, but only the value obtained at the current (final) time t is used to estimate the current state: $\widehat{\mathrm{x}}(t) = \mathrm{x}^*_{[t_0,t]}(t)$. This implies that, to have an estimator for all t, we have to continuously calculate the optimal trajectory $\mathrm{x}^*(t)$ and take its final value. Unfortunately, given $t_0 < t_1 < t_2$, in general $\mathrm{x}^*_{[t_0,t_1]}(\tau) \neq \mathrm{x}^*_{[t_0,t_2]}(\tau)$ for $t_0 \leq \tau \leq t_1$, which means that the optimal trajectory in $[t_0, t_1]$ for $\mathrm{x}^*_{[t_0,t_1]}(\tau)$ is different from the optimal trajectory in $[t_0, t_1]$ for $\mathrm{x}^*_{[t_0,t_2]}(\tau)$. Therefore, the collected trajectories for the previous intervals of time, cannot be used as starting trajectories for the actual one (Figure 2.1).

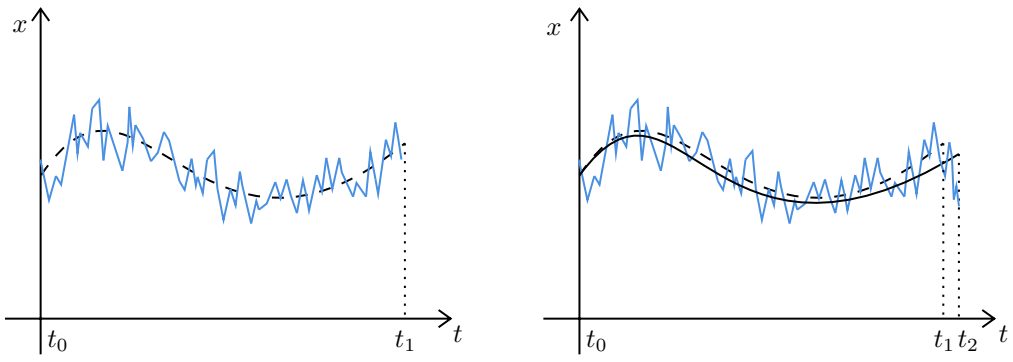


Figure 2.1: Optimal trajectories x^* for the time intervals $[t_0, t_1]$ (dashed) and $[t_0, t_2]$ (continuous). In blue is the measurement. Adding of the measurement between t_1 and t_2 modifies the optimal trajectory also in the interval $[t_0, t_1]$.

In order to improve the computational efficiency of the filter, it is convenient to rewrite the issue as an optimal control problem. Substituting (2.1) into (2.4), we obtain

$$\begin{aligned} J_t = &+ \frac{1}{2}[\mathrm{x}(t_0) - \mu]\Lambda^{-1}[\mathrm{x}(t_0) - \mu] + \frac{1}{2}\int_{t_0}^{t} v'(\tau) R^{-1}(\tau) v(\tau) \\ &+ [\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)]' Q^{-1}(\tau)[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)] \mathrm{d}\tau. \end{aligned} \tag{2.5}$$

In this way, the new purpose is to deal $v(t)$ as an optimal control variable and find it in order to minimize the functional (2.5) knowing that it is subject to the differential constraint

$$\dot{\mathrm{x}}(\tau) = f(\mathrm{x}(\tau), \tau) + v(\tau). \tag{2.6}$$

The Hamiltonian associated with this optimal control problem is then

$$\begin{aligned} H(\mathrm{x}, p, v, t) =& \frac{1}{2} v'(\tau) R^{-1}(\tau) v(\tau) + p'(f(\mathrm{x}(\tau), \tau) + v(\tau)) \\ &+ \frac{1}{2}[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)]' Q^{-1}(\tau)[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)] \end{aligned} \tag{2.7}$$

where x, p, v and t are independent variables. Since the Hamiltonian function is quadratic in v, the optimal choice of v is given by

$$\nabla_v H = 0 \Rightarrow v = -R(\tau)p \tag{2.8}$$

and thus, the optimal Hamiltonian is

$$\begin{aligned} H^*(\mathrm{x}, p, t) =& \frac{1}{2} p'(\tau) R(\tau) p(\tau) + \frac{1}{2}[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)]' Q^{-1}(\tau)[\mathrm{y}(\tau) - h(\mathrm{x}(\tau), \tau)] \\ &+ p'(f(\mathrm{x}(\tau), \tau) + v(\tau)). \end{aligned} \tag{2.9}$$

Hamilton's equations related to Hamiltonian (2.9) are

$$\begin{aligned} \dot{\mathrm{x}}(\tau) =& \nabla_p H^*(\mathrm{x}, p, \tau) = f(\mathrm{x}, \tau) - R(\tau)p \\ \dot{p}(\tau) =& -\nabla_{\mathrm{x}} H^*(\mathrm{x}, p, \tau) = -f_{\mathrm{x}}(\mathrm{x}, \tau)p + h_{\mathrm{x}}(\mathrm{x}, \tau) Q^{-1}(\tau)[\mathrm{y}(\tau) - h(\mathrm{x}, \tau)] \end{aligned} \tag{2.10}$$

with boundary conditions

$$p(t_0) = \Lambda^{-1}[\mu - \mathrm{x}(t_0)], \quad p(t) = 0, \tag{2.11}$$

where f_{x}, h_{x} are the Jacobian matrices of f and h, respectively. The optimal estimate $\widehat{\mathrm{x}}(t)$ is the final value of the optimal trajectory $\mathrm{x}^*(t)$. With this approach, one has to find the optimal trajectory for every t, and then evaluate it at the current time to find the optimal estimate. The procedure could be long and computationally speaking it could be very wasteful. Thus, another viewpoint that considers the Hamilton-Jacobi theory is preferable.

We define the value function

$$V(\mathrm{x}, t; \mu, t_0) = \min_{v(t), t_0 \leq \tau \leq t} J_t. \tag{2.12}$$

The Hamilton-Jacobi equation for this function is

$$\frac{\partial}{\partial t} V(\mathrm{x}, t; \mu, t_0) + H^*[\mathrm{x}, \nabla_{\mathrm{x}}(\mathrm{x}, t; \mu, t_0), t] = 0 \tag{2.13}$$

with the boundary condition

$$V(\mathrm{x}, t_0; \mu, t_0) = \frac{1}{2}[\mathrm{x}(t_0) - \mu]\Lambda^{-1}[\mathrm{x}(t_0) - \mu]. \tag{2.14}$$

Since we have no final conditions on the state variable x (i.e. $\mathrm{x}(t)$ is free), the vector of adjoint variables satisfies $p(t) = 0$. This are the adjoint variables of the Hamiltonian function (2.13), and thus, the final condition brings to

$$\nabla_{\mathrm{x}} V(\mathrm{x}, t; \mu, t_0) = 0. \tag{2.15}$$

With this approach, one obtains the solution of V satisfying (2.13), and then, one computes its root in order to find the optimal estimate $\widehat{\mathrm{x}}(t)$.

The procedure described so far allows to identify $\widehat{\mathrm{x}}(t)$ solving the partial differential equation (2.13). The next step consists in write a system of differential equations that is satisfied by $\widehat{\mathrm{x}}(t)$. The total time derivative of (2.15) is

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{dt}}\Big\{[\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)]_{\mathrm{x}=\widehat{\mathrm{x}}(t)}\Big\}\\&\quad=\left\{\nabla_{\mathrm{x}}\left[\frac{\partial V}{\partial t}(\mathrm{x},t;\mu,t_0)+\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)\left(\frac{\mathrm{d}\widehat{\mathrm{x}}(t)}{\mathrm{dt}}\right)\right]\right\}_{\mathrm{x}=\widehat{\mathrm{x}}(t)}.\end{aligned}\tag{2.16}$$

Substituting the Hessian matrix $\Pi(\mathrm{x},t;\mu,t_0)$

$$(\Pi)_{ij}=\frac{\partial^2}{\partial \mathrm{x}_i\partial \mathrm{x}_j}V(\mathrm{x},t;\mu,t_0)\tag{2.17}$$

and (2.13) into (2.16), we obtain

$$\begin{aligned}&\frac{\mathrm{d}}{\mathrm{dt}}\Big\{[\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)]_{\mathrm{x}=\widehat{\mathrm{x}}(t)}\Big\}\\&\quad=\left\{-\nabla_{\mathrm{x}}H^*[\mathrm{x},\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0),t]+\Pi(\mathrm{x},t;\mu,t_0)\left(\frac{\mathrm{d}\widehat{\mathrm{x}}(t)}{\mathrm{dt}}\right)\right\}_{\mathrm{x}=\widehat{\mathrm{x}}(t)}.\end{aligned}\tag{2.18}$$

The first addend on the right-hand side can be expanded using the chain rule

$$\begin{aligned}&\nabla_{\mathrm{x}}H^*[(\mathrm{x},\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0),t)]\\&=[\nabla_{\mathrm{x}}H^*(\mathrm{x},p,t)]_{p=\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)}+\Pi(\mathrm{x},t;\mu,t_0)[\nabla_pH(\mathrm{x},p,t)]_{p=\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)}.\end{aligned}\tag{2.19}$$

From Hamilton's equations (2.10)

$$\begin{aligned}&\nabla_{\mathrm{x}}H^*[(\mathrm{x},\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0),t)]\\&\quad=-f_{\mathrm{x}}(\mathrm{x},t)\cdot\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)+h_{\mathrm{x}}(\mathrm{x},t)Q^{-1}(t)[\mathrm{y}(t)-h(\mathrm{x},t)]\\&\qquad+\Pi(\mathrm{x},t;\mu,t_0)f(\mathrm{x},t)-\Pi(\mathrm{x},t;\mu,t_0)R(t)\cdot\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0).\end{aligned}\tag{2.20}$$

Rewriting equation (2.15) as

$$[\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0)]_{\mathrm{x}=\widehat{\mathrm{x}}(t)}=0\tag{2.21}$$

and $K(\mathrm{x},t)=h_{\mathrm{x}}(\mathrm{x},t)$, substituting these into (2.20), and evaluating along $\widehat{\mathrm{x}}(t)$, we end up with

$$\begin{aligned}&\{\nabla_{\mathrm{x}}H^*[(\mathrm{x},\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0),t)]\}_{\mathrm{x}=\widehat{\mathrm{x}}(t)}\\&\quad=\Pi(\widehat{\mathrm{x}}(t),t;\mu,t_0)f(\widehat{\mathrm{x}}(t),t)+K(\widehat{\mathrm{x}}(t),t)Q^{-1}(t)[\mathrm{y}(t)-h(\widehat{\mathrm{x}}(t),t)].\end{aligned}\tag{2.22}$$

From equation (2.21), equation (2.18) becomes

$$\{\nabla_{\mathrm{x}}H^*[\mathrm{x},\nabla_{\mathrm{x}}V(\mathrm{x},t;\mu,t_0),t]\}_{\mathrm{x}=\widehat{\mathrm{x}}(t)}=\Pi(\widehat{\mathrm{x}}(t),t;\mu,t_0)\left(\frac{\mathrm{d}\widehat{\mathrm{x}}(t)}{\mathrm{dt}}\right).\tag{2.23}$$

Assuming non-singularity of the matrix Π, and substituting into equation (2.22) we finally obtain

$$\frac{\mathrm{d}\widehat{\mathrm{x}}(t)}{\mathrm{dt}} = f(\widehat{\mathrm{x}}(t), t) + \Pi^{-1}(\widehat{\mathrm{x}}(t), t; \mu, t_0) K(\widehat{\mathrm{x}}(t), t) Q^{-1}(t)[\mathrm{y}(t) - h(\widehat{\mathrm{x}}(t), t)]. \tag{2.24}$$

This equation, together with the initial condition

$$\widehat{\mathrm{x}}(t_0) = \mu, \tag{2.25}$$

is the maximum-likelihood nonlinear filter. Matrix $\Pi^{-1}(\widehat{\mathrm{x}}(t), t; \mu, t_0)$ corresponds to the error covariance matrix in Kalman-Bucy theory and can be obtained with an ordinary matrix differential equation of Riccati type with initial condition

$$\Pi^{-1}(t_0, t_0) = \Lambda. \tag{2.26}$$

2.2 Second-order optimal minimum energy filters on Lie groups

In this section, we recall a second-order optimal minimum energy filter constructed on Lie groups. This filter represents the result of the work in [35] and constitutes the base of the rest of this thesis.

Let us better specify some concepts that are useful for understanding the filter.

Definition 2.1. (Connection function). A left-invariant affine connection is an affine connection ∇ on G such that $L_g^*(\nabla_X Y) = \nabla_{L_{g^*}X} L_g^* Y$ for all $g \in G$ and $X, Y \in \mathfrak{X}(TG)$. Such a left-invariant affine connection is fully characterized by its bilinear connection function $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ through the identity $\nabla_{gX}(gY) = g\omega(X, Y)$.

Definition 2.2. (Hessian operator). Given a twice differentiable function $f : G \to \mathbb{R}$ we can define the Hessian operator $\mathrm{Hess}\, f(g) : T_gG \to T_g^*G$ at a point $g \in G$ by $\mathrm{Hess}\, f(g)(gX)(gY) = d(df(g)(gY))(gX) - df(g)(\nabla_{gX}(gY))$ for all gX, $gY \in T_gG$.

Definition 2.3. (Exponential functor). Given three vector spaces U, V, W and a linear map $\phi : U \to V$, the exponential functor $(\cdot)^W$ lifts the map ϕ to the linear map $\phi^W : \mathcal{L}(W, U) \to \mathcal{L}(W, V)$ defined by $\phi^W(\psi) = \phi \circ \psi$.

The optimal filter described in this section proposes the same ideas as the general maximum-likelihood nonlinear filter described in Section 2.1. In particular, it is stated as a minimum energy filter constructed on Lie groups.

Since the geometric structure that underlies the dynamics of our system is a Lie group G, we can write it considering the following deterministic system:

$$\dot{g}(t) = g(t)\left[\lambda(g(t), u(t), t) + B\delta(t)\right], \quad g(t_0) = g_0 \tag{2.27}$$

where $g(t) \in G$ is the state, $u(t) \in \mathbb{R}^m$ is an external input, $\lambda : G \times \mathbb{R}^m \times \mathbb{R} \to \mathfrak{g}$ is the left-trivialized dynamics, $\delta(t) \in \mathbb{R}^d$ is unknown model error, $B : \mathbb{R}^d \to \mathfrak{g}$ is a linear map and g_0 is the unknown initial condition at time t_0.

The known measurement output $\mathrm{y} \in \mathbb{R}^p$ is given by the following measurement equation:

$$\mathrm{y}(t) = h(g(t), t) + D\varepsilon(t) \tag{2.28}$$

where $h : G \times \mathbb{R} \to \mathbb{R}^p$ is the nominal output map, $\varepsilon \in \mathbb{R}^p$ is the unknown measurement error and $D : \mathbb{R}^p \to \mathbb{R}^p$ is an invertible linear map. This equation depicts the knowledge the system acquires of its state through sensors.

In order to settle the filter as an optimal control problem, it is necessary to define a cost functional to be minimized. We choose

$$J(\delta, \varepsilon, g_0; t, t_0) = \min_{(g(\cdot), \delta(\cdot))} m(g(t_0), t, t_0) + \int_{t_0}^{t} \ell(\delta(\tau), \varepsilon(\tau), t, \tau) d\tau \tag{2.29}$$

where

$$m(g_0, t, t_0) := 1/2 e^{-\alpha(t-t_0)} m_0(g_0), \tag{2.30}$$

$$\ell(\delta, \varepsilon, t, \tau) := 1/2 e^{-\alpha(t-t_0)} (\mathcal{R}(\delta) + \mathcal{Q}(\varepsilon)). \tag{2.31}$$

Equation (2.30) represents an initial cost where $m_0 : G \to \mathbb{R}$ is a bounded smooth function with a unique global minimum on G that encodes the a-priori information about the state at time t_0. Equation (2.31), instead, represents an incremental cost. The two quadratic forms $\mathcal{R} : \mathbb{R}^d \to \mathbb{R}$ and $\mathcal{Q} : \mathbb{R}^p \to \mathbb{R}$ measure the instantaneous energy of model error $\mathcal{R}(\delta)$ and measurement error $\mathcal{Q}(\varepsilon)$.

The filter takes as input the measurement $\mathrm{y}(\tau)$ and the input $u(\tau)$ and produces the filter estimate $\widehat{g}(t)$. The error signals δ and ε are modelled as unknown deterministic functions of time. Together with the initial conditions g_0, these three signals are unknown in the optimization problem. Each choice of this triple corresponds to a different state trajectory $g(\tau)$. The principle of minimum energy filter consists in finding the "best" trajectory that minimizes the cost functional induced by the signals $(\delta(\tau), \varepsilon(\tau), g_0)$ for $\tau \in [t_0, t]$.

The filter estimation $\widehat{g}(t)$ coincides with the optimal minimum trajectory at time t, that means $\widehat{g}(t) := g^*_{[t_0,t]}(t)$. The minimum energy filter should be posed and solved on the interval $[t_0, t]$. Based on the same idea presented in Section (2.1), it is not necessary to solve the optimization problem for each new time t and then evaluate the optimal estimate $\widehat{g}(t)$, because we can exploit the dynamic programming principle and apply the Hamilton-Jacobi-Bellman equation in terms of the value function. The new approach based on Lie group theory provides a new formulation that takes care of the symmetry and geometry of the system. Here we state the second-order optimal minimum energy filter designed for Lie groups systems.

Theorem 2.1. *Consider the system define by* (2.27) *and* (2.28) *with the energy cost functional* (2.29)-(2.31). *Then the second-order-optimal minimum-energy filter* $\widehat{g}$ *is given by*

$$\widehat{g}^{-1}\dot{\widehat{g}} = \lambda_t(\widehat{g}, u) + K(t) r_t(\widehat{g}), \quad \widehat{g}(t_0) = \widehat{g}_0 \tag{2.32}$$

where $K(t) : \mathfrak{g}^* \to \mathfrak{g}$ is a (time-varying) second-order-optimal symmetric gain operator satisfying the perturbed Riccati operator (2.35) *given below,*

$$\widehat{g}_0 = \arg\min_{g \in G} m_0(g), \tag{2.33}$$

and the residual $r_t(\widehat{g}) \in \mathfrak{g}^*$ is computed as

$$r_t(\widehat{g}) = T_e L^*_{\widehat{g}}[\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{g}))\big) \circ \mathrm{d}h_t(\widehat{g})]. \tag{2.34}$$

The perturbed Riccati equation for K is

$$\begin{aligned} \dot{K} = &- \alpha \cdot K + A \circ K + K \circ A^* - K \circ E \circ K \\ &+ B \circ R^{-1} \circ B^* - \omega_{Kr} \circ K - K \circ \omega^*_{Kr} \end{aligned} \tag{2.35}$$

with initial condition $K(t_0) = X_0^{-1}$. The operators $X_0 : \mathfrak{g} \to \mathfrak{g}^*$, $A(t) : \mathfrak{g} \to \mathfrak{g}$, and $E(t) : \mathfrak{g} \to \mathfrak{g}^*$ are given by

$$X_0 = T_e L^*_{\widehat{g}_0} \circ \mathrm{Hess}\, m_0(\widehat{g}) \circ T_e L_{\widehat{g}_0} \tag{2.36}$$

$$A(t) = \mathrm{d}_1 \lambda_t(\widehat{g}, u) \circ T_e L_{\widehat{g}} - \mathrm{ad}_{\lambda_t(\widehat{g},u)} - T_{\lambda_t(\widehat{g},u)} \tag{2.37}$$

$$\begin{aligned} E(t) = -T_e L^*_{\widehat{g}} \circ \big[&\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{g}))\big)^{T_{\widehat{g}G}} \circ \mathrm{Hess}\, h_t(\widehat{g}) \\ &- (\mathrm{d}h_t(\widehat{g}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h_t(\widehat{g})\big] \circ T_e L_{\widehat{g}}. \end{aligned} \tag{2.38}$$

The symbol $\circ$ denotes the composition between maps, Kr is a shorthand notation for $K(t)r_t(\widehat{g})$ and ω_{Kr} is the connection form that depends on the chosen affine connection. R and Q are two symmetric positive definite matrices representative of the quadratic forms $\mathcal{R}$ and $\mathcal{Q}$.

Proof. See Appendix C. □

As can be seen, the theorem describes a second-order filter since the gain K, solved by a Riccati equation (2.35), is a second-order approximation of the analytical solution that appears considering the Hamilton-Jacobi-Bellman equation (see Appendix C).

The residual r_t considers the difference between the real measurements and the estimated one. Through this operator, the estimation errors, that belong to $\mathbb{R}^p$, are mapped onto the dual of the Lie algebra $\mathfrak{g}$. Operator A represents the coefficient of the linear part of the Riccati equation. The first term $\mathrm{d}_1\lambda(\widehat{g}, u) \circ T_e L_{\widehat{g}}$ is the differential of the left-trivialized dynamics with respect to the group elements that, together with the adjoint operator $\mathrm{ad}_{\lambda(\widehat{g},u)}$, provide a linearization for the trivialized dynamics, while the last term $T_{\lambda(\widehat{g},u)}$ is the torsion that takes care of the choice of the connection function adopted. The operator E represents the second-order term of the Riccati equation that does not depend on the gain, while the operator ω_{K_r} represents the second-order term that depends on the gain.

We notice that the residual r_t, the operators K, A, E, the starting condition X_0 and the connection function ω are all defined on the Lie algebra of the Lie group or on its dual.

The filter does not require any hypothesis on the model and measurement errors, which are treated as deterministic but unknown functions. This implies that it is not necessary to provide a statistical description for the errors as instead is required in stochastic filters. In real life, model errors represent uncertainty related to model dynamics (e.g. uncertainty on masses, inertias,...), while measurement errors affect the precision of the sensors. In the thesis, we will simulate them as Gaussian variables in the simulations to get a realistic representation. In the case of dynamics constructed on tangent bundle, we will add model errors only on the velocity evolution of the system and not on the kinematics ([35]).

Chapter 3

Second-order optimal filter applied to a free rigid body [1]

In this chapter, we study the second-order optimal filter applied to the case of a planar free rigid body. We will focus in particular on the study of the geometric structure that underlies the dynamics and the impacts of different measurement equations. The last part of the chapter is devoted to a comparison with the extended Kalman filter.

3.1 Free rigid body dynamics in TSE(2)

The free rigid body that we consider is shown in Figure 3.1.

We consider an inertial frame of reference $\Sigma_I = \{e_x, e_y\}$ integral to the ground and a body reference frame $\Sigma_b = \{e_1^b, e_2^b\}$ attached to the rigid body with the origin of the axes settled on the center of mass (x, y). We indicate with θ the angle between e_x and e_1^b that represents the orientation of the body. We will refer to the triple (θ, x, y) as the pose of the body. This pose is an element $g \in \mathrm{SE}(2)$ and admits a well-known matrix representation given by

$$g = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix}.$$

SE(2) is a matrix Lie group, called special Euclidean group, that features the rigid motion of the Euclidean space on $\mathbb{R}^2$, comprising all translations and rotations. Its Lie algebra is indicated with $\mathfrak{se}(2)$, and an element $\eta^g \in \mathfrak{se}(2)$

[1] This chapter is based on the following publications:
▷ Rigo D., Segala C., Sansonetto N., & Muradore R. (2022). *Second-order-optimal filter on Lie groups for planar rigid bodies.* IEEE Transactions on Automatic Control.
▷ Rigo D., Sansonetto N., & Muradore R. (2021, December). *A comparison between the Extended Kalman Filter and a Minimum-Energy Filter in the TSE(2) case.* In 2021 60th IEEE Conference on Decision and Control (CDC) (pp. 6175-6180). IEEE.

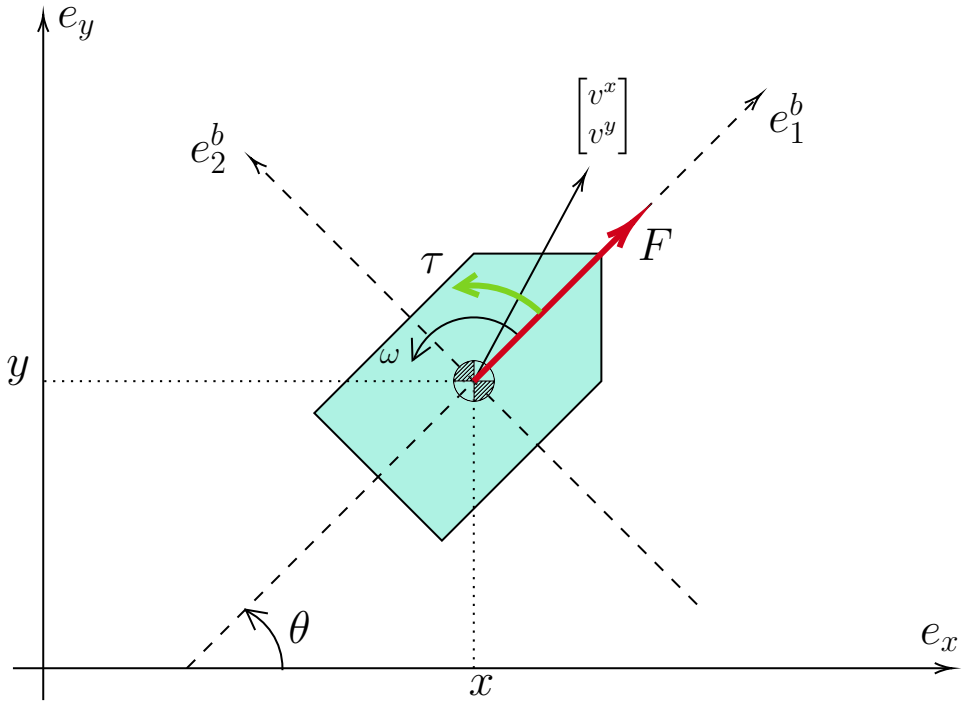


Figure 3.1: Planar free rigid body.

admits the matrix representation

$$\eta^g = \begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix}.$$

We introduce the Lie algebra isomorphisms $\vee : \mathfrak{se}(2) \to \mathbb{R}^3$ and $\wedge : \mathbb{R}^3 \to \mathfrak{se}(2)$ as

$$\begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix}^\vee = \begin{bmatrix} \eta^\theta \\ \eta^x \\ \eta^y \end{bmatrix}, \quad \begin{bmatrix} \eta^\theta \\ \eta^x \\ \eta^y \end{bmatrix}^\wedge = \begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix},$$

between the Lie algebra $(\mathbb{R}^3, \star)$ and the matrix Lie algebra $(\mathfrak{se}(2), [\cdot,\cdot])$, where $\star : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is the Lie bracket operation defined as

$$\begin{bmatrix} \eta_1^\theta \\ \eta_1^x \\ \eta_1^y \end{bmatrix} \star \begin{bmatrix} \eta_2^\theta \\ \eta_2^x \\ \eta_2^y \end{bmatrix} = \begin{bmatrix} 0 \\ \eta_1^y \eta_2^\theta - \eta_1^\theta \eta_2^y \\ \eta_1^\theta \eta_2^x - \eta_1^x \eta_2^\theta \end{bmatrix}, \tag{3.1}$$

and $[\cdot,\cdot]$ is the usual matrix commutator (see e.g. [25]). The adjoint operator $\mathrm{ad}_{\eta^g} \in \mathcal{L}(\mathfrak{se}(2); \mathfrak{se}(2))$ of $\eta^g \in \mathfrak{g}$ admits the matrix representation

$$\mathrm{ad}_{\eta^g} = \begin{bmatrix} 0 & 0 & 0 \\ v^y & 0 & -\omega \\ -v^x & \omega & 0 \end{bmatrix},$$

while its dual $\mathrm{ad}^*_{\eta^g} \in \mathcal{L}(\mathfrak{se}(2)^*; \mathfrak{se}(2)^*)$ satisfies $\mathrm{ad}^*_{\eta^g} = (\mathrm{ad}_{\eta^g})^T$.

The space of velocities has the same structure as the Lie algebra $\mathfrak{g}$, and thus a generic element Ω can be written as

$$\Omega = \begin{bmatrix} 0 & -\omega & v^x \\ \omega & 0 & v^y \\ 0 & 0 & 0 \end{bmatrix}.$$

With reference to the rigid body in Figure 3.1, ω represents the angular velocity while the couple (v^x, v^y) denotes the linear velocity of the body written in body coordinate.

Given the dual basis $\{\hat{e}^1, \hat{e}^2, \hat{e}^3\}$ on $\mathfrak{se}(2)^*$, we denote by $\mathbb{I}$ the constant inertia tensor

$$\mathbb{I} = J\hat{e}^1 \otimes \hat{e}^1 + m\hat{e}^2 \otimes \hat{e}^2 + m\hat{e}^3 \otimes \hat{e}^3 \tag{3.2}$$

with matrix representation (in the standard basis for $\mathbb{R}^3$, see [7])

$$\mathbb{I} = \mathrm{diag}(J, m, m) \tag{3.3}$$

where J is the inertia along the axis passing to the center of mass and orthogonal to the plane and with m the mass of the rigid body. We denote by $u = (\tau/J, F/m, 0)^T$ the control inputs that are functions of time such that τ and F act respectively as a torque applied around the center of mass and a force applied along the body first axis as shown in Figure 3.1.

The dynamics of the body evolves on TSE(2), the tangent bundle of $G := \mathrm{SE}(2)$ that we can identify with $\overline{G} = \mathrm{SE}(2) \times \mathfrak{se}(2)$ via left translation, [7], and is given by the Euler-Poincarè equation

$$g^{-1}\dot{g} = \Omega \tag{3.4}$$

$$\dot{\Omega} = \mathbb{I}^{\sharp}\mathrm{ad}^*_{\Omega}\mathbb{I}^{\flat}\Omega + u^{\wedge} \tag{3.5}$$

with $(g, \Omega) \in \overline{G}$. In component-wise form, the system can be written as

$$\begin{cases} \dot{\theta} &= \omega \\ \dot{x} &= v^x \cos\theta - v^y \sin\theta \\ \dot{y} &= v^x \sin\theta + v^y \cos\theta \\ \dot{\omega} &= \tau/J \\ \dot{v}_1 &= \omega v^y + F/m \\ \dot{v}_2 &= -\omega v^x. \end{cases} \tag{3.6}$$

To take into account unmodelled dynamics in (3.4)-(3.5) we consider the unknown error δ (modelled as a normalized Gaussian white noise) and the mapping

$$\begin{aligned} B : \mathbb{R}^3 &\to \mathfrak{se}(2) \times \mathfrak{se}(2) \\ \delta &\mapsto (0_{3\times 3}, (B_2\delta)^{\wedge}) \end{aligned} \tag{3.7}$$

with $B_2 \in \mathbb{R}^{3\times 3}$. The full dynamic equations thus become

$$g^{-1}\dot{g} = \Omega \tag{3.8}$$

$$\dot{\Omega} = \mathbb{I}^{\sharp}\mathrm{ad}^{*}_{\Omega}\mathbb{I}^{\flat}\Omega + u^{\wedge} + (B_2\delta)^{\wedge} \tag{3.9}$$

where the model error δ does not affect the reconstruction equation $g^{-1}\dot{g} = \Omega$. In extended form, we can summarize the contribution of the model errors with the vector $\xi = (0, 0, 0, \xi_\omega, \xi_{v^x}, \xi_{v^y})$, and rewrite the dynamics as

$$\begin{cases} \dot{\theta} &= \omega \\ \dot{x} &= v^x \cos\theta - v^y \sin\theta \\ \dot{y} &= v^x \sin\theta + v^y \cos\theta \\ \dot{\omega} &= \tau/J + \xi_\omega \\ \dot{v}_1 &= \omega v^y + F/m + \xi_{v_x} \\ \dot{v}_2 &= -\omega v^x + \xi_{v_y}. \end{cases} \tag{3.10}$$

3.2 The $\mathrm{SE}(2) \times \mathfrak{se}(2)$ structure

In this section, we present some mathematical tools to better face the computations in $\mathrm{SE}(2) \times \mathfrak{se}(2)$.

Let (g, Ω) be an element of $\overline{G} = \mathrm{SE}(2) \times \mathfrak{se}(2)$, we represent it in matrix form as

$$\begin{bmatrix} g & 0 & 0 \\ 0 & I & \Omega \\ 0 & 0 & I \end{bmatrix}.$$

The group operation is

$$(g, \Omega) \cdot (f, \Psi) = (gf, \Omega + \Psi),$$

that in matrix form reads

$$\begin{bmatrix} g & 0 & 0 \\ 0 & I & \Omega \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} f & 0 & 0 \\ 0 & I & \Psi \\ 0 & 0 & I \end{bmatrix} = \begin{bmatrix} gf & 0 & 0 \\ 0 & I & \Omega + \Psi \\ 0 & 0 & I \end{bmatrix}, \tag{3.11}$$

and lets to define the left translation on $\overline{G}$ as

$$L_{(g,\Omega)}(f, \Psi) = (gf, \Omega + \Psi).$$

The unit element of this group is $e = (I, 0)$ and for each (g, Ω) its inverse is $(g^{-1}, -\Omega)$.

The Lie algebra of the (product) group $\overline{G}$ is the (product) algebra $\overline{\mathfrak{g}} = \mathfrak{se}(2) \times \mathfrak{se}(2)$, whose generic element (η^g, η^Ω) can be represented as

$$\begin{bmatrix} \eta^g & 0 & 0 \\ 0 & 0 & \eta^\Omega \\ 0 & 0 & 0 \end{bmatrix}.$$

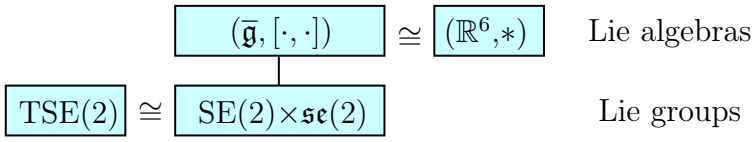


Figure 3.2: Geometric scheme for TSE(2).

$\overline{\mathfrak{g}}$ with the matrix commutator is isomorphic (as Lie algebra) with $(\mathbb{R}^6, *)$ where

$$\begin{bmatrix} \eta_1^\theta \\ \eta_1^x \\ \eta_1^y \\ \eta_1^\omega \\ \eta_1^{v^x} \\ \eta_1^{v^y} \end{bmatrix} * \begin{bmatrix} \eta_2^\theta \\ \eta_2^x \\ \eta_2^y \\ \eta_2^\omega \\ \eta_2^{v^x} \\ \eta_2^{v^y} \end{bmatrix} = \begin{bmatrix} 0 \\ \eta_1^y\eta_2^\theta - \eta_1^\theta\eta_2^y \\ \eta_1^\theta\eta_2^x - \eta_1^x\eta_2^\theta \\ 0 \\ 0 \\ 0 \end{bmatrix}.$$

The geometric scheme for TSE(2) is summarized in Figure 3.2.

We define the "product" between an element of $\overline{G}$ and an element of its Lie algebra as

$$(g, \Omega) * (\eta^g, \eta^\Omega) = (g\eta^g, \eta^\Omega),$$

that in matrix form reads:

$$\begin{bmatrix} g & 0 & 0 \\ 0 & I & \Omega \\ 0 & 0 & I \end{bmatrix} \begin{bmatrix} \eta^g & 0 & 0 \\ 0 & 0 & \eta^\Omega \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} g\eta^g & 0 & 0 \\ 0 & 0 & \eta^\Omega \\ 0 & 0 & 0 \end{bmatrix}.$$

The tangent map is

$$T_{(I,0)}L_{(g,\Omega)}(\eta^g, \eta^\Omega) = (g\eta^g, \eta^\Omega).$$

Eventually, we define the Lie algebra operation

$$(\eta^g, \eta^\Omega) \star (\eta^f, \eta^\Psi) = \Big(\eta^g\eta^f, 0\Big),$$

that in matrix form is

$$\begin{bmatrix} \eta^g & 0 & 0 \\ 0 & 0 & \eta^\Omega \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \eta^f & 0 & 0 \\ 0 & 0 & \eta^\Psi \\ 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} \eta^g\eta^f & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}.$$

The adjoint representation of the Lie algebra into itself is

$$\mathrm{ad}_{(\eta^g, \eta^\Omega)}(\eta^f, \eta^\Psi) = (\mathrm{ad}_{\eta^g}\eta^f, 0)$$

and its matrix form is represented by the 6×6 matrix

$$\mathrm{ad}_{(\eta^g, \eta^v)} = \begin{bmatrix} \mathrm{ad}_{\eta^g} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}. \tag{3.12}$$

In order to feature the notion of differentiation of vector fields on $\overline{G}$, it is necessary to introduce an affine connection. A left-invariant affine connection

∇ on $\overline{G}$ is characterized by its bilinear connection function $\omega : \mathfrak{g} \times \mathfrak{g} \to \mathfrak{g}$ through the identity $\nabla_{\overline{g}X}(\overline{g}Y) = \overline{g}\omega(X, Y)$ for all X, $Y \in \overline{\mathfrak{g}}$ (see, e.g., [27], [13]). Thus $\nabla_X Y \in \overline{\mathfrak{g}}$, and, since this map is $\mathbb{R}$-linear, it is a multiplication in $\overline{\mathfrak{g}}$. We choose skew-symmetric connection function of the form $\nabla_X Y = \lambda[X, Y]$, $\lambda \in \mathbb{R}$ (see e.g. [9], [30]). For example, the choices $\lambda = 0, \frac{1}{2}, 1$ define the (-), (0), (+) Cartan-Schouten connections, that have negative, null and positive torsion, respectively.

Recall that given the control system $\dot{\overline{g}} = f(\overline{g}, u, t)$, the left-trivialization of f is defined as $\lambda(\overline{g}, u, t) \coloneqq \overline{g}^{-1} f(\overline{g}, u, t)$. In our case we obtain from (3.4)-(3.5)

$$\dot{\overline{g}} = f(\overline{g}, u, t) = \left(g\Omega, \mathbb{I}^\sharp \mathrm{ad}^*_\Omega \mathbb{I}^\flat \Omega + u^\wedge\right)$$

and

$$\begin{aligned}\lambda(\overline{g}, u, t) &= (g^{-1}, -\Omega)\left(g\Omega, \mathbb{I}^\sharp \mathrm{ad}^*_\Omega \mathbb{I}^\flat \Omega + u^\wedge\right) \\ &= \left(\Omega, \mathbb{I}^\sharp \mathrm{ad}^*_\Omega \mathbb{I}^\flat \Omega + u^\wedge\right) \in \mathfrak{se}(2) \times \mathfrak{se}(2).\end{aligned}$$

3.3 Free rigid body optimal filters

The second-order optimal filter minimizes the cost functional (2.29)-(2.31) where the initial cost map is given by

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 \tag{3.13}$$

where $\|\cdot\|_F^2$ stands for the Frobenius norm, and where the matrix representation of the form $\mathcal{R}$ is

$$R = B^T B, \quad B = \begin{bmatrix} 0_{3\times 3} \\ B_2 \end{bmatrix}. \tag{3.14}$$

The estimation of $\overline{G} \ni \overline{g} \coloneqq (g, \Omega)$ is provided using the measurement equation

$$\mathrm{y}(t) = h(\overline{g}(t), t) + D\varepsilon(t), \tag{3.15}$$

where $h(\overline{g}(t), t)$ is the output map and ε represents the unknown measurement error (modelled as a normalized Gaussian white noise). We design three different filters using three different measurement equations: velocity, velocity and pose, velocity and two GPS-like antennas. For all of these scenarios we consider the Cartan-Schouten (0)-connection form $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$ (see e.g. [22]).

In what follows we will use $\eta^{\overline{g}} = (\eta^g, \eta^\Omega) = (\eta^\theta, \eta^x, \eta^y, \eta^\omega, \eta^{v^x}, \eta^{v^y})^T \in \mathbb{R}^6$ to indicate the vector form of an element of the Lie algebra $\overline{\mathfrak{g}}$ and $\overline{g}\eta^{\overline{g}} = (g\eta^g, \eta^\Omega) = T_e L_{\overline{g}}(\eta^{\overline{g}}) = (\theta', x', y', \omega', v^{x\prime}, v^{y\prime})^T \in \mathbb{R}^6$ for the vector form of an element of the tangent space $T_{\overline{g}}\overline{G}$ where $\theta' = \eta^\theta$, $x' = \eta^x \cos\theta - \eta^y \sin\theta$, $y' = \eta^x \sin\theta + \eta^y \cos\theta$, $\omega' = \eta^\omega$, $v^{x\prime} = \eta^{v^x}$, $v^{y\prime} = \eta^{v^y}$. We will denote with $\widehat{\cdot}$ an estimated variable.

We are now ready to design three different filters corresponding to the three different measurement equations.

First case

In the first case, the measurement equation is given by

$$h_1(\overline{g}(t), t) = (\Omega)^\vee = \begin{bmatrix} \omega \\ v^x \\ v^y \end{bmatrix} \tag{3.16}$$

that represents the velocities provided, for example, by an IMU and an odometer. For simplicity, the D matrix in (3.15) takes the form

$$D_1 = \text{diag}\{d_3, d_4, d_4\},\ d_i \in \mathbb{R}^+, \tag{3.17}$$

while the Q matrix associated to the quadratic form $\mathcal{Q}$ in (2.31) becomes

$$Q_1 = \text{diag}\{q_3, q_4, q_4\},\ q_i \in \mathbb{R}^+. \tag{3.18}$$

Proposition 3.1. *Consider the dynamic system* (3.8)-(3.9) *with measurement equation* (2.28) *where the output map h and the linear map D are given by* (3.16) *and* (3.17)*, respectively. Consider the cost functional* (2.29)-(2.31) *where the initial cost m_0 is given by* (3.13) *and the matrix representation of the forms $\mathcal{R}$, $\mathcal{Q}$ are given by* (3.14) *and* (3.18)*, respectively. Then the second-order optimal filter is*

$$\begin{aligned} \widehat{g}^{-1}\dot{\widehat{g}} &= \widehat{\Omega} + (K_{11}r^g + K_{12}r^\Omega)^\wedge \\ \dot{\widehat{\Omega}} &= \mathbb{I}^\sharp \text{ad}^*_{\widehat{\Omega}} \mathbb{I}^\flat \widehat{\Omega} + u^\wedge + (K_{21}r^g + K_{22}r^\Omega)^\wedge \end{aligned} \tag{3.19}$$

where the residual $r_t = \begin{bmatrix} r^g & r^\Omega \end{bmatrix}$, with $r^g,\ r^\Omega \in \mathbb{R}^3$, is

$$r_t = \begin{bmatrix} r^g \\ r^\Omega \end{bmatrix}^T = \begin{bmatrix} 0_{3\times 1} \\ \text{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\}\ (y - \widehat{y}) \end{bmatrix}^T, \tag{3.20}$$

and

$$\widehat{y} = (\widehat{\Omega})^\vee. \tag{3.21}$$

The second-order optimal gain $K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$ is the solution of the perturbed matrix Riccati differential equation

$$\begin{aligned} \dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K, r_t)K - KW(K, r_t)^T \end{aligned} \tag{3.22}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ -\widehat{v}^y & 0 & \widehat{\omega} & 0 & 1 & 0 \\ \widehat{v}^x & -\widehat{\omega} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widehat{v}^y & 0 & \widehat{\omega} \\ 0 & 0 & 0 & -\widehat{v}^x & -\widehat{\omega} & 0 \end{bmatrix}, \tag{3.23}$$

$$E = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & \operatorname{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \end{bmatrix}, \tag{3.24}$$

$$BR^{-1}B^T = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & B_2R^{-1}B_2^T \end{bmatrix}, \tag{3.25}$$

$$W(K, r_t) = \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^\Omega)^\wedge} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}. \tag{3.26}$$

The initial conditions for the estimation and the operator K are

$$\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{3.27}$$

$$K(t_0) = \operatorname{diag}\{1/2, 1, 1, 1/2, 1, 1\}. \tag{3.28}$$

Proof. The proof is split into different sections to make it more readable.

Computation of r_t

The expression for the residual $r_t(\widehat{\overline{g}}) \in \overline{\mathfrak{g}}^*$ is defined as

$$r_t(\widehat{\overline{g}}) = T_eL^*_{\widehat{\overline{g}}}[\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}}))\big) \circ \mathrm{d}h_t(\widehat{\overline{g}})]. \tag{3.29}$$

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in T_{\widehat{\overline{g}}}\overline{G}$, we have that the differential of h_t in $\widehat{\overline{g}}$ applied to $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\mathrm{d}h_t(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}(\widehat{\Omega}(s))^\vee = (\eta^\Omega)^\vee,$$

and we can write the operator $\mathrm{d}h_t(\widehat{\overline{g}})$ as

$$\mathrm{d}h_t(\widehat{\overline{g}}) = \begin{bmatrix} 0_{3\times3} & I_{3\times3} \end{bmatrix}. \tag{3.30}$$

From (3.21) and the definition of the matrices (3.17) and (3.18) it follows that

$$(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \left[\operatorname{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T. \tag{3.31}$$

Using (3.30) and (3.31) we obtain

$$\begin{aligned} r_t(\widehat{\overline{g}}) &= \left[\operatorname{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T \begin{bmatrix} 0_{3\times3} & I_{3\times3} \end{bmatrix} \\ &= \left[0_{1\times3} \quad \left(\operatorname{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \ (\mathrm{y} - \widehat{\mathrm{y}})\right)\right]^T. \end{aligned}$$

Computation of A

The operator A represents the coefficients of the linear part of the Riccati equation and its formula $A(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}$ is given by:

$$A(t) = \mathrm{d}_1\lambda_t(\widehat{\overline{g}}, u) \circ T_eL_{\widehat{\overline{g}}} - \mathrm{ad}_{\lambda_t(\widehat{\overline{g}},u)} - T_{\lambda_t(\widehat{\overline{g}},u)}. \tag{3.32}$$

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in T_{\widehat{\overline{g}}}\overline{G}$, we have

$$\begin{aligned} \mathrm{d}_1\lambda(\widehat{\overline{g}},u)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\lambda_g(s)\\ \lambda_\Omega(s)\end{bmatrix} \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\widehat{\omega}(s)\\ \widehat{v}^x(s)\\ \widehat{v}^y(s)\\ \tau/J\\ \widehat{\omega}\widehat{v}^y+F/m\\ -\widehat{\omega}\widehat{v}^x\end{bmatrix} = \begin{bmatrix}\widehat{\omega}'\\ \widehat{v}^{x\prime}\\ \widehat{v}^{y\prime}\\ 0\\ \widehat{\omega}'\widehat{v}^y+\widehat{\omega}\widehat{v}^{y\prime}\\ -\widehat{\omega}'\widehat{v}^x-\widehat{\omega}\widehat{v}^{x\prime}\end{bmatrix} \end{aligned} \tag{3.33}$$

and thus

$$\mathrm{d}_1\lambda(\widehat{\overline{g}},u)\circ T_eL_{\widehat{\overline{g}}} = \begin{bmatrix}0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1\\ 0&0&0&0&0&0\\ 0&0&0&\widehat{v}^y&0&\widehat{\omega}\\ 0&0&0&-\widehat{v}^x&-\widehat{\omega}&0\end{bmatrix}. \tag{3.34}$$

The adjoint matrix representation (3.12) implies

$$\mathrm{ad}_{\lambda_t(\widehat{\overline{g}},u)} = \begin{bmatrix}0&0&0&0_{1\times3}\\ \widehat{v}^y&0&-\widehat{\omega}&0_{1\times3}\\ -\widehat{v}^x&\widehat{\omega}&0&0_{1\times3}\\ 0_{3\times1}&0_{3\times1}&0_{3\times1}&0_{3\times3}\end{bmatrix}. \tag{3.35}$$

Considering the Cartan-Schouten (0)-connection form $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$, the torsion function T vanishes (see [22]), thus, in matrix form, it is given by

$$T_{\lambda_t(\widehat{\overline{g}},u)} = \begin{bmatrix}0_{3\times3}&0_{3\times3}\\ 0_{3\times3}&0_{3\times3}\end{bmatrix}. \tag{3.36}$$

Using (3.34), (3.35) and (3.36) we obtain the matrix (3.23).

Computation of E

The function $E(t)$ in (2.38) extended to the group $\overline{G}$ is $E(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ and takes the form

$$\begin{aligned} E(t) = -T_eL^*_{\widehat{\overline{g}}}\circ\big[\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}}\circ \mathrm{Hess}\,h_t(\widehat{\overline{g}}) \\ -(\mathrm{d}h_t(\widehat{\overline{g}}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ\mathrm{d}h_t(\widehat{\overline{g}})\big]\circ T_eL_{\widehat{\overline{g}}}. \end{aligned} \tag{3.37}$$

The dual operators $(\mathrm{d}h_t(\widehat{\overline{g}}))^*$ and $T_eL^*_{\widehat{\overline{g}}}$ are $(\mathrm{d}h_t(\widehat{\overline{g}}))^T$ and $(T_eL_{\widehat{\overline{g}}})^T$ respectively. We can find the compositions

$$(\mathrm{d}h_t(\widehat{\overline{g}}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ\mathrm{d}h_t(\widehat{\overline{g}}) = \mathrm{diag}\left\{0,0,0,\frac{q_3}{d_3^2},\frac{q_4}{d_4^2},\frac{q_4}{d_4^2}\right\} \tag{3.38}$$

and

$$(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \left[\mathrm{diag}\left\{\tfrac{q_3}{d_3^2}, \tfrac{q_4}{d_4^2}, \tfrac{q_4}{d_4^2}\right\} \ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T. \tag{3.39}$$

The Hessian of h_t in $\widehat{\overline{g}}$ applied to $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is the null operator since it is the differential of (3.30), which is constant. Combining (3.38) and (3.39) with the tangent operator and its dual, we end up with

$$E = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & \mathrm{diag}\left\{\tfrac{q_3}{d_3^2}, \tfrac{q_4}{d_4^2}, \tfrac{q_4}{d_4^2}\right\} \end{bmatrix}.$$

Computation of W

From the adjoint matrix form (3.12) and recalling that we consider the Cartan-Schouten (0)-connection ([33, 34]), we have

$$W(K, r_t) = \frac{1}{2}\mathrm{ad}_{((K_{11}r^g + K_{12}r^\Omega)^\wedge,\ (K_{21}r^g + K_{22}r^\Omega)^\wedge)} = \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^\Omega)^\wedge} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}.$$

Initial condition

The initial condition for the filter is given by (2.33) while the initial condition for the gain is $K(t_0) = X_0^{-1}$ where the operators $X_0 : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ satisfies (2.36).

We rewrite m_0 as

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 = \frac{1}{2}\mathrm{tr}\left[(I_{9\times9} - \overline{g}^{-1}\overline{g}_0)^T(I_{9\times9} - \overline{g}^{-1}\overline{g}_0)\right]. \tag{3.40}$$

From (3.11) it easily follows

$$I_{9\times9} - \overline{g}^{-1}\overline{g}_0 = \begin{bmatrix} I_{3\times3} - g^{-1}g_0 & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & \Omega_0 - \Omega \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \end{bmatrix} \tag{3.41}$$

and thus

$$\begin{aligned} &(I_{9\times9} - \overline{g}^{-1}\overline{g}_0)^T(I_{9\times9} - \overline{g}^{-1}\overline{g}_0) \\ &\quad = \begin{bmatrix} (I_{3\times3} - g^{-1}g_0)^T(I_{3\times3} - g^{-1}g_0) & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} & (\Omega_0 - \Omega)^T(\Omega_0 - \Omega) \end{bmatrix}. \end{aligned} \tag{3.42}$$

Computing the trace we obtain

$$\begin{aligned} m_0(\overline{g}) =& \frac{1}{2}[4(1 - \cos(\theta - \theta_0)) + (x - x_0)^2 + (y - y_0)^2 \\ &+ 2(\omega - \omega_0)^2 + (v^x - v_0^x)^2 + (v^y - v_0^y)^2] \end{aligned} \tag{3.43}$$

and from (2.33) it follows that $\widehat{\overline{g}}(t_0) = \overline{g}_0$.

The Hessian of the function m_0 at a point $\overline{g} \in \overline{G}$ is defined as

$$\mathrm{Hess}\, m_0(\overline{g})(\overline{g}X)(\overline{g}Y) = \mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) - \mathrm{d}m_0(\overline{g})(\nabla_{\overline{g}X}(\overline{g}Y))$$

for all $\overline{g}X$, $\overline{g}Y \in T_{\overline{g}}\overline{G}$. The differential of m_0 is given by

$$\mathrm{d}m_0(\overline{g}) = \begin{bmatrix} 2\sin(\theta-\theta_0) & (x-x_0) & (y-y_0) & 2(\omega-\omega_0) & (v^x-v_0^x) & (v^y-v_0^y) \end{bmatrix} \tag{3.44}$$

while, given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{g}\eta^{g_1}, \eta^{\Omega_1})$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) \in T_{\overline{\widehat{g}}\overline{G}}$, the affine connection yields

$$\nabla_{(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})}(\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) = \frac{1}{2}\begin{bmatrix} 0 \\ \widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3.45}$$

Combining (3.44) and (3.45) we obtain

$$\begin{aligned} &\mathrm{d}m_0(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) \\ &\quad = \Big[(x-x_0)(\widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2') + (y-y_0)(-\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2')\Big] \end{aligned} \tag{3.46}$$

that evaluating in $\widehat{\overline{g}}_0$ produces

$$\mathrm{d}m_0(\widehat{\overline{g}}_0)(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) = 0. \tag{3.47}$$

The double differential takes the form

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{2\cos(\theta-\theta_0), 1, 1, 2, 1, 1\} \tag{3.48}$$

that, evaluating in $\widehat{\overline{g}}_0$, produces

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{2, 1, 1, 2, 1, 1\} \tag{3.49}$$

and thus, from (3.47) and (3.49)

$$\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) = \mathrm{diag}\{2, 1, 1, 2, 1, 1\}. \tag{3.50}$$

From

$$K(t_0) = X_0^{-1} = T_eL_{\widehat{\overline{g}}_0}^* \circ \mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) \circ T_eL_{\widehat{\overline{g}}_0} \tag{3.51}$$

we obtain the initial condition of $K(t_0)$

$$\begin{aligned} K(t_0) &= (T_eL_{\widehat{\overline{g}}_0})^{-1}(\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0))^{-1}(T_eL_{\widehat{\overline{g}}_0}^*)^{-1} \\ &= \mathrm{diag}\{1/2, 1, 1, 1/2, 1, 1\}. \end{aligned} \tag{3.52}$$

This computation ends the proof. □

Second case

In the second case, we consider as measurement equation

$$h_2(\overline{g}(t),t) = \begin{bmatrix} \theta(t) \\ x(t) \\ y(t) \\ \omega(t) \\ v^x(t) \\ v^y(t) \end{bmatrix} \tag{3.53}$$

in which we measure both the pose (θ, x, y) and the velocity components (ω, v^x, v^y). The D and Q matrices in this case take the forms

$$D_2 = \mathrm{diag}\{d_1, d_2, d_2, d_3, d_4, d_4\},\ d_i \in \mathbb{R}^+, \tag{3.54}$$

$$Q_2 = \mathrm{diag}\{q_1, q_2, q_2, q_3, q_4, q_4\},\ q_i \in \mathbb{R}^+. \tag{3.55}$$

Proposition 3.2. *Consider the dynamic system* (3.8)-(3.9) *with measurement equation* (2.28) *where the output map h and the linear map D are given by* (3.53) *and* (3.54)*, respectively. Consider the cost functional* (2.29)-(2.31) *where the initial cost m_0 is given by* (3.13) *and the matrix representation of the forms $\mathcal{R}$, $\mathcal{Q}$ are given by* (3.14) *and* (3.55)*, respectively. Then the second-order optimal filter is*

$$\begin{aligned} \widehat{g}^{-1}\dot{\widehat{g}} &= \widehat{\Omega} + (K_{11}r^g + K_{12}r^\Omega)^\wedge \\ \dot{\widehat{\Omega}} &= \mathbb{I}^\sharp \mathrm{ad}^*_{\widehat{\Omega}} \mathbb{I}^\flat \widehat{\Omega} + u^\wedge + (K_{21}r^g + K_{22}r^\Omega)^\wedge \end{aligned} \tag{3.56}$$

where the residual r_t is

$$r_t = \begin{bmatrix} \frac{q_1}{d_1^2}(\mathrm{y}_1 - \widehat{\mathrm{y}}_1) \\ \frac{q_2}{d_2^2}\cos(\theta)(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) + \frac{q_2}{d_2^2}\sin(\theta)(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) \\ -\frac{q_2}{d_2^2}\sin(\theta)(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) + \frac{q_2}{d_2^2}\cos(\theta)(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) \\ \frac{q_3}{d_3^2}(\mathrm{y}_4 - \widehat{\mathrm{y}}_4) \\ \frac{q_4}{d_4^2}(\mathrm{y}_5 - \widehat{\mathrm{y}}_5) \\ \frac{q_4}{d_4^2}(\mathrm{y}_6 - \widehat{\mathrm{y}}_6) \end{bmatrix}^T.$$

The second-order optimal gain $K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$ is the solution of the perturbed matrix Riccati differential equation

$$\begin{aligned} \dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K, r_t)K - KW(K, r_t)^T \end{aligned} \tag{3.57}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ -\widehat{v}^y & 0 & \widehat{\omega} & 0 & 1 & 0 \\ \widehat{v}^x & -\widehat{\omega} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widehat{v}^y & 0 & \widehat{\omega} \\ 0 & 0 & 0 & -\widehat{v}^x & -\widehat{\omega} & 0 \end{bmatrix}, \tag{3.58}$$

$$E = \begin{bmatrix} F & 0_{3\times 3} \\ 0_{3\times 3} & \operatorname{diag}\left\{\frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \end{bmatrix}$$

with

$$F = \begin{bmatrix} \frac{q_1}{d_1^2} & -\frac{q_2}{2d_2^2}(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) & \frac{q_2}{2d_2^2}(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) \\ \frac{q_2}{2d_2^2}(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) & \frac{q_2}{d_2^2} & 0 \\ -\frac{q_2}{2d_2^2}(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) & 0 & \frac{q_2}{d_2^2} \end{bmatrix},$$

$$BR^{-1}B^T = \begin{bmatrix} 0_{3\times 3} & 0_{3\times 3} \\ 0_{3\times 3} & B_2 R^{-1} B_2^T \end{bmatrix},$$

$$W(K, r_t) = \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^\Omega)^\wedge} & 0_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix}.$$

The initial conditions for the estimation and the operator K are

$$\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{3.59}$$
$$K(t_0) = \operatorname{diag}\{1/2, 1, 1, 1/2, 1, 1\}. \tag{3.60}$$

Proof. **Computation of r_t**

The extension of (2.34) to the group $\overline{G}$ is given by (3.29). Given $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in T_{\widehat{\overline{g}}}\overline{G}$, the differential of h_t in $\widehat{\overline{g}}$ applied to $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\mathrm{d}h_t(\widehat{\overline{g}})(T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} (\widehat{g}(s))^\vee \\ (\widehat{\Omega}(s))^\vee \end{bmatrix} = \begin{bmatrix} (\eta^{\widehat{g}})^\vee \\ (\eta^{\widehat{\Omega}})^\vee \end{bmatrix},$$

and we can write the operator $\mathrm{d}h_t(\widehat{\overline{g}})$ as

$$\mathrm{d}h_t(\widehat{\overline{g}}) = I_{6\times 6}. \tag{3.61}$$

From the definition of the matrices (3.54) and (3.55) it follows that

$$\begin{aligned} &(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \\ &= \left[\operatorname{diag}\left\{\frac{q_1}{d_1^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} (\mathrm{y} - \widehat{\mathrm{y}})\right]^T. \end{aligned} \tag{3.62}$$

Evaluating $T_e L^*_{\widehat{\overline{g}}}$ on (3.62) the result follows.

Computation of A

As in the proof of Proposition 3.1.

Computation of E

We can find the compositions

$$(\mathrm{d}h_t(\widehat{\overline{g}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h_t(\widehat{\overline{g}}) = \mathrm{diag}\left\{\frac{q_1}{d_1^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \tag{3.63}$$

and

$$(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \left[\mathrm{diag}\left\{\frac{q_1}{d_1^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T. \tag{3.64}$$

Given the twice differentiable function $h_t : \overline{G} \to \mathbb{R}^6$ and $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{g}\eta^{g_1}, \eta^{\Omega_1})$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) \in T_{\overline{g}\overline{G}}$, the Hessian operator $\mathrm{Hess}\, h_t(\overline{g}) : T_{\overline{g}}\overline{G} \to T_{\overline{g}}\overline{G}^*$ is defined by

$$\begin{aligned}\mathrm{Hess}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})) =& \mathrm{d}(\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) \\ &- \mathrm{d}h(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))\end{aligned} \tag{3.65}$$

and we have that

$$\mathrm{d}(\mathrm{d}h_t(\widehat{\overline{g}})(\widehat{g}\eta^{g_2}, \eta^{\Omega_2}))(\widehat{g}\eta^{g_1}, \eta^{\Omega_1}) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\left[(\widehat{g}\eta^{g_2}, \eta^{\Omega_2})\right] = \left[0_{1\times 6}\right]. \tag{3.66}$$

Moreover, since we are working with Cartan-Schouten (0)-connection, we get

$$\nabla_{(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})}(\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) = \frac{1}{2}\widehat{\overline{g}}\ \mathrm{ad}_{(\eta^{g_1}, \eta^{\Omega_1})}(\eta^{g_2}, \eta^{\Omega_2}) = \frac{1}{2}\ (\widehat{g}\ \mathrm{ad}_{\eta^{g_1}}\eta^{g_2}, 0) \tag{3.67}$$

where

$$\widehat{g}\ \mathrm{ad}_{\eta^{g_1}}\eta^{g_2} = \frac{1}{2}\begin{bmatrix} 0 \\ (\widehat{v}_1^y\widehat{\omega}_2 - \widehat{\omega}_1\widehat{v}_2^y)\cos\widehat{\theta} - (-\widehat{v}_1^x\widehat{\omega}_2 + \widehat{\omega}_1\widehat{v}_2^x)\sin\widehat{\theta} \\ (\widehat{v}_1^y\widehat{\omega}_2 - \widehat{\omega}_1\widehat{v}_2^y)\sin\widehat{\theta} + (-\widehat{v}_1^x\widehat{\omega}_2 + \widehat{\omega}_1\widehat{v}_2^x)\cos\widehat{\theta} \end{bmatrix} \tag{3.68}$$

and so

$$\mathrm{d}h_t(\widehat{\overline{g}})(\nabla_{(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})}(\widehat{g}\eta^{g_2}, \eta^{\Omega_2})) = \frac{1}{2}\begin{bmatrix} 0 \\ \widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3.69}$$

Thus, the Hessian evaluated in $(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})$ and $(\widehat{g}\eta^{g_2}, \eta^{\Omega_2})$ takes the form

$$\mathrm{Hess}h_t(\widehat{\overline{g}})(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})(\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) = -\frac{1}{2}\begin{bmatrix} 0 \\ \widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \\ 0 \end{bmatrix}. \tag{3.70}$$

From (3.64) and (3.70) it follows that

$$\left((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}}))\right)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \mathrm{Hess}\, h_t(\widehat{\overline{g}}) = \begin{bmatrix} F & 0_{3\times 3} \\ 0_{3\times 3} & 0_{3\times 3} \end{bmatrix} \tag{3.71}$$

with

$$F = \begin{bmatrix} 0 & -\frac{q_2}{2d_2^2}(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) & \frac{q_2}{2d_2^2}(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) \\ \frac{q_2}{2d_2^2}(\mathrm{y}_3 - \widehat{\mathrm{y}}_3) & 0 & 0 \\ -\frac{q_2}{2d_2^2}(\mathrm{y}_2 - \widehat{\mathrm{y}}_2) & 0 & 0 \end{bmatrix}$$

Combining (3.63) and (3.71) the result follows.

Computation of W

As in the proof of Proposition 3.1.

Initial condition

As in the proof of Proposition 3.1.

□

Third case

In the third case, we consider a measurement equation given by

$$h_3(\overline{g}(t), t) = \begin{bmatrix} x(t) + \ell\, \cos(\theta(t)) \\ y(t) + \ell\, \sin(\theta(t)) \\ x(t) - \ell\, \cos(\theta(t)) \\ y(t) - \ell\, \sin(\theta(t)) \\ \omega(t) \\ v^x(t) \\ v^y(t) \end{bmatrix} \tag{3.72}$$

which models two GPS-like systems at distance ℓ from the center of mass (first four rows), an IMU and an odometer (last three rows). The GPS measurements are the positions of two antennas attached to the body at distance ℓ from the center of gravity as shown in Figure 3.3. The D and Q matrices become

$$D_3 = \mathrm{diag}\{d_2, d_2, d_2, d_2, d_3, d_4, d_4\},\ d_i \in \mathbb{R}^+, \tag{3.73}$$

$$Q_3 = \mathrm{diag}\{q_2, q_2, q_2, q_2, q_3, q_4, q_4\},\ q_i \in \mathbb{R}^+. \tag{3.74}$$

Proposition 3.3. *Consider the dynamic system* (3.8)-(3.9) *with measurement equation* (2.28) *where the output map h and the linear map D are given by* (3.72) *and* (3.73)*, respectively. Consider the cost functional* (2.29)-(2.31) *where the initial cost m_0 is given by* (3.13) *and the matrix representation of the forms $\mathcal{R}$, $\mathcal{Q}$ are given by* (3.14) *and* (3.74)*, respectively. Then the second-order optimal filter is*

$$\begin{aligned} \widehat{g}^{-1}\dot{\widehat{g}} &= \widehat{\Omega} + (K_{11} r^g + K_{12} r^\Omega)^\wedge \\ \dot{\widehat{\Omega}} &= \mathbb{I}^\sharp \mathrm{ad}^*_{\widehat{\Omega}} \mathbb{I}^\flat \widehat{\Omega} + u^\wedge + (K_{21} r^g + K_{22} r^\Omega)^\wedge \end{aligned} \tag{3.75}$$

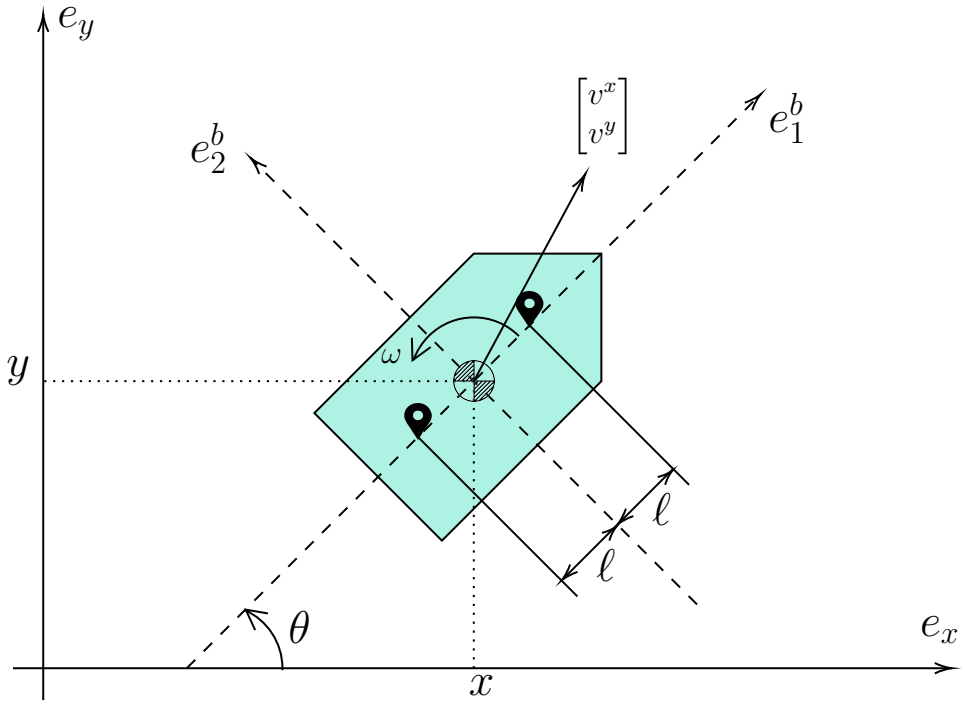


Figure 3.3: Planar rigid body with two antennas.

where the residual r_t *is*

$$r_t = \begin{bmatrix} r^g \\ r^\Omega \end{bmatrix} = \begin{bmatrix} -(\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell \sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell \cos\widehat{\theta} \\ (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} \\ -(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} \\ \widetilde{\mathrm{y}}_5 \\ \widetilde{\mathrm{y}}_6 \\ \widetilde{\mathrm{y}}_7 \end{bmatrix}^T, \tag{3.76}$$

and

$$\mathbb{R}^7 \ni \widetilde{\mathrm{y}} = \left[\mathrm{diag}\left\{\frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\}\right](\mathrm{y} - \widehat{\mathrm{y}}). \tag{3.77}$$

The second-order optimal gain $K = \begin{bmatrix} K_{11} & K_{12} \\ K_{21} & K_{22} \end{bmatrix}$ *is the solution of the perturbed matrix Riccati differential equation*

$$\begin{aligned} \dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K, r_t)K - KW(K, r_t)^T \end{aligned} \tag{3.78}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ -\widehat{v}^y & 0 & \widehat{\omega} & 0 & 1 & 0 \\ \widehat{v}^x & -\widehat{\omega} & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widehat{v}^y & 0 & \widehat{\omega} \\ 0 & 0 & 0 & -\widehat{v}^x & -\widehat{\omega} & 0 \end{bmatrix}, \tag{3.79}$$

$$E = \begin{bmatrix} E_1 & 0_{3\times3} \\ 0_{3\times3} & I_{3\times3} \end{bmatrix}, \quad E_1 = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & 2 & 0 \\ a_{3,1} & 0 & 2 \end{bmatrix}, \tag{3.80}$$

with

$$\begin{aligned}
a_{1,1} =& (\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell\, \cos\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell\, \sin\widehat{\theta} + 2\ell^2, \\
a_{1,2} =& -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} + \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta}, \\
a_{1,3} =& +\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} + \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta}, \\
a_{2,1} =& +\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} - \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta}, \\
a_{3,1} =& -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} - \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta},
\end{aligned}$$

$$BR^{-1}B^T = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & B_2R^{-1}B_2^T \end{bmatrix},$$

$$W(K, r_t) = \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^\Omega)^\wedge} & 0_{3\times3} \\ 0_{3\times3} & 0_{3\times3} \end{bmatrix}. \tag{3.81}$$

The initial conditions for the estimation and the operator K are

$$\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{3.82}$$

$$K(t_0) = \mathrm{diag}\{1/2, 1, 1, 1/2, 1, 1\}. \tag{3.83}$$

Proof. **Computation of r_t**

The extension of (2.34) to the group $\overline{G}$ is given by (3.29). Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in T_{\widehat{\overline{g}}}\overline{G}$, the differential of h in $\widehat{\overline{g}}$ applied to $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} \widehat{x}(s) + \ell\, \cos\widehat{\theta}(s) \\ \widehat{y}(s) + \ell\, \sin\widehat{\theta}(s) \\ \widehat{x}(s) - \ell\, \cos\widehat{\theta}(s) \\ \widehat{y}(s) - \ell\, \sin\widehat{\theta}(s) \\ \widehat{\omega}(s) \\ \widehat{v}^x(s) \\ \widehat{v}^y(s) \end{bmatrix} = \begin{bmatrix} \widehat{x}' - \ell\,\widehat{\theta}'\sin\widehat{\theta} \\ \widehat{y}' + \ell\,\widehat{\theta}'\cos\widehat{\theta} \\ \widehat{x}' + \ell\,\widehat{\theta}'\sin\widehat{\theta} \\ \widehat{y}' - \ell\,\widehat{\theta}'\cos\widehat{\theta} \\ \widehat{\omega}' \\ \widehat{v}^{x\prime} \\ \widehat{v}^{y\prime} \end{bmatrix}, \tag{3.84}$$

and thus we can write the operator $\mathrm{d}h_t(\widehat{\overline{g}})$ as

$$\mathrm{d}h_t(\widehat{\overline{g}}) = \begin{bmatrix} -\ell\, \sin\widehat{\theta} & 1 & 0 & 0_{1\times3} \\ +\ell\, \cos\widehat{\theta} & 0 & 1 & 0_{1\times3} \\ +\ell\, \sin\widehat{\theta} & 1 & 0 & 0_{1\times3} \\ -\ell\, \cos\widehat{\theta} & 0 & 1 & 0_{1\times3} \\ 0_{3\times1} & 0_{3\times1} & 0_{3\times1} & I_{3\times3} \end{bmatrix}. \tag{3.85}$$

From the definition of the matrices (3.73) and (3.74) it follows that

$$\begin{aligned}&(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \\ &= \left[\mathrm{diag}\left\{\frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\}\ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T,\end{aligned} \tag{3.86}$$

composing (3.85) and (3.86), and evaluating $T_eL^*_{\widehat{\overline{g}}}$, we obtain (2.34).

Computation of A

As in the proof of Proposition 3.1.

Computation of E

The function $E(t)$ in (2.38) extended to the group $\overline{G}$ is $E(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ and takes the form (3.37). Now we can find the compositions

$$(\mathrm{d}h_t(\widehat{\overline{g}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h_t(\widehat{\overline{g}}) = \mathrm{diag}\left\{2\ell^2\frac{q_2}{d_2^2}, 2\frac{q_2}{d_2^2}, 2\frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\} \tag{3.87}$$

and

$$(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}})) = \left[\mathrm{diag}\left\{\frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}, \frac{q_4}{d_4^2}\right\}\ (\mathrm{y} - \widehat{\mathrm{y}})\right]^T. \tag{3.88}$$

Given the twice differentiable function $h_t : \overline{G} \to \mathbb{R}^7$ and $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{g}\eta^{g_1}, \eta^{\Omega_1})$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) \in T_{\overline{g}\overline{G}}$, the Hessian operator $\mathrm{Hess}\, h_t(\overline{g}) : T_{\overline{g}}\overline{G} \to T_{\overline{g}}\overline{G}^*$ is defined by

$$\begin{aligned}\mathrm{Hess}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})) =& \mathrm{d}(\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) \\ &- \mathrm{d}h(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))).\end{aligned} \tag{3.89}$$

We have

$$\begin{aligned}&\mathrm{d}(\mathrm{d}h_t(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) = \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\widehat{x}_2' - \widehat{\theta}_2'\ell\,\sin\widehat{\theta}(s) \\ \widehat{y}_2' + \widehat{\theta}_2'\ell\,\cos\widehat{\theta}(s) \\ \widehat{x}_2' + \widehat{\theta}_2'\ell\,\sin\widehat{\theta}(s) \\ \widehat{y}_2' - \widehat{\theta}_2'\ell\,\cos\widehat{\theta}(s) \\ 0_{3\times1}\end{bmatrix} = \begin{bmatrix}-\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} \\ -\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} \\ 0_{3\times1}\end{bmatrix}.\end{aligned} \tag{3.90}$$

Since we are working with the Cartan-Schouten (0)-connection, we end up with

$$\nabla_{(\widehat{g}\eta^{g_1}, \eta^{\Omega_1})}(\widehat{g}\eta^{g_2}, \eta^{\Omega_2}) = \frac{1}{2}\widehat{\overline{g}}\ \mathrm{ad}_{(\eta^{g_1}, \eta^{\Omega_1})}(\eta^{g_2}, \eta^{\Omega_2}) = \frac{1}{2}\ (\widehat{g}\ \mathrm{ad}_{\eta^{g_1}}\eta^{g_2}, 0) \tag{3.91}$$

and so

$$\mathrm{d}h_t(\widehat{\overline{g}})(\nabla_{(\widehat{g}\eta^{g_1},\eta^{\Omega_1})}(\widehat{g}\eta^{g_2},\eta^{\Omega_2})) = \frac{1}{2}\begin{bmatrix} +\widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ +\widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0_{3\times 1} \end{bmatrix}. \tag{3.92}$$

The Hessian evaluated in $(\widehat{g}\eta^{g_1},\eta^{\Omega_1})$ and $(\widehat{g}\eta^{g_2},\eta^{\Omega_2})$ takes the form

$$\mathrm{Hess}h_t(\widehat{\overline{g}})(\widehat{g}\eta^{g_1},\eta^{\Omega_1})(\widehat{g}\eta^{g_2},\eta^{\Omega_2}) = \begin{bmatrix} -\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ 0_{3\times 1} \end{bmatrix}. \tag{3.93}$$

From (3.88) and (3.93) it follows that

$$\begin{gathered} \left((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\overline{g}}))\right)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \mathrm{Hess}\, h_t(\widehat{\overline{g}}) = \\ \begin{bmatrix} a_{1,1} & -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & 0_{1\times 3} \\ \frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & 0 & 0 & 0_{1\times 3} \\ -\frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & 0 & 0 & 0_{1\times 3} \\ 0_{3\times 1} & 0_{3\times 1} & 0_{3\times 1} & 0_{3\times 3} \end{bmatrix}, \end{gathered} \tag{3.94}$$

where $a_{1,1} = -\widetilde{\mathrm{y}}_1\ell\,\cos\widehat{\theta} - \widetilde{\mathrm{y}}_2\ell\,\sin\widehat{\theta} + \widetilde{\mathrm{y}}_3\ell\,\cos\widehat{\theta} + \widetilde{\mathrm{y}}_4\ell\,\sin\widehat{\theta}$. In conclusion, combining (3.87) and (3.94) with $T_eL^*_{\widehat{\overline{g}}}$ and $T_eL_{\widehat{\overline{g}}}$ the result follows.

Computation of W

As in the proof of Proposition 3.1.

Initial condition

As in the proof of Proposition 3.1.

□

3.4 Simulations and discussions

In this section we explain how the second-order optimal minimum-energy filters on $\mathrm{TSE}(2) \simeq \mathrm{SE}(2) \times \mathfrak{se}(2)$ works.

The inertia tensor of the planar rigid body under consideration (in Figure 3.1 and Figure 3.3) is diagonal

$$\mathbb{I} = \begin{bmatrix} J & 0 & 0 \\ 0 & m & 0 \\ 0 & 0 & m \end{bmatrix},$$

where the inertia along the z axis is $J = 6.125\,\mathrm{kg\,m^2}$ and the mass is $m = 125\,\mathrm{kg}$. The matrix representation of $\mathbb{I}^\sharp$ is equal to the inverse of $\mathbb{I}$ and $\mathbb{I}^\flat = \mathbb{I}$, as well. We consider the control inputs $\tau = 50\sin(\frac{1}{50}t)$ and $F = \sin(t)$ with total simulation time $T = 60\,\mathrm{s}$. These inputs are used to generate the trajectory g.

To solve all the differential equations we use a Runge-Kutta 4^{th} order method with integration step $T_s = 1\mathrm{ms}$. The initial conditions are $\theta(0) = \pi/6$ rad, $x(0) = 1\,\mathrm{m}$, $y(0) = 2\,\mathrm{m}$, $\omega(0) = 0\,\mathrm{rad/s}$, $v^x(0) = 0\,\mathrm{m/s}$ and $v^y(0) = 0\,\mathrm{m/s}$. The matrices B_2, R and B related to the model uncertainty (2.27), (2.35) are

$$
\begin{aligned}
B_2 &= \mathrm{diag}\{0.1, 0.1, 0.1\},\\
B &= \begin{bmatrix} 0_{3\times 3} \\ B_2 \end{bmatrix},\\
R &= B^T B.
\end{aligned}
$$

Case 1: only velocity measurement

The measured output is obtained by adding a Gaussian white noise with zero mean and standard deviation depending on D_1. For the matrices D_1 and Q_1 in (3.17) and (3.18), we choose $d_3 = 0.1$, $d_4 = 0.1$ and $q_i = d_i^2$, $i = 3, 4$. The results of the numerical simulation are shown in Figure 3.5. Comparing the actual and filtered poses, it can be seen that without measurements on the group elements we have little biases on x, y, and θ. This is not surprising because the state of the system (3.4)-(3.5) is not observable with the measurement equation (3.16) (see e.g. [16], [40]).

Case 2: positions and velocity measurements

In the case of measurement equation (3.53), the measured output is obtained by adding a Gaussian white noise with zero mean and standard deviation, depending on D_2, to the pose obtained integrating (3.8) and to the velocities obtained integrating (3.9). For the matrices D_2 and Q_2 in (3.54) and (3.55) respectively, we chose $d_1 = 1$, $d_2 = 0.5$, $d_3 = 0.1$, $d_4 = 0.1$ and $q_i = d_i^2$, $i = 1, ..., 4$. The initial conditions of the filter are as before. The results of the numerical simulation are shown in Figure 3.6. The filter works well both on the poses and on the velocities. The system is now fully observable.

Case 3: antennas and velocity measurements

The measurement (3.72) is obtained by adding a Gaussian white noise with zero mean and standard deviation depending on D_3 to the velocities computed integrating (3.9) and to the "antennas" components $x_1 = x + \ell\cos(\theta)$, $y_1 = x + \ell\sin(\theta)$, $x_2 = x - \ell\cos(\theta)$, $y_2 = x - \ell\sin(\theta)$ where we chose $\ell = 0.4\,\mathrm{m}$. For the matrix D_3 in (3.73) we chose $d_2 = 0.5$, $d_3 = 0.1$, $d_4 = 0.1$. For the matrix Q_3 in (3.74) we choose $q_i = d_i^2$, $i = 2, ..., 4$. The initial conditions of the filter, the actual and measured velocities, are as before, while the "antennas" outputs are shown in Figure 3.4. The results of the numerical simulation are given in Figure 3.4. Also in this case, the system is observable.

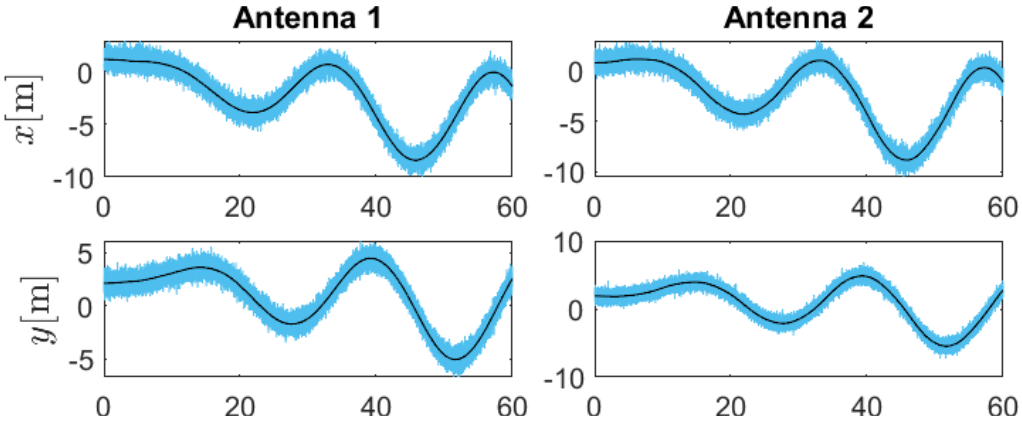


Figure 3.4: Nominal (black) and measured (blue) GPS antennas for the third case.

The three filters proposed in this chapter estimate the pose and velocity of a planar rigid body. Its dynamics is governed by the Euler-Poincaré equation for rigid body that fits well with the formulation via Lie groups. In particular, the underlying geometric structure is represented by the tangent bundle TSE(2) of the special Euclidean group SE(2). The application of the second-order filter is made more feasible thanks to the Lie algebra isomorphism between $\mathfrak{g}$ and $\mathbb{R}^3$. The choice of the measurement equation affects the accuracy of the filter. The first one has as measurement inputs the angular and linear velocity. As can be seen, this filter does not provide good estimations for the pose since the system is not observable. Anyway, the accuracy of the velocity is good. This type of filter can be considered when only measurements on the velocity are available, or when the other measurements on the pose are not accessible. The second filter has as input all the dynamic variables and provides good estimations for all of the dynamic variables. This represents the optimal case when all the sensors provide direct measurements. The third case has as input the angular and linear velocity and the position of two antennas integral to the body. This filter is based on GPS-like and odometer and IMU-like measurements and is the most relevant for applications.

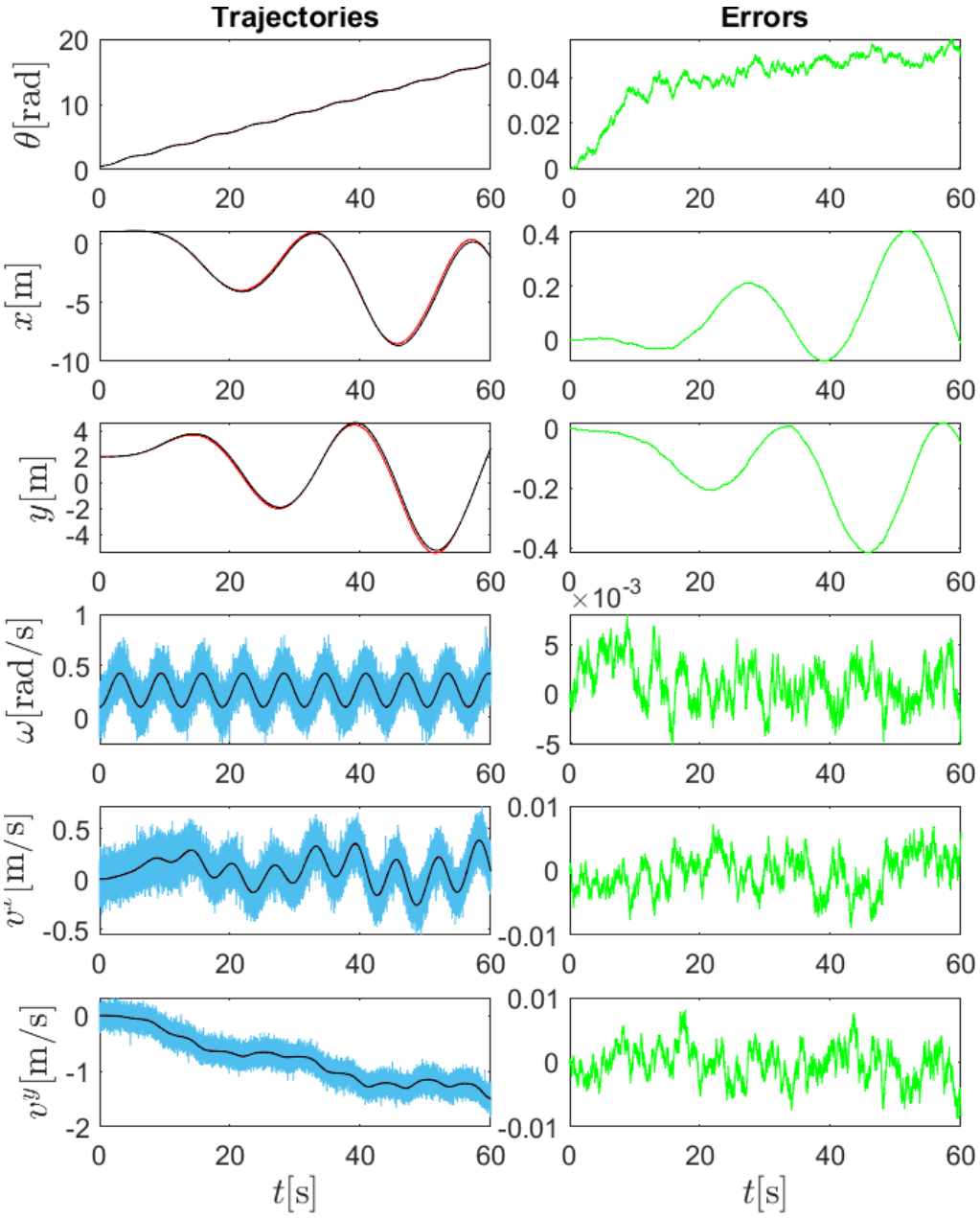


Figure 3.5: Nominal (black), measured (blue), filtered (red) trajectories on the left and their errors (green) on the right for the first case.

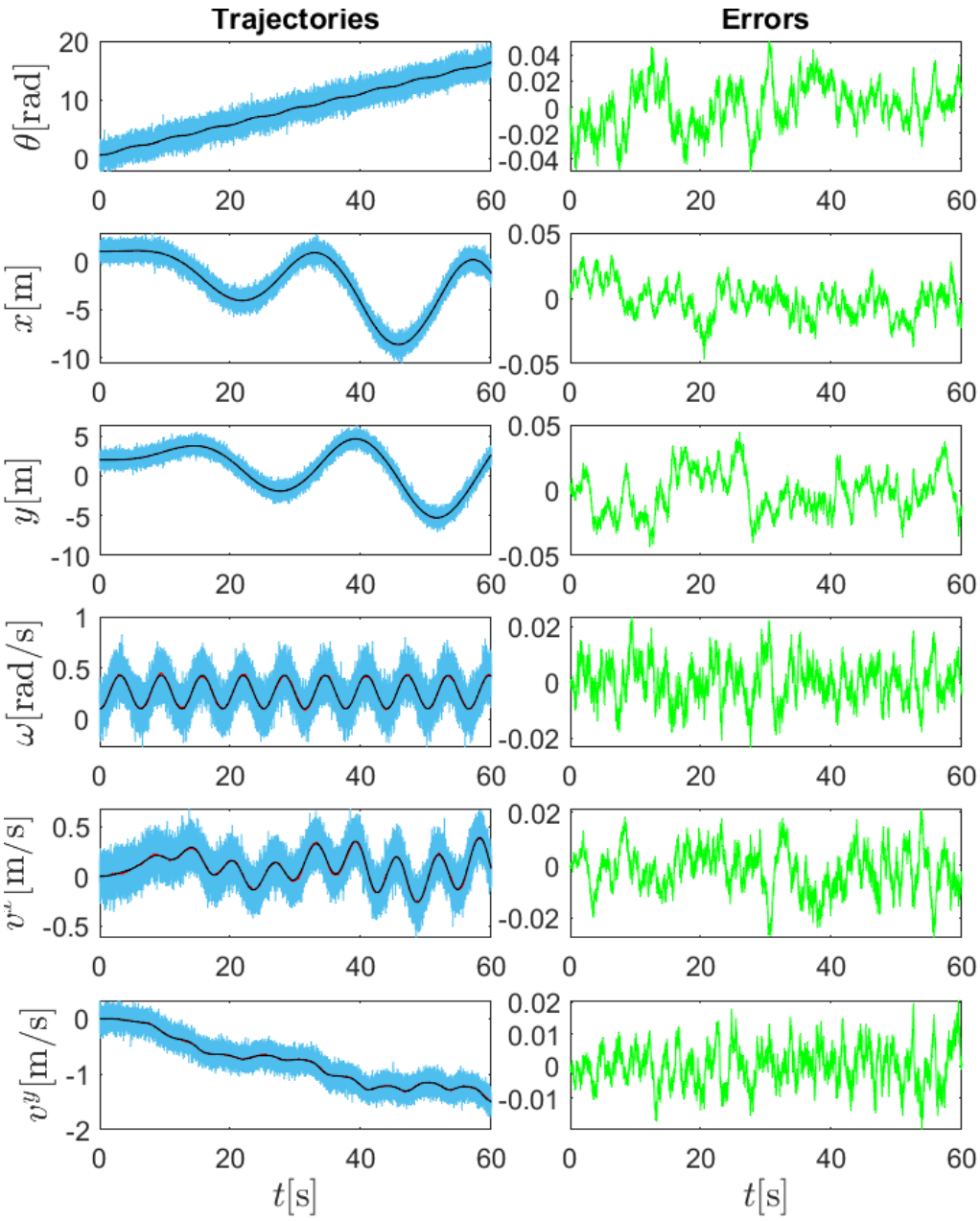


Figure 3.6: Nominal (black), measured (blue), filtered (red) trajectories on the left and their errors (green) on the right for the second case.

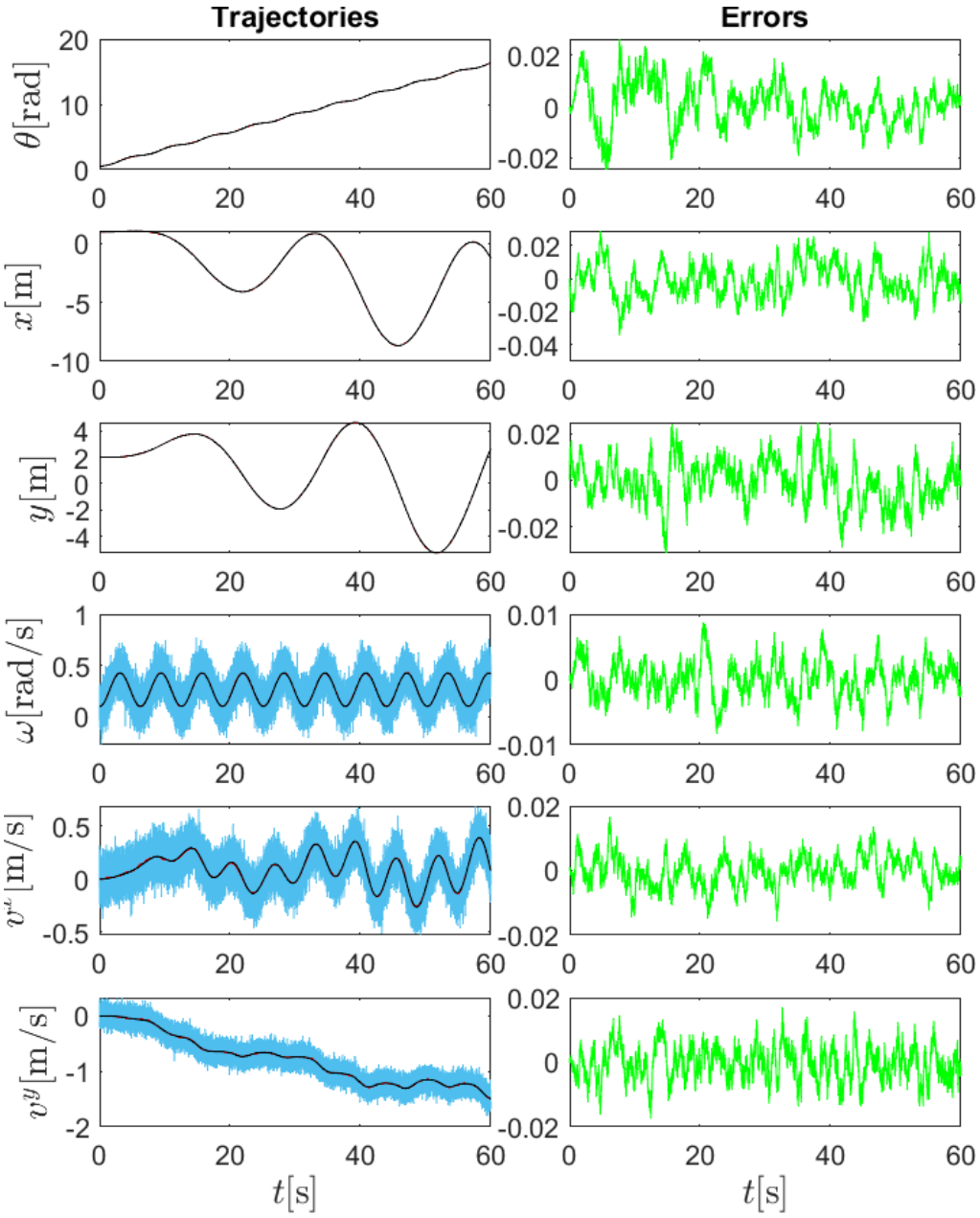


Figure 3.7: Nominal (black), measured (blue), filtered (red) trajectories on the left and their errors (green) on the right for the third case.

3.5 A comparison with the extended Kalman filter

In this section, we compare the second-order filter, derived for the two antennas case, with the extended Kalman filter. There exist two main versions of the extended Kalman filter: the continuous and the discrete. We treat the continuous version since the second-order filter constructed on Lie group is a continuous version too.

The extended Kalman filter does not require any particular geometric structure for the dynamics. Thus, we can consider the full dynamic equation (3.10), written as

$$\dot{\mathrm{x}} = f(\mathrm{x}, u) + \xi$$

where $\mathrm{x} = (\theta, x, y, \omega, v^x, v^y)^T$ and where the model error $\xi = (0, 0, 0, \xi_\omega, \xi_{v^x}, \xi_{v^y})^T$ is modelled as a Gaussian white noise with variance R_K: $\xi \sim \mathcal{N}(0, R_K(t))$. We consider the case of two GPS-like and odometer/IMU sensors studied in Section 3.3 (Figure 3.3). The measurement equation is thus given by

$$\mathrm{y}(t) = h_3(\mathrm{x}(t), t) + \varepsilon$$

where h_3 is provided by equation (3.72) and ε is the measurement error modelled as a Gaussian white noise with variance Q_K: $\varepsilon \sim \mathcal{N}(0, Q_K(t))$. From these definitions and from equations (3.7) and (3.73), we have

$$R_K = \begin{bmatrix} 0_{3\times 3} & 0_{3\times 3} \\ 0_{3\times 3} & B_2^T B_2 \end{bmatrix}, \quad Q_K = D_3^T D_3.$$

We now design the continuous time extended Kalman filter for the free rigid body case. For a rigorous treatment on Kalman filter and extended Kalman filter see e.g. [1], [17], [28].

Proposition 3.4. *For the dynamic model described by* (3.10) *and* (3.72)*, the estimation* $\widehat{\mathrm{x}}(t)$ *and the gain* $P(t)$ *are given by*

- *initial conditions:*

$$\widehat{\mathrm{x}}(t_0) = \mathbb{E}[\mathrm{x}(t_0)], \;\; K(t_0) = \mathrm{Var}[\mathrm{x}(t_0)], \tag{3.95}$$

- *predict-update:*

$$\dot{\widehat{\mathrm{x}}}(t) = f(\widehat{\mathrm{x}}(t), u(t)) + P(t)(\mathrm{y}(t) - h_3(\widehat{\mathrm{x}}(t))), \tag{3.96}$$

$$\dot{K}(t) = F(t)K(t) + K(t)F(t)^T - P(t)H(t)K(t) + R_K(t), \tag{3.97}$$

$$P(t) = K(t)H(t)^T Q_K(t)^{-1}, \tag{3.98}$$

where

$$F(t) = \left.\frac{\partial f}{\partial \mathrm{x}}\right|_{\mathrm{x}=\widehat{\mathrm{x}}(t)} = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 0 \\ -\widehat{v}^x \sin\widehat{\theta} - \widehat{v}^y \cos\widehat{\theta} & 0 & 0 & 0 & \cos\widehat{\theta} & -\sin\widehat{\theta} \\ \widehat{v}^x \cos\widehat{\theta} - \widehat{v}^y \sin\widehat{\theta} & 0 & 0 & 0 & \sin\widehat{\theta} & -\cos\widehat{\theta} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \widehat{v}^y & 0 & \widehat{\omega} \\ 0 & 0 & 0 & -\widehat{v}^x & -\widehat{\omega} & 0 \end{bmatrix}$$

and

$$H(t) = \left.\frac{\partial h_3}{\partial \mathrm{x}}\right|_{\mathrm{x}=\widehat{\mathrm{x}}(t)} = \begin{bmatrix} -\ell\sin\widehat{\theta} & 1 & 0 & 0 & 0 & 0 \\ \ell\cos\widehat{\theta} & 0 & 1 & 0 & 0 & 0 \\ \ell\sin\widehat{\theta} & 1 & 0 & 0 & 0 & 0 \\ -\ell\cos\widehat{\theta} & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}.$$

We will refer to the second-order optimal filter derived in Proposition 3.3 as $\mathcal{F}_L$, and to the extended Kalman filter derived in Proposition 3.4 as $\mathcal{F}_K$. The formulations of the two Propositions are similar: both try to estimate the state variables considering the correction of the dynamic equation obtained by adding a "gain" that has to satisfy a Riccati-like equation. The main difference between the two approaches regards the fact that $\mathcal{F}_K$ does not take into consideration the geometry of the system. Moreover, $\mathcal{F}_L$ allows more freedom in the construction of the filter. Firstly, in the choice of the connection function ω that describes the derivation on the tangent space, that is not necessary for $\mathcal{F}_K$ since the geometric differential structure is neglected. The presence of the initial term (2.30) in the cost function (2.29) enables us to have more choices in the starting initial condition of $\mathcal{F}_L$. Finally, the presence of the forms $\mathcal{R}$ and $\mathcal{Q}$ in the incremental cost (2.31) of (2.29) let us weigh the contribution of the variances of the model and measurement errors. It should also be taken into consideration that the construction of the Kalman filter started from the hypothesis of linearity of the dynamics and Gaussianity of the errors, the second-order filter, instead, does not require them.

In order to better investigate the relation of the filters $\mathcal{F}_K$ and $\mathcal{F}_L$, it could be useful the following Proposition.

Proposition 3.5. *Under the hypothesis of linear dynamics and quadratic cost, the second-order optimal filter and the extended Kalman filter are equivalent.*

Proof. We consider the state variables $\mathrm{x} = (x_1, ..., x_n)^T$ and the dynamic linear system $\dot{\mathrm{x}} = F\mathrm{x}$, in extended form

$$\begin{cases} \dot{x}_1 &= F_{11}x_1 + \cdots + F_{1n}x_n \\ &\vdots \\ \dot{x}_n &= F_{n1}x_1 + \cdots + F_{nn}x_n \end{cases} \tag{3.99}$$

where F_{ij}, $i, j = 1, ..., n$ are the components of the matrix F. This system evolves in $\mathbb{R}^n$, that is a Lie group with the addition operation. A matrix Lie group representation for an element $\mathrm{x} \in \mathbb{R}^n$ is given by $g \in G$ where G is the matrix Lie group

$$G := \left\{ \begin{bmatrix} I_{n\times n} & \mathrm{x} \\ 0_{1\times n} & 1 \end{bmatrix} \in \mathrm{GL}(n+1), \mathrm{x} \in \mathbb{R}^n \right\}. \tag{3.100}$$

The matrix multiplication produces

$$\begin{bmatrix} I_{n\times n} & \mathrm{x}_1 \\ 0_{1\times n} & 1 \end{bmatrix} \begin{bmatrix} I_{n\times n} & \mathrm{x}_2 \\ 0_{1\times n} & 1 \end{bmatrix} = \begin{bmatrix} I_{n\times n} & \mathrm{x}_1 + \mathrm{x}_2 \\ 0_{1\times n} & 1 \end{bmatrix} \tag{3.101}$$

from which we can derive the inverse of g

$$g^{-1} = \begin{bmatrix} I_{n\times n} & -\mathrm{x} \\ 0_{1\times n} & 1 \end{bmatrix}. \tag{3.102}$$

The Lie algebra $\mathfrak{g}$ of G is given by

$$\mathfrak{g} := \left\{ \begin{bmatrix} 0_{n\times n} & \eta^{\mathrm{x}} \\ 0_{1\times n} & 0 \end{bmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}, \eta^{\mathrm{x}} \in \mathbb{R}^n \right\} \tag{3.103}$$

with the Lie bracket $[\cdot,\cdot]_{\mathfrak{g}} = 0_{(n+1)\times(n+1)}$. We define the Lie algebra isomorphism $\wedge : \mathfrak{g} \to \mathbb{R}^n$ from the Lie algebra $\mathfrak{g}$ with Lie bracket $[\cdot,\cdot]_{\mathfrak{g}}$ to the Lie algebra $\mathbb{R}^n$ with Lie bracket $[\cdot,\cdot]_{\mathbb{R}^n} = 0_{n,1}$. It follows that, given an element $\eta^g \in \mathfrak{g}$, the matrix representation of the adjoint operator is null

$$\mathrm{ad}_{\eta^g} = 0_{(n+1)\times(n+1)}. \tag{3.104}$$

The tangent map and its dual are the identity matrix.

The dynamics is described by

$$g^{-1}\dot{g} = \lambda_t \tag{3.105}$$

and since we consider a linear system, the left-trivialized dynamics has the vector representation (in the Lie algebra $\mathbb{R}^n$)

$$\lambda_t = \begin{bmatrix} \lambda_{x_1} \\ \vdots \\ \lambda_{x_n} \end{bmatrix} = \begin{bmatrix} F_{11}x_1 + \cdots + F_{1n}x_n \\ \vdots \\ F_{n1}x_1 + \cdots + F_{nn}x_n \end{bmatrix} = F\mathrm{x}. \tag{3.106}$$

The differential of the left-trivialized dynamics λ_t is $\mathrm{d}\lambda_t = F$.

For the measurement equation we choose

$$h_t = \begin{bmatrix} H_{11}x_1 + \cdots + H_{1n}x_n \\ \vdots \\ H_{p1}x_1 + \cdots + H_{pn}x_n \end{bmatrix} = H\mathrm{x} \tag{3.107}$$

where $H \in \mathbb{R}^{p\times n}$. The differential of the measurement equation h_t is $\mathrm{d}h_t = H$.

Now with these considerations, we can prove the equivalence of the filters. To emphasize the additional property of the group operation, we will use as argument of the functions the wordings x. Moreover, we will drop the time dependence when not needed.

The hypothesis of quadratic cost leads us to rewrite Equation (2.29) as

$$J(\delta, \varepsilon, \mathrm{x}_0; t, t_0) = \int_{t_0}^{t} \ell(\delta(\tau), \varepsilon(\tau), t, \tau) d\tau \tag{3.108}$$

which means that the initial cost equation (2.30) evaluated in x_0 is equal to zero

$$m(\mathrm{x}_0, t, t_0) = 1/2 e^{-\alpha(t-t_0)} m_0(\mathrm{x}_0) = 0. \tag{3.109}$$

We choose the starting quadratic cost

$$m_0(\mathrm{x}_0, t, t_0) = \frac{1}{2}(\mathrm{x} - \mathbb{E}[\mathrm{x}_0])^T \mathrm{Var}[\mathrm{x}_0]^{-1}(\mathrm{x} - \mathbb{E}[\mathrm{x}_0]). \tag{3.110}$$

This equation attains its minimum at $\mathbb{E}[\mathrm{x}(t_0)])$ which proves that

$$\widehat{\mathrm{x}}_0 = \mathbb{E}[\mathrm{x}(t_0)]. \tag{3.111}$$

Its Hessian at t_0 is

$$\mathrm{Hess}\, m_0(\mathrm{x}(t_0), t, t_0) = \mathrm{Var}[\mathrm{x}(t_0)]^{-1}. \tag{3.112}$$

The starting condition for the operator K of the second-order filter is the inverse of X_0 in equation (2.36), and thus

$$K(t_0) = \mathrm{Var}[\mathrm{x}(t_0)]. \tag{3.113}$$

The incremental cost in (2.31), with the choices $\alpha = 0$ and the matrix form for $\mathcal{R}$ and $\mathcal{Q}$ equal to the identity matrix, becomes

$$\ell(\delta, \varepsilon, t, \tau) = \langle \delta, \delta \rangle + \langle \varepsilon, \varepsilon \rangle. \tag{3.114}$$

From the construction of the Lie group G and the derivation of the various maps and operators, the residual in (2.34) takes the form

$$\begin{aligned} r_t(\widehat{\mathrm{x}}) &= T_e L^*_{\widehat{\mathrm{x}}}[\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\mathrm{x}}))\big) \circ \mathrm{d}h_t(\widehat{\mathrm{x}})] \\ &= [\big((D^{-1})^T D^{-1}(\mathrm{y} - h_t(\widehat{\mathrm{x}}))\big)^T H]^T \\ &= H^T Q_K^{-1}(\mathrm{y} - h_t(\widehat{\mathrm{x}})) \end{aligned} \tag{3.115}$$

where with $Q_K = D^T D$ we define the variance of the measurement error.

The operator A is defined in (2.37) and takes the form

$$A(t) = \mathrm{d}_1 \lambda_t(\widehat{\mathrm{x}}, u) \circ T_e L_{\widehat{\mathrm{x}}} - \mathrm{ad}_{\lambda_t(\widehat{\mathrm{x}}, u)} - T_{\lambda_t(\widehat{\mathrm{x}}, u)} = F \tag{3.116}$$

since the dynamics is linear and both the adjoint operator and torsion are the null matrix.

The operator E is defined in (2.38) and takes the form

$$\begin{aligned} E = -T_e L^*_{\widehat{\mathrm{x}}} \circ \big[\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{\mathrm{x}}))\big)^{T_{\widehat{\mathrm{x}}}G} \circ \mathrm{Hess}\, h_t(\widehat{\mathrm{x}}) \\ -(\mathrm{d}h_t(\widehat{\mathrm{x}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h_t(\widehat{\mathrm{x}})\big] \circ T_e L_{\widehat{\mathrm{x}}}. \end{aligned} \tag{3.117}$$

The Hessian operator in this case is a 3-dimensional tensor with all the entries equal to zero since the measurement equation is linear. From the definition of $Q_K = D^T D$ it follows that

$$E = H^T Q_K^{-1} H. \tag{3.118}$$

The connection function is the null operator since the Lie bracket operator vanishes.

With the new operators, the Riccati equation (2.35) becomes

$$\begin{aligned}\dot{K} =& -\alpha \cdot K + A \circ K + K \circ A^* - K \circ E \circ K \\ &+ B \circ R^{-1} \circ B^* - \omega_{Kr} \circ K - K\omega_{Kr}^* \\ =& FK + KF^T - KH^T Q_K^{-1} HK + R_K\end{aligned} \tag{3.119}$$

where $R_K = B^T B$ is the matrix variance of model error.

From (3.105) and (3.106) it follows that we can rewrite (2.32) as

$$\begin{aligned}\dot{\widehat{\mathrm{x}}} &= F\widehat{\mathrm{x}} + Kr_t(\widehat{\mathrm{x}}) \\ &= F\widehat{\mathrm{x}} + KH^T Q_K^{-1}(\mathrm{y} - h_t(\widehat{\mathrm{x}})).\end{aligned} \tag{3.120}$$

From the equations (3.119) and (3.120) and defining $P = KH^T Q_K^{-1}$, we obtain the Kalman filter form. □

3.6 Comparison and simulations

In these simulations, we show how the two filters $\mathcal{F}_L$ and $\mathcal{F}_K$ work when we set them in the same conditions. With this, we mean that we choose the parameters of the optimal problem filter $\mathcal{F}_L$ (initial conditions, operators R and Q, etc.) as done in Proposition 3.5. For the inertia tensor of the planar rigid body under consideration (3.2) we set $J = 6.125\,\mathrm{kg\,m^2}$ and $m = 125\,\mathrm{kg}$. We consider the control inputs $u_\omega = 10\sin(\frac{1}{2}t)$ and $u_v = \frac{1}{2}\sin(\frac{1}{5}t)$, with total simulation time $T = 60\,\mathrm{s}$. The initial conditions are $\theta(0) = \pi/6$ rad, $x(0) = 1\,\mathrm{m}$, $y(0) = 2\,\mathrm{m}$, $\omega(0) = 0\,\mathrm{rad/s}$, $v^x(0) = 0\,\mathrm{m/s}$ and $v^y(0) = 0\,\mathrm{m/s}$. The forgetting factor α in (2.35) will take the value $\alpha = 0$. The matrices R and Q related to the incremental cost function are the identity matrices. The matrix B_2 in (3.9) related to model error is $B_2 = \mathrm{diag}\{b_\omega, b_{v^x}, b_{v^y}\}$ and thus $R_K = \begin{bmatrix} 0_{3\times3} & 0_{3\times3} \\ 0_{3\times3} & B_2^T B_2 \end{bmatrix}$ where we put $b_\omega = 0.1$, $b_{v^x} = 0.1$, $b_{v^y} = 0.1$. The matrix D_3 related to measurement error is given by (3.73) and $Q_K = D_3^T D_3$, where we choose $d_2 = 0.5$, $d_3 = 0.1$, $d_4 = 0.1$. We consider two time steps of $T_s = 1\,\mathrm{ms}$ and $T_s = 10\,\mathrm{ms}$. In Figure 3.8 and 3.9 are shown the trajectories for the dynamic variables filtered with both filters. As can be noticed, $\mathcal{F}_L$ and $\mathcal{F}_K$ results overlap almost perfectly. This is due to the fact that, even if the dynamics is nonlinear, the integration step is sufficiently small to consider the dynamics almost linear. This aspect is clear in Table 3.1, where the mean, the standard deviation and the root mean square of the difference between the reference and filtered trajectories

$$\Delta\theta := \theta - \widehat{\theta}, \quad \Delta x := x - \widehat{x}, \quad \cdots$$

are shown.

It is known that the Kalman filter produces the optimal estimations for linear systems with quadratic costs and Gaussian errors. Proposition 3.5 and the

fact that in the simulations the extended Kalman filter and the second-order filter on Lie groups present the same behavior show that, in these cases, the results of the two filters are comparable. The major difference concerns the fact the second-order one allows to have more freedom in choosing the parameters (matrices R, B, connection function...). Another important difference regards that the differential equation for the gain operator K is defined on the whole tangent space in the Kalman case, while it evolves on the Lie algebra in the geometric filter case. The corrected Lie algebra components are then mapped onto the tangent space through the left-trivialized dynamics. This suggests that in the case of fast dynamics, with a longer sampling time, the second-order filter could perform better since it considers the geometric structure of the system between samples. As future works, it could be worthwhile to design a discrete version of this filter and compare it with the discrete Kalman filter.

Table 3.1: Mean, standard deviation and root mean square values of the errors for $T_s = 1\,\text{ms}$ and $T_s = 10\,\text{ms}$.

T_s		EKF			TSE(2) 2nd order		
		μ	σ	rms	μ	σ	rms
1ms	$\Delta\theta\,[\text{mrad}]$	1.8	6.6	6.9	1.9	6.8	7.0
	$\Delta x\,[\text{mm}]$	-0.2	7.1	7.1	0.2	7.4	7.4
	$\Delta y\,[\text{mm}]$	0.3	7.1	7.1	0.0	7.6	7.6
	$\Delta\omega\,\left[\frac{\text{mrad}}{\text{s}}\right]$	0.0	1.9	1.9	0.1	1.9	1.9
	$\Delta v^x\,\left[\frac{\text{mm}}{\text{s}}\right]$	0.2	3.5	3.5	0.1	3.4	3.4
	$\Delta v^y\,\left[\frac{\text{mm}}{\text{s}}\right]$	-0.2	2.4	2.4	-0.2	2.6	2.6
10ms	$\Delta\theta\,[\text{mrad}]$	-6.8	16.4	17.8	-5.9	17.6	18.6
	$\Delta x\,[\text{mm}]$	4.9	31.6	32.0	5.7	32.0	32.5
	$\Delta y\,[\text{mm}]$	-3.6	23.2	23.5	-5.2	26.0	26.5
	$\Delta\omega\,\left[\frac{\text{mrad}}{\text{s}}\right]$	0.5	5.8	5.8	0.7	5.8	5.8
	$\Delta v^x\,\left[\frac{\text{mm}}{\text{s}}\right]$	2.5	10.3	10.6	2.2	9.5	9.8
	$\Delta v^y\,\left[\frac{\text{mm}}{\text{s}}\right]$	-0.1	6.7	6.8	-0.0	7.7	7.7

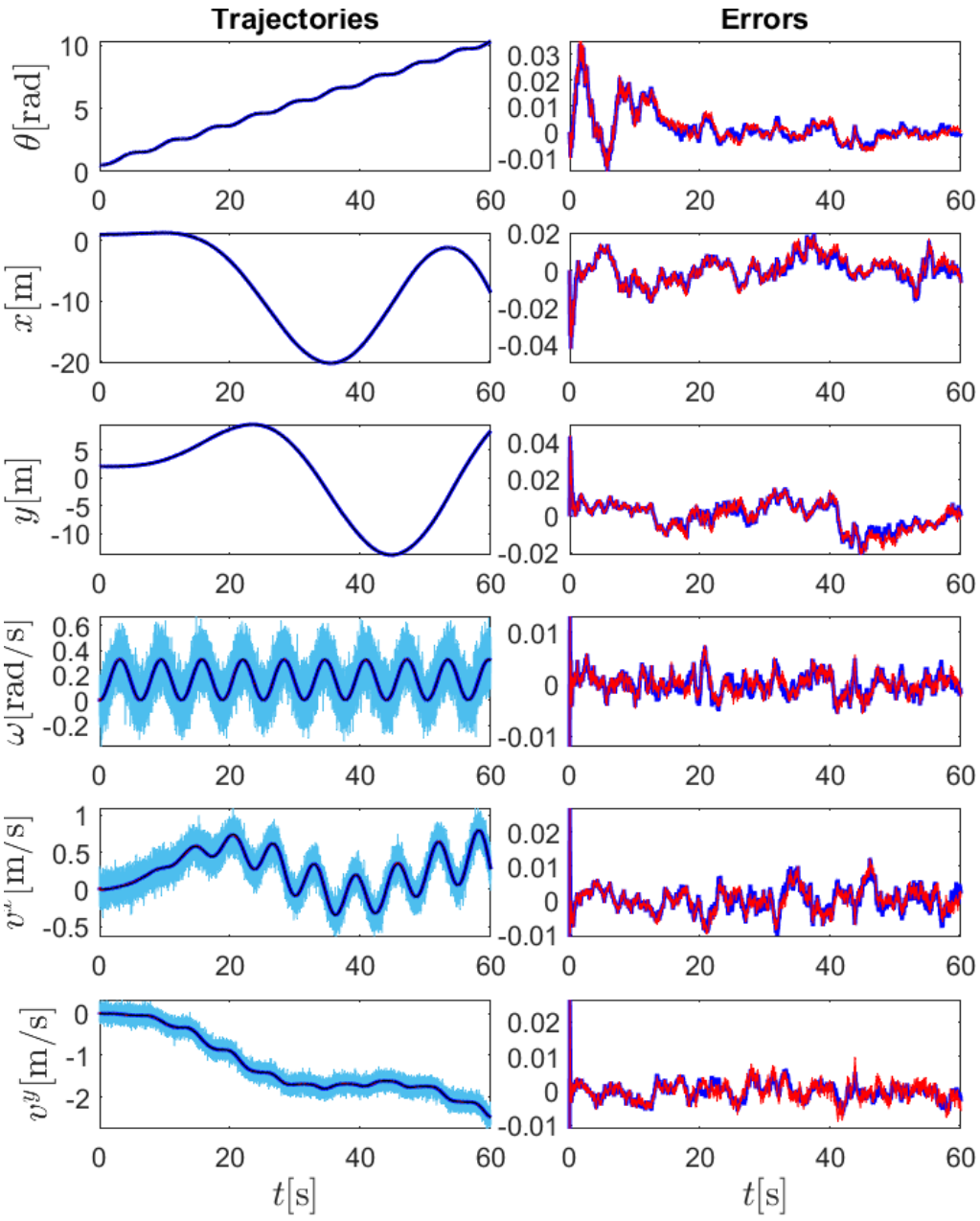


Figure 3.8: Measured (sky-blue), nominal (black), filtered with the Second-order filter (red), and filtered with extended Kalman filter (blue) trajectories on the left and their errors on the right with time step $T_s = 1\,\text{ms}$.

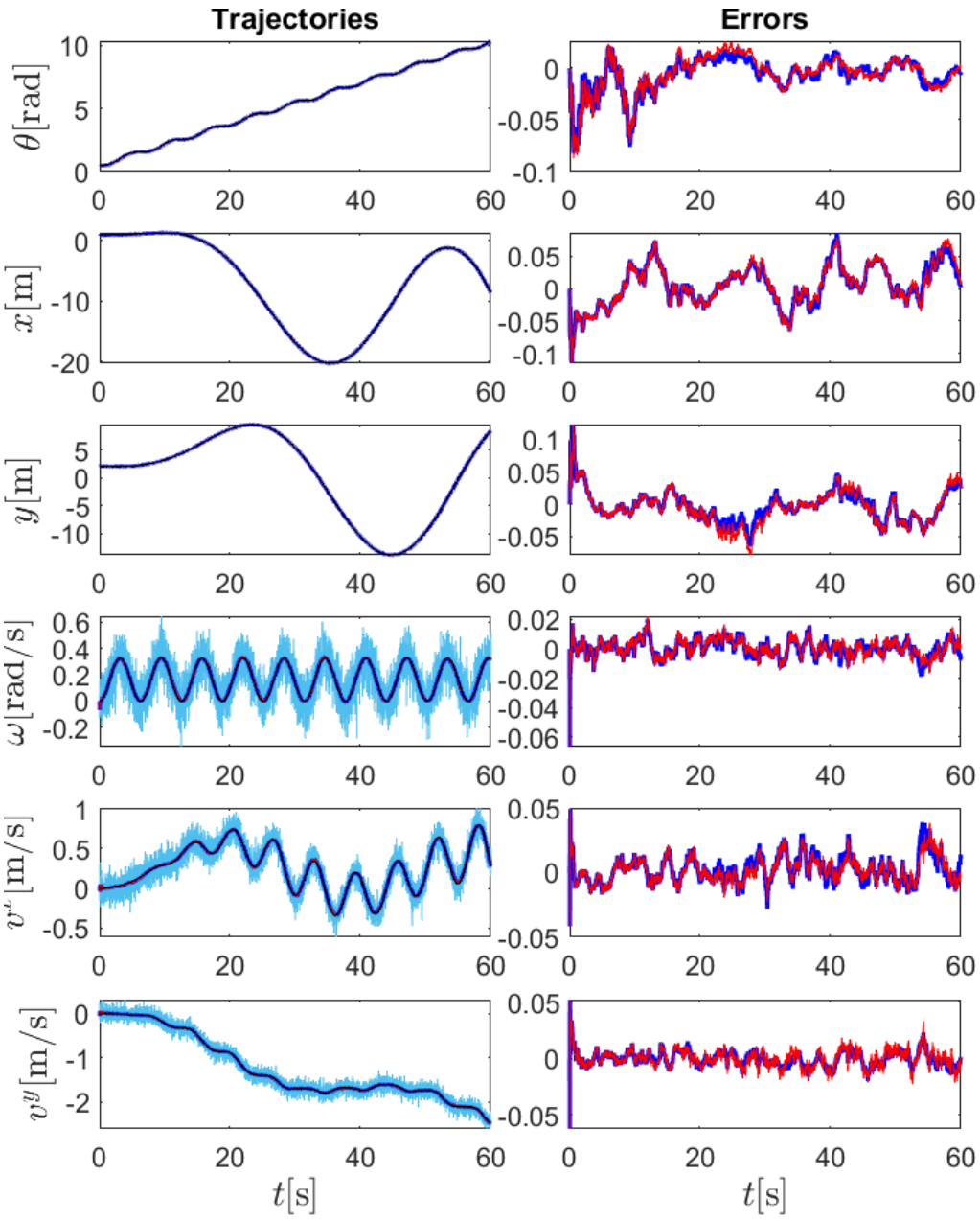


Figure 3.9: Measured (sky-blue), nominal (black), filtered with the Second-order filter (red), and filtered with extended Kalman filter (blue) trajectories on the left and their errors on the right with time step $T_s = 10\,\text{ms}$.

Chapter 4

Second-order optimal filter applied to the Chaplygin sleigh[1]

In this chapter, we apply the second-order filter to the Chaplygin sleigh which is a planar rigid body with nonholonomic constraint (see e.g. [3], [29], [37]). The main difficulty is represented by the fact that the presence of nonholonomic constraint modifies the state space and some considerations have to be taken. The recognition of the geometric structure is obtained through the use of Hamel's equations. A relevant importance in this type of filter is represented by the choice of the affine connection ([9], [27], [30]).

4.1 The Chaplygin sleigh

The Chaplygin sleigh is a nonholonomic system that models a planar rigid body supported at three points, two of which slide freely while the third is a blade at distance a from the center of mass and that cannot move perpendicularly. Let $\Sigma_b = \left\{e_{\parallel}^b, e_{\perp}^b\right\}$ be the right-handed body frame attached to the Chaplygin sleigh centered at the contact point between the blade and the ground with the $e_{\parallel}^b$-axis aligned with the blade, and let $\Sigma_I = \{e_x, e_y\}$ be an inertial frame (also called space frame) fixed in space as shown in Figure 4.1.

The configuration space is SE(2) with coordinates $q = (\theta, x, y)$, where (x, y) denote the position of the contact point of the blade in the plane, and θ is the angle that the blade forms with the horizontal axis e_x. The velocity components, also named quasi velocities or Hamel's velocities, with respect to the body frame Σ_b are $(\omega, v, v_\perp)$, where ω is the angular velocity, v and $v_\perp$ are the (linear) velocities of the body along the $e_{\parallel}^b$ and $e_{\perp}^b$ axes, respectively. The nonholonomic constraint imposes that the orthogonal component $v_\perp$ of the velocity with respect to the blade vanishes, namely $v_\perp = 0$. In the inertial frame, where $(\dot\theta, \dot x, \dot y)$ are the components of the velocity, the constraint reads

[1] This chapter is based on the work
▷ Rigo D., Sansonetto N., & Muradore R. *Second-order-optimal filtering on* $\mathrm{SE}(2) \times \mathbb{R}^2$ *for the Chaplygin sleigh* (accepted) presented at System & Control Letters.

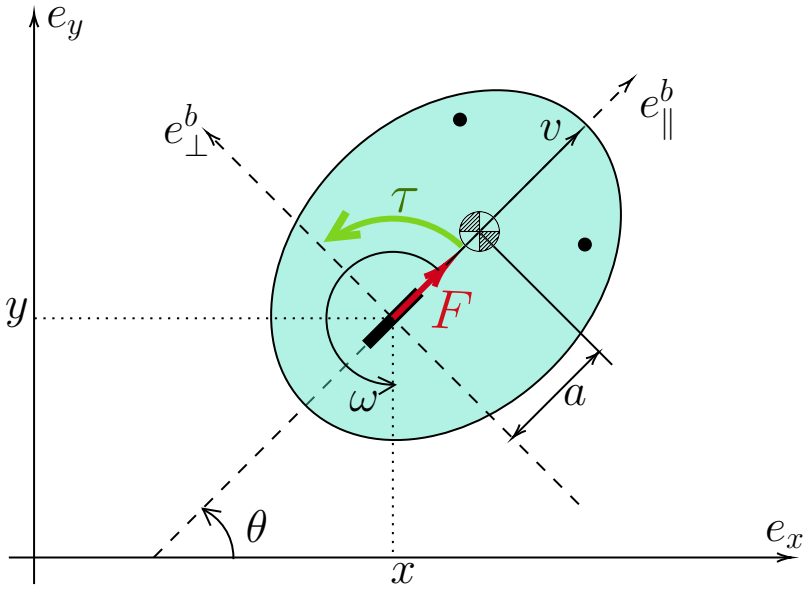


Figure 4.1: Planar rigid body where the blade is indicated with a thicker segment and the two passive supporting wheels as •.

as

$$\dot{x}\sin\theta - \dot{y}\cos\theta = 0. \tag{4.1}$$

The velocities $(\dot{\theta}, \dot{x}, \dot{y})$ and quasi-velocities $(\omega, v, v_\perp)$ are related by

$$\dot{x} = v\cos\theta, \quad \dot{y} = v\sin\theta, \quad \dot{\theta} = \omega. \tag{4.2}$$

The nonholonomic constraint (4.1) defines a constant rank-2 distribution $\mathcal{D}$ on the configuration space locally generated by

$$X_v = \cos\theta\frac{\partial}{\partial x} + \sin\theta\frac{\partial}{\partial y}, \quad X_\omega = \frac{\partial}{\partial\theta}, \tag{4.3}$$

called constrained manifold. The state space, that is the constrained manifold, can be therefore identified with $\mathrm{SE}(2)\times\mathbb{R}^2$ and parameterized by $(\theta, x, y, \omega, v)$.

The center of mass (x_c, y_c) is settled at distance a from the contact point (x, y) according to the equations

$$(x_c, y_c) = (x + a\cos\theta, y + a\sin\theta). \tag{4.4}$$

We denote by J the inertia along the axis passing to the center of mass and orthogonal to the plane and with m the mass of the rigid body. The control inputs $u(t) = (\tau(t), F(t))$ are functions of time that act respectively as a torque applied around the contact point and a force applied to the center of the body frame along $e^b_\parallel$ as shown in Figure 4.1.

In order to put in evidence the geometric structure of the dynamics, it is useful to consider Hamel's approach to the equations of motion of nonholonomic system (see [3], [4]). Hamel's equations with external input are then

$$\begin{aligned} m\dot{v} &= ma\omega^2 + F, \\ (J + ma^2)\dot{\omega} &= -mav\omega + \tau. \end{aligned} \tag{4.5}$$

Equations (4.5) together with (4.2) define the motion of the Chaplygin sleigh with external forces. We stress the fact that the constrained manifold $\mathcal{D}$ can be identified with $\mathrm{SE}(2) \times \mathbb{R}^2$, and therefore it can be endowed with a Lie group structure.

Since we assume that our model is not perfectly accurate, we add a model error that consists of an additive term that takes into account unmodelled dynamics and uncertainty on the parameters' values. According to [35], it only affects the evolution of the velocity of the system and not the kinematics. This error is denoted by $\xi(t) = (\xi_\omega(t), \xi_v(t))^T$ and is modelled as a Gaussian white noise with zero mean and diagonal and positive definite variance Σ.

The controlled Chaplygin sleigh equations are then (see e.g. [3])

$$\begin{cases} \dot{\theta}(t) &= \omega(t) \\ \dot{x}(t) &= v(t)\cos\theta(t) \\ \dot{y}(t) &= v(t)\sin\theta(t) \\ \dot{\omega}(t) &= -\frac{ma}{J+ma^2}\omega(t)v(t) + \frac{1}{J+ma^2}\tau(t) + \xi_\omega(t) \\ \dot{v}(t) &= a\omega(t)^2 + \frac{1}{m}F(t) + \xi_v(t). \end{cases} \tag{4.6}$$

The measurement equation is given by

$$\mathrm{y}(t) = h(\overline{g}(t)) + D\varepsilon(t) \tag{4.7}$$

where

$$h(\overline{g}(t)) = \begin{bmatrix} x(t) + \ell\ \cos(\theta(t)) \\ y(t) + \ell\ \sin(\theta(t)) \\ x(t) - \ell\ \cos(\theta(t)) \\ y(t) - \ell\ \sin(\theta(t)) \\ \omega(t) \\ v(t) \end{bmatrix}, \tag{4.8}$$

represents a GPS-like system that provides the position of two antennas on the rigid body located at distance ℓ to the origin of the body frame Σ_b (see Figure 4.2), and an INS-like system that measures the angular velocity ω and the linear velocity v along the $e_\parallel^b$ axis. The measurement noise ε is modelled as a Gaussian white noise with zero mean and diagonal and positive definite variance Λ.

4.2 The $\mathrm{SE}(2) \times \mathbb{R}^2$ structure

We recall that, given the Lie algebra $\mathfrak{se}(2)$ of SE(2), we introduce the Lie algebra isomorphism $^\wedge : \mathbb{R}^3 \to \mathfrak{se}(2)$

$$\Omega = \begin{bmatrix} \eta^\theta \\ \eta^x \\ \eta^y \end{bmatrix} \longmapsto \Omega^\wedge = \eta^g = \begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix} \tag{4.9}$$

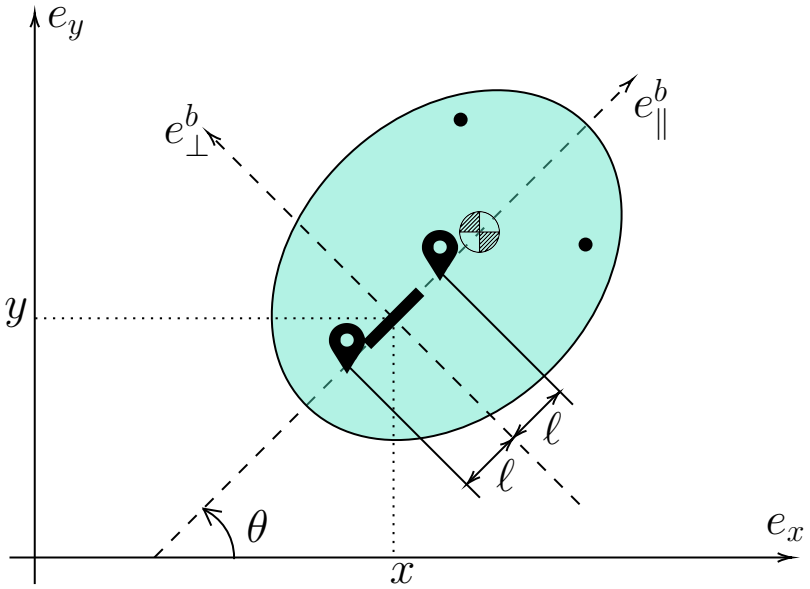


Figure 4.2: Planar rigid body with two antennas 📍.

from the Lie algebra $(\mathbb{R}^3, \star)$ to the matrix Lie algebra $(\mathfrak{se}(2), [\cdot,\cdot])$, where $\star : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is the Lie bracket operation defined as

$$\begin{bmatrix} \eta_1^\theta \\ \eta_1^x \\ \eta_1^y \end{bmatrix} \star \begin{bmatrix} \eta_2^\theta \\ \eta_2^x \\ \eta_2^y \end{bmatrix} = \begin{bmatrix} 0 \\ \eta_1^y \eta_2^\theta - \eta_1^\theta \eta_2^y \\ \eta_1^\theta \eta_2^x - \eta_1^x \eta_2^\theta \end{bmatrix}, \tag{4.10}$$

and $[\cdot,\cdot]$ is the usual matrix commutator.

The tangent bundle TSE(2) is isomorphic to $\mathrm{SE}(2) \times \mathfrak{se}(2)$ via left translation; since we impose the nonholonomic constraint $v_\perp = 0$, a sub-bundle of it describes the admissible velocities. The velocity pair $V = (\omega, v)^T$ is an element of $\mathbb{R}^2$ that is a Lie group with respect to the sum operation, with abelian Lie algebra $\mathbb{R}^2$. The elements of such Lie algebra are denoted by $\eta^V = (\eta^\omega, \eta^v)$.

An element $\overline{g} = (g, V) \in \overline{G} := \mathrm{SE}(2) \times \mathbb{R}^2$ can be represented by a 6×6 matrix

$$\overline{g} = \begin{bmatrix} g & 0_{3\times 2} & 0_{3\times 1} \\ 0_{2\times 3} & I_{2\times 2} & V \\ 0_{1\times 3} & 0_{1\times 2} & 1 \end{bmatrix}, \tag{4.11}$$

and the group operation is defined by

$$(g, V) \cdot (f, W) = (gf, V + W) \tag{4.12}$$

with unit element $e = (I_{3\times 3}, 0_{2\times 1})$ and inverse $(g, V)^{-1} = (g^{-1}, -V)$. The Lie algebra $\overline{\mathfrak{g}} = \mathfrak{se}(2) \times \mathbb{R}^2$ of $\overline{G}$ can be identified, up to Lie algebra isomorphism, with $\mathbb{R}^5$. We exploit this isomorphism to work on $\mathbb{R}^5$ and to report the operators that appear in the optimal filter equations in their matrix forms.

Let $\eta^{\overline{g}} = (\eta^g, \eta^V)$ be an element of $\overline{\mathfrak{g}}$, the matrix form of the $\mathrm{ad}_{\eta^{\overline{g}}}$ operator

is given by the 5×5 matrix

$$\mathrm{ad}_{\eta^{\overline{g}}} = \begin{bmatrix} 0 & 0 & 0 & 0_{1\times 2} \\ \eta^y & 0 & -\eta^\theta & 0_{1\times 2} \\ -\eta^x & \eta^\theta & 0 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix}. \tag{4.13}$$

Finally, the matrix representation of the tangent map is

$$T_e L_{\overline{g}} = \begin{bmatrix} 1 & 0 & 0 & 0_{1\times 2} \\ 0 & \cos\theta & -\sin\theta & 0_{1\times 2} \\ 0 & \sin\theta & \cos\theta & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & I_{2\times 2} \end{bmatrix} \tag{4.14}$$

whose dual map matrix representation $T_e L^*_{\overline{g}}$ satisfies $T_e L^*_{\overline{g}} = (T_e L_{\overline{g}})^T$.

4.3 Chaplygin sleigh optimal filter

We rewrite equation (4.6) using the geometric structure outlined in Section 4.2 as

$$\dot{\overline{g}}(t) = \overline{g}(t)(\lambda(\overline{g}(t), u(t)) + B\delta(t)), \quad \overline{g}(t_0) = \overline{g}_0, \tag{4.15}$$

where $\overline{g}(t) = (g(t), V(t)) \in \overline{G}$ is the state, $u(t) \in \mathbb{R}^2$ is the input, $\delta(t)$ is a Gaussian white noise with zero mean and unit variance, $B : \mathbb{R}^2 \to \overline{\mathfrak{g}}$ is a linear map with matrix representation $B \in \mathbb{R}^{5\times 2}$

$$B = \begin{bmatrix} 0_{3\times 2} \\ B_2 \end{bmatrix}, B_2 = \begin{bmatrix} b_\omega & 0 \\ 0 & b_v \end{bmatrix} \tag{4.16}$$

such that $B_2\delta(t)$ is a zero mean white Gaussian noise with variance $\Sigma = B^T B$, and $\lambda : G \times \mathbb{R}^2 \to \overline{\mathfrak{g}}$ is the left-trivialized dynamics (4.18). The function λ can be rewritten splitting it into its SE(2) and $\mathbb{R}^2$ components obtaining

$$g^{-1}\dot{g} = {\lambda_g}^\wedge, \quad \dot{V} = \lambda_V + B_2\delta \tag{4.17}$$

where $\lambda = ({\lambda_g}^\wedge, \lambda_V) \in \overline{\mathfrak{g}}$. In particular, the expressions for λ_g and λ_V are

$$\lambda_g(\overline{g}, u) = \begin{bmatrix} \omega \\ v \\ 0 \end{bmatrix}, \quad \lambda_V(\overline{g}, u) = \begin{bmatrix} -\frac{ma}{J+ma^2}\omega v + \frac{1}{J+ma^2}\tau \\ a\omega^2 + \frac{1}{m}F \end{bmatrix}. \tag{4.18}$$

The measurement equation (4.7) can be then rewritten as

$$\mathrm{y}(t) = h(\overline{g}(t)) + D\overline{\varepsilon}(t) \tag{4.19}$$

where $h : \overline{G} \to \mathbb{R}^6$ is the output map (4.8), $\overline{\varepsilon} \in \mathbb{R}^6$ is the measurement error and $D : \mathbb{R}^6 \to \mathbb{R}^6$ is an invertible linear map with the property that $\varepsilon = D\overline{\varepsilon}$ (i.e., $\Lambda = D^T D$, $\overline{\varepsilon}$ is a Gaussian white noise with zero mean and unit variance). We choose for the matrix D in (4.7) the matrix form

$$D = \mathrm{diag}\{d_2, d_2, d_2, d_2, d_3, d_4\}. \tag{4.20}$$

For the cost functional to minimize (2.29)-(2.31), we consider $\alpha = 0$ and the initial cost

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 \tag{4.21}$$

where $\|\cdot\|_F^2$ stands for the Frobenius norm, while for the matrix representation of the forms $\mathcal{R}$ and $\mathcal{Q}$ we choose

$$R = \mathrm{diag}\{r_\omega, r_v\}, \tag{4.22}$$
$$Q = \mathrm{diag}\{q_2, q_2, q_2, q_2, q_3, q_4\}. \tag{4.23}$$

We define the weighted output error $\widetilde{\mathrm{y}}$ as

$$\widetilde{\mathrm{y}} = \left[\mathrm{diag}\left(\frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}, \frac{q_4}{d_4^2}\right)\right](\mathrm{y} - \widehat{\mathrm{y}}) \in \mathbb{R}^6. \tag{4.24}$$

In this section, we use the Cartan-Schouten (0)-connection given by $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$ [34].

The following proposition is an extension of the theorems in [35], [32] to the nonholonomic case, and represents the second-order-optimal filter tailored for the Chaplygin sleigh case.

Proposition 4.1. *Consider the dynamic system* (4.15) *with measurement equation* (4.19) *where the output map h and the linear map D are given by* (4.8) *and* (4.20)*, respectively and where the operator B takes the form* (4.16)*. Consider the cost functional* (2.29)-(2.31) *where the initial cost m_0 is given by* (4.21) *and the matrix representation of the forms $\mathcal{R}$, $\mathcal{Q}$ are given by* (4.22) *and* (4.23)*, respectively. Then the second-order optimal filter is*

$$\widehat{g}^{-1}\dot{\widehat{g}} = (\widehat{\omega}, \widehat{v}, 0)^\wedge + (K_{11}r^g + K_{12}r^V)^\wedge \tag{4.25}$$
$$\dot{\widehat{V}} = \begin{bmatrix} -\frac{ma}{J+ma^2}\widehat{\omega}\widehat{v} + \frac{\tau}{J+ma^2} \\ a\widehat{\omega}^2 + \frac{F}{m} \end{bmatrix} + (K_{21}r^g + K_{22}r^V) \tag{4.26}$$

where the residual r is

$$r = \begin{bmatrix} r^g \\ r^V \end{bmatrix}^T = \begin{bmatrix} -(\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell\,\sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell\,\cos\widehat{\theta} \\ (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} \\ -(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} \\ \widetilde{\mathrm{y}}_5 \\ \widetilde{\mathrm{y}}_6 \end{bmatrix}^T. \tag{4.27}$$

The optimal gain $K = (K_{11}, K_{12}; K_{21}, K_{22}) : (\mathbb{R}^5)^* \to \mathbb{R}^5$ (with $K_{11} \in \mathbb{R}^{3\times3}$, $K_{12} \in \mathbb{R}^{3\times2}$, $K_{21} \in \mathbb{R}^{2\times3}$ and $K_{22} \in \mathbb{R}^{2\times2}$) is the solution of the perturbed matrix Riccati differential equation

$$\begin{aligned} \dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K, r)K - KW(K, r)^T \end{aligned} \tag{4.28}$$

where

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & \widehat{\omega} & 0 & 1 \\ \widehat{v} & -\widehat{\omega} & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{ma}{J+ma^2}\widehat{v} & -\frac{ma}{J+ma^2}\widehat{\omega} \\ 0 & 0 & 0 & 2a\widehat{\omega} & 0 \end{bmatrix}, \tag{4.29}$$

$$E = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0 & 0 \\ a_{2,1} & 2\frac{q_2}{d_2^2} & 0 & 0 & 0 \\ a_{3,1} & 0 & 2\frac{q_2}{d_2^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{q_3}{d_3^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{q_4}{d_4^2} \end{bmatrix} \tag{4.30}$$

with

$$\begin{aligned} a_{1,1} =& (\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell\,\cos\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell\,\sin\widehat{\theta} + 2\frac{q_2}{d_2^2}\ell^2, \\ a_{1,2} =& -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} + \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta}, \\ a_{1,3} =& +\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} + \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta}, \\ a_{2,1} =& +\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} - \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta}, \\ a_{3,1} =& -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} - \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta}, \end{aligned}$$

and

$$W(K, r) = \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^V)^\wedge} & 0_{3\times 2} \\ 0_{2\times 3} & 0_{2\times 2} \end{bmatrix}. \tag{4.31}$$

The initial conditions for the equations (4.25)-(4.26) *and* (4.28) *are*

$$\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{4.32}$$

$$K(t_0) = \mathrm{diag}\{1/2, 1, 1, 1/2, 1, 1\}. \tag{4.33}$$

Proof. In what follows we will use $\eta^{\overline{g}} = (\eta^\theta, \eta^x, \eta^y, \eta^\omega, \eta^v)^T \in \mathbb{R}^5$ to indicate the vector form of an element of the Lie algebra $\overline{\mathfrak{g}}$ and with $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}}) = (\theta', x', y', \omega', v')^T$ the vector form of an element of the tangent space $T_{\widehat{\overline{g}}}\overline{G}$.

Computation of r

According to [35], [32], the residual $r \in (\mathbb{R}^5)^*$ is given by

$$r(\widehat{\overline{g}}) = T_e L^*_{\widehat{\overline{g}}} \left[\left((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}})) \right) \circ \mathrm{d}h(\widehat{\overline{g}}) \right]. \tag{4.34}$$

From the definition of the matrices D and Q it follows that

$$(D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}})) = \widetilde{\mathrm{y}}^T. \tag{4.35}$$

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^5$, the differential of h in $\widehat{\overline{g}}$ applied to $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\begin{aligned} &\mathrm{d}h\big(\widehat{\overline{g}}\big)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} \widehat{x}(s) + \ell\, \cos\widehat{\theta}(s) \\ \widehat{y}(s) + \ell\, \sin\widehat{\theta}(s) \\ \widehat{x}(s) - \ell\, \cos\widehat{\theta}(s) \\ \widehat{y}(s) - \ell\, \sin\widehat{\theta}(s) \\ \widehat{\omega}(s) \\ \widehat{v}(s) \end{bmatrix} = \begin{bmatrix} \widehat{x}' - \ell\, \widehat{\theta}' \sin\widehat{\theta} \\ \widehat{y}' + \ell\, \widehat{\theta}' \cos\widehat{\theta} \\ \widehat{x}' + \ell\, \widehat{\theta}' \sin\widehat{\theta} \\ \widehat{y}' - \ell\, \widehat{\theta}' \cos\widehat{\theta} \\ \widehat{\omega}' \\ \widehat{v}' \end{bmatrix}, \end{aligned} \tag{4.36}$$

and we can write the operator $\mathrm{d}h\big(\widehat{\overline{g}}\big)$ as

$$\mathrm{d}h\big(\widehat{\overline{g}}\big) = \begin{bmatrix} -\ell\, \sin\widehat{\theta} & 1 & 0 & 0 & 0 \\ +\ell\, \cos\widehat{\theta} & 0 & 1 & 0 & 0 \\ +\ell\, \sin\widehat{\theta} & 1 & 0 & 0 & 0 \\ -\ell\, \cos\widehat{\theta} & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}. \tag{4.37}$$

Using (4.37) and (4.35) we obtain

$$\begin{aligned} &\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big) \circ \mathrm{d}h(\widehat{\overline{g}}) = \\ &= \begin{bmatrix} -(\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell\, \sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell\, \cos\widehat{\theta} \\ \widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3 \\ \widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4 \\ \widetilde{\mathrm{y}}_5 \\ \widetilde{\mathrm{y}}_6 \end{bmatrix}^T. \end{aligned} \tag{4.38}$$

Evaluating $T_eL^*_{\widehat{\overline{g}}}$ on (4.38) we finally get (4.27).

Computation of A

The expression for the operator $A : \mathbb{R}^5 \to \mathbb{R}^5$ is

$$A = \mathrm{d}_1\lambda(\widehat{\overline{g}}, u) \circ T_eL_{\widehat{\overline{g}}} - \mathrm{ad}_{\lambda(\widehat{\overline{g}},u)} - T_{\lambda(\widehat{\overline{g}},u)}. \tag{4.39}$$

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^5$, the differential of λ is

$$\begin{aligned} &\mathrm{d}_1\lambda(\widehat{\overline{g}}, u)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} \lambda_g(s) \\ \lambda_V(s) \end{bmatrix} = \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} \widehat{\omega}(s) \\ \widehat{v}(s) \\ 0 \\ -\frac{ma\widehat{\omega}(s)\widehat{v}(s)}{J+ma^2} + \frac{\tau}{J+ma^2} \\ a\widehat{\omega}^2(s) + \frac{F}{m} \end{bmatrix} = \begin{bmatrix} \widehat{\omega}' \\ \widehat{v}' \\ 0 \\ -\frac{ma(\widehat{\omega}\widehat{v}' + \widehat{v}\widehat{\omega}')}{J+ma^2} \\ 2a\widehat{\omega}\widehat{\omega}' \end{bmatrix} \end{aligned} \tag{4.40}$$

and thus

$$\mathrm{d}_1\lambda(\widehat{\overline{g}},u)\circ T_eL_{\widehat{\overline{g}}}=\begin{bmatrix}0&0&0&1&0\\0&0&0&0&1\\0&0&0&0&0\\0&0&0&-\frac{ma\widehat{v}}{J+ma^2}&-\frac{ma\widehat{\omega}}{J+ma^2}\\0&0&0&2a\widehat{\omega}&0\end{bmatrix}. \tag{4.41}$$

The adjoint matrix representation (4.13) implies

$$\mathrm{ad}_{\lambda(\widehat{\overline{g}},u)}=\begin{bmatrix}0&0&0&0_{1\times2}\\0&0&-\widehat{\omega}&0_{1\times2}\\-\widehat{v}&\widehat{\omega}&0&0_{1\times2}\\0_{2\times1}&0_{2\times1}&0_{2\times1}&0_{2\times2}\end{bmatrix}. \tag{4.42}$$

Since we consider the Cartan-Schouten (0)-connection form $\omega^{(0)}=\frac{1}{2}\mathrm{ad}$, the torsion function $T_{\lambda(\widehat{\overline{g}},u)}$ vanishes (see [22]), thus, in matrix form, it is given by

$$T_{\lambda(\widehat{\overline{g}},u)}=\begin{bmatrix}0_{3\times3}&0_{3\times2}\\0_{2\times3}&0_{2\times2}\end{bmatrix}. \tag{4.43}$$

Using (4.41), (4.42) and (4.43) we obtain (4.29).

Computation of E

The operator $E:\mathbb{R}^5\to(\mathbb{R}^5)^*$ takes the form

$$\begin{aligned}E=&-T_eL^*_{\widehat{\overline{g}}}\circ\big[\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}}\circ\mathrm{Hess}\,h(\widehat{\overline{g}})\\&-(\mathrm{d}h(\widehat{\overline{g}}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ\mathrm{d}h(\widehat{\overline{g}})\big]\circ T_eL_{\widehat{\overline{g}}}.\end{aligned} \tag{4.44}$$

From (4.37) and the definitions of the matrices D and Q we can find the composition

$$\begin{aligned}&(\mathrm{d}h(\widehat{\overline{g}}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ\mathrm{d}h(\widehat{\overline{g}})=\\&=\mathrm{diag}\left(2\frac{q_2}{d_2^2}\ell^2,2\frac{q_2}{d_2^2},2\frac{q_2}{d_2^2},\frac{q_3}{d_3^2},\frac{q_4}{d_4^2}\right).\end{aligned} \tag{4.45}$$

The function

$$\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h(\widehat{\overline{g}}))\big):\mathbb{R}^6\to(\mathbb{R}^6)^*$$

in lifted through the exponential functor $(\cdot)^{T_{\widehat{\overline{g}}}\overline{G}}$ to the linear map

$$\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}}:\mathcal{L}(T_{\widehat{\overline{g}}}\overline{G},\mathbb{R}^6)\to\mathcal{L}(T_{\widehat{\overline{g}}}\overline{G},(\mathbb{R}^6)^*)$$

defined as

$$\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}(\xi)}=\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h(\widehat{\overline{g}}))\big)\circ\xi.$$

Let $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{\theta}_1', \widehat{x}_1', \widehat{y}_1', \widehat{\omega}_1', \widehat{v}_1')^T$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{\theta}_2', \widehat{x}_2', \widehat{y}_2', \widehat{\omega}_2', \widehat{v}_2')^T \in T_{\widehat{\overline{g}}}\overline{G}$ be two vector fields, then the Hessian matrix is defined by

$$\begin{aligned}\operatorname{Hess} h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})) =& \mathrm{d}(\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) \\ &- \mathrm{d}h(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))),\end{aligned} \tag{4.46}$$

from the choice of Cartan-Schouten (0)-connection, we get

$$\nabla_{T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}}(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \frac{1}{2}T_eL_{\widehat{\overline{g}}}(\mathrm{ad}_{\eta^{\overline{g}_1}}\eta^{\overline{g}_2}). \tag{4.47}$$

The Hessian evaluated in $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}$ and $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}$ is therefore

$$\operatorname{Hess}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \begin{bmatrix} -\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\cos\widehat{\theta} - \frac{1}{2}\widehat{\theta}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\theta}_1'\widehat{y}_2' \\ +\widehat{\theta}_1'\widehat{\theta}_2'\ell\,\sin\widehat{\theta} + \frac{1}{2}\widehat{\theta}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \end{bmatrix}. \tag{4.48}$$

From (4.35) and (4.48) it follows that

$$\begin{aligned}&\left((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\right)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \operatorname{Hess} h(\widehat{\overline{g}}) = \\ &\begin{bmatrix} a_{1,1} & -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & 0_{1\times 2} \\ \frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & 0 & 0 & 0_{1\times 2} \\ -\frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & 0 & 0 & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix},\end{aligned} \tag{4.49}$$

$$a_{1,1} = -\widetilde{\mathrm{y}}_1\ell\,\cos\widehat{\theta} - \widetilde{\mathrm{y}}_2\ell\,\sin\widehat{\theta} + \widetilde{\mathrm{y}}_3\ell\,\cos\widehat{\theta} + \widetilde{\mathrm{y}}_4\ell\,\sin\widehat{\theta}.$$

In conclusion, combining (4.45) and (4.49) with $T_eL^*_{\widehat{\overline{g}}}$ and $T_eL_{\widehat{\overline{g}}}$, the matrix (4.30) is obtained.

Computation of W

From the adjoint matrix form (4.13) and the Cartan-Schouten (0)-connection (see, e.g., [33],[34]), we have

$$\begin{aligned}W(K, r) &= \frac{1}{2}\mathrm{ad}_{((K_{11}r^g + K_{12}r^V)^\wedge, (K_{21}r^g + K_{22}r^V)^\wedge)} \\ &= \begin{bmatrix} \frac{1}{2}\mathrm{ad}_{(K_{11}r^g + K_{12}r^V)^\wedge} & 0_{3\times 2} \\ 0_{2\times 3} & 0_{2\times 2} \end{bmatrix}.\end{aligned} \tag{4.50}$$

Initial condition

The initial condition for the filter is given by (2.33) while the initial condition for the gain is $K(t_0) = X_0^{-1}$ where the operator $X_0 : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ satisfies (2.36).

We rewrite m_0 as

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 = \frac{1}{2}\mathrm{tr}\left[(I_{6\times 6} - \overline{g}^{-1}\overline{g}_0)^T(I_{6\times 6} - \overline{g}^{-1}\overline{g}_0)\right]. \quad (4.51)$$

From (4.12) it follows

$$I_{6\times 6} - \overline{g}^{-1}\overline{g}_0 = \begin{bmatrix} I_{3\times 3} - g^{-1}g_0 & 0_{3\times 2} & 0_{3\times 1} \\ 0_{2\times 3} & 0_{2\times 2} & V_0 - V \\ 0_{1\times 3} & 0_{1\times 2} & 0_{1\times 1} \end{bmatrix} \quad (4.52)$$

and thus

$$\begin{aligned}(I_{6\times 6} - \overline{g}^{-1}\overline{g}_0)^T&(I_{6\times 6} - \overline{g}^{-1}\overline{g}_0) \\ &= \begin{bmatrix} (I_{3\times 3} - g^{-1}g_0)^T(I_{3\times 3} - g^{-1}g_0) & 0_{3\times 2} & 0_{3\times 1} \\ 0_{2\times 3} & 0_{2\times 2} & 0_{2\times 1} \\ 0_{1\times 3} & 0_{1\times 2} & (V_0 - V)^T(V_0 - V) \end{bmatrix}. \end{aligned} \quad (4.53)$$

Computing the trace we obtain

$$\begin{aligned} m_0(\overline{g}) =& \frac{1}{2}[4(1 - \cos(\theta - \theta_0)) + (x - x_0)^2 + (y - y_0)^2 \\ &+ 2(\omega - \omega_0)^2 + (v - v_0)^2] \end{aligned} \quad (4.54)$$

and from (2.33) it follows that $\widehat{\overline{g}}(t_0) = \overline{g}_0$.

The Hessian of the function m_0 at a point $\overline{g} \in \overline{G}$ is defined as

$$\mathrm{Hess}\, m_0(\overline{g})(\overline{g}X)(\overline{g}Y) = \mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) - \mathrm{d}m_0(\overline{g})(\nabla_{\overline{g}X}(\overline{g}Y))$$

for all $\overline{g}X$, $\overline{g}Y \in T_{\overline{g}}\overline{G}$. The differential of m_0 is given by

$$\mathrm{d}m_0(\overline{g}) = \begin{bmatrix} 2\sin(\theta - \theta_0) & (x - x_0) & (y - y_0) & 2(\omega - \omega_0) & (v - v_0) \end{bmatrix} \quad (4.55)$$

while, given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{g}\eta^{g_1}, \eta^{V_1})$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{g}\eta^{g_2}, \eta^{V_2}) \in T_{\overline{g}\overline{G}}$, the affine connection yields

$$\nabla_{(\widehat{g}\eta^{g_1}, \eta^{V_1})}(\widehat{g}\eta^{g_2}, \eta^{V_2}) = \frac{1}{2}\begin{bmatrix} 0 \\ \widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ 0 \end{bmatrix}. \quad (4.56)$$

Combining (4.55) and (4.56) we obtain

$$\begin{aligned} &\mathrm{d}m_0(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) \\ &\quad = \left[(x - x_0)(\widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2') + (y - y_0)(-\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2')\right] \end{aligned} \quad (4.57)$$

that evaluating in $\widehat{\overline{g}}_0$ produces

$$\mathrm{d}m_0(\widehat{\overline{g}}_0)(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) = 0. \quad (4.58)$$

The double differential takes the form

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{2\cos(\theta-\theta_0), 1, 1, 2, 1\} \tag{4.59}$$

that, evaluating in $\widehat{\overline{g}}_0$, produces

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{2, 1, 1, 2, 1\} \tag{4.60}$$

and thus, from (4.58) and (4.60)

$$\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) = \mathrm{diag}\{2, 1, 1, 2, 1, 1\}. \tag{4.61}$$

From

$$K(t_0) = X_0^{-1} = (T_e L^*_{\widehat{\overline{g}}_0} \circ \mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) \circ T_e L_{\widehat{\overline{g}}_0})^{-1} \tag{4.62}$$

we obtain the initial condition of $K(t_0)$

$$\begin{aligned} K(t_0) &= (T_e L_{\widehat{\overline{g}}_0})^{-1} (\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0))^{-1} (T_e L^*_{\widehat{\overline{g}}_0})^{-1} \\ &= \mathrm{diag}\{1/2, 1, 1, 1/2, 1\}. \end{aligned} \tag{4.63}$$

This computation ends the proof. □

Remark 1. If one considers the case where only GPS measurements (the first four lines in (4.8)) are available, then the residual r and the matrix E become:

$$r = \begin{bmatrix} -(\widetilde{\mathrm{y}}_1 - \widetilde{\mathrm{y}}_3)\ell\, \sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 - \widetilde{\mathrm{y}}_4)\ell\, \cos\widehat{\theta} \\ (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\theta} \\ -(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\theta} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\theta} \\ 0_{2\times 1} \end{bmatrix}^T, \tag{4.64}$$

$$E = \begin{bmatrix} a_{1,1} & a_{1,2} & a_{1,3} & 0_{1\times 2} \\ a_{2,1} & 2\frac{q_2}{d_2^2} & 0 & 0_{1\times 2} \\ a_{3,1} & 0 & 2\frac{q_2}{d_2^2} & 0_{1\times 2} \\ 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 1} & 0_{2\times 2} \end{bmatrix}. \tag{4.65}$$

The state is still observable using a computation similar to [40].

4.4 Optimal filter with nonholonomic constraint

The aim of this section is to investigate under which conditions the filter preserves the nonholonomic constraint (4.1), that the formulation in Section 4.3 cannot guarantee. The preservation of the nonholonomic constraint produces more feasible trajectories that will be more suitable for control implementations.

We first observe that, if the dynamics does not present model errors, the nonholonomic information is enclosed by the left trivialized dynamics, in particular by its "group" part

$$\lambda_g(\overline{g}, u) = \begin{bmatrix} \omega \\ v \\ 0 \end{bmatrix}, \tag{4.66}$$

where the last entry is equal to 0, because it represents $v_\perp$, that is the orthogonal projection of the linear velocity with respect to the first axis direction. In order to preserve the constraint, it is necessary to "force" the filter to keep such value equal to 0 on the right-hand side of the equations (4.25)-(4.26). It is not possible to act directly on the filter parameters in order to impose the third component of the residual (4.27) be always 0, thus, it is necessary to operate on the gain operator K structure whose dynamics is governed by the Riccati equation (4.28).

Whenever the third row and column of the operator K in (4.28) are 0, the product KEK has the third row and column equal to 0, whatever the components of E. Moreover, since the system has only model errors that apply to the dynamics that model the evolution of the velocity and not to the kinematics, the term $BR^{-1}B^T$ has non-null components only on the last 2×2 submatrix. The analysis can be therefore limited to the study of the operators A and W and on the choice of the connection function. Continuing working with skew-symmetric connection functions (see e.g. [9], [30]), a choice of the connection could be $\nabla_X Y = \lambda[X, Y]$ with $\lambda \in \mathbb{R}$. A generic version of the operator A provided in (4.29) suggests, for our purposes, to set $\lambda = 0$, which corresponds to the Cartan Schouten (-)-connection $\omega^{(-)} = 0$ and produces

$$A = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\frac{ma}{J+ma^2}\widehat{v} & -\frac{ma}{J+ma^2}\widehat{\omega} \\ 0 & 0 & 0 & 2a\widehat{\omega} & 0 \end{bmatrix}. \tag{4.67}$$

With this choice, the operator W is represented by

$$W = 0_{5\times 5}, \tag{4.68}$$

and then the products $W(K, r)K$ and $K^T W(K, r)$ have the third rows and columns equal to 0. The new matrix representation for the operator E becomes:

$$E = \begin{bmatrix} a_{1,1} & 0 & 0 & 0 & 0 \\ 0 & 2\frac{q_2}{d_2^2} & 0 & 0 & 0 \\ 0 & 0 & 2\frac{q_2}{d_2^2} & 0 & 0 \\ 0 & 0 & 0 & \frac{q_3}{d_3^2} & 0 \\ 0 & 0 & 0 & 0 & \frac{q_4}{d_4^2} \end{bmatrix} \tag{4.69}$$

with

$$a_{1,1} = (\widetilde{y}_1 - \widetilde{y}_3)\ell \cos\widehat{\theta} + (\widetilde{y}_2 - \widetilde{y}_4)\ell \sin\widehat{\theta} + 2\frac{q_2}{d_2^2}\ell^2.$$

The preservation of the nonholonomic constraint is then granted by choosing the initial conditions of the operator K as

$$K(t_0) = \begin{bmatrix} 1_{2\times 2} & 0_{2\times 1} & 0_{2\times 2} \\ 0_{1\times 2} & 0_{1\times 1} & 0_{1\times 2} \\ 0_{2\times 2} & 0_{2\times 1} & 1_{2\times 2} \end{bmatrix}. \tag{4.70}$$

4.5 Simulations and discussions

In this section, we show how the second-order optimal filter applied to the Chaplygin sleigh works in simulated scenarios. We set $J = 6.125\ \mathrm{kg\,m^2}$, $m = 125\ \mathrm{kg}$, $a = 0.3\ \mathrm{m}$ and $\ell = 0.2\ \mathrm{m}$. We consider the control inputs $\tau(t) = 10\sin((1/2)t)$ and $F(t) = (1/2)\sin((1/5)t)$. The total simulation time is $T = 20\,\mathrm{s}$. The initial conditions are $\theta(0) = \pi/6\ \mathrm{rad}$, $x(0) = 1\,\mathrm{m}$, $y(0) = 2\,\mathrm{m}$, $\omega(0) = 0\,\mathrm{rad/s}$ and $v(0) = 0\,\mathrm{m/s}$. For the matrix D we choose $d_2 = 0.4$, $d_3 = d_4 = 0.2$, while for the matrix B_2 we choose $b_\omega = b_v = 10$. To solve all the differential equations we used a Runge-Kutta 4^{th} order method with steps $T_s = 10\,\mathrm{ms}$ and $T_s = 1\,\mathrm{ms}$.

Figures 4.3 and 4.4 show how the filters work with a sample time of 10ms in the cases of Cartan-Scouten (-)-connection ($\lambda = 0$) and Cartan-Scouten (0)-connection ($\lambda = 1/2$) using both GPS and INS devices and only GPS. The estimation errors are calculated with respect to the nominal values

$$\Delta\theta := \theta - \widehat{\theta}, \quad \Delta x := x - \widehat{x}, \quad \cdots .$$

Figures 4.5 and 4.6 show the same trajectories and errors in the case of a sample time of 1ms. Table 4.1 and Table 4.2 compare the mean, the standard deviation and the root mean square values of the errors obtained by applying the second-order filter in the GPS+INS and GPS cases, with both connection functions, using $T_s = 10\,\mathrm{ms}$ and $T_s = 1\,\mathrm{ms}$, respectively.

The filter designed on the Chaplygin sleigh using the Cartan-Schouten (0)-connection provides a good estimation both for the pose and the velocity, even if the two antennas do not measure directly the position of the contact point of the blade nor the orientation. The filter designed using the Cartan-Schouten (-)-connection has less accurate estimations but preserves the nonholonomic constraint. With no INS measurements, the accuracy of the velocities for the Cartan-Schouten (-)-connection gets worse, but it allows to gain precision on the positions. Instead, for the Cartan-Schouten (0)-connection the addition of the INS measurements improves all the estimations. In both cases, a finer sampling of measurements greatly improves the accuracy of the estimates as can be seen by comparing the $T_s = 10\,\mathrm{ms}$ and $T_s = 1\,\mathrm{ms}$ cases.

Table 4.1: Mean, standard deviation and root mean square values of the errors for $T_s = 10\,\text{ms}$ for the cases $\lambda = 1/2$ and $\lambda = 0$ (in bracket).

	GPS+INS			GPS		
	μ	σ	rms	μ	σ	rms
$\Delta\theta$ [mrad]	28.7	41.1	50.1	-2.1	180.5	180.4
	(71.1)	(32.7)	(78.2)	(-6.8)	(194.7)	(194.8)
Δx [mm]	-0.3	19.2	19.2	4.9	66.8	67.0
	(-37.8)	(122.6)	(128.3)	(-13.8)	(77.9)	(79.1)
Δy [mm]	-20.5	30.9	37.1	-11.6	59.3	60.4
	(-281.0)	(177.6)	(332.3)	(-51.9)	(103.2)	(115.5)
$\Delta\omega$ [mm/rad]	12.3	94.2	95.0	-2.5	208.8	208.8
	(12.3)	(94.2)	(95.0)	(4.4)	(203.8)	(203.8)
Δv [mm/rad]	4.9	99.4	99.5	0.0	249.5	249.4
	(4.8)	(99.4)	(99.4)	(15.4)	(251.9)	(252.3)

Table 4.2: Mean, standard deviation and root mean square values of the errors for $T_s = 1\,\text{ms}$ for the cases $\lambda = 1/2$ and $\lambda = 0$ (in bracket).

	GPS+INS			GPS		
	μ	σ	rms	μ	σ	rms
$\Delta\theta$ [mrad]	-0.1	6.6	6.6	-6.1	44.4	44.8
	(4.9)	(9.2)	(10.4)	(-4.0)	(57.0)	(57.1)
Δx [mm]	8.3	9.5	12.6	5.5	20.9	21.7
	(-0.4)	(18.0)	(18.0)	(-8.8)	(31.9)	(33.1)
Δy [mm]	-6.1	7.3	9.5	-5.1	16.3	17.1
	(-28.6)	(22.0)	(36.1)	(-33.5)	(36.6)	(49.6)
$\Delta\omega$ [mm/rad]	1.5	31.2	31.2	-1.1	54.2	54.2
	(1.5)	(31.2)	(31.2)	(-0.4)	(57.1)	(57.1)
Δv [mm/rad]	-1.8	33.9	33.9	2.8	77.1	77.2
	(-1.9)	(33.9)	(33.9)	(1.0)	(80.0)	(80.0)

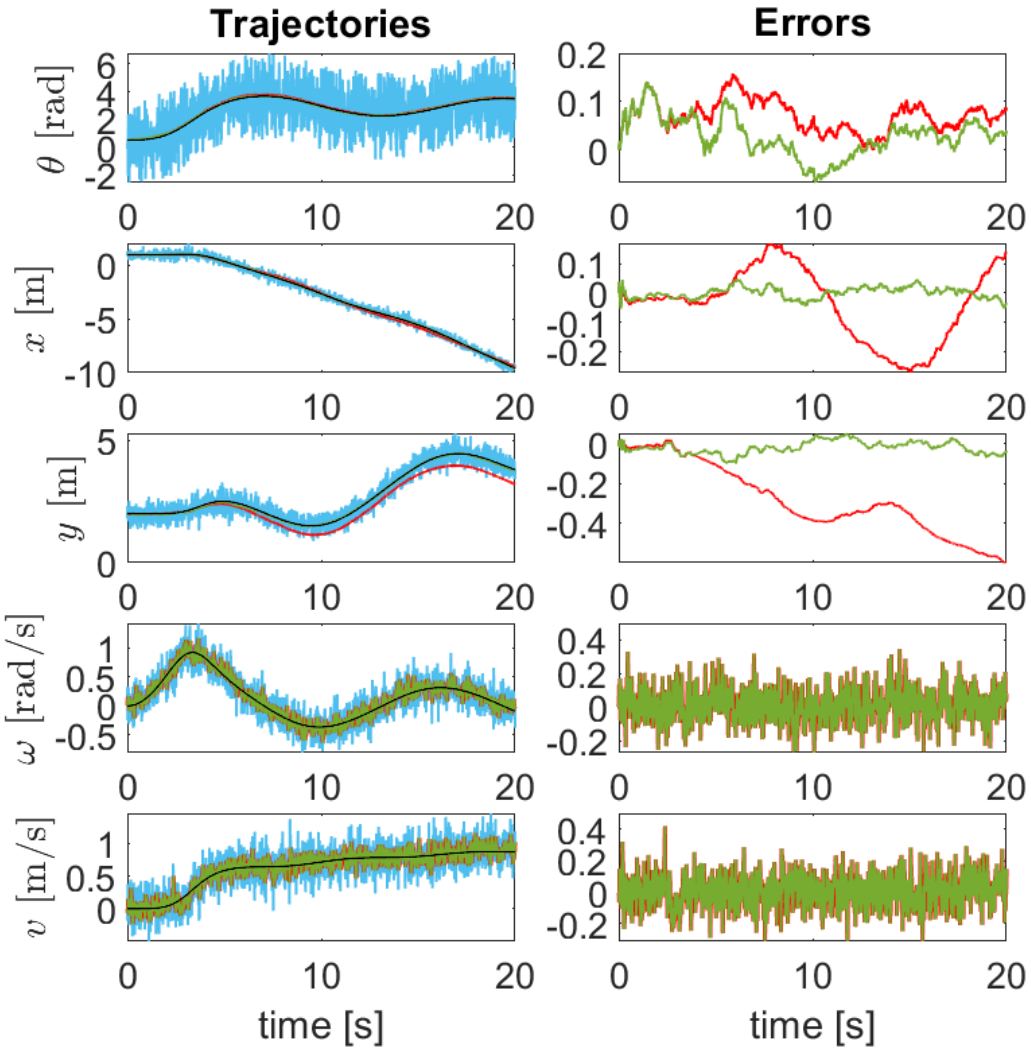


Figure 4.3: Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with GPS and INS in the case $T_s = 10$ms.

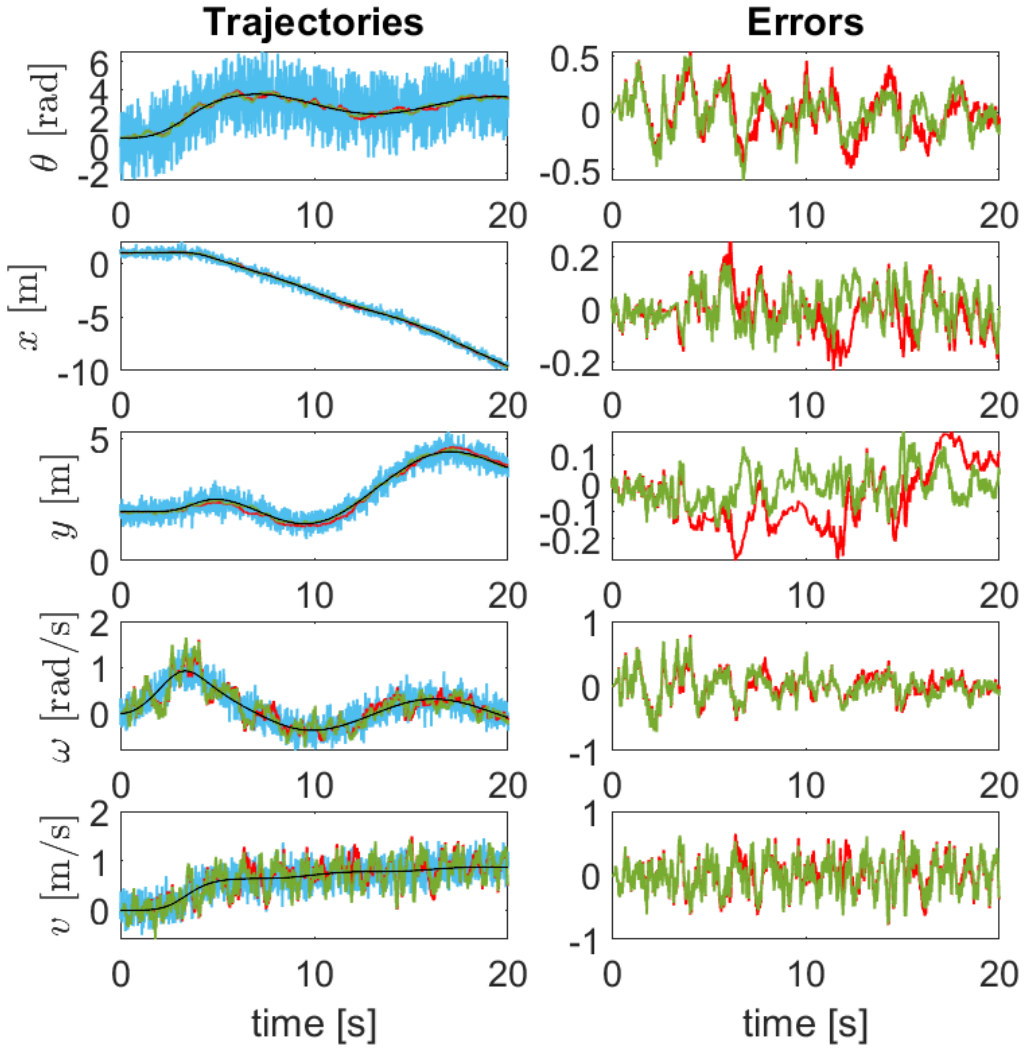


Figure 4.4: Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with only GPS in the case $T_s = 10$ms.

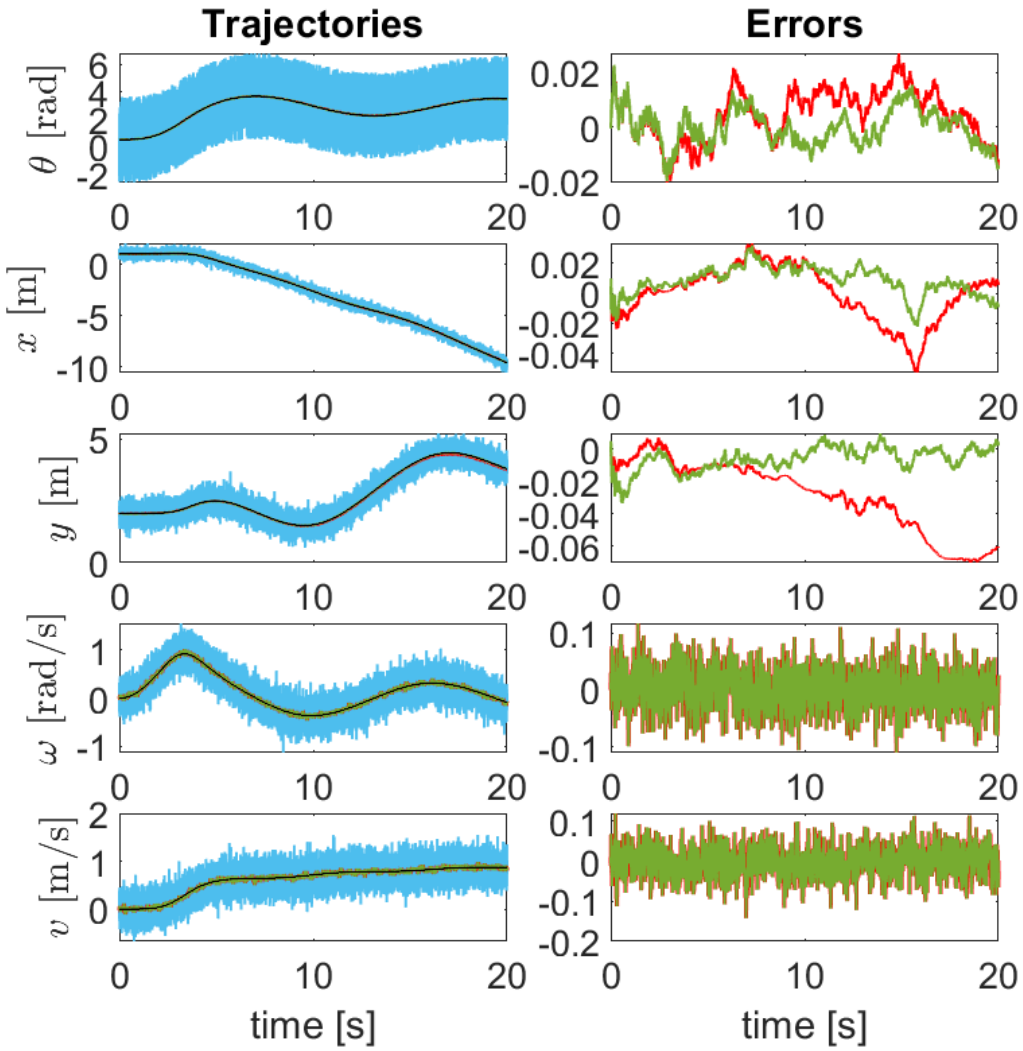


Figure 4.5: Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with GPS and INS in the case $T_s = 1$ms.

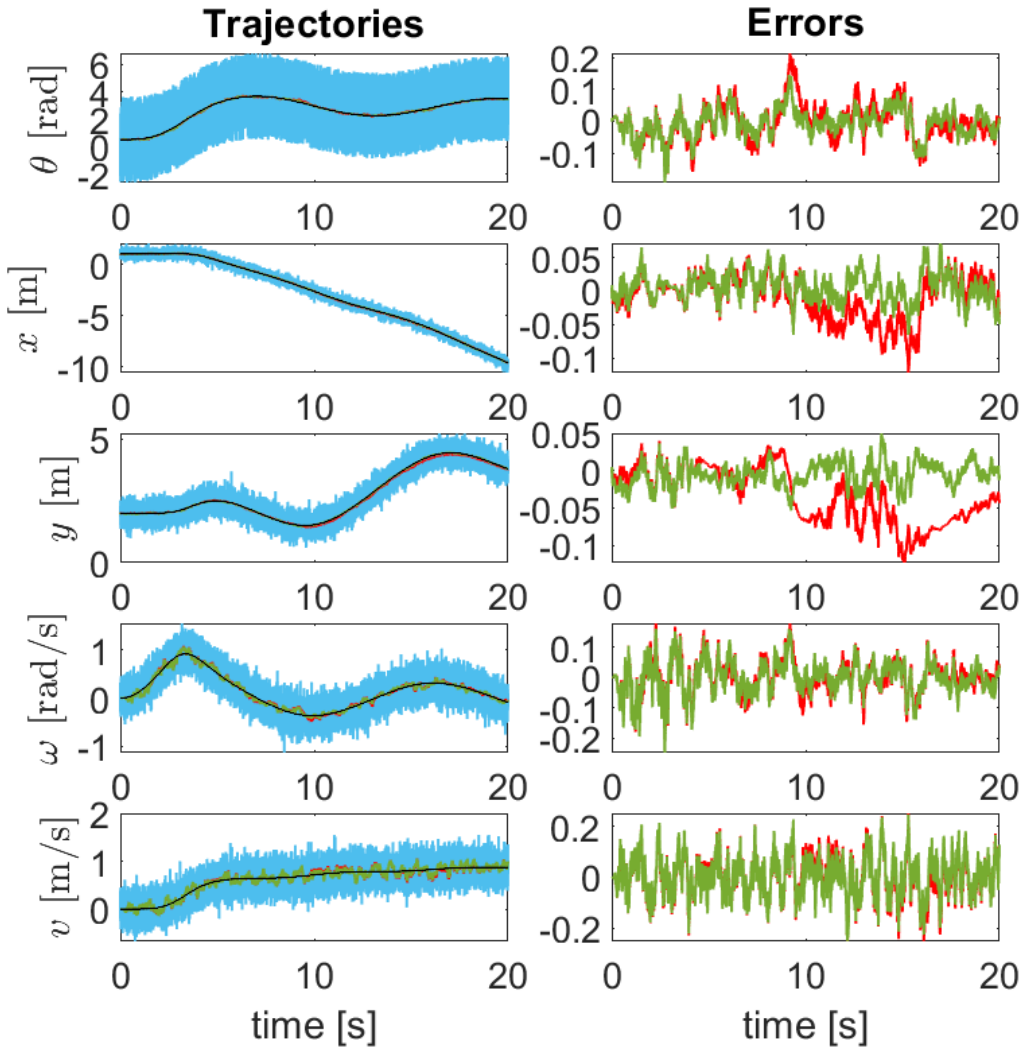


Figure 4.6: Nominal (black), measured (blue), filtered with $\lambda = 1/2$ (green) and filtered with $\lambda = 0$ (red) trajectories on the left and their corresponding errors on the right with only GPS in the case $T_s = 1$ms.

Chapter 5

Second-order optimal filter applied to an n-articulated vehicle system[1]

In this chapter, we design the second-order optimal filter for an articulated n-trailer vehicle when masses and inertias parameters are (i) known at each instant of time, (ii) unknown but time-invariant, treated as constants, (iii) unknown and time-varying and thus treated as state variables and filtered. The nature of the nonholonomic and hooking constraints allows us to estimate the other dynamic variables like the poses of the trailers and their velocities.

5.1 The dynamics of an articulated n-trailer vehicle with different masses and inertias

The system that we consider consists of a leading car $\mathcal{B}_0$, that pulls n trailers $\mathcal{B}_1$,...,$\mathcal{B}_n$. The leading car and each trailer are connected as shown in Figure 5.1 for the case $n = 2$. We denote with (x_0, y_0) the midpoint of the wheels' axis of the leading car and with (x_j, y_j) the midpoint of the wheels' axis of the trailers with respect to the inertial frame $\Sigma_I = \{e_x, e_y\}$. We consider $n + 1$ right-handed body frames $\Sigma_j = \{e^j_{\parallel}, e^j_{\perp}\}$, $j = 0, ..., n$ attached to $\mathcal{B}_j$, centered at (x_j, y_j) with the $e^j_{\parallel}$-axes aligned with the wheels and the $e^j_{\perp}$-axes perpendicular to them. The orientation of each body $\mathcal{B}_j$ is determined by the angle θ_j between the axis $e^j_{\parallel}$ aligned with the wheels and the e_x axis of the inertial frame. The configuration space of the system is the $n + 3$ dimensional manifold $Q = \mathrm{SE}(2) \times \mathbb{T}^n$, where $\mathrm{SE}(2)$ is the special Euclidean group on $\mathbb{R}^2$ that describes the pose of the leading car, and $\mathbb{T}^n$ is the $n-$dimensional torus that identifies the angles between the bodies $\mathcal{B}_0, ..., \mathcal{B}_n$. The leading car and the trailers are subjected to nonholonomic constraints that do not allow them to move in the orthogonal direction with respect to the axis $e^0_{\parallel}$, $e^1_{\parallel}$,..., $e^n_{\parallel}$ i.e.,

$$\dot{x}_j \sin\theta_j - \dot{y}_j \cos\theta_j = 0, \quad j = 0, 1, ..., n. \tag{5.1}$$

[1]This chapter is based on the work

▷ Rigo D., Sansonetto N., & Muradore R. *Second-order optimal filtering for planar rigid body with trailers and unknown parameters* (submitted).

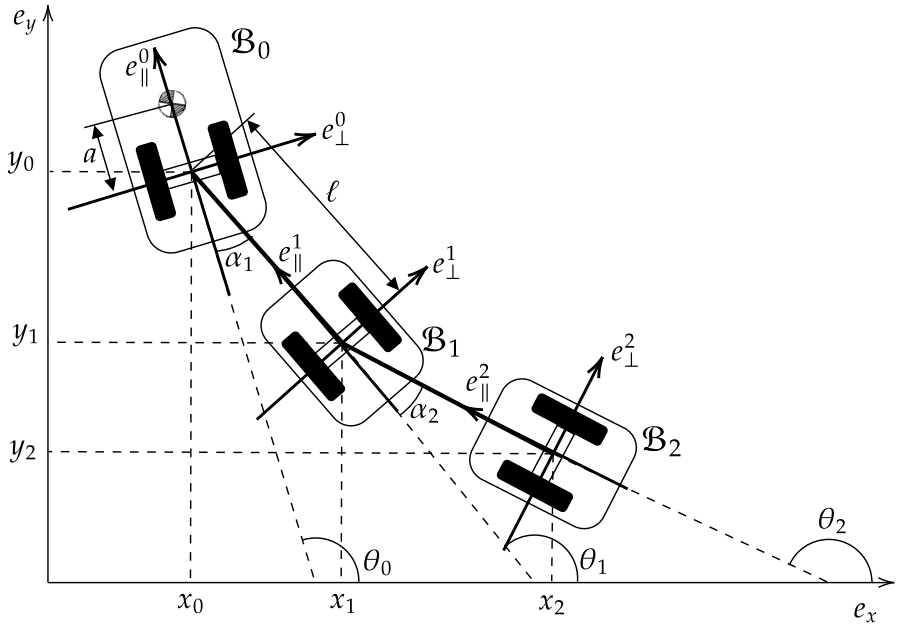


Figure 5.1: Articulated vehicle with 2 trailers.

These constraints define a rank 2 distribution $\mathcal{D}$ on Q (see [6]). We denote the angles between subsequent bodies in the convoy by

$$\alpha_j = \theta_{j-1} - \theta_j, \quad j = 1, ..., n.$$

These new angle coordinates are a base of $\mathbb{T}^n$ and are useful since they do not change under the SE(2) action given by (see [6])

$$\begin{aligned} &g \cdot (x, y, \theta_0, \theta_1 ..., \theta_n) \\ &\quad = (x\cos\varphi - y\sin\varphi + r, x\sin\varphi + y\cos\varphi + s, \theta + \varphi, \theta_1 + \varphi ..., \theta + \varphi) \end{aligned}$$

where

$$g = \begin{bmatrix} \cos\varphi & -\sin\varphi & r \\ \sin\varphi & \cos\varphi & s \\ 0 & 0 & 1 \end{bmatrix} \in \mathrm{SE}(2).$$

We denote with (x_c, y_c) the center of mass of $\mathcal{B}_0$ at distance a from (x_0, y_0):

$$x_c = x_0 + a\cos\theta_0, \quad y_c = y_0 + a\sin\theta_0. \tag{5.2}$$

The leading car is modelled as a Chaplygin sleigh presented in Section 4. Each trailer is connected to the previous one with a rigid link of length ℓ and the hooks of the convoy define the $2n$ holonomic constraints

$$x_j + \ell\cos\theta_j - x_{j-1} = 0, \quad y_j + \ell\cos\theta_j - y_{j-1} = 0, \quad j = 1, ..., n, \tag{5.3}$$

that can be read also as

$$x_j = x_0 - \ell\sum_{k=1}^{j}\cos\theta_k, \quad y_j = y_0 - \ell\sum_{k=1}^{j}\sin\theta_k, \quad j = 1, ..., n. \tag{5.4}$$

From this we can rewrite the nonholonomic constraints (5.1) as

$$\dot{x}\sin\theta_j - \dot{y}\cos\theta_j + \ell\sum_{k=1}^{j}\cos(\theta_j-\theta_k)\dot{\theta}_k = 0, \quad j = 0,...,n. \tag{5.5}$$

We assume that the leading car has a total mass m_0 and moment of inertia around its center of mass J_0, and the $j-$th trailer has the center of mass lying in the midpoint of the axis (x_j, y_j), total mass of m_j and moment of inertia J_j around (x_j, y_j). The velocity components of the leading car are $(v, v_\perp, \omega)$, where v and $v_\perp$ are the (linear) velocities of the body along the $e_\parallel^0$ and $e_\perp^0$ axes, respectively, and ω is the angular velocity. The nonholonomic constraint imposes that the orthogonal component $v_\perp$ vanishes. The linear and angular velocities of the trailers can be obtained by exploiting the hooking and nonholonomic constraints, so v and ω are enough to describe the whole dynamics of the articulated $n-$trailer. The state space of the system is then given by $\mathrm{SE}(2)\times\mathbb{T}^n\times\mathbb{R}^2$. The control inputs $u(t) = (F(t), \tau(t))$ consist of a linear force F applied to the center of the body frame along $e_\parallel^0$ and a torque τ applied around (x_0, y_0). The following theorem is an extension of Theorem 1 presented in [6] in the case of different masses and inertias.

Theorem 5.1. *The reduced equations of motion of the n-trailer vehicle with masses m_0, m_1,...,m_n and inertias J_0, J_1,...,J_n and control inputs F and τ are given by*

$$\begin{cases} \dot{x}_0 &= v\cos\theta_0 \\ \dot{y}_0 &= v\sin\theta_0 \\ \dot{\theta}_0 &= \omega \\ \dot{\alpha}_1 &= \omega - \frac{v\sin\alpha_1}{\ell} \\ \dot{\alpha}_k &= \frac{v}{\ell}\left(\prod_{j=1}^{k-2}\cos\alpha_k\right)(\sin\alpha_{k-1} - \cos\alpha_{k-1}\sin\alpha_k) \quad k = 2,...,n \\ \dot{v} &= -\frac{1}{2R(\alpha)}\left(\sum_{k=1}^{n} A_k\frac{\partial R(\alpha)}{\partial\alpha_k}\right)v^2 + \frac{Q(\alpha)}{\ell^2 R(\alpha)}v\omega + \frac{m_0 a}{R(\alpha)}\omega^2 + \frac{F}{R(\alpha)} \\ \dot{\omega} &= -\frac{m_0 a v\omega}{J_0+m_0a^2} + \frac{\tau}{J_0+m_0a^2} \end{cases} \tag{5.6}$$

where the coefficients A_k are

$$A_k = \frac{1}{\ell}\left(\prod_{j=1}^{k-2}\cos\alpha_j\right)(\sin\alpha_{k-1} - \cos\alpha_{k-1}\sin\alpha_k) \tag{5.7}$$

and

$$\begin{aligned} Q(\alpha) := &\left(-\frac{J_1}{2} + \ell^2\sum_{j=1}^{n} m_j\left(\prod_{k=2}^{j}\cos^2\alpha_k\right)\right. \\ &\left. + \frac{1}{2}\sum_{j=2}^{n}\left(J_j\left(\prod_{k=2}^{j-1}\cos^2\alpha_k\right)\sin^2\alpha_j\right)\right)\cos\alpha_1\sin\alpha_1, \end{aligned} \tag{5.8}$$

$$R(\alpha) := m_0 + \sum_{j=1}^{n} \left(m_j \prod_{k=1}^{j} \cos^2 \alpha_k \right) + \frac{1}{\ell^2} \sum_{j=1}^{n} \left(J_j \left(\prod_{k=1}^{j-1} \cos^2 \alpha_k \right) \sin^2 \alpha_j \right). \tag{5.9}$$

Proof. The construction of the kinematic system for (5.6) (i.e., the equations for $\dot{x}_0$, $\dot{y}_0$, $\dot{\theta}_0$, $\dot{\alpha_1}$, and $\dot{\alpha_k}$) follows exactly the one proposed in [6], thus we will focus on the dynamic part.

Differentiating the equations (5.2) and (5.4), and adding the contributions of all the bodies we obtain the Lagrangian function of the system

$$\begin{aligned}\mathcal{L} =& \frac{1}{2} \Bigg((J_0 + m_0 a^2)\dot{\theta}_0^2 + \Big(m_0 + \sum_{j=1}^{n} m_j \Big)(\dot{x}_0^2 + \dot{y}_0^2) \\ &+ 2m_0 a \dot{\theta}_0 (\dot{y}_0 \cos\theta_0 - \dot{x}_0 \sin\theta_0) + 2\ell \sum_{j=1}^{n} m_j \sum_{k=1}^{n} \dot{\theta}_k (\dot{x}_0 \sin\theta_k - \dot{y}_0 \cos\theta_k) \\ &+ \sum_{k=1}^{n} \Big(J_k + \sum_{j=k}^{n} m_j \ell^2 \Big) \dot{\theta}_k^2 + 2\ell^2 \sum_{i=1}^{n} \sum_{k=i+1}^{n} \sum_{j=k}^{n} \dot{\theta}_i \dot{\theta}_k \cos(\theta_i - \theta_k) \Bigg).\end{aligned} \tag{5.10}$$

The nonholonomic constraints in (5.5) do not impose any restriction on the value of $\dot{\theta}_0 = \omega$, thus we can apply the forced Euler-Lagrange equation to the Lagrangian (5.10) finding

$$\dot{\omega} = -\frac{m_0 a v \omega}{J_0 + m_0 a^2} + \frac{\tau}{J_0 + m_0 a^2}.$$

where we set τ as the external torque.

To find the equations of motion of v, we exploit the constrained Lagrangian $\mathcal{L}_c$ that is the restriction of $\mathcal{L}$ to the constraint distribution $\mathcal{D}$. First of all, we notice that

$$\dot{x}_0 = v\cos\theta_0,\ \dot{y}_0 = v\sin\theta_0,\ \dot{\theta}_0 = \omega,\ \dot{\theta}_k = \frac{v\sin\alpha_k}{\ell} \prod_{j=1}^{k-1} \cos\alpha_j,\ k = 1, ..., n$$

and that, if $j \geq 1$, we have

$$\dot{x}_j^2 + \dot{y}_j^2 = v^2 \prod_{k=1}^{j} \cos\alpha_k.$$

It follows that the kinetic energy of the $j-$th trailer is given by

$$\mathcal{K}_j = \frac{1}{2}\Big(J_j \dot{\theta}_j^2 + m_j(\dot{x}_j^2 + \dot{y}_j^2) \Big) = \frac{v^2}{2} \left(\prod_{k=1}^{j-1} \cos^2\alpha_k \right) \left(\frac{J_j}{\ell^2} \sin^2\alpha_j + m_j \cos^2\alpha_j \right),$$

while for $j = 0$, the leading car, we have

$$\mathcal{K}_0 = \frac{1}{2}\Big(J_0 \dot{\theta}^2 + m_0(\dot{x}_c^2 + \dot{y}_c^2) \Big) = \frac{1}{2}\Big((J_0 + m_0 a^2)\omega^2 + m_0 v^2 \Big).$$

Adding up all the contributions of the body in the convoy, we get the constrained Lagrangian

$$\mathcal{L}_c = \frac{1}{2}\left(R(\alpha)v^2 + (J_0 + m_0a^2)\omega^2\right)$$

where $R(\alpha)$ is given by (5.9). Following the approach in [6] and [14] (in the case of an external force F) we obtain the equations of motion

$$\frac{d}{dt}\left(\frac{\partial \mathcal{L}_c}{\partial v}\right) = -\mathcal{C}_{12}^1\omega\frac{\partial \mathcal{L}_c}{\partial v} - \mathcal{C}_{12}^2\omega\frac{\partial \mathcal{L}_c}{\partial \omega} + \sum_{k=1}^{n} A_k\frac{\partial \mathcal{L}_c}{\partial \alpha_k} + F$$

where

$$\mathcal{C}_{12}^1 = \frac{Q(\alpha)}{\ell^2 R(\alpha)}, \quad \mathcal{C}_{12}^2 = -\frac{m_0 a}{J_0 + m_0a^2}$$

with $Q(\alpha)$ and $R(\alpha)$ given by (5.8) and (5.9), respectively, and thus

$$\frac{d}{dt}\left(R(\alpha)v\right) = -\frac{Q(\alpha)}{\ell^2}v\omega + m_0a\omega^2 + \frac{1}{2}\sum_{k=1}^{n} A_k\frac{\partial R(\alpha)}{\partial \alpha_k}v^2 + F. \tag{5.11}$$

From the derivative of the product, the left-hand side becomes

$$\frac{d}{dt}(R(\alpha)v) = v\sum_{k=1}^{n}\frac{\partial R(\alpha)}{\partial \alpha_k}\dot{\alpha}_k + R(\alpha)\dot{v}. \tag{5.12}$$

Then from (5.6) and (5.7), the derivative of the angles can be written as

$$\dot{\alpha}_1 = \omega + vA_1, \quad \dot{\alpha}_k = vA_k, \quad k = 2, ..., n, \tag{5.13}$$

while from (5.8) and (5.9) we have

$$\frac{\partial R(\alpha)}{\partial \alpha_1} = -\frac{2Q(\alpha)}{\ell^2}. \tag{5.14}$$

Substituting (5.12), (5.13) and (5.14) into (5.11) we finally obtain

$$\dot{v} = -\frac{1}{2R(\alpha)}\left(\sum_{k=1}^{n} A_k\frac{\partial R(\alpha)}{\partial \alpha_k}\right)v^2 + \frac{Q(\alpha)}{\ell^2 R(\alpha)}v\omega + \frac{m_0a}{R(\alpha)}\omega^2 + \frac{F}{R(\alpha)}.$$

□

5.2 The $\mathrm{SE}(2) \times \mathbb{T}^n \times \mathbb{R}^2$ structure

From here on, we will indicate the pose (θ_0, x_0, y_0) as (θ, x, y) in order to not burden the notation.

The state space of the $n-$trailer system is given by the Lie group $\overline{G} = \mathrm{SE}(2) \times \mathbb{T}^n \times \mathbb{R}^2$. An element $\overline{g} \in \overline{G}$ admits the following block-diagonal matrix representation

$$\overline{g} = \begin{bmatrix} g & & & & \\ & A_1 & & & \\ & & \ddots & & \\ & & & A_n & \\ & & & & V \end{bmatrix}$$

with

$$g = \begin{bmatrix} \cos\theta & -\sin\theta & x \\ \sin\theta & \cos\theta & y \\ 0 & 0 & 1 \end{bmatrix}, A_k = \begin{bmatrix} \cos\alpha_k & -\sin\alpha_k \\ \sin\alpha_k & \cos\alpha_k \end{bmatrix}, V = \begin{bmatrix} 1 & 0 & v \\ 0 & 1 & \omega \\ 0 & 0 & 1 \end{bmatrix},$$

where x, y, θ, α_k, v, ω are defined in Section 5.1. The group operation is given by the matrix multiplication

$$\overline{g}_1 \cdot \overline{g}_2 = \begin{bmatrix} g^1 g^2 & & & & \\ & A_1^1 A_1^2 & & & \\ & & \ddots & & \\ & & & A_n^1 A_n^2 & \\ & & & & V^1 V^2 \end{bmatrix}. \tag{5.15}$$

The unit element of this group is the identity matrix and the inverse operation is the matrix inversion. The generic element $\eta^{\overline{g}}$ of the Lie algebra $\overline{\mathfrak{g}}$ of $\overline{G}$ can be represented as

$$\eta^{\overline{g}} = \begin{bmatrix} \eta^g & & & & \\ & \eta^{A_1} & & & \\ & & \ddots & & \\ & & & \eta^{A_n} & \\ & & & & \eta^V \end{bmatrix}$$

where

$$\eta^g = \begin{bmatrix} 0 & -\eta^\theta & \eta^x \\ \eta^\theta & 0 & \eta^y \\ 0 & 0 & 0 \end{bmatrix}, \eta^{A_k} = \begin{bmatrix} 0 & -\eta^{\alpha_k} \\ \eta^{\alpha_k} & 0 \end{bmatrix}, \eta^V = \begin{bmatrix} 0 & 0 & \eta^v \\ 0 & 0 & \eta^\omega \\ 0 & 0 & 0 \end{bmatrix},$$

and zero on the other entries. For convenience, we define the Lie algebra isomorphism $\wedge : \mathbb{R}^{n+5} \to \overline{\mathfrak{g}}$ as

$$\begin{bmatrix} \eta^x \\ \eta^y \\ \eta^\theta \\ \eta^{\alpha_1} \\ \vdots \\ \eta^{\alpha_n} \\ \eta^v \\ \eta^\omega \end{bmatrix}^\wedge = \eta^{\overline{g}}$$

from the Lie algebra $(\mathbb{R}^{n+5}, \star)$ to the Lie algebra $(\overline{\mathfrak{g}}, [\cdot,\cdot])$, where $[\cdot,\cdot]$ is the usual matrix commutator and where $\star : \mathbb{R}^{n+5} \times \mathbb{R}^{n+5} \to \mathbb{R}^{n+5}$ is the Lie bracket operation defined as

$$\begin{bmatrix} \eta_1^x \\ \eta_1^y \\ \eta_1^\theta \\ \eta_1^{\alpha_1} \\ \vdots \\ \eta_1^{\alpha_n} \\ \eta_1^v \\ \eta_1^\omega \end{bmatrix} \star \begin{bmatrix} \eta_2^x \\ \eta_2^y \\ \eta_2^\theta \\ \eta_2^{\alpha_1} \\ \vdots \\ \eta_2^{\alpha_n} \\ \eta_2^v \\ \eta_2^\omega \end{bmatrix} = \begin{bmatrix} \eta_1^y\eta_2^\theta - \eta_1^\theta\eta_2^y \\ \eta_1^\theta\eta_2^x - \eta_1^x\eta_2^\theta \\ 0 \\ 0 \\ \vdots \\ 0 \\ 0 \\ 0 \end{bmatrix},$$

(see, e.g., [25]). Given $\eta^{\overline{g}_1} \in \overline{\mathfrak{g}}$, the adjoint operator $\mathrm{ad}_{\eta^{\overline{g}_1}} \in \mathcal{L}(\overline{\mathfrak{g}}; \overline{\mathfrak{g}})$ admits the matrix representation

$$\mathrm{ad}_{\eta^{\overline{g}_1}} = \begin{bmatrix} 0 & -\eta^{\theta_1} & \eta^{y1} & 0_{1\times(n+2)} \\ \eta^{\theta_1} & 0 & -\eta^{x1} & 0_{1\times(n+2)} \\ 0 & 0 & 0 & 0_{1\times(n+2)} \\ 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times(n+2)} \end{bmatrix}. \tag{5.16}$$

It is useful to notice that the tangent map for left matrix multiplication is also left matrix multiplication: $T_eL_{\overline{g}}(\eta^{\overline{g}}) = \overline{g}\eta^{\overline{g}} \in T_{\overline{g}}\overline{G}$ (see [7]). Given $\overline{g} \in \overline{G}$ the left multiplication tangent map admits the following matrix representation

$$T_eL_{\overline{g}} = \begin{bmatrix} \cos\theta & -\sin\theta & 0 & 0_{1\times(n+2)} \\ \sin\theta & \cos\theta & 0 & 0_{1\times(n+2)} \\ 0 & 0 & 1 & 0_{1\times(n+2)} \\ 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & I_{(n+2)\times(n+2)} \end{bmatrix}$$

while its dual operator matrix representation satisfies $T_eL^*_{\overline{g}} = (T_eL_{\overline{g}})^T$.

Using this geometric construction, we can rewrite the dynamics (5.6) as

$$\dot{\overline{g}}(t) = \overline{g}(t)\lambda(\overline{g}(t), u(t), t)^\wedge, \quad \overline{g}(t_0) = \overline{g}_0, \tag{5.17}$$

where $\overline{g}(t) \in \overline{G}$ is the state, $u(t) \in \mathbb{R}^2$ is the input and $\lambda : \overline{G} \times \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}^{n+5}$ is the left-trivialized dynamics

$$\lambda = \begin{bmatrix} \dot{x}\cos\theta + \dot{y}\sin\theta \\ -\dot{x}\sin\theta + \dot{y}\cos\theta \\ \dot{\theta} \\ \dot{\alpha}_1 \\ \vdots \\ \dot{\alpha}_k \\ \vdots \\ \dot{v} \\ \dot{\omega} \end{bmatrix} = \begin{bmatrix} v \\ 0 \\ \omega \\ \omega - \frac{v\sin\alpha_1}{\ell} \\ \vdots \\ \frac{v}{\ell}\left(\prod_{j=1}^{k-2}\cos\alpha_k\right)(\sin\alpha_{k-1} - \cos\alpha_{k-1}\sin\alpha_k) \\ \vdots \\ -\frac{1}{2R(\alpha)}\left(\sum_{k=1}^{n} A_k\frac{\partial R(\alpha)}{\partial\alpha_k}\right)v^2 + \frac{Q(\alpha)}{\ell^2 R(\alpha)}v\omega + \frac{m_0 a}{R(\alpha)}\omega^2 \\ -\frac{m_0 a v\omega}{J_0 + m_0 a^2} \end{bmatrix}.$$

5.3 n-articulated vehicle system optimal filter

In order to take into account the inaccuracy of the dynamics we add the unknown error $\delta \in \mathbb{R}^2$ (modelled as a normalized Gaussian white noise with zero mean and unitary variance) and the map $B : \mathbb{R}^2 \to \overline{\mathfrak{g}}$ to (5.6) to obtain the system

$$\dot{\overline{g}}(t) = \overline{g}(t)(\lambda(\overline{g}(t), u(t), t)^\wedge + B\delta(t)), \quad \overline{g}(t_0) = \overline{g}_0. \tag{5.18}$$

Since the kinematic reconstruction equation is supposed to be correct, we choose B in such a way that it admits the matrix representation

$$B = \begin{bmatrix} 0_{n+3,1} & 0_{n+3,1} \\ b_v & 0 \\ 0 & b_\omega \end{bmatrix} \tag{5.19}$$

meaning that we add a correction only on the dynamic part. The measurement function is

$$\mathbf{y}(t) = h(\overline{g}(t), t) + D\varepsilon(t) = \begin{bmatrix} \mathrm{y}_1 \\ \mathrm{y}_2 \end{bmatrix} \tag{5.20}$$

where $h : \overline{G} \times \mathbb{R} \to \mathbb{R}^p$ is the nominal output map defined as

$$h(\overline{g}(t), t) = \begin{bmatrix} x(t) \\ y(t) \end{bmatrix} \tag{5.21}$$

that represents a GPS antenna settled at (x, y), $\varepsilon \in \mathbb{R}^2$ is the unknown measurement error (modelled as a normalized Gaussian white noise with zero mean and diagonal and unitary variance) and $D : \mathbb{R}^2 \to \mathbb{R}^2$ is an invertible linear map

$$D := \mathrm{diag}\left\{d_1, d_2\right\}. \tag{5.22}$$

The knowledge of the position of the leading car is sufficient to observe the whole system as can be seen by noting that the co-distribution matrix

$$H = \begin{bmatrix} dh_i \\ dL_f h_i \\ \vdots \\ dL_f^{n-1} h_i \end{bmatrix} \in \mathbb{R}^{pn \times n} \tag{5.23}$$

has full rank, where $L_f h$ is the Lie derivative of h with respect to f (see, e.g., [16] [40]). The tests of the observations for the case $n = 1$ and $n = 2$ were done by the use of Matlab's symbolic calculus.

For the cost functional to minimize (2.29)-(2.31), we consider $\alpha = 0$, the initial cost

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 \tag{5.24}$$

where $\|\cdot\|_F^2$ stands for the Frobenius norm, and the matrix representations for the two forms $\mathcal{R}$ and $\mathcal{Q}$ as

$$R := \mathrm{diag}\left\{r_v, r_\omega\right\} \tag{5.25}$$

$$Q := \mathrm{diag}\,\{q_v, q_\omega\} \tag{5.26}$$

respectively.

We are now ready to design the second-order optimal filter for the dynamic system (5.18) which minimizes the cost functional (2.29)-(2.31) using the measurement equation (5.20). We consider the Cartan-Shouten (0)-connection form $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$. This choice is justified by the fact that this connection has null torsion (see [22]) and works better than the others as highlighted in [35].

Proposition 5.1. *Consider the dynamic system* (5.18) *with measurement equation* (5.20) *where the output map h and the linear map D are given by* (5.21) *and* (5.22)*, respectively and where the operator B takes the form* (5.19)*. Consider the cost functional* (2.29)-(2.31) *where the initial cost m_0 is given by* (5.24) *and the matrix representation of the forms $\mathcal{R}$, $\mathcal{Q}$ are given by* (5.25) *and* (5.26)*, respectively. Then the second-order optimal filter is*

$$\widehat{\overline{g}}^{-1}\dot{\widehat{\overline{g}}} = \Big(\lambda_t(\widehat{\overline{g}}, u) + K(t) r_t(\widehat{\overline{g}})\Big)^\wedge, \quad \widehat{\overline{g}}(t_0) = \widehat{\overline{g}}_0$$

where the residual r is

$$r_t(\widehat{\overline{g}}) = \begin{bmatrix} \widetilde{\mathrm{y}}_1 \cos\widehat{\theta} + \widetilde{\mathrm{y}}_2 \sin\widehat{\theta} \\ -\widetilde{\mathrm{y}}_1 \sin\widehat{\theta} + \widetilde{\mathrm{y}}_2 \cos\widehat{\theta} \\ 0_{(n+3)\times 1} \end{bmatrix}^T . \tag{5.27}$$

The optimal gain $K : (\mathbb{R}^{n+5})^* \to \mathbb{R}^{n+5}$ is the solution of the perturbed matrix Riccati differential equation

$$\dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T - W(K,r)K - KW(K,r)^T$$

where the matrix forms of the operators $A(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}$, $E(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ and $W(K,t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}$ are given by

$$A = \begin{bmatrix} 0 & \widehat{\omega} & 0 & 0 & \cdots & 0 & 1 & 0 \\ -\widehat{\omega} & 0 & +\widehat{v} & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 \\ 0 & 0 & 0 & \frac{\partial \dot{\alpha}_1}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\alpha}_1}{\partial \alpha_n} & \frac{\partial \dot{\alpha}_1}{\partial v} & \frac{\partial \dot{\alpha}_1}{\partial \omega} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \frac{\partial \dot{\alpha}_n}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\alpha}_n}{\partial \alpha_n} & \frac{\partial \dot{\alpha}_n}{\partial v} & \frac{\partial \dot{\alpha}_n}{\partial \omega} \\ 0 & 0 & 0 & \frac{\partial \dot{v}}{\partial \alpha_1} & \cdots & \frac{\partial \dot{v}}{\partial \alpha_n} & \frac{\partial \dot{v}}{\partial v} & \frac{\partial \dot{v}}{\partial \omega} \\ 0 & 0 & 0 & \frac{\partial \dot{\omega}}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\omega}}{\partial \alpha_n} & \frac{\partial \dot{\omega}}{\partial v} & \frac{\partial \dot{\omega}}{\partial \omega} \end{bmatrix}, \tag{5.28}$$

$$E = \begin{bmatrix} E_{11} & E_{12} & E_{13} & 0_{1\times(n+2)} \\ E_{21} & E_{22} & E_{23} & 0_{1\times(n+2)} \\ E_{31} & E_{32} & E_{33} & 0_{1\times(n+2)} \\ 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times(n+2)} \end{bmatrix} \tag{5.29}$$

with

$$E_{11} = \frac{q_1}{d_1^2}\cos^2\widehat{\theta} + \frac{q_2}{d_2^2}\sin^2\widehat{\theta}, \tag{5.30}$$

$$E_{12} = -\frac{q_1}{d_1^2}\cos\widehat{\theta}\sin\widehat{\theta} + \frac{q_2}{d_2^2}\cos\widehat{\theta}\sin\widehat{\theta}, \tag{5.31}$$

$$E_{13} = \frac{1}{2}\widetilde{\mathrm{y}}_2\cos\widehat{\theta} - \frac{1}{2}\widetilde{\mathrm{y}}_1\sin\widehat{\theta}, \tag{5.32}$$

$$E_{21} = -\frac{q_1}{d_1^2}\cos\widehat{\theta}\sin\widehat{\theta} + \frac{q_2}{d_2^2}\cos\widehat{\theta}\sin\widehat{\theta}, \tag{5.33}$$

$$E_{22} = \frac{q_1}{d_1^2}\sin^2\widehat{\theta} + \frac{q_2}{d_2^2}\cos^2\widehat{\theta}, \tag{5.34}$$

$$E_{23} = -\frac{1}{2}\widetilde{\mathrm{y}}_2\sin\widehat{\theta} - \frac{1}{2}\widetilde{\mathrm{y}}_1\cos\widehat{\theta}, \tag{5.35}$$

$$E_{31} = -\frac{1}{2}\widetilde{\mathrm{y}}_2\cos\widehat{\theta} + \frac{1}{2}\widetilde{\mathrm{y}}_1\sin\widehat{\theta}, \tag{5.36}$$

$$E_{32} = \frac{1}{2}\widetilde{\mathrm{y}}_2\sin\widehat{\theta} + \frac{1}{2}\widetilde{\mathrm{y}}_1\cos\widehat{\theta}, \tag{5.37}$$

$$E_{33} = 0, \tag{5.38}$$

and

$$W(K, r) = \frac{1}{2}\mathrm{ad}_{(Kr)^\wedge}. \tag{5.39}$$

The initial conditions for the estimator and the gain are

$$\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{5.40}$$

$$K(t_0) = \mathrm{diag}\{1, 1, 1/2, 1, ..., 1, 1/2, 1\}. \tag{5.41}$$

Proof. In what follows we will use $\eta^{\overline{g}} = (\eta^x, \eta^y, \eta^\theta, \eta^{\alpha_1}, ..., \eta^{\alpha_n}, \eta^v, \eta^\omega)^T \in \mathbb{R}^{n+5}$ to indicate the vector form of an element of the Lie algebra $\overline{\mathfrak{g}}$ and with $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) = (x', y', \theta', \alpha_1', ..., \alpha_n', v', \omega')^T$ the vector form of an element of the tangent space $T_{\widehat{\overline{g}}}\overline{G}$.

Computation of r

According to [35], the residual $r \in (\mathbb{R}^5)^*$ is given by

$$r(\widehat{\overline{g}}) = T_eL^*_{\widehat{\overline{g}}}\left[\left((D^{-1})^* \circ Q \circ D^{-1}(\mathbf{y} - h(\widehat{\overline{g}}))\right) \circ \mathrm{d}h(\widehat{\overline{g}})\right].$$

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^{n+5}$, the differential of h in $\widehat{\overline{g}}$ applied to $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\mathrm{d}h\big(\widehat{\overline{g}}\big)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0}\begin{bmatrix}\widehat{x}(s)\\ \widehat{y}(s)\end{bmatrix} = \begin{bmatrix}\widehat{x}'\\ \widehat{y}'\end{bmatrix},$$

and we can write the operator $\mathrm{d}h\big(\widehat{\overline{g}}\big)$ as

$$\mathrm{d}h\big(\widehat{\overline{g}}\big) = \begin{bmatrix}1 & 0 & 0_{1\times(n+3)}\\ 0 & 1 & 0_{1\times(n+3)}\end{bmatrix}. \tag{5.42}$$

Defining the weighted errors

$$\widetilde{\mathbf{y}} = (D^{-1})^* \circ Q \circ D^{-1}(\mathbf{y} - h\big(\widehat{\overline{g}}\big)) = \begin{bmatrix}\widetilde{\mathrm{y}}_1 & \widetilde{\mathrm{y}}_2\end{bmatrix}^T \in \mathbb{R}^2 \tag{5.43}$$

and using (5.42) we obtain

$$\left((D^{-1})^* \circ Q \circ D^{-1}(\mathbf{y} - h(\widehat{\overline{g}}))\right) \circ \mathrm{d}h(\widehat{\overline{g}}) = \begin{bmatrix}\widetilde{\mathrm{y}}_1 & \widetilde{\mathrm{y}}_2 & 0_{1\times(n+3)}\end{bmatrix}. \tag{5.44}$$

Evaluating $T_eL^*_{\widehat{\overline{g}}}$ on (5.44) we get (5.27).

Computation of A

The formula for the operator $A : \mathbb{R}^{n+5} \to \mathbb{R}^{n+5}$ is

$$A = \mathrm{d}_1\lambda(\widehat{\overline{g}}, u) \circ T_e L_{\widehat{\overline{g}}} - \mathrm{ad}_{\lambda(\widehat{\overline{g}},u)} - T_{\lambda(\widehat{\overline{g}},u)}.$$

Given $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^{n+5}$, the differential of λ is

$$\begin{aligned}
&\mathrm{d}_1\lambda(\widehat{\overline{g}}, u)(T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \lambda(s) = \\
&= \begin{bmatrix}
\widehat{v}' \\
0 \\
\widehat{\omega}' \\
\widehat{\omega}' - \frac{\widehat{v}' \sin\widehat{\alpha}'_1}{\ell} \\
\vdots \\
\frac{\widehat{v}'}{\ell}\left(\prod_{j=1}^{k-2}\cos\widehat{\alpha}'_k\right)(\sin\widehat{\alpha}'_{k-1} - \cos\widehat{\alpha}'_{k-1}\sin\widehat{\alpha}'_k) \\
\vdots \\
-\frac{1}{2R(\widehat{\alpha}')}\left(\sum_{k=1}^{n}\widehat{A}'_k\frac{\partial R(\widehat{\alpha}')}{\partial\widehat{\alpha}'_k}\right)\widehat{v}'^2 + \frac{Q(\widehat{\alpha}')}{\ell^2 R(\widehat{\alpha}')}\widehat{v}'\widehat{\omega}' + \frac{m_0 a}{R(\widehat{\alpha}')}\widehat{\omega}'^2 \\
-\frac{m_0 a\widehat{v}'\widehat{\omega}'}{J_0 + m_0 a^2}
\end{bmatrix}.
\end{aligned} \tag{5.45}$$

The adjoint matrix representation (5.16) implies

$$\mathrm{ad}_{\lambda(\widehat{\overline{g}},u)} = \begin{bmatrix}
0 & -\widehat{\omega} & 0 & 0_{1\times(n+2)} \\
\widehat{\omega} & 0 & -\widehat{v} & 0_{1\times(n+2)} \\
0 & 0 & 0 & 0_{1\times(n+2)} \\
0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times(n+2)}
\end{bmatrix}. \tag{5.46}$$

Since we consider the Cartan-Schouten (0)-connection form $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$, the torsion function $T_{\lambda(\widehat{\overline{g}},u)}$ vanishes (see [22]), thus, in matrix form, it is given by

$$T_{\lambda(\widehat{\overline{g}},u)} = \left[0_{(n+5)\times(n+5)}\right]. \tag{5.47}$$

Using the differential form of (5.45), the adjoint matrix (5.46) and the torsion (5.47), we obtain (5.28).

Computation of E

The operator $E : \mathbb{R}^{n+5} \to (\mathbb{R}^{n+5})^*$ takes the form

$$\begin{aligned}
E = -T_e L^*_{\widehat{\overline{g}}} \circ \big[&\big((D^{-1})^* \circ Q \circ D^{-1}(\mathbf{y} - h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \mathrm{Hess}\, h(\widehat{\overline{g}}) \\
&-(\mathrm{d}h(\widehat{\overline{g}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h(\widehat{\overline{g}})\big] \circ T_e L_{\widehat{\overline{g}}}.
\end{aligned}$$

From (5.42) and the definitions of the matrices D and Q, we can find the composition

$$\begin{aligned}
&(\mathrm{d}h(\widehat{\overline{g}}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h(\widehat{\overline{g}}) = \\
&= \mathrm{diag}\left(\frac{q_1}{d_1^2}\cos^2\widehat{\theta} + \frac{q_2}{d_2^2}\sin^2\widehat{\theta}, \frac{q_1}{d_1^2}\sin^2\widehat{\theta} + \frac{q_2}{d_2^2}\cos^2\widehat{\theta}, 0, ..., 0\right).
\end{aligned} \tag{5.48}$$

Let $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{\theta}'_1, \widehat{x}'_1, \widehat{y}'_1, \widehat{\omega}'_1, \widehat{v}'_1)^T$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{\theta}'_2, \widehat{x}'_2, \widehat{y}'_2, \widehat{\omega}'_2, \widehat{v}'_2)^T \in T_{\widehat{\overline{g}}}\overline{G}$ be two vector fields, then the Hessian matrix is defined by

$$\begin{aligned}\operatorname{Hess} h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})) =& \mathrm{d}(\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) \\ &- \mathrm{d}h(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))),\end{aligned}$$

and from the choice of Cartan-Schouten (0)-connection, we get

$$\nabla_{T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}}(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \frac{1}{2}T_eL_{\widehat{\overline{g}}}(\mathrm{ad}_{\eta^{\overline{g}_1}}\eta^{\overline{g}_2}).$$

The Hessian evaluated in $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}$ and $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}$ is therefore

$$\operatorname{Hess} h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \begin{bmatrix} -\frac{1}{2}\widehat{\theta}'_2\widehat{y}'_1 + \frac{1}{2}\widehat{\theta}'_1\widehat{y}'_2 \\ +\frac{1}{2}\widehat{\theta}'_2\widehat{x}'_1 - \frac{1}{2}\widehat{\theta}'_1\widehat{x}'_2 \\ 0_{(n+3)\times 1} \end{bmatrix}. \tag{5.49}$$

From (5.43) and (5.49) it follows that

$$\begin{aligned}&\left((D^{-1})^* \circ Q \circ D^{-1}(\mathbf{y} - h(\widehat{\overline{g}}))\right)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \operatorname{Hess} h(\widehat{\overline{g}}) = \\ &\begin{bmatrix} 0 & 0 & -\widetilde{\mathrm{y}}_2 & 0_{1\times(n+2)} \\ 0 & 0 & \widetilde{\mathrm{y}}_1 & 0_{1\times(n+2)} \\ \widetilde{\mathrm{y}}_2 & -\widetilde{\mathrm{y}}_1 & 0 & 0_{1\times(n+2)} \\ 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times 1} & 0_{(n+2)\times(n+2)} \end{bmatrix}.\end{aligned} \tag{5.50}$$

In conclusion, combining (5.48) and (5.50) with $T_eL^*_{\widehat{\overline{g}}}$ and $T_eL_{\widehat{\overline{g}}}$, the matrix (5.29) is obtained.

Computation of W

From the adjoint matrix form (5.16) and recalling that we consider the Cartan-Schouten (0)-connection (see, e.g., [33],[34]), we have

$$W(K, r) = \frac{1}{2}\mathrm{ad}_{(Kr)^\wedge}.$$

Initial condition

The initial condition for the filter is given by (2.33) while the initial condition for the gain is $K(t_0) = X_0^{-1}$ where the operators $X_0 : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ satisfies (2.36).

We rewrite m_0 as

$$\begin{aligned}m_0(\overline{g}) &= \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 \\ &= \frac{1}{2}\mathrm{tr}\left[(I_{(n+5)\times(n+5)} - \overline{g}^{-1}\overline{g}_0)^T(I_{(n+5)\times(n+5)} - \overline{g}^{-1}\overline{g}_0)\right].\end{aligned} \tag{5.51}$$

From (5.15) it follows

$$
\begin{aligned}
&I_{(n+5)\times(n+5)} - \overline{g}^{-1}\overline{g}_0 \\
&\quad = \begin{bmatrix}
I_{3\times 3} - g^{-1}g_0 & & & & & & \\
& I_{2\times 2} - A_1^{-1}A_{1(0)} & & & & & \\
& & \ddots & & & & \\
& & & I_{2\times 2} - A_n^{-1}A_{n(0)} & & & \\
& & & & I_{3\times 3} - V^{-1}V_0
\end{bmatrix}
\end{aligned} \tag{5.52}
$$

and thus

$$
(I_{(n+5)\times(n+5)} - \overline{g}^{-1}\overline{g}_0)^T(I_{(n+5)\times(n+5)} - \overline{g}^{-1}\overline{g}_0) = \begin{bmatrix}
M_g & & & & \\
& M_{A_1} & & & \\
& & \ddots & & \\
& & & M_{A_n} & \\
& & & & M_V
\end{bmatrix} \tag{5.53}
$$

where

$$
\begin{aligned}
M_g &= (I_{3\times 3} - g^{-1}g_0)^T(I_{3\times 3} - g^{-1}g_0), \\
M_{A_k} &= (I_{2\times 2} - A_n^{-1}A_{n(0)})^T(I_{2\times 2} - A_n^{-1}A_{n(0)}), \quad k = 1, ..., n, \\
M_V &= (I_{3\times 3} - V^{-1}V_0)^T(I_{3\times 3} - V^{-1}V_0).
\end{aligned}
$$

Computing the trace we obtain

$$
\begin{aligned}
m_0(\overline{g}) =& \frac{1}{2}[4(1 - \cos(\theta - \theta_0)) + (x - x_0)^2 + (y - y_0)^2 \\
&+ \sum_{k=1}^{n}[2 - 2\cos(\alpha_k - \alpha_{k(0)})] + 2(\omega - \omega_0)^2 + (v - v_0)^2]
\end{aligned} \tag{5.54}
$$

and from (2.33) it follows that $\widehat{\overline{g}}(t_0) = \overline{g}_0$.

The Hessian of the function m_0 at a point $\overline{g} \in \overline{G}$ is defined as

$$
\operatorname{Hess} m_0(\overline{g})(\overline{g}X)(\overline{g}Y) = \mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) - \mathrm{d}m_0(\overline{g})(\nabla_{\overline{g}X}(\overline{g}Y))
$$

for all $\overline{g}X$, $\overline{g}Y \in T_{\overline{g}}\overline{G}$. The differential of m_0 is given by

$$
\mathrm{d}m_0(\overline{g}) = \begin{bmatrix}
(x - x_0) \\
(y - y_0) \\
2\sin(\theta - \theta_0) \\
\vdots \\
\sin(\alpha_k - \alpha_{k(0)}) \\
\vdots \\
2(\omega - \omega_0) \\
(v - v_0)
\end{bmatrix}^T \tag{5.55}
$$

while, given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = \widehat{\overline{g}}\eta^{\overline{g}_1}$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = \widehat{\overline{g}}\eta^{\overline{g}_2} \in T_{\overline{g}}\overline{G}$, the affine connection yields

$$\nabla_{\widehat{\overline{g}}\eta^{\overline{g}_1}}\widehat{\overline{g}}\eta^{\overline{g}_2} = \frac{1}{2}\begin{bmatrix} \widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2' \\ -\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2' \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{5.56}$$

Combining (5.55) and (5.56) we obtain

$$\mathrm{d}m_0(\widehat{\overline{g}})(\nabla_{\widehat{\overline{g}}\eta^{\overline{g}_1}}\widehat{\overline{g}}\eta^{\overline{g}_2}) = \Big[(x-x_0)(\widehat{\theta}_2'\widehat{y}_1' - \widehat{\theta}_1'\widehat{y}_2') + (y-y_0)(-\widehat{\theta}_2'\widehat{x}_1' + \widehat{\theta}_1'\widehat{x}_2')\Big] \tag{5.57}$$

that evaluating in $\widehat{\overline{g}}_0$ produces

$$\mathrm{d}m_0(\widehat{\overline{g}}_0)(\nabla_{\widehat{\overline{g}}\eta^{\overline{g}_1}}\widehat{\overline{g}}\eta^{\overline{g}_2}) = 0. \tag{5.58}$$

The double differential takes the form

$$\begin{aligned} &\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) \\ &\quad = \mathrm{diag}\{1, 1, 2\cos(\theta-\theta_0), \cos(\alpha_1 - \alpha_{1(0)}), ..., \cos(\alpha_n - \alpha_{n(0)}), 2, 1\} \end{aligned} \tag{5.59}$$

that, evaluating in $\widehat{\overline{g}}_0$, produces

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{1, 1, 2, 1, ..., 1, 2, 1\} \tag{5.60}$$

and thus, from (5.58) and (5.60)

$$\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) = \mathrm{diag}\{1, 1, 2, 1, ..., 1, 2, 1\}. \tag{5.61}$$

From

$$K(t_0) = X_0^{-1} = (T_eL^*_{\widehat{\overline{g}}_0} \circ \mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) \circ T_eL_{\widehat{\overline{g}}_0})^{-1} \tag{5.62}$$

we obtain the initial condition of $K(t_0)$

$$\begin{aligned} K(t_0) &= (T_eL_{\widehat{\overline{g}}_0})^{-1}(\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0))^{-1}(T_eL^*_{\widehat{\overline{g}}_0})^{-1} \\ &= \mathrm{diag}\{1, 1, 1/2, 1, ..., 1, 1/2, 1\}. \end{aligned} \tag{5.63}$$

□

5.4 The system with uncertain masses and inertias

In this section, we aim to provide the second-order optimal filter for the convoy system in case of unknown but time-invariant masses and inertias. To do this,

we treat masses and inertias as state variables that are constant along the motion. The new dynamic equations become:

$$\begin{cases} \dot{x} & = v\cos\theta \\ \dot{y} & = v\sin\theta \\ \dot{\theta} & = \omega \\ \dot{\alpha}_1 & = \omega - \frac{v\sin\alpha_1}{\ell} \\ \dot{\alpha}_k & = \frac{v}{\ell}\left(\prod_{j=1}^{k-2}\cos\alpha_k\right)(\sin\alpha_{k-1} - \cos\alpha_{k-1}\sin\alpha_k) \quad k = 2, ..., n \\ \dot{v} & = -\frac{1}{2R(\alpha)}\left(\sum_{k=1}^{n} A_k \frac{\partial R(\alpha)}{\partial \alpha_k}\right) v^2 + \frac{Q(\alpha)}{\ell^2 R(\alpha)} v\omega + \frac{m_0 a}{R(\alpha)}\omega^2 + \frac{F}{R(\alpha)} \\ \dot{\omega} & = -\frac{m_0 a v\omega}{J_0 + m_0 a^2} + \frac{\tau}{J_0 + m_0 a^2} \\ \dot{m}_k & = 0 \quad k = 0, ..., n \\ \dot{J}_k & = 0 \quad k = 0, ..., n \end{cases} \tag{5.64}$$

where the coefficients A_k, $Q(\alpha)$ and $R(\alpha)$ are given as before by (5.7), (5.8) and (5.9), respectively. The geometric structure that underlies the dynamics (5.64) is featured by the Lie group $\overline{G} = \mathrm{SE}(2) \times \mathbb{T}^n \times \mathbb{R}^2 \times \mathbb{R}^{2(n+1)}$. As done in Section 5.2, we can describe the group operation, the Lie algebra, the adjoint operator and the tangent map accordingly. Evaluating the rank of the matrix (5.23) on this system, we can check the observability. The dynamics with model error (modelled as a normalized Gaussian) is given by

$$\dot{\overline{g}}(t) = \overline{g}(t)(\lambda(\overline{g}(t), u(t), t)^\wedge + B\delta(t)), \quad \overline{g}(t_0) = \overline{g}_0. \tag{5.65}$$

with $B : \mathbb{R}^{2+2(n+1)} \to \overline{\mathfrak{g}}$. The matrix representation of R and B related to the new filter are

$$R = \mathrm{diag}\left(r_v, r_\omega, r_{m_0}, ..., r_{m_n}, r_{J_0}, ..., r_{J_n}\right), r_i \in \mathbb{R}^+, \tag{5.66}$$

$$B = \begin{bmatrix} 0_{n+3,2n+4)} \\ B_2 \end{bmatrix}, \tag{5.67}$$

$$B_2 = \mathrm{diag}\left(b_v, b_\omega, b_{m_0}, ..., b_{m_n}, b_{J_0}, ..., b_{J_n}\right), b_i \in \mathbb{R}^+, \tag{5.68}$$

since the new model error has now $2 + 2(n+1)$ components. The parameters $b_{m_0}, ..., b_{m_n}, b_{J_0}, ..., b_{J_n}$ are needed to estimate masses and inertias and their values are a trade-off between promptness and accuracy.

Proposition 5.2. *Consider the dynamic system* (5.65) *with measurement equation* (5.20) *where the output map* h *and the linear map* D *are given by* (5.21) *and* (5.22)*, respectively and where the operator* B *takes the form* (5.67)*. Consider the cost functional* (2.29)-(2.31) *where the initial cost* m_0 *is given by* (5.24) *and the matrix representation of the forms* $\mathcal{R}$*,* $\mathcal{Q}$ *are given by* (5.66) *and* (5.26)*, respectively. Then the second-order optimal filter is*

$$\widehat{\overline{g}}^{-1}\dot{\widehat{\overline{g}}} = (\lambda_t(\widehat{\overline{g}}, u) + K(t)r_t(\widehat{\overline{g}}))^\wedge, \quad \widehat{\overline{g}}(t_0) = \widehat{\overline{g}}_0$$

where the residual r *is*

$$r_t(\widehat{\overline{g}}) = \begin{bmatrix} \widetilde{\mathrm{y}}_1\cos\widehat{\theta} + \widetilde{\mathrm{y}}_2\sin\widehat{\theta} \\ -\widetilde{\mathrm{y}}_1\sin\widehat{\theta} + \widetilde{\mathrm{y}}_2\cos\widehat{\theta} \\ 0_{(3n+5)\times 1} \end{bmatrix}^T.$$

The optimal gain $K : (\mathbb{R}^{3n+7})^* \to \mathbb{R}^{3n+7}$ *is the solution of the perturbed matrix Riccati differential equation*

$$\dot{K} = -\alpha K + AK + KA^T - KEK + BR^{-1}B^T - W(K,r)K - KW(K,r)^T$$

where the matrix forms of the operators $A(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}$, $E(t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ *and* $W(K,t) : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}$ *are given by*

$$A = \begin{bmatrix} 0 & \widehat{\omega} & 0 & 0 & \cdots & 0 & 1 & 0 & 0 & \cdots & 0 \\ -\widehat{\omega} & 0 & \widehat{v} & 0 & \cdots & 0 & 0 & 0 & 0 & \cdots & 0 \\ 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 0 & \frac{\partial \dot{\alpha}_1}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\alpha}_1}{\partial \alpha_n} & \frac{\partial \dot{\alpha}_1}{\partial v} & \frac{\partial \dot{\alpha}_1}{\partial \omega} & \frac{\partial \dot{\alpha}_1}{\partial m_0} & \cdots & \frac{\partial \dot{\alpha}_1}{\partial J_n} \\ \vdots & \vdots & \vdots & \vdots & & \vdots & \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & 0 & \frac{\partial \dot{\alpha}_n}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\alpha}_n}{\partial \alpha_n} & \frac{\partial \dot{\alpha}_n}{\partial v} & \frac{\partial \dot{\alpha}_n}{\partial \omega} & \frac{\partial \dot{\alpha}_n}{\partial m_0} & \cdots & \frac{\partial \dot{\alpha}_n}{\partial J_n} \\ 0 & 0 & 0 & \frac{\partial \dot{v}}{\partial \alpha_1} & \cdots & \frac{\partial \dot{v}}{\partial \alpha_n} & \frac{\partial \dot{v}}{\partial v} & \frac{\partial \dot{v}}{\partial \omega} & \frac{\partial \dot{v}}{\partial m_0} & \cdots & \frac{\partial \dot{v}}{\partial J_n} \\ 0 & 0 & 0 & \frac{\partial \dot{\omega}}{\partial \alpha_1} & \cdots & \frac{\partial \dot{\omega}}{\partial \alpha_n} & \frac{\partial \dot{\omega}}{\partial v} & \frac{\partial \dot{\omega}}{\partial \omega} & \frac{\partial \dot{\omega}}{\partial m_0} & \cdots & \frac{\partial \dot{\omega}}{\partial J_n} \end{bmatrix},$$

$$E = \begin{bmatrix} E_{11} & E_{12} & E_{13} & 0_{1\times(3n+4)} \\ E_{21} & E_{22} & E_{23} & 0_{1\times(3n+4)} \\ E_{31} & E_{32} & E_{33} & 0_{1\times(3n+4)} \\ 0_{(3n+4)\times 1} & 0_{(3n+4)\times 1} & 0_{(3n+4)\times 1} & 0_{(3n+4)\times(3n+4)} \end{bmatrix}$$

with

$$\begin{aligned} E_{11} &= \frac{q_1}{d_1^2}\cos^2\widehat{\theta} + \frac{q_2}{d_2^2}\sin^2\widehat{\theta}, \\ E_{12} &= -\frac{q_1}{d_1^2}\cos\widehat{\theta}\sin\widehat{\theta} + \frac{q_2}{d_2^2}\cos\widehat{\theta}\sin\widehat{\theta}, \\ E_{13} &= \frac{1}{2}\widetilde{\mathrm{y}}_2\cos\widehat{\theta} - \frac{1}{2}\widetilde{\mathrm{y}}_1\sin\widehat{\theta}, \\ E_{21} &= -\frac{q_1}{d_1^2}\cos\widehat{\theta}\sin\widehat{\theta} + \frac{q_2}{d_2^2}\cos\widehat{\theta}\sin\widehat{\theta}, \\ E_{22} &= \frac{q_1}{d_1^2}\sin^2\widehat{\theta} + \frac{q_2}{d_2^2}\cos^2\widehat{\theta}, \\ E_{23} &= -\frac{1}{2}\widetilde{\mathrm{y}}_2\sin\widehat{\theta} - \frac{1}{2}\widetilde{\mathrm{y}}_1\cos\widehat{\theta}, \\ E_{31} &= -\frac{1}{2}\widetilde{\mathrm{y}}_2\cos\widehat{\theta} + \frac{1}{2}\widetilde{\mathrm{y}}_1\sin\widehat{\theta}, \\ E_{32} &= \frac{1}{2}\widetilde{\mathrm{y}}_2\sin\widehat{\theta} + \frac{1}{2}\widetilde{\mathrm{y}}_1\cos\widehat{\theta}, \\ E_{33} &= 0, \end{aligned}$$

and

$$W(K,r) = \frac{1}{2}\mathrm{ad}_{(Kr)^\wedge}.$$

Proof. The proof is similar to the one for Proposition 5.1 and is not reported for the sake of space. □

5.5 Simulations and discussions

In this section, we show how the second-order optimal filter applied to the convoy works. We will consider the cases $n = 1$, $n = 2$ and compare three different filters:

- $\mathcal{F}_1$, where we consider perfectly known masses and inertias at every instant of time;
- $\mathcal{F}_2$, where the masses and inertias change but they are known only at the initial configuration;
- $\mathcal{F}_3$, where we consider masses and inertias as state variables and we filter them (Section 5.4).

We set the initial conditions at $x(0) = y(0) = 0\,\text{m}$, $\theta(0) = \alpha_1(0) = \alpha_2(0) = 0$ rad, $v(0) = 0\,\text{m/s}$ and $\omega(0) = 0\,\text{rad/s}$. We consider the control inputs

$$F(t) = 1.5(t/(1+t)) + 1.5\sin(t/5),$$
$$\tau(t) = 1.5(t/(1+t)) + 1.5\sin(t/9).$$

The total simulation time is $T = 200\,\text{s}$, while we put $\ell = 0.8\,\text{m}$ and $a = 0.2\,\text{m}$.

We start with known values for masses and inertias, and at time 100 s they change their values, according to Table 5.1. The variations simulate the unknown change of weights due to the addition of new luggage or new packages to a convoy in an airport or warehouse. For the matrix D and Q we choose $d_x = d_y = 0.5$ and $q_x = q_y = 0.25$ respectively, while for the matrix B we choose $b_{m_0} = b_{m_1} = b_{m_2} = b_{J_0} = b_{J_1} = b_{J_2} = 50$, $b_v = b_\omega = 0.1$.

Table 5.1: Values of masses and inertias.

t [s]	$0 - 100$	$100 - 200$
m_0 [kg]	100	130
m_1 [kg]	12	16
m_2 [kg]	10	12
J_0 [kg m^2]	7	9
J_1 [kg m^2]	4	5
J_2 [kg m^2]	3	4

To solve the differential equations of the dynamic systems (5.6) and (5.64) we used a 4-th order Runge-Kutta method, while to solve the equations of the filters we used a forward Euler method with step $T_s = 10$ms. The 2D trajectories of the leading car on the plane are shown in Figure 5.2a ($n = 1$) and Figure 5.2b ($n = 2$), while the error time-series of the state variables for the filters $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ are shown in Figure 5.3 ($n = 1$) and Figure 5.4 ($n = 2$). We can observe that in the first 100 s of the simulations all filters work well, with slightly better results for the filters that know the exact values of the parameters. The drastic change of masses and inertias values produces a variation of the trajectories, and while the filter $\mathcal{F}_2$ (with constant

masses and inertia) fails to align with correct values, the filter $\mathcal{F}_3$ (that filters also the parameters), after a settling period, produces excellent results and comparable with the filter $\mathcal{F}_1$ (that know the exact values at each instant of time). To better highlight the difference among the filters, we report in Table 5.2 the mean, the standard deviation and the root mean square values of the errors

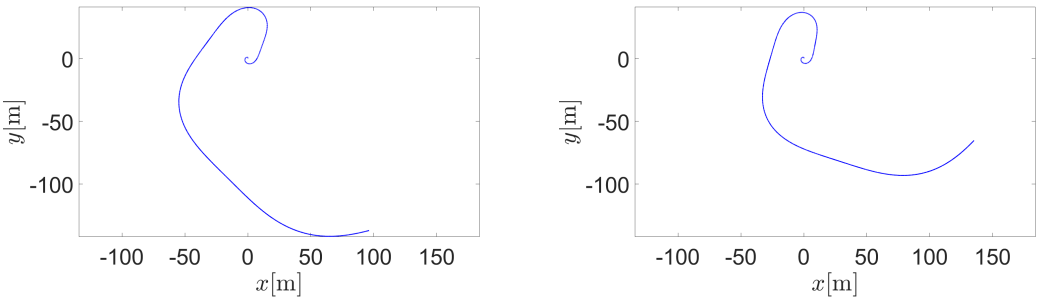
$$\Delta x := x - \widehat{x}, \quad \Delta y := y - \widehat{y}, \quad \cdots$$

in the two time intervals $I_1 = [0, 100)$ and $I_2 = [100, 200]$. These results highlight how the filter $\mathcal{F}_3$ produces better estimates compared with $\mathcal{F}_2$, comparable with $\mathcal{F}_1$.

We can observe that all the state variables can be estimated, even if the measurement equation provides only the position (x, y) of the first leading car. The observability of pose and velocity is possible thanks to the nonholonomic and hooking constraints.

Table 5.2: Mean, standard deviation and root mean square value of the errors for $n = 1$ and $n = 2$.

T		$\mathcal{F}_1$			$\mathcal{F}_2$			$\mathcal{F}_3$		
n		μ	σ	rms	μ	σ	rms	μ	σ	rms
	Δx [mm]	18.0	24.4	30.4	18.0	24.4	30.4	16.5	36.7	40.2
	Δy [mm]	-10.5	24.3	26.5	-10.5	24.3	26.5	-9.6	35.0	36.3
I_1	$\Delta\theta$ [mrad]	-3.5	26.9	27.1	-3.5	26.9	27.1	-5.8	40.6	41.0
$n = 1$	$\Delta\alpha_1$ [mrad]	-3.5	18.4	18.7	-3.5	18.4	18.7	-4.4	27.4	27.7
	$\Delta v \left[\frac{\text{mm}}{\text{s}}\right]$	0.1	4.9	4.9	0.1	4.9	4.9	0.6	11.5	11.5
	$\Delta\omega \left[\frac{\text{mrad}}{\text{s}}\right]$	-0.5	2.4	2.4	-0.5	2.4	2.4	-1.8	7.7	7.9
	Δx [mm]	1.4	28.0	28.1	202.8	205.2	288.5	3.7	35.4	35.6
	Δy [mm]	0.4	28.0	28.0	151.5	230.7	276.0	-1.9	34.2	34.2
I_2	$\Delta\theta$ [mrad]	0.1	1.7	1.7	44.0	30.9	53.8	1.5	5.2	5.4
$n = 1$	$\Delta\alpha_1$ [mrad]	0.0	0.0	0.0	2.3	1.6	2.8	0.2	0.4	0.4
	$\Delta v \left[\frac{\text{mm}}{\text{s}}\right]$	-0.3	5.3	5.3	11.7	11.0	16.1	0.5	10.0	10.0
	$\Delta\omega \left[\frac{\text{mrad}}{\text{s}}\right]$	0.0	0.1	0.1	7.3	4.7	8.7	0.5	1.1	1.2
	Δx [mm]	19.4	24.5	31.3	19.4	24.5	31.3	16.9	37.3	41.0
	Δy [mm]	-11.1	26.2	28.5	-11.1	26.2	28.5	-11.3	36.1	37.8
	$\Delta\theta$ [mrad]	-3.9	30.6	30.9	-3.9	30.6	30.9	-6.7	47.2	47.7
I_1	$\Delta\alpha_1$ [mrad]	-3.1	20.3	20.6	-3.1	20.3	20.6	-4.8	32.0	32.4
$n = 2$	$\Delta\alpha_2$ [mrad]	-13.5	242.7	243.1	-13.5	242.7	243.1	17.2	202.0	202.7
	$\Delta v \left[\frac{\text{mm}}{\text{s}}\right]$	0.7	9.5	9.5	0.7	9.5	9.5	0.6	14.5	14.5
	$\Delta\omega \left[\frac{\text{mrad}}{\text{s}}\right]$	-0.4	2.8	2.8	-0.4	2.8	2.8	-2.0	9.0	9.2
	Δx [mm]	2.6	26.9	27.0	104.3	234.3	256.4	4.5	35.2	35.5
	Δy [mm]	0.5	29.3	29.3	217.7	207.9	301.0	-0.6	34.6	34.6
	$\Delta\theta$ [mrad]	0.1	1.7	1.7	47.1	33.0	57.5	1.8	5.3	5.6
I_2	$\Delta\alpha_1$ [mrad]	-0.0	0.1	0.1	2.6	1.8	3.2	0.2	0.4	0.5
$n = 2$	$\Delta\alpha_2$ [mrad]	-0.0	0.1	0.1	2.6	1.8	3.2	0.2	0.4	0.5
	$\Delta v \left[\frac{\text{mm}}{\text{s}}\right]$	0.5	5.9	6.0	12.1	8.8	15.0	0.6	10.2	10.2
	$\Delta\omega \left[\frac{\text{mrad}}{\text{s}}\right]$	-0.0	0.1	0.1	7.8	5.1	9.3	0.6	1.2	1.3



(a) 2D Trajectory of the leading car for $n = 1$ (b) 2D Trajectory of the leading car for $n = 2$

Figure 5.2

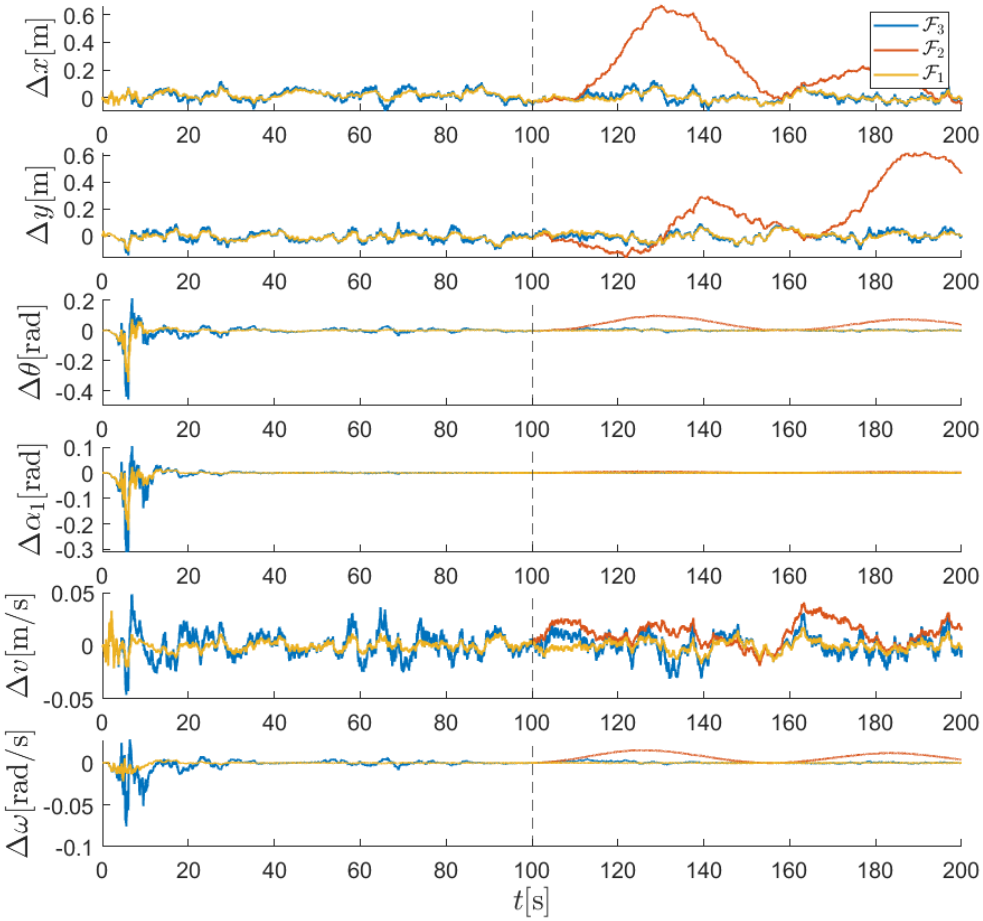


Figure 5.3: Trajectories errors for the filters $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ for $n = 1$.

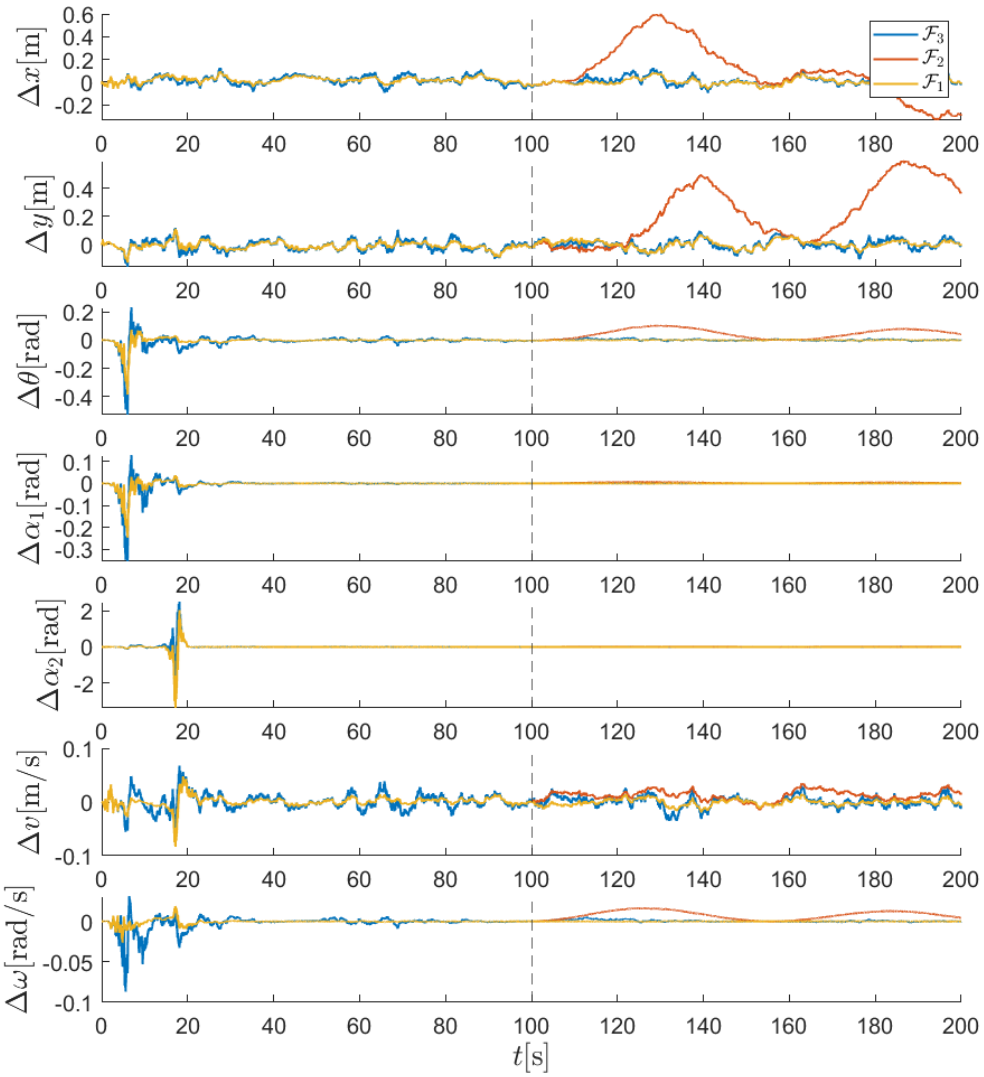


Figure 5.4: Trajectories errors for the filters $\mathcal{F}_1$, $\mathcal{F}_2$ and $\mathcal{F}_3$ for $n = 2$.

Chapter 6

Second-order optimal filter applied to a truck semi-trailer system in a real case model[1]

In this chapter, we apply the second-order optimal filter to a scale model of a truck semi-trailer system engaged in a parking maneuver (Figure 6.1). The measured trajectories are obtained by adding noises to the poses provided by a motion capture optical system in order to reproduce a real-case scenario. We apply the second-order filter simulating different measurement equations that correspond to the cases where GPS and LIDAR sensors are available or not ([12], [39], [41]).

(a) Truck and semi-trailer in a real parking area

(b) Truck and semi-trailer model with a reduction ratio of 1:13.3 in Automotive lab

Figure 6.1

[1]This chapter is based on the work
▷ Rigo D., Saccon A., Sansonetto N., & Muradore R. *State Estimation for a Tractor Semi-trailer System using a Minimum-Energy Filter* (submitted).
This work was done at Eindhoven University of Technology under the supervision of prof. Alessandro Saccon within the Dynamics and Control group and with the assistance of the Automotive Lab people.

6.1 Kinematics of truck semi-trailer systems

The system under study is an articulated vehicle composed of a leading truck and a semi-trailer (Figure 6.2). It moves in a maneuvering area and its goal is to park in a docking station. We consider an inertial frame of reference $\Sigma_I = \{e_x, e_y\}$ attached to this docking station with the origin in the final target, and a right-handed body frame $\Sigma_b = \left\{e_{\|}^b, e_{\perp}^b\right\}$ on the truck, centered in the midpoint of the rear wheels (x_1, y_1).

The configuration space of the truck is SE(2) with coordinates (x_1, y_1, α) with respect to the inertial frame, where α is the angle that the truck forms with the inertial frame. At distance ℓ_1 from (x_1, y_1), the pair (x_0, y_0) represents the midpoint of the front wheels.

Also the configuration space of the trailer is given by SE(2) with coordinates (x_2, y_2, β). The pair (x_2, y_2) represents the position of the midpoint of the rear wheels, and β is the angle the semi-trailer forms with respect to the inertial frame. The semi-trailer is hooked to the truck through an articulation point (x_{1_c}, y_{1_c}) at distance (ℓ_{1_c}) from (x_1, y_1), while ℓ_2 is the distance between (x_2, y_2) and the articulation point.

The truck is modelled as a rigid body with a nonholonomic constraint given by

$$\dot{x}_1 \sin\alpha - \dot{y}_1 \cos\alpha = 0 \tag{6.1}$$

while the nonholonomic constraint for the semi-trailer is given by

$$\dot{x}_2 \sin\beta - \dot{y}_2 \cos\beta = 0. \tag{6.2}$$

They do not allow orthogonal components of the velocities.

The lateral velocity of the front wheels, expressed in the chassis frame, is equal to $V_1 \tan\theta$, where V_1 is the vehicle forward velocity at the rear axle and θ is the front wheels' steering angle. Similarly, the lateral velocity of the attachment point (x_{1_c}, y_{1_c}) is equal to $V_2 \tan\alpha$, where V_2 is the velocity at the rear axle of the trailer.

The equations of motion for the truck semi-trailer system are then given by

$$\begin{cases} \dot{x} &= V_1 \cos\alpha \\ \dot{y} &= V_1 \sin\alpha \\ \dot{\alpha} &= \frac{1}{\ell_1} V_1 \tan\theta \\ \dot{\beta} &= \frac{1}{\ell_2}\left(V_1 \sin\gamma_1 - \dot{\alpha}\ell_{1c}\cos\gamma_1\right) \end{cases} \tag{6.3}$$

where $V_2 = V_1 \cos\gamma_1 - \dot{\alpha}\ell_{1_c}\sin\gamma_1$ and $\gamma_1 = \alpha - \beta$ is the angle between the truck and the trailer.

6.2 The $\mathrm{SE}(2) \times \mathrm{SO}(2)$ structure

In order to apply the second-order filter to the system (6.3) it is necessary to investigate its geometric structure. The state space that underlies the kine-

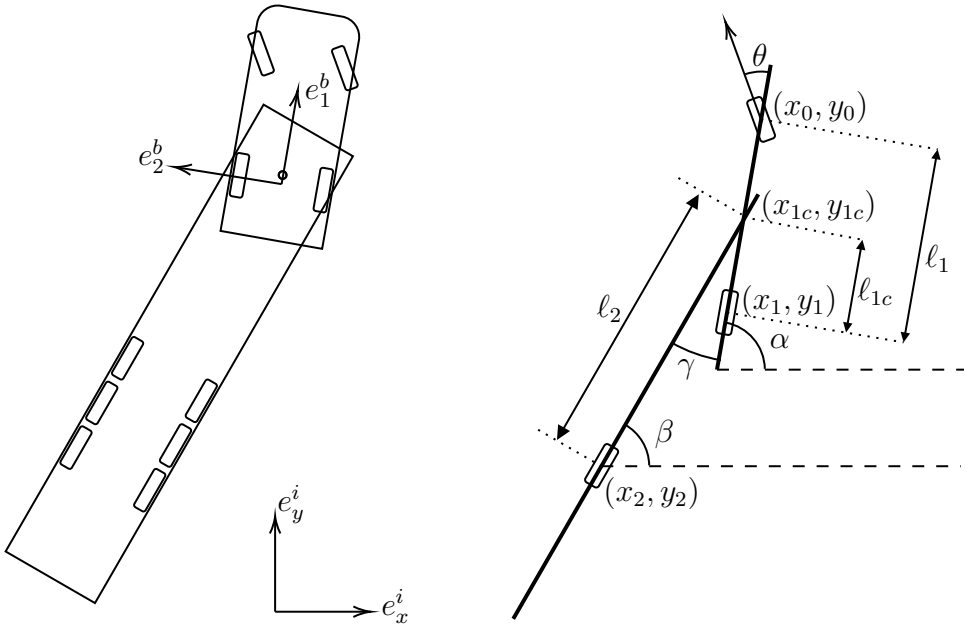


Figure 6.2: Scheme of truck semi-trailer model.

matics (6.3) is the Lie group

$$\overline{G} = \mathrm{SE}(2) \times \mathrm{SO}(2) \tag{6.4}$$

whose generic element $\overline{g} \in \overline{G}$ admits the matrix representation

$$\overline{g} = \begin{bmatrix} \cos\alpha & -\sin\alpha & x & 0 & 0 \\ \sin\alpha & \cos\alpha & y & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & \cos\beta & -\sin\beta \\ 0 & 0 & 0 & \sin\beta & \cos\beta \end{bmatrix}. \tag{6.5}$$

Given the Lie algebra $\overline{\mathfrak{g}} = \mathfrak{se}(2) \times \mathfrak{so}(2)$, we introduce the Lie algebra isomorphism $^\wedge : \mathbb{R}^4 \to \mathfrak{se}(2) \times \mathfrak{so}(2)$

$$\begin{bmatrix} \eta^x \\ \eta^y \\ \eta^\alpha \\ \eta^\beta \end{bmatrix}^\wedge \cong \begin{bmatrix} 0 & -\eta^\alpha & \eta^x & 0 & 0 \\ \eta^\alpha & 0 & \eta^y & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -\eta^\beta \\ 0 & 0 & 0 & \eta^\beta & 0 \end{bmatrix} \tag{6.6}$$

from the Lie algebra $(\mathbb{R}^4, \star)$ to the matrix Lie algebra $(\mathfrak{se}(2), [\cdot,\cdot])$, where $\star : \mathbb{R}^3 \times \mathbb{R}^3 \to \mathbb{R}^3$ is the Lie bracket operation defined as

$$\begin{bmatrix} \eta_1^x \\ \eta_1^y \\ \eta_1^\alpha \\ \eta_1^\beta \end{bmatrix} \star \begin{bmatrix} \eta_2^x \\ \eta_2^y \\ \eta_2^\alpha \\ \eta_2^\beta \end{bmatrix} = \begin{bmatrix} -\eta_1^\alpha \eta_2^y + \eta_2^\alpha \eta_1^y \\ \eta_1^\alpha \eta_2^x - \eta_2^\alpha \eta_1^x \\ 0 \\ 0 \end{bmatrix} \tag{6.7}$$

and $[\cdot,\cdot]$ is the usual matrix commutator (see, e.g., [25]).

The left-trivialization dynamics of (6.3), obtained from $\lambda^\wedge = \overline{g}^{-1}\dot{\overline{g}}$, is given by $\lambda = (\lambda_x, \lambda_y, \lambda_\alpha, \lambda_\beta)$ where

$$\begin{aligned}
\lambda_x &= V_1 \\
\lambda_y &= 0 \\
\lambda_\alpha &= \frac{1}{\ell} V_1 \tan\theta \\
\lambda_\beta &= \frac{1}{\ell} V_1 \sin(\alpha - \beta) + \frac{\ell_{1_c}}{\ell_2} \frac{1}{\ell} V_1 \tan\theta \cos(\alpha - \beta).
\end{aligned} \tag{6.8}$$

The tangent map and the adjoint representation are given by

$$T_e L_{\overline{g}} = \begin{bmatrix} \cos\alpha & -\sin\alpha & 0 & 0 \\ \sin\alpha & \cos\alpha & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{6.9}$$

$$\mathrm{ad}_{(\eta^\Omega)^\wedge} = \begin{bmatrix} 0 & -\eta^\alpha & \eta^y & 0 \\ \eta^\alpha & 0 & -\eta^x & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \tag{6.10}$$

respectively.

The choice of the connection function ω on the Lie algebra is related to the choice of a left-invariant affine connection ∇ on the Lie group. We use the so-called Cartan-Schouten (0)-connection, characterized by $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$, that has the following matrix representation

$$\omega^{(0)} = \frac{1}{2} \begin{bmatrix} 0 & -\eta^\alpha & \eta^y & 0 \\ \eta^\alpha & 0 & -\eta^x & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.11}$$

This decision is justified by the fact that it has null torsion and results in better estimations (see [22]).

6.3 Truck semi-trailer optimal filter

To take into account unmodelled kinematics in (6.3) we consider the unknown error δ (modelled as a normalized Gaussian white noise) and the mapping

$$B : \mathbb{R}^4 \to \overline{\mathfrak{g}}. \tag{6.12}$$

With these definitions, we rewrite the kinematics as

$$\dot{\overline{g}}(t) = \overline{g}(t)\left[\lambda(\overline{g}(t), u(t), t) + B\delta(t)\right], \quad \overline{g}(t_0) = \overline{g}_0. \tag{6.13}$$

For the first part of the maneuver, when the truck is far away from the docking station, we assume to measure the position of two GPS devices settled

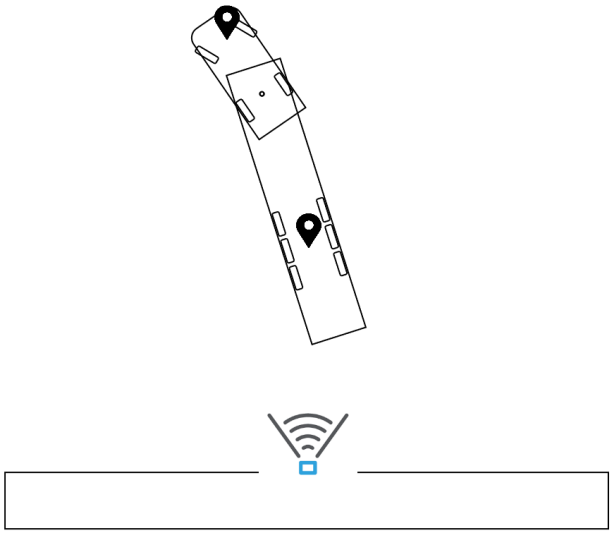


Figure 6.3: LIDAR e GPS devices

in (x_0, y_0) and (x_2, y_2). Since the use of GPS is not sufficient during the maneuver that requires great precision (especially when reversing), we combine them with other sensors. In the second part of the maneuver, when the vehicle is sufficiently close to the final target, we simulate a LIDAR device settled at the docking station, with the purpose of helping the truck during the reversing (see Figure 6.3). This LIDAR provides the pose of the trailer through a laser scan and thus the position of (x_2, y_2) and the angle β. This is justified because, when reversing during the parking maneuver, the LIDAR can spot only the semi-trailer, therefore it can improve only its pose. Summarizing, the two measurement equations

$$\mathrm{y}_i(t) = h_i(\overline{g}(t), t) + D_i \varepsilon_i(t), \quad i = 1, 2, \tag{6.14}$$

have the following output maps:

$$h_1(\overline{g}(t), t) = \begin{bmatrix} x_{0_{\mathrm{GPS}}} \\ y_{0_{\mathrm{GPS}}} \\ x_{2_{\mathrm{GPS}}} \\ y_{2_{\mathrm{GPS}}} \end{bmatrix} = \begin{bmatrix} x + \ell_1 \cos\alpha|_{\mathrm{GPS}} \\ y + \ell_1 \sin\alpha|_{\mathrm{GPS}} \\ x + \ell_{1_c} \cos\alpha - \ell_2 \cos\beta|_{\mathrm{GPS}} \\ y + \ell_{1_c} \sin\alpha - \ell_2 \sin\beta|_{\mathrm{GPS}} \end{bmatrix}, \tag{6.15}$$

$$h_2(\overline{g}(t), t) = \begin{bmatrix} x_{0_{\mathrm{GPS}}} \\ y_{0_{\mathrm{GPS}}} \\ x_{2_{\mathrm{LIDAR}}} \\ y_{2_{\mathrm{LIDAR}}} \\ \beta_{2_{\mathrm{LIDAR}}} \end{bmatrix} = \begin{bmatrix} x + \ell_1 \cos\alpha|_{\mathrm{GPS}} \\ y + \ell_1 \sin\alpha|_{\mathrm{GPS}} \\ x + \ell_{1_c} \cos\alpha - \ell_2 \cos\beta|_{\mathrm{LIDAR}} \\ y + \ell_{1_c} \sin\alpha - \ell_2 \sin\beta|_{\mathrm{LIDAR}} \\ \beta|_{\mathrm{LIDAR}} \end{bmatrix}. \tag{6.16}$$

The second-order optimal filter minimizes the cost functional (2.29)-(2.31) with the initial cost map given by

$$m_0(\overline{g}) = \frac{1}{2} \| I - \overline{g}^{-1}(t) \overline{g}_0 \|_F^2 \tag{6.17}$$

where $\|\cdot\|_F^2$ stands for the Frobenius norm, and the matrix representations of the quadratic forms $\mathcal{R}$ in (2.31) and of the operator B in (6.12) are given by

$$R := \text{diag}\{r_x, r_y, r_\alpha, r_\beta\}, \tag{6.18}$$
$$B := \text{diag}\{b_x, b_y, b_\alpha, b_\beta\}. \tag{6.19}$$

The matrix representations of the quadratic forms $\mathcal{Q}$ in (2.31) and of the operator D in (6.14) in the first part of the maneuver are given by

$$Q := \text{diag}\{q_1, q_1, q_1, q_1\}, \tag{6.20}$$
$$D := \text{diag}\{d_1, d_1, d_1, d_1\}. \tag{6.21}$$

while in the second part are

$$Q := \text{diag}\{q_1, q_1, q_2, q_2, q_3\}, \tag{6.22}$$
$$D := \text{diag}\{d_1, d_1, d_2, d_2, d_3\}. \tag{6.23}$$

For the first filter we define the weighted output error $\widetilde{\mathrm{y}}$ as

$$\widetilde{\mathrm{y}} = \left[\text{diag}\left(\frac{q_1}{d_1^2}, \frac{q_1}{d_1^2}, \frac{q_1}{d_1^2}, \frac{q_1}{d_1^2}\right)\right](\mathrm{y} - \widehat{\mathrm{y}}) \in \mathbb{R}^4 \tag{6.24}$$

while for the second as

$$\widetilde{\mathrm{y}} = \left[\text{diag}\left(\frac{q_1}{d_1^2}, \frac{q_1}{d_1^2}, \frac{q_2}{d_2^2}, \frac{q_2}{d_2^2}, \frac{q_3}{d_3^2}\right)\right](\mathrm{y} - \widehat{\mathrm{y}}) \in \mathbb{R}^5. \tag{6.25}$$

Proposition 6.1. *Consider the dynamic system* (6.13) *with output map* h *and with linear map* D *given by* (6.15) *and* (6.21)*, respectively, and where the operator* B *takes the form* (6.19)*. Consider the cost functional* (2.29)-(2.31) *where the initial cost* m_0 *is given by* (6.17) *and the matrix representation of the forms* $\mathcal{R}$*,* $\mathcal{Q}$ *are given by* (6.18) *and* (6.20)*, respectively. Then the second-order optimal filter is*

$$\widehat{\overline{g}}^{-1}\dot{\widehat{\overline{g}}} = \lambda_t(\widehat{\overline{g}}, u) + K(t) r_t(\widehat{\overline{g}}), \quad \widehat{\overline{g}}(t_0) = \widehat{\overline{g}}_0 \tag{6.26}$$

where the residual r *is*

$$r_t = \begin{bmatrix} (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\alpha} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\alpha} \\ -(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\alpha} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\alpha} \\ -\widetilde{\mathrm{y}}_1\ell_1\sin\widehat{\alpha} + \widetilde{\mathrm{y}}_2\ell_1\cos\widehat{\alpha} - \widetilde{\mathrm{y}}_3\ell_{1_c}\sin\widehat{\alpha} + \widetilde{\mathrm{y}}_4\ell_{1_c}\cos\widehat{\alpha} \\ \widetilde{\mathrm{y}}_3\ell_2\sin\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2\cos\widehat{\beta} \end{bmatrix}^T. \tag{6.27}$$

The optimal gain $K : (\mathbb{R}^4)^* \to \mathbb{R}^4$ *is the solution of the perturbed matrix Riccati differential equation*

$$\begin{aligned} \dot{K} = -cK + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K,r)K - KW(K,r)^T \end{aligned} \tag{6.28}$$

with

$$A(t) = \begin{bmatrix} 0 & a_{12} & 0 & 0 \\ a_{21} & 0 & a_{23} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \tag{6.29}$$

where

$$\begin{aligned} a_{12} &= \frac{1}{\ell} V_1 \tan\theta \\ a_{21} &= -\frac{1}{\ell} V_1 \tan\theta \\ a_{23} &= -V_1 \\ a_{43} &= \frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) - \frac{\ell_{1_C}}{\ell_2} \frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}) \\ a_{44} &= -\frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) + \frac{\ell_{1_C}}{\ell_2} \frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}), \end{aligned}$$

and

$$E(t) = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{14} \\ E_{21} & E_{22} & E_{23} & E_{24} \\ E_{31} & E_{32} & E_{33} & E_{34} \\ E_{41} & E_{42} & E_{43} & E_{44} \end{bmatrix} \tag{6.30}$$

where

$$\begin{aligned} E_{11} &= e_{11}\cos^2\widehat{\alpha} + e_{21}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{12}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{22}\sin^2\widehat{\alpha} \\ E_{12} &= -e_{11}\cos\widehat{\alpha}\sin\widehat{\alpha} - e_{21}\sin^2\widehat{\alpha} + e_{12}\cos^2\widehat{\alpha} + e_{22}\cos\widehat{\alpha}\sin\widehat{\alpha} \\ E_{13} &= e_{13}\cos\widehat{\alpha} + e_{23}\sin\widehat{\alpha} \\ E_{14} &= e_{14}\cos\widehat{\alpha} + e_{24}\sin\widehat{\alpha} \\ E_{21} &= -e_{11}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{21}\cos^2\widehat{\alpha} - e_{12}\sin^2\widehat{\alpha} + e_{22}\cos\widehat{\alpha}\sin\widehat{\alpha} \\ E_{22} &= e_{11}\sin^2\widehat{\alpha} - e_{21}\cos\widehat{\alpha}\sin\widehat{\alpha} - e_{12}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{22}\cos^2\widehat{\alpha} \\ E_{23} &= -e_{13}\sin\widehat{\alpha} + e_{23}\cos\widehat{\alpha} \\ E_{24} &= -e_{14}\sin\widehat{\alpha} + e_{24}\cos\widehat{\alpha} \\ E_{31} &= e_{31}\cos\widehat{\alpha} + e_{32}\sin\widehat{\alpha} \\ E_{32} &= -e_{31}\sin\widehat{\alpha} + e_{32}\cos\widehat{\alpha} \\ E_{33} &= e_{33} \\ E_{34} &= e_{34} \\ E_{41} &= e_{41}\cos\widehat{\alpha} + e_{42}\sin\widehat{\alpha} \\ E_{42} &= -e_{41}\sin\widehat{\alpha} + e_{42}\cos\widehat{\alpha} \\ E_{43} &= e_{34} \\ E_{44} &= e_{44} \\ e_{11} &= 2\frac{q_1}{d_1^2} \\ e_{12} &= 0 \end{aligned}$$

$$
\begin{aligned}
e_{13} &= -\frac{q_1}{d_1^2}\ell_1 \sin\widehat{\alpha} - \frac{q_1}{d_1^2}\ell_{1_c}\sin\widehat{\alpha} - \frac{1}{2}(\widehat{\mathrm{y}}_2 + \widehat{\mathrm{y}}_4)\\
e_{14} &= \frac{q_1}{d_1^2}\ell_2 \sin\widehat{\beta}\\
e_{21} &= 0\\
e_{22} &= 2\frac{q_1}{d_1^2}\\
e_{23} &= \frac{q_1}{d_1^2}\ell_1 \cos\widehat{\alpha} + \frac{q_1}{d_1^2}\ell_{1_c}\cos\widehat{\alpha} + \frac{1}{2}(\widehat{\mathrm{y}}_1 + \widehat{\mathrm{y}}_3)\\
e_{24} &= -\frac{q_1}{d_1^2}\ell_2 \cos\widehat{\beta}\\
e_{31} &= -\frac{q_1}{d_1^2}\ell_1 \sin\widehat{\alpha} - \frac{q_1}{d_1^2}\ell_{1_c}\sin\widehat{\alpha} + \frac{1}{2}(\widehat{\mathrm{y}}_2 + \widehat{\mathrm{y}}_4)\\
e_{32} &= \frac{q_1}{d_1^2}\ell_1 \cos\widehat{\alpha} + \frac{q_1}{d_1^2}\ell_{1_c}\cos\widehat{\alpha} - \frac{1}{2}(\widehat{\mathrm{y}}_1 + \widehat{\mathrm{y}}_3)\\
e_{33} &= \frac{q_1}{d_1^2}\ell_1^2 + \frac{q_1}{d_1^2}\ell_{1_c}^2 + \widetilde{\mathrm{y}}_1\ell_1\,\cos\widehat{\alpha} + \widetilde{\mathrm{y}}_2\ell_1\,\sin\widehat{\alpha} - \widetilde{\mathrm{y}}_3\ell_{1_c}\,\cos\widehat{\alpha} - \widetilde{\mathrm{y}}_4\ell_{1_c}\,\sin\widehat{\alpha}\\
e_{34} &= -\frac{q_1}{d_1^2}\ell_{1_c}\ell_2 \sin\widehat{\alpha}\sin\widehat{\beta} - \frac{q_1}{d_1^2}\ell_{1_c}\ell_2\cos\widehat{\alpha}\cos\widehat{\beta}\\
e_{41} &= \frac{q_1}{d_1^2}\ell_2 \sin\widehat{\beta}\\
e_{42} &= -\frac{q_1}{d_1^2}\ell_2 \cos\widehat{\beta}\\
e_{43} &= -\frac{q_1}{d_1^2}\ell_{1_c}\ell_2 \sin\widehat{\alpha}\sin\widehat{\beta} - \frac{q_1}{d_1^2}\ell_{1_c}\ell_2\cos\widehat{\alpha}\cos\widehat{\beta}\\
e_{44} &= \frac{q_1}{d_1^2}\ell_2^2 - \widetilde{\mathrm{y}}_3\ell_2\cos\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2\sin\widehat{\beta}
\end{aligned}
$$

and with

$$
W(K,r) = \frac{1}{2}\mathrm{ad}_{(Kr)^\wedge}. \tag{6.31}
$$

The initial conditions for the equations (6.26) *and* (6.28) *are*

$$
\widehat{\overline{g}}(t_0) = \overline{g}_0 \tag{6.32}
$$

$$
K(t_0) = \mathrm{diag}\{1, 1, 1/2, 1\}. \tag{6.33}
$$

Proof. In what follows we will use $\eta^{\overline{g}} = (\eta^x, \eta^y, \eta^\alpha, \eta^\beta)^T \in \mathbb{R}^4$ to indicate the vector form of an element of the Lie algebra $\overline{\mathfrak{g}}$ and with $T_e L_{\widehat{\overline{g}}}(\eta^{\overline{g}}) = (x', y', \alpha', \beta')^T$ the vector form of an element of the tangent space $T_{\widehat{\overline{g}}}\overline{G}$.

Computation of r

Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^4$, the differential of h in $\widehat{\overline{g}}$ applied to $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})$ is

$$\begin{aligned} \mathrm{d}h\big(\,\widehat{\overline{g}}\,\big)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) =& \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} \widehat{x}(s) + \ell_1 \cos\widehat{\alpha}(s) \\ \widehat{y}(s) + \ell_1 \sin\widehat{\alpha}(s) \\ \widehat{x}(s) + \ell_{1_c} \cos\widehat{\alpha}(s) - \ell_2 \cos\widehat{\beta}(s) \\ \widehat{y}(s) + \ell_{1_c} \sin\widehat{\alpha}(s) - \ell_2 \sin\widehat{\beta}(s) \end{bmatrix} \\ &= \begin{bmatrix} \widehat{x}' - \ell_1\,\widehat{\alpha}'\sin\widehat{\alpha} \\ \widehat{y}' + \ell_1\,\widehat{\alpha}'\cos\widehat{\alpha} \\ \widehat{x}' - \ell_{1_c}\,\widehat{\alpha}'\sin\widehat{\alpha} + \ell_2\,\widehat{\beta}'\sin\widehat{\beta} \\ \widehat{y}' + \ell_{1_c}\,\widehat{\alpha}'\cos\widehat{\alpha} - \ell_2\,\widehat{\beta}'\cos\widehat{\beta} \end{bmatrix}, \end{aligned} \tag{6.34}$$

and we can write the operator $\mathrm{d}h\big(\,\widehat{\overline{g}}\,\big)$ as

$$\mathrm{d}h\big(\,\widehat{\overline{g}}\,\big) = \begin{bmatrix} 1 & 0 & -\ell_1 \sin\widehat{\alpha} & 0 \\ 0 & 1 & \ell_1 \cos\widehat{\alpha} & 0 \\ 1 & 0 & -\ell_{1_c} \sin\widehat{\alpha} & \ell_2 \sin\widehat{\beta} \\ 0 & 1 & \ell_{1_c} \cos\widehat{\alpha} & -\ell_2 \cos\widehat{\beta} \end{bmatrix}. \tag{6.35}$$

Thus, we obtain

$$\begin{aligned} &\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big) \circ \mathrm{d}h(\widehat{\overline{g}}) = \\ &= \begin{bmatrix} (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) \\ (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) \\ -\widetilde{\mathrm{y}}_1\ell_1 \sin\widehat{\alpha} + \widetilde{\mathrm{y}}_2\ell_1 \cos\widehat{\alpha} - \widetilde{\mathrm{y}}_3\ell_{1_c} \sin\widehat{\alpha} + \widetilde{\mathrm{y}}_4\ell_{1_c} \cos\widehat{\alpha} \\ \widetilde{\mathrm{y}}_3\ell_2 \sin\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2 \cos\widehat{\beta} \end{bmatrix}. \end{aligned} \tag{6.36}$$

Evaluating $T_eL^*_{\widehat{\overline{g}}}$ on (6.36) we finally get (6.27).

Computation of A

The expression for the operator $A : \mathbb{R}^4 \to \mathbb{R}^4$ is (2.37). Given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}}) \in \mathbb{R}^4$, the differential of λ is

$$\begin{aligned} &\mathrm{d}_1\lambda(\widehat{\overline{g}}, u)(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}})) = \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \lambda(s) \\ &= \frac{\mathrm{d}}{\mathrm{d}s}\Big|_{s=0} \begin{bmatrix} V_1 \\ 0 \\ \frac{1}{\ell} V_1 \tan\theta \\ \frac{1}{\ell} V_1 \sin\left(\widehat{\alpha}(s) - \widehat{\beta}(s)\right) + \frac{\ell_{1_c}}{\ell_2}\frac{1}{\ell} V_1 \tan\theta \cos\left(\widehat{\alpha}(s) - \widehat{\beta}(s)\right) \end{bmatrix} \end{aligned} \tag{6.37}$$

and thus

$$\mathrm{d}_1\lambda(\widehat{\overline{g}}, u) \circ T_eL_{\widehat{\overline{g}}} = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \tag{6.38}$$

where

$$\begin{aligned} a_{43} &= \frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) - \frac{\ell_{1_C}}{\ell_2} \frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}) \\ a_{44} &= -\frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) + \frac{\ell_{1_C}}{\ell_2} \frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}). \end{aligned} \tag{6.39}$$

The adjoint matrix representation (6.10) implies

$$\mathrm{ad}_{\lambda(\widehat{\bar{g}},u)} = \begin{bmatrix} 0 & -\frac{1}{\ell} V_1 \tan\theta & 0 & 0 \\ \frac{1}{\ell} V_1 \tan\theta & 0 & V_1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}. \tag{6.40}$$

Since we consider the Cartan-Schouten (0)-connection form $\omega^{(0)} = \frac{1}{2}\mathrm{ad}$, the torsion function $T_{\lambda(\widehat{\bar{g}},u)}$ vanishes (see [22]), thus, in matrix form, it is given by

$$T_{\lambda(\widehat{\bar{g}},u)} = 0_{4\times 4}. \tag{6.41}$$

Using (6.38), (6.40) and (6.41) we obtain (6.29).

Computation of E

The operator $E : \mathbb{R}^4 \to (\mathbb{R}^4)^*$ takes the form (2.38). From (6.35) and the definitions of the matrices D and Q we can find the composition

$$(\mathrm{d}h(\,\widehat{\bar{g}}\,))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ \mathrm{d}h(\,\widehat{\bar{g}}\,) = \begin{bmatrix} \epsilon_{11} & \epsilon_{12} & \epsilon_{13} & \epsilon_{14} \\ \epsilon_{21} & \epsilon_{22} & \epsilon_{23} & \epsilon_{24} \\ \epsilon_{31} & \epsilon_{32} & \epsilon_{33} & \epsilon_{34} \\ \epsilon_{41} & \epsilon_{42} & \epsilon_{43} & \epsilon_{44} \end{bmatrix} \tag{6.42}$$

where

$$\begin{aligned} \epsilon_{11} &= 2\frac{q_1}{d_1^2} \\ \epsilon_{12} &= \epsilon_{21} = 0 \\ \epsilon_{13} &= \epsilon_{31} = -\frac{q_1}{d_1^2}\ell_1 \sin\widehat{\alpha} - \frac{q_1}{d_1^2}\ell_{1_c} \sin\widehat{\alpha} \\ \epsilon_{14} &= \epsilon_{41} = \frac{q_1}{d_1^2}\ell_2 \sin\widehat{\beta} \\ \epsilon_{22} &= 2\frac{q_1}{d_1^2} \\ \epsilon_{23} &= \epsilon_{32} = \frac{q_1}{d_1^2}\ell_1 \cos\widehat{\alpha} + \frac{q_1}{d_1^2}\ell_{1_c} \cos\widehat{\alpha} \\ \epsilon_{24} &= \epsilon_{42} = -\frac{q_1}{d_1^2}\ell_2 \cos\widehat{\beta} \\ \epsilon_{33} &= \frac{q_1}{d_1^2}\ell_1^2 + \frac{q_1}{d_1^2}\ell_{1_c}^2 \\ \epsilon_{34} &= \epsilon_{43} = -\frac{q_1}{d_1^2}\ell_{1_c}\ell_2 \sin\widehat{\alpha}\sin\widehat{\beta} - \frac{q_1}{d_1^2}\ell_{1_c}\ell_2 \cos\widehat{\alpha}\cos\widehat{\beta} \end{aligned}$$

$$\epsilon_{44} = \frac{q_1}{d_1^2}\ell_2^2.$$

The function

$$\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big) : \mathbb{R}^4 \to (\mathbb{R}^4)^*$$

in lifted through the exponential functor $(\cdot)^{T_{\widehat{\overline{g}}}\overline{G}}$ to the linear map

$$\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}} : \mathcal{L}(T_{\widehat{\overline{g}}}\overline{G}, \mathbb{R}^4) \to \mathcal{L}(T_{\widehat{\overline{g}}}\overline{G}, (\mathbb{R}^4)^*)$$

defined as

$$\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}(\xi)} = \big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big) \circ \xi.$$

Let $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = (\widehat{x}_1', \widehat{y}_1', \widehat{\alpha}_1', \widehat{\beta}_1')^T$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = (\widehat{x}_2', \widehat{y}_2', \widehat{\alpha}_2', \widehat{\beta}_2')^T \in T_{\widehat{\overline{g}}}\overline{G}$ be two vector fields, then the Hessian matrix is defined by

$$\begin{aligned}\operatorname{Hess} h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})) =& \mathrm{d}(\mathrm{d}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2})))(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})) \\ &- \mathrm{d}h(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))),\end{aligned} \tag{6.43}$$

from the choice of Cartan-Schouten (0)-connection, we get

$$\nabla_{T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}}(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \frac{1}{2}T_eL_{\widehat{\overline{g}}}(\mathrm{ad}_{\eta^{\overline{g}_1}}\eta^{\overline{g}_2}). \tag{6.44}$$

The Hessian evaluated in $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1}$ and $T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}$ is therefore

$$\begin{aligned}&\mathrm{Hess}h(\widehat{\overline{g}})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_1})(T_eL_{\widehat{\overline{g}}}\eta^{\overline{g}_2}) = \\ &\qquad = \begin{bmatrix} -\widehat{\alpha}_1'\widehat{\alpha}_2'\ell_1\cos\widehat{\alpha} - \frac{1}{2}\widehat{\alpha}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\alpha}_1'\widehat{y}_2' \\ -\widehat{\alpha}_1'\widehat{\alpha}_2'\ell_1\sin\widehat{\alpha} + \frac{1}{2}\widehat{\alpha}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\alpha}_1'\widehat{x}_2' \\ -\widehat{\alpha}_1'\widehat{\alpha}_2'\ell_{1_c}\cos\widehat{\alpha} + \widehat{\beta}_1'\widehat{\beta}_2'\ell_2\cos\widehat{\beta} - \frac{1}{2}\widehat{\alpha}_2'\widehat{y}_1' + \frac{1}{2}\widehat{\alpha}_1'\widehat{y}_2' \\ -\widehat{\alpha}_1'\widehat{\alpha}_2'\ell_{1_c}\sin\widehat{\alpha} + \widehat{\beta}_1'\widehat{\beta}_2'\ell_2\sin\widehat{\beta} + \frac{1}{2}\widehat{\alpha}_2'\widehat{x}_1' - \frac{1}{2}\widehat{\alpha}_1'\widehat{x}_2' \end{bmatrix}.\end{aligned} \tag{6.45}$$

It follows that

$$\begin{aligned}&\big((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h(\widehat{\overline{g}}))\big)^{T_{\widehat{\overline{g}}}\overline{G}} \circ \operatorname{Hess} h(\widehat{\overline{g}}) = \\ &\begin{bmatrix} 0 & 0 & \frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & 0 \\ 0 & 0 & -\frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & 0 \\ -\frac{1}{2}(\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4) & \frac{1}{2}(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3) & e_{33} & 0 \\ 0 & 0 & 0 & \widetilde{\mathrm{y}}_3\ell_2\cos\widehat{\beta} + \widetilde{\mathrm{y}}_4\ell_2\sin\widehat{\beta} \end{bmatrix},\end{aligned} \tag{6.46}$$

$$e_{3,3} = -\widetilde{\mathrm{y}}_1\ell_1\cos\widehat{\alpha} - \widetilde{\mathrm{y}}_2\ell_1\sin\widehat{\alpha} + \widetilde{\mathrm{y}}_3\ell_{1_c}\cos\widehat{\alpha} + \widetilde{\mathrm{y}}_4\ell_{1_c}\sin\widehat{\alpha}.$$

In conclusion, combining (6.42) and (6.46) with $T_eL_{\widehat{\overline{g}}}^*$ and $T_eL_{\widehat{\overline{g}}}$, the matrix (6.30) is obtained.

Computation of W

Equation (6.31) follows from the choice of the Cartan-Schouten (0)-connection.

Initial condition

The initial condition for the filter is given by (2.33) while the initial condition for the gain is $K(t_0) = X_0^{-1}$ where the operators $X_0 : \overline{\mathfrak{g}} \to \overline{\mathfrak{g}}^*$ satisfies (2.36).

We rewrite m_0 as

$$m_0(\overline{g}) = \frac{1}{2}\|I - \overline{g}^{-1}(t)\overline{g}_0\|_F^2 = \frac{1}{2}\mathrm{tr}\left[(I_{5\times 5} - \overline{g}^{-1}\overline{g}_0)^T(I_{5\times 5} - \overline{g}^{-1}\overline{g}_0)\right]. \quad (6.47)$$

Computing the trace we obtain

$$m_0(\overline{g}) = \frac{1}{2}[4(1-\cos(\alpha-\alpha_0)) + (x-x_0)^2 + (y-y_0)^2 + 2(1-\cos(\beta-\beta_0))] \quad (6.48)$$

and from (2.33) it follows that $\widehat{\overline{g}}(t_0) = \overline{g}_0$.

The Hessian of the function m_0 at a point $\overline{g} \in \overline{G}$ is defined as

$$\mathrm{Hess}\, m_0(\overline{g})(\overline{g}X)(\overline{g}Y) = \mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) - \mathrm{d}m_0(\overline{g})(\nabla_{\overline{g}X}(\overline{g}Y))$$

for all $\overline{g}X$, $\overline{g}Y \in T_{\overline{g}}\overline{G}$. The differential of m_0 is given by

$$\mathrm{d}m_0(\overline{g}) = \begin{bmatrix} 2\sin(\alpha-\alpha_0) & (x-x_0) & (y-y_0) & \sin(\beta-\beta_0)\end{bmatrix} \quad (6.49)$$

while, given $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1}) = \widehat{\overline{g}}\eta^{\overline{g}_1}$, $T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}) = \widehat{\overline{g}}\eta^{\overline{g}_2} \in T_{\overline{g}\overline{G}}$, the affine connection yields

$$\nabla_{(\widehat{\overline{g}}\eta^{\overline{g}_1})}(\widehat{\overline{g}}\eta^{\overline{g}_2}) = \frac{1}{2}\begin{bmatrix} \widehat{\alpha}_2'\widehat{y}_1' - \widehat{\alpha}_1'\widehat{y}_2' \\ -\widehat{\alpha}_2'\widehat{x}_1' + \widehat{\alpha}_1'\widehat{x}_2' \\ 0 \\ 0 \end{bmatrix}. \quad (6.50)$$

Combining (6.49) and (6.50) we obtain

$$\begin{aligned} &\mathrm{d}m_0(\widehat{\overline{g}})(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) \\ &\quad = \left[(x-x_0)(\widehat{\alpha}_2'\widehat{y}_1' - \widehat{\alpha}_1'\widehat{y}_2') + (y-y_0)(-\widehat{\alpha}_2'\widehat{x}_1' + \widehat{\alpha}_1'\widehat{x}_2')\right] \end{aligned} \quad (6.51)$$

that evaluating in $\widehat{\overline{g}}_0$ produces

$$\mathrm{d}m_0(\widehat{\overline{g}}_0)(\nabla_{T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_1})}(T_eL_{\widehat{\overline{g}}}(\eta^{\overline{g}_2}))) = 0. \quad (6.52)$$

The double differential takes the form

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{1, 1, 2\cos(\alpha-\alpha_0), \cos(\beta-\beta_0)\} \quad (6.53)$$

that, evaluating in $\widehat{\overline{g}}_0$, produces

$$\mathrm{d}(\mathrm{d}m_0(\overline{g})(\overline{g}Y))(\overline{g}X) = \mathrm{diag}\{1, 1, 2, 1\} \quad (6.54)$$

and thus, from (6.52) and (6.54)

$$\mathrm{Hess}\, m_0(\widehat{\overline{g}}_0) = \mathrm{diag}\{1, 1, 2, 1\}. \quad (6.55)$$

From

$$K(t_0) = X_0^{-1} = T_e L^*_{\widehat{\overline{g}}_0} \circ \operatorname{Hess} m_0(\widehat{\overline{g}}_0) \circ T_e L_{\widehat{\overline{g}}_0} \tag{6.56}$$

we obtain the initial condition of $K(t_0)$

$$\begin{aligned} K(t_0) &= (T_e L_{\widehat{\overline{g}}_0})^{-1} (\operatorname{Hess} m_0(\widehat{\overline{g}}_0))^{-1} (T_e L^*_{\widehat{\overline{g}}_0})^{-1} \\ &= \operatorname{diag}\{1, 1, 1/2, 1\}. \end{aligned} \tag{6.57}$$

This computation ends the proof. □

For the second case (reversing maneuver with LIDAR sensor) we have the following Proposition. The proof is similar to the previous one and is omitted.

Proposition 6.2. *Consider the dynamic system* (6.13) *with output map* h *and the linear map* D *are given by* (6.16) *and* (6.23)*, respectively and where the operator* B *takes the form* (6.19)*. Consider the cost functional* (2.29)-(2.31) *where the initial cost* m_0 *is given by* (6.17) *and the matrix representation of the forms* $\mathcal{R}$*,* $\mathcal{Q}$ *are given by* (6.18) *and* (6.22)*, respectively. Then the second-order optimal filter is*

$$\widehat{\overline{g}}^{-1}\dot{\widehat{\overline{g}}} = \lambda_t(\widehat{\overline{g}}, u) + K(t) r_t(\widehat{\overline{g}}), \quad \widehat{\overline{g}}(t_0) = \widehat{\overline{g}}_0 \tag{6.58}$$

where the residual r *is*

$$r_t = \begin{bmatrix} (\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\cos\widehat{\alpha} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\sin\widehat{\alpha} \\ -(\widetilde{\mathrm{y}}_1 + \widetilde{\mathrm{y}}_3)\sin\widehat{\alpha} + (\widetilde{\mathrm{y}}_2 + \widetilde{\mathrm{y}}_4)\cos\widehat{\alpha} \\ -\widetilde{\mathrm{y}}_1\ell_1\sin\widehat{\alpha} + \widetilde{\mathrm{y}}_2\ell_1\cos\widehat{\alpha} - \widetilde{\mathrm{y}}_3\ell_{1_c}\sin\widehat{\alpha} + \widetilde{\mathrm{y}}_4\ell_{1_c}\cos\widehat{\alpha} \\ \widetilde{\mathrm{y}}_3\ell_2\sin\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2\cos\widehat{\beta} \\ \widetilde{\mathrm{y}}_5 \end{bmatrix}^T . \tag{6.59}$$

The optimal gain $K =: (\mathbb{R}^4)^* \to \mathbb{R}^4$ *is the solution of the perturbed matrix Riccati differential equation*

$$\begin{aligned} \dot{K} = -cK + AK + KA^T - KEK + BR^{-1}B^T \\ -W(K,r)K - KW(K,r)^T \end{aligned} \tag{6.60}$$

with

$$A(t) = \begin{bmatrix} 0 & a_{12} & 0 & 0 \\ a_{21} & 0 & a_{23} & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & a_{43} & a_{44} \end{bmatrix} \tag{6.61}$$

where

$$\begin{aligned} a_{12} &= \frac{1}{\ell} V_1 \tan\theta \\ a_{21} &= -\frac{1}{\ell} V_1 \tan\theta \\ a_{23} &= -V_1 \\ a_{43} &= \frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) - \frac{\ell_{1_C}}{\ell_2}\frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}) \\ a_{44} &= -\frac{1}{\ell_2} V_1 \cos(\widehat{\alpha} - \widehat{\beta}) + \frac{\ell_{1_C}}{\ell_2}\frac{1}{\ell} V_1 \tan\theta \sin(\widehat{\alpha} - \widehat{\beta}), \end{aligned}$$

and

$$E(t) = \begin{bmatrix} E_{11} & E_{12} & E_{13} & E_{14} \\ E_{21} & E_{22} & E_{23} & E_{24} \\ E_{31} & E_{32} & E_{33} & E_{34} \\ E_{41} & E_{42} & E_{43} & E_{44} \end{bmatrix} \tag{6.62}$$

where

$$E_{44} = \frac{q_3}{d_3^2}\ell_2^2 \sin^2\widehat{\beta} + \frac{q_4}{d_4^2}\ell_2^2 \cos^2\widehat{\beta} + \frac{q_5}{d_5^2} - \widetilde{\mathrm{y}}_3\ell_2 \cos\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2 \sin\widehat{\beta}$$

$$\begin{aligned}
E_{11} &= e_{11}\cos^2\widehat{\alpha} + e_{21}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{12}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{22}\sin^2\widehat{\alpha} \\
E_{12} &= -e_{11}\cos\widehat{\alpha}\sin\widehat{\alpha} - e_{21}\sin^2\widehat{\alpha} + e_{12}\cos^2\widehat{\alpha} + e_{22}\cos\widehat{\alpha}\sin\widehat{\alpha} \\
E_{13} &= e_{13}\cos\widehat{\alpha} + e_{23}\sin\widehat{\alpha} \\
E_{14} &= e_{14}\cos\widehat{\alpha} + e_{24}\sin\widehat{\alpha} \\
E_{21} &= -e_{11}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{21}\cos^2\widehat{\alpha} - e_{12}\sin^2\widehat{\alpha} + e_{22}\cos\widehat{\alpha}\sin\widehat{\alpha} \\
E_{22} &= e_{11}\sin^2\widehat{\alpha} - e_{21}\cos\widehat{\alpha}\sin\widehat{\alpha} - e_{12}\cos\widehat{\alpha}\sin\widehat{\alpha} + e_{22}\cos^2\widehat{\alpha} \\
E_{23} &= -e_{13}\sin\widehat{\alpha} + e_{23}\cos\widehat{\alpha} \\
E_{24} &= -e_{14}\sin\widehat{\alpha} + e_{24}\cos\widehat{\alpha} \\
E_{31} &= e_{31}\cos\widehat{\alpha} + e_{32}\sin\widehat{\alpha} \\
E_{32} &= -e_{31}\sin\widehat{\alpha} + e_{32}\cos\widehat{\alpha} \\
E_{33} &= e_{33} \\
E_{34} &= e_{34} \\
E_{41} &= e_{41}\cos\widehat{\alpha} + e_{42}\sin\widehat{\alpha} \\
E_{42} &= -e_{41}\sin\widehat{\alpha} + e_{42}\cos\widehat{\alpha} \\
E_{43} &= e_{34} \\
E_{44} &= e_{44} \\
e_{11} &= \frac{q_1}{d_1^2} + \frac{q_2}{d_2^2} \\
e_{12} &= 0 \\
e_{13} &= -\frac{q_1}{d_1^2}\ell_1 \sin\widehat{\alpha} - \frac{q_2}{d_2^2}\ell_{1_c} \sin\widehat{\alpha} - \frac{1}{2}(\widehat{\mathrm{y}}_2 + \widehat{\mathrm{y}}_4) \\
e_{14} &= \frac{q_2}{d_2^2}\ell_2 \sin\widehat{\beta} \\
e_{21} &= 0 \\
e_{22} &= \frac{q_1}{d_1^2} + \frac{q_2}{d_2^2} \\
e_{23} &= \frac{q_1}{d_1^2}\ell_1 \cos\widehat{\alpha} + \frac{q_2}{d_2^2}\ell_{1_c} \cos\widehat{\alpha} + \frac{1}{2}(\widehat{\mathrm{y}}_1 + \widehat{\mathrm{y}}_3) \\
e_{24} &= -\frac{q_2}{d_2^2}\ell_2 \cos\widehat{\beta} \\
e_{31} &= -\frac{q_1}{d_1^2}\ell_1 \sin\widehat{\alpha} - \frac{q_2}{d_2^2}\ell_{1_c} \sin\widehat{\alpha} + \frac{1}{2}(\widehat{\mathrm{y}}_2 + \widehat{\mathrm{y}}_4)
\end{aligned}$$

$$
\begin{aligned}
e_{32} &= \frac{q_1}{d_1^2}\ell_1 \cos\widehat{\alpha} + \frac{q_2}{d_2^2}\ell_{1_c}\cos\widehat{\alpha} - \frac{1}{2}(\widehat{\mathrm{y}}_1 + \widehat{\mathrm{y}}_3)\\
e_{33} &= \frac{q_1}{d_1^2}\ell_1^2 + \frac{q_2}{d_2^2}\ell_{1_c}^2 + \widetilde{\mathrm{y}}_1\ell_1\ \cos\widehat{\alpha} + \widetilde{\mathrm{y}}_2\ell_1\ \sin\widehat{\alpha} - \widetilde{\mathrm{y}}_3\ell_{1_c}\ \cos\widehat{\alpha} - \widetilde{\mathrm{y}}_4\ell_{1_c}\ \sin\widehat{\alpha}\\
e_{34} &= -\frac{q_2}{d_2^2}\ell_{1_c}\ell_2 \sin\widehat{\alpha}\sin\widehat{\beta} - \frac{q_2}{d_2^2}\ell_{1_c}\ell_2\cos\widehat{\alpha}\cos\widehat{\beta}\\
e_{41} &= \frac{q_2}{d_2^2}\ell_2\sin\widehat{\beta}\\
e_{42} &= -\frac{q_2}{d_2^2}\ell_2\cos\widehat{\beta}\\
e_{43} &= -\frac{q_2}{d_2^2}\ell_{1_c}\ell_2 \sin\widehat{\alpha}\sin\widehat{\beta} - \frac{q_2}{d_2^2}\ell_{1_c}\ell_2\cos\widehat{\alpha}\cos\widehat{\beta}\\
e_{44} &= \frac{q_2}{d_2^2}\ell_2^2 + \frac{q_3}{d_3^2} - \widetilde{\mathrm{y}}_3\ell_2\cos\widehat{\beta} - \widetilde{\mathrm{y}}_4\ell_2\sin\widehat{\beta}
\end{aligned}
$$

and

$$W(K,r) = \frac{1}{2}\mathrm{ad}_{(Kr)^\wedge}. \tag{6.63}$$

The initial conditions for the equations (6.58) *and* (6.28) *are*

$$\widehat{\overline{g}}(t_0) = \quad \overline{g}_0 \tag{6.64}$$

$$K(t_0) = \quad I_{5\times 5}. \tag{6.65}$$

6.4 Laboratory setting

The experimental validations were conducted on truck and semi-trailer scaled models at the Automotive Lab, Eindhoven University of Technology, within the project TruckLab (Figure 6.4). The operating space in the laboratory is $7\,\mathrm{m} \times 7\,\mathrm{m}$ including a docking station and 3 tractors and semi-trailers. The laboratory is equipped with motion capture cameras Prime^{x}13 and, together with markers attached to the scaled vehicles, allows to have a position accuracy of about $\pm 20\,\mathrm{mm}$; so we can consider such measurements as the (true) reference trajectories. The scaled vehicles operate on ROS (Robot Operating System) and are configured with Turtlebot3 Waffle Pi software architecture. The scaled model vehicles are a faithful reproduction of real truck semi-trailer vehicles, with a reduction ratio of 1:13.3. The model dimensions in Figure 6.2 are listed in Table 6.1.

element	cm
ℓ_1	28
ℓ_{1_c}	5.5
ℓ_2	56.7

Table 6.1: Model datasheet

The steering angle is measured with a combination of odometer and IMU measurements. The steering wheels of the scale reproductions have a maximum

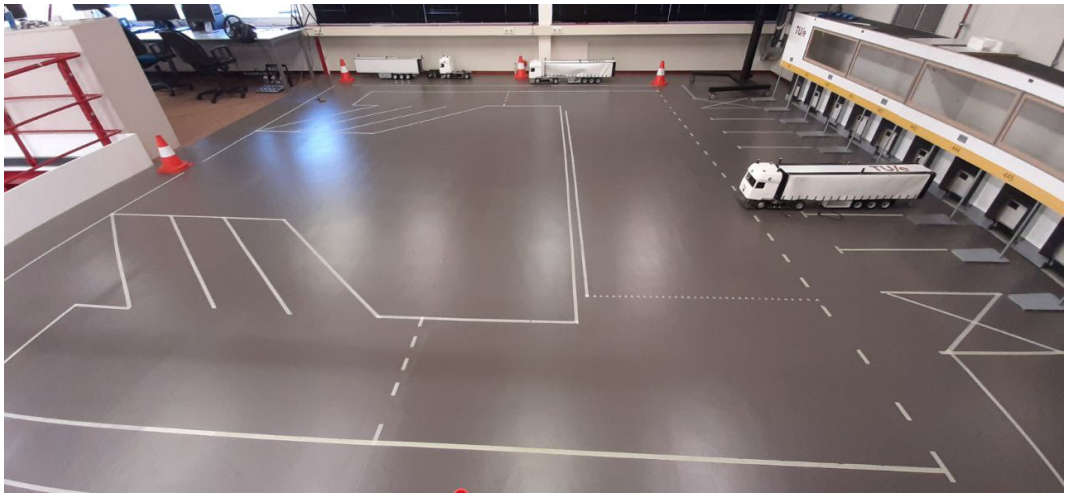


Figure 6.4: Automotive Laboratory (Eindhoven University of Technology)

steering angle θ of $\pm 38\,\text{deg}$. Thus, we impose the inequality $-38\,\text{deg} \leq \theta \leq +38\,\text{deg}$.

Another constraint is represented by the so-called jackknifing, which is a condition where the articulation angle between the tractor and the semi-trailer becomes very large. This problem can arise when driving forward and applying a large steering angle, or, when driving backward (in this case the vehicle configuration is unstable), a small constant steering input in the articulation angle will grow until the cabin collides with the semi-trailer. This condition results in the inequality $-100\,\text{deg} \leq \gamma_1 \leq +100\,\text{deg}$.

6.5 Simulations and discussions

To simulate noisy measurements given by the GPS and LIDAR devices, we add Gaussian white noises to the reference trajectories provided by the optic cameras. The standard deviations of these measurement errors are reported in Table 6.2.

measure	standard deviation
$x_{0_{\text{GPS}}}$	5 m
$y_{0_{\text{GPS}}}$	5 m
$x_{2_{\text{GPS}}}$	5 m
$y_{2_{\text{GPS}}}$	5 m
$x_{2_{\text{LIDAR}}}$	0.10 m
$y_{2_{\text{LIDAR}}}$	0.10 m
β_{LIDAR}	0.02 rad

Table 6.2: Measurement errors

The linear velocity of the truck is obtained by adding to the reference velocity a Gaussian white noise with standard deviation of 0.1m/s.

The matrix representations of the quadratic forms $\mathcal{R}$ and the linear operator B are

$$R := \text{diag}\{1, 1, 1, 1\},$$
$$B := \text{diag}\{0.1, 0.1, 0.1, 0.1\}.$$

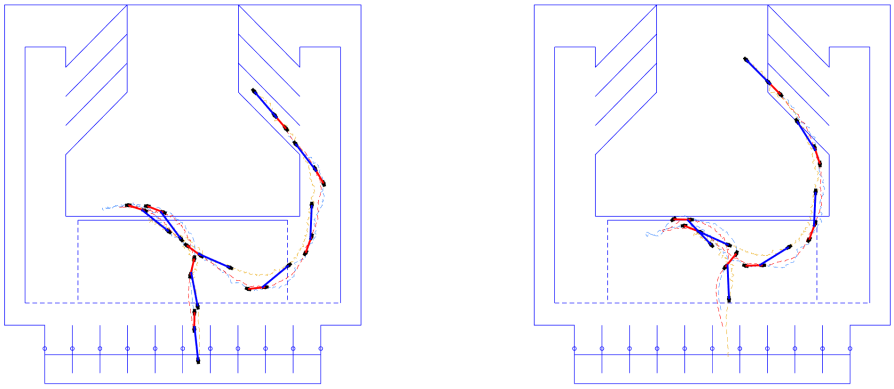
The matrix representations of the quadratic forms $\mathcal{Q}$ and the linear operator D in the first part of the maneuver are

$$Q := \text{diag}\{1, 1, 1, 1\},$$
$$D := \text{diag}\{0.5, 0.5, 0.5, 0.5\},$$

while for the second part, they are

$$Q := \text{diag}\{1, 1, 1, 1, 1\},$$
$$D := \text{diag}\{0.5, 0.5, 0.5, 0.5, 0.5\}.$$

To solve the differential equations, we use a forward Euler method with a sample time of $T_s = 10\text{ms}$.



(a) Parking maneuver for the dataset 1

(b) Parking maneuver for the dataset 2

Figure 6.5

In Figure 6.5a and Figure 6.5b we show the maneuvers of the truck semi-trailer system for two datasets both consisting in pose, steering angle and linear velocity trajectories. In Figure 6.6 and 6.7 we report the graphs of the reference, measured (corrupted by noises) and filtered trajectories. The first vertical line corresponds to the instant when the vehicle starts the reversing maneuver, while the second one to the instant when the filter uses also the LIDAR sensor. As can be seen, the filter performs well even if the noises are large. The final errors due to filter approximation are acceptable if compared with the parking space, the dimension of the vehicles and the sensors' accuracy.

The addition of the LIDAR measurements significantly improves the precision of the filter and allows to have better estimations in the final part of the maneuvering. Another advantage of considering a LIDAR sensor is highlighted by its ductility of being able to be settled both on the semi-trailer or on the docking station, having the characteristic of detecting nearby physical objects. For this reason, the choice of using a single LIDAR on the docking station for several vehicles can be cost-effective.

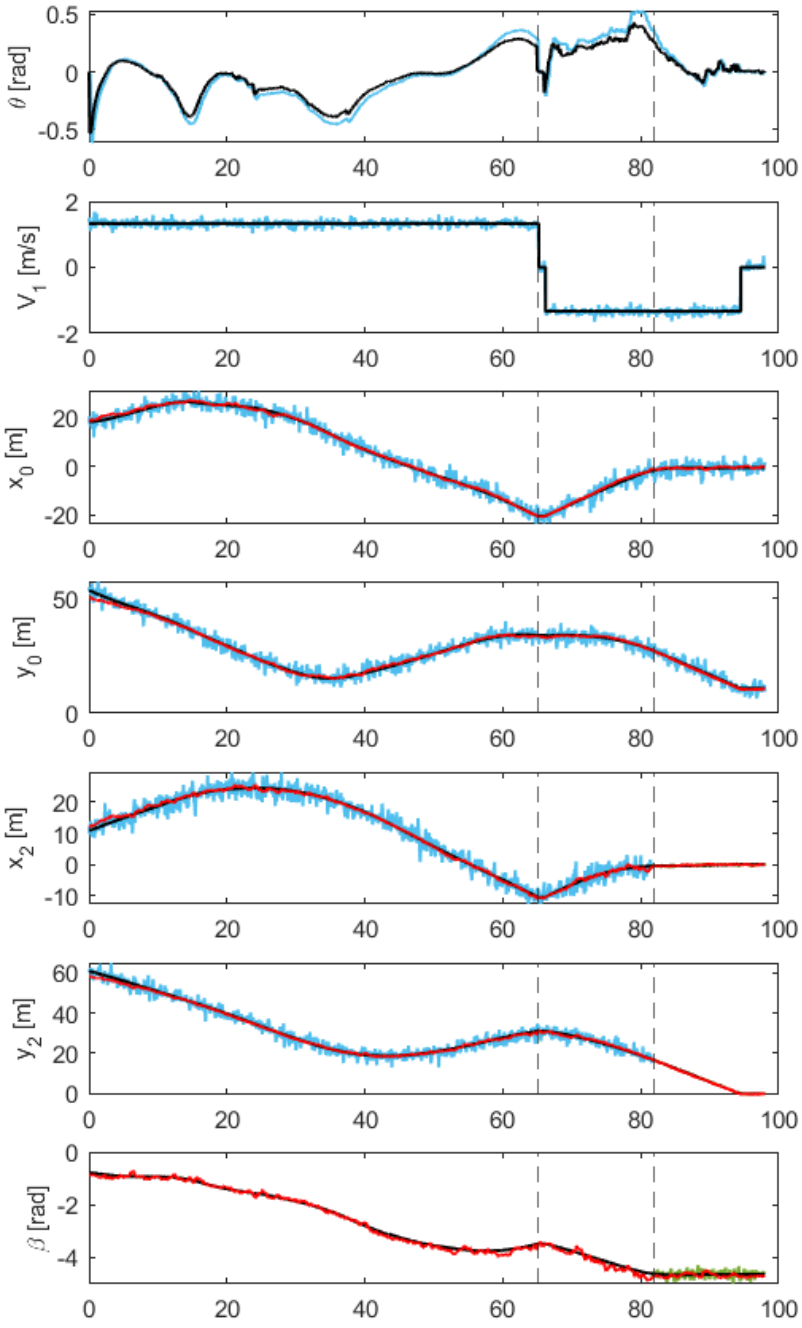


Figure 6.6: References (black), measured with GPS (blue) measured with LIDAR (green) and filtered (red) inputs and trajectories for dataset 1

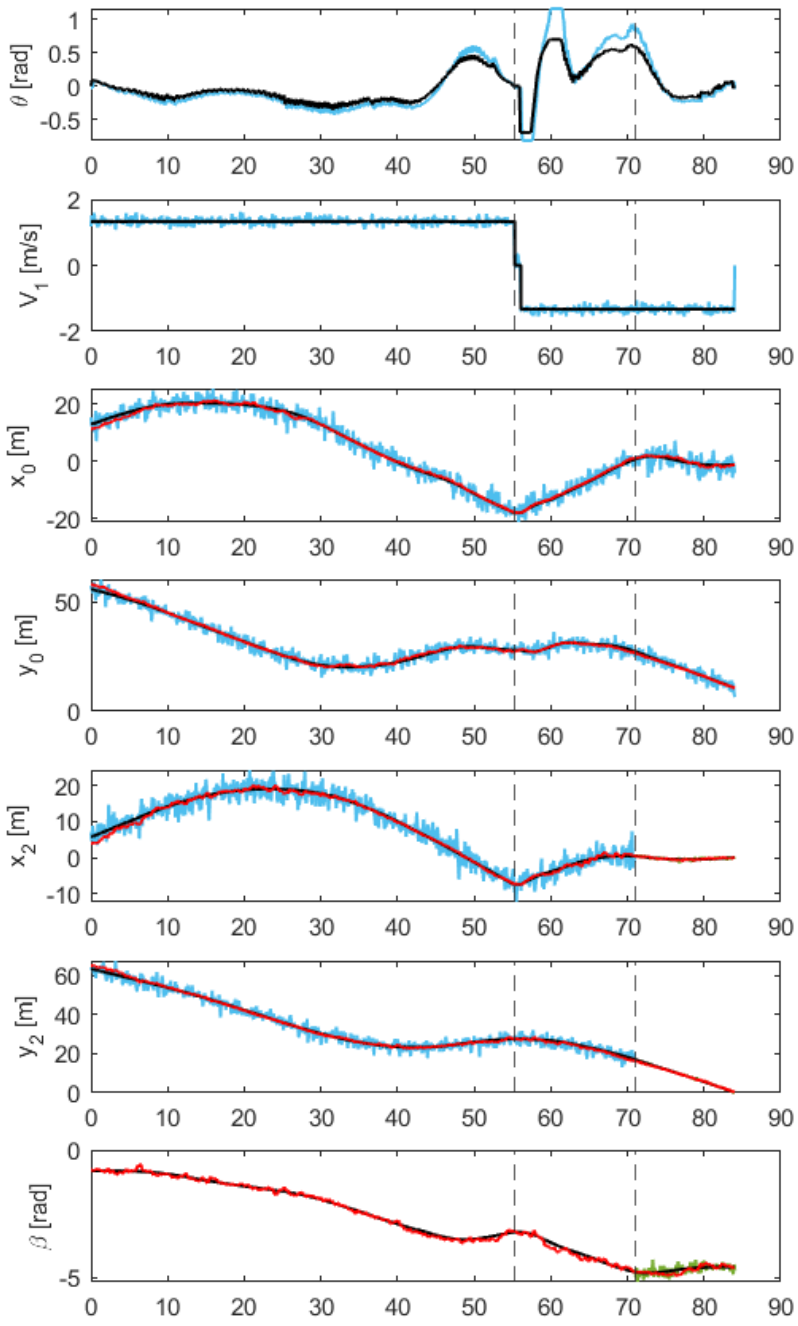


Figure 6.7: References (black), measured with GPS (blue) measured with LIDAR (green) and filtered (red) inputs and trajectories for dataset 2

Chapter 7

Conclusions and future works

In this thesis, we dealt with the design of second-order optimal filters based on the Lie group theory and applied to planar rigid bodies. This filter is part of the so-called minimum energy filters since its optimality is achieved considering a quadratic cost function in the model and measurement errors. The dynamics evolves on Lie groups, and this lets us consider its trivialization and take into account the symmetries of the system. The advantages of this filter are many, such as (1) it does not require any hypothesis on the errors, treating them as deterministic functions, (2) the Lie group structure lets use global coordinates, and so we only need to use a single chart. It is a second-order optimal filter since, for the exact derivation of the filter, it would be necessary to solve an infinite dimension Hamilton-Jacobi-Bellman (HJB) equation, and thus, our filter relies on a second-order approximation of the value function.

We first applied the filter to the case of a free planar rigid body whose underlying dynamics evolves on the tangent bundle of the special Euclidean group on the plane. This case occurs for example in many marine vehicle systems such as hovercraft, or in planar rigid bodies without constraints. We started our analysis by investigating the matrix Lie group structure that underlies the dynamics described by Euler-Poincaré equation for a free rigid body. An important aspect of this work is the description of the Lie algebra homomorphism between $\mathfrak{se}(2)$ and $\mathbb{R}^3$. Particular attention was paid to the choice of the measurement equation and its relation with the observability of the system. After that, we compared this filter with the extended Kalman filter. We highlighted that under the hypothesis of linearity and Gaussianity of the errors the two filters are equivalent. The filter on Lie groups proved to have almost the same behavior of the extended Kalman filter and thus the former can be seen as an extension of the latter.

Since in real life most vehicles are subject to constraints originated by the use of wheels (such as bicycles or cars) or blades (such as sleighs), we studied the design of second-order filters applied to nonholonomic systems. In particular, we examine the Chaplygin sleigh, which is a nonholonomic system that models a planar rigid body supported at three points, two of which slide freely while the third is a blade at distance a from the center of mass and that cannot move perpendicularly. From a geometric point of view, the presence

of a nonholonomic constraint changes the state space and thus, it is no more described as the tangent space of a base manifold, but by a distribution of it. We derived the Lie group matrix representation and the related concepts in light of this distribution. Particular attention was devoted to the left-trivialized dynamics. In general, the filter does not preserve the nonholonomic constraint due to the presence of non-null components on the gain-residual term. Thus, we investigated the conditions that ensure the preservation of the nonholonomic constraint by properly choosing the affine connection which guarantees that the orthogonal velocity is equal to zero by design.

After that, we shifted our focus to articulated vehicle systems. These systems model multi-body vehicles composed of rigid bodies connected by hooking constraints such as luggage carriers in airports or cars with trailers in warehouses. We considered the case of a leading car pulling n trailers with different masses and inertias. The use of Hamel's coordinates to write the dynamic equations allows to derive the Lie group structure for this system. We designed three second-order optimal filters where the masses and inertias are (i) known at each instant of time, (ii) unknown but time-invariant, (iii) unknown and time-varying and thus treated as dynamic variables and estimated. The latter can be applied for example in case of consumption of fuel, or if some weights are loaded or unloaded from the trailers (e.g. in an airport or warehouse).

In the last case we studied the application of the filter to a scaled model of a truck semi-trailer system in a parking area. This work was done in cooperation with the Dynamics and Control group and with the Automotive Lab at Eindhoven University of Technology. The measurements are obtained by adding artificial noise to the ones obtained using motion capture cameras. We consider different types of measurements in order to have better estimations when the system is moving reversing, in which case the dynamics become unstable.

In the thesis, we investigated many aspects that characterize the filter, such as the choice of the measurement equations, the choice of the connection function, and the presence of nonholonomic and hooking constraints.

In the future, we plan to extend these results to bodies whose dynamics no longer evolve on the plane but in space, such as drones, airplanes or satellites. Another important research topic could be the study of stability. This result will be not trivial since the residual r is within the Riccati equation, and it makes the study of the convergence of the Riccati-like equation and the analysis of the stability of the non-linear time-invariant system quite challenging. Finally, an important theoretical result could be the formulation of the discrete version of the filter, as done for the Kalman filter. The current version requires measurements at each instant of time: this hypothesis is not feasible in real-life applications. A discrete version would be better tailored to the sampling time of the sensing system and a continuous time. Discrete measurement equations would allow to estimate the state between measurements.

Appendix A

Review of differential geometry

In this appendix, we report the main topics of differential geometry theory. For a complete treatment see e.g. [5], [7], [8], [11], [13], [22].

Before giving the standard definitions of differential geometry, we provide the notion of tensor that will be useful in the discussion.

Definition A.1. (Tensor product of vector spaces). Let V and W be two vector spaces. The tensor product of V and W denoted by $V \otimes W$ is a vector space with a bilinear map $\otimes : V \times W \to V \otimes W$ such that, given a vector space Z and a bilinear map $\varphi : V \times W \to Z$, then there exists a unique linear map $\overline{\varphi} : V \otimes W \times Z$ such that $\varphi = \overline{\varphi} \circ \otimes$, namely $\varphi(v, w) = \overline{\varphi}(v \otimes w)$ for every $v \in V$ and $w \in W$.

It is possible to extend this definition to an arbitrary product of vector spaces $V_1 \otimes V_2 \otimes \cdots \otimes V_n$.

Definition A.2. (Tensor of type (r, s)). Let V be a vector space and V^* its dual. A tensor of type (r, s) is a multilinear map

$$T^r_s : \underbrace{V^* \times \cdots \times V^*}_{r} \times \underbrace{V \times \cdots \times V}_{s} \to \mathbb{R}.$$

A tensor $T^r_s \in \mathcal{L}(V^*, ..., V^*, V, ..., V; \mathbb{R})$ is an element of the product tensor space

$$\otimes^r_s(V) := V^{\otimes r} \otimes (V^*)^{\otimes s} = \underbrace{V \otimes ... \otimes V}_{r} \otimes \underbrace{V^* \otimes ... \otimes V^*}_{s}.$$

The study of differential geometry revolves around the notion of differentiable manifold. Before giving its definition, it is necessary to introduce the concepts of charts, atlases and differentiable structures. These concepts let to describe a manifold putting it in relation to open spaces in $\mathbb{R}^n$.

Definition A.3. (Charts, atlases, differentiable structures). Let S be a set.

- A chart for S is a pair $(\mathcal{U}, \phi)$ where $\mathcal{U}$ is a subset of S and $\phi : \mathcal{U} \to \mathbb{R}^n$ is an injection for which $\phi(\mathcal{U})$ is an open subset of $\mathbb{R}^n$.

- For $r \in \mathbb{N} \cup \{\infty\}$, a C^r-atlas for S is a collection $\mathcal{A} = \{(\mathcal{U}_a, \phi_a)\}_{a \in A}$ of charts with the properties that $S = \bigcup_{a \in A} \mathcal{A}_a$ and, whenever $\mathcal{U}_a \cap \mathcal{U}_b \neq \emptyset$, $\phi_a(\mathcal{U}_a \cap \mathcal{U}_b)$ and $\phi_b(\mathcal{U}_a \cap \mathcal{U}_b)$ are open subsets of $\mathbb{R}^n$ and the map $\phi_b \circ \phi_a^{-1} : \phi_a(\mathcal{U}_a \cap \mathcal{U}_b) \to \phi_b(\mathcal{U}_a \cap \mathcal{U}_b)$ is a C^r-diffeomorphism.

- Two C^r-atlases $\mathcal{A}_1$ and $\mathcal{A}_2$ are equivalent if $\mathcal{A}_1 \cup \mathcal{A}_2$ is also a C^r-atlas. A C^r-differentiable structure on S is an equivalence class of atlases under this equivalence relation.

Definition A.4. (Manifolds). A C^r-differentiable manifold M is a set S with a C^r-differentiable structure. If all charts take values in $\mathbb{R}^n$, then n is the dimension of the manifold.

Roughly speaking, a manifold represents a set that "locally looks like" an open set of $\mathbb{R}^n$. Using the definition of charts and atlases, it is possible to define a set of coordinates $x^1, .., x^n$, which allows locating the points on the manifold.

Definition A.5. (Differentiable function). Let M be an $n-$dimensional manifold and $\mathcal{A}$ an atlas. A function $f : M \to \mathbb{R}$ is said to be differentiable of class C^k if the composition $f \circ \phi^{-1}$ is differentiable of class C^k for every chart $(\mathcal{U}, \phi)$ in the atlas $\mathcal{A}$.

The set of all differentiable functions from M to $\mathbb{R}$ will be denoted by $C^\infty(M)$.

Definition A.6. (Differentiable map between manifolds). Let M and N be two manifolds, and let $f : M \to N$ be a continuous map between them. Then f is said to be differentiable of class C^k if for any pair of charts $(\mathcal{U}, \phi)$ of M and $(\mathcal{V}, \psi)$ of N, the composition $\psi \circ f \circ \phi^{-1}$ if a differentiable map of class C^k. If f and f^{-1} are bijections of class C^k, then f is diffeomorphism of class C^k.

In order to define the notion of tangent vector and thus of tangent space at a point x of a manifold, it is useful to introduce the notion of differentiable curve.

Definition A.7. (Differentiable curve). A curve at $x \in M$ is a function $\gamma : I \to M$, where I is an interval of $\mathbb{R}$ containing 0 and such that $I(0) = x$.

Two curves γ_1, γ_2 are said to be equivalent at x if in a coordinate chart $(\mathcal{U}, \phi)$ with $x \in \mathcal{U}$, the local representative of γ_1 and γ_2 have the same derivative at 0. This equivalence is independent of the chart chosen. We will write $\gamma_1 \sim_x \gamma_2$ to indicate that γ_1 and γ_2 are equivalent at x. This is an equivalence relation $[\gamma]_x$.

Definition A.8. (Tangent vector). Let M be a manifold. A tangent vector v at $x \in M$ is the equivalence class $[\gamma]_x$.

Definition A.9. (Tangent space). The tangent space of M at x is the set of all tangent vectors of M at x and will be denoted by T_xM.

The tangent space T_xM is a vector space with basis $\left\{\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}\right\}$ where $x^1, .., x^n$ is the set of coordinates.

Definition A.10. (Tangent bundle). The tangent bundle TM is the set of all tangent space T_xM of all $x \in M$:

$$TM = \bigcup_{x \in M} T_xM.$$

It can be proven that, given an $n-$dimensional differentiable manifold M, the tangent bundle TM is a differentiable manifold with dimension $2n$. The canonical projection, $\pi : TM \to M$ is the mapping defined by $\pi(v) = x$ when $v \in T_xM$, thus $\pi^{-1}(x) = T_xM$.

Definition A.11. (Tangent map). Let M and N be two differentiable manifolds and let $\phi : M \to N$ be a C^k-map. Then the tangent map $T\phi : TM \to TN$ maps the vector $v = [\gamma]_x \in T_xM$ into the vector $T_{\phi(x)}\phi(v) = [\phi \circ \gamma]_{\phi(x)} \in T_{\phi(x)}N$.

Definition A.12. (Submersion, immersion, embedding). Let M and N be two manifolds and $f : M \to N$ be a smooth map. Then f is

- a submersion if its differential is surjective at each point;
- an immersion if its differential is injective at each point;
- an embedding if it is an immersion and it is a homeomorphism onto $f(M)$.

Definition A.13. (Submanifold). If $N \subset M$ and the inclusion $i : N \to M$ is an embedding, then N is a submanifold of M.

Definition A.14. (Vector fields). A vector field X of class C^k on M is a function that assigns to each point x of M a vector $v \in T_xM$ whose components in the frames of any local coordinates $(\mathcal{U}, \phi)$ are functions of class C^k on the domain $\mathcal{U}$ of the coordinates.

The set of all vector fields on a manifold M is denoted by $\mathfrak{X}(M)$.

The tangent space in a point x forms a vector space in which tangent vectors live. A related concept is that of cotangent space, namely the dual of the tangent space where are defined the linear applications on vectors.

Definition A.15. (Cotangent vector). Let M be a differentiable manifold and $x \in M$. A cotangent vector of M at x is a linear operator $\alpha : T_xM \to \mathbb{R}$ that associates to each vector v in T_xM a real number.

Definition A.16. (Cotangent space). Let M be a differentiable manifold. The cotangent space of M at x, denoted by T_x^*M, is the set of all cotangent vectors of M at x.

The cotangent space T_x^*M is a covector space with basis $\left\{\mathrm{d}x^1, ..., \mathrm{d}x^n\right\}$ where

$$\mathrm{d}x^i \frac{\partial}{\partial x^j} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

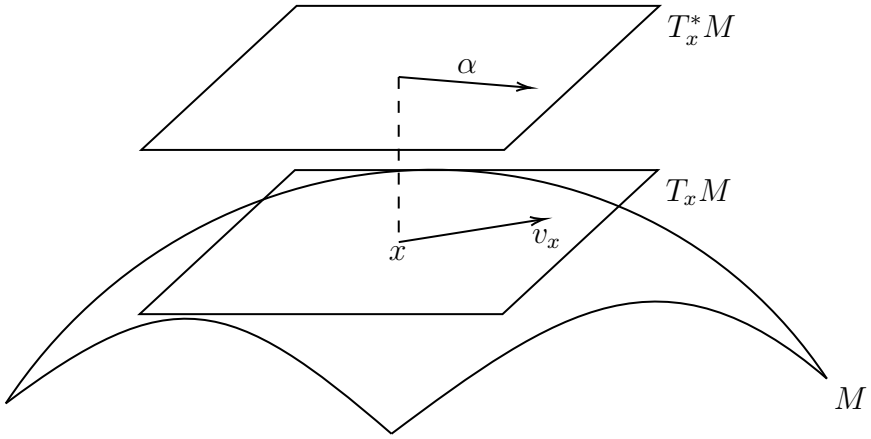


Figure A.1: Tangent and cotangent space in a point x. To view it, the cotangent bundle is depicted as a vector space and cotangent vectors as vectors.

Definition A.17. (Cotangent bundle). The cotangent bundle T^*M is the set of all cotangent space T_x^*M of all $x \in M$:

$$T^*M = \bigcup_{x \in M} T_x^*M.$$

The tangent and cotangent bundle, together with differentiable functions, let us to define the pull-back and the push-forward.

Definition A.18. (Pull-back of a real-function). Let M and N be two differentiable manifolds, consider a diffeomorphism $\phi : M \to N$ and a differentiable function $f : N \to \mathbb{R}$. Then the pull-back of f under ϕ is defined by $\phi^* f = f \circ \phi$.

Definition A.19. (Pull-back and push-forward of a vector field). Let M and N be two differentiable manifolds, and consider a diffeomorphism $\phi : M \to N$. Then:

- the pull-back of $Y \in \mathfrak{X}(N)$ is the vector field $\phi^* Y \in \mathfrak{X}(M)$ defined as $\phi^* Y = T\phi^{-1} \circ Y \circ \phi$;

- the push-forward of $X \in \mathfrak{X}(M)$ is the vector field $\phi_* X \in \mathfrak{X}(N)$ defined as $\phi_* Y = T\phi \circ X \circ \phi^{-1}$.

Via a vector field on a manifold M it is possible to construct a curve whose tangent vector in each point is equal to the value of the vector field. A wider concept is that of a flow.

Definition A.20. (Integral curve). An integral curve $\gamma : I \to M$ is a differentiable curve at $x \in M$ for a vector field X, if $\gamma'(t) = X(\gamma(t))$ and $\gamma(0) = x$.

Definition A.21. (Flow). Let $X \in \mathfrak{X}(M)$. A flow of X is a differentiable map $\Phi^X : U \times I \to M$, where $I \in \mathbb{R}$ is an interval containing 0 and U is an open subset of M, such that, for any $x \in U$, the map $\Phi^X(x,t)$ is an integral curve of X with $\Phi^X(x,0) = x$. We will use $\Phi_t^X(x)$ for $\Phi^X(x,t)$.

It is possible to define the concept of a derivation of a function with respect to a vector field. In a similar way one could define the derivative of a vector field with respect to a vector field. These are called Lie derivatives.

Definition A.22. (Lie derivative of a function). Let M be a differentiable manifold, X be a vector field with flow Φ_t^X, $f : M \to \mathbb{R}$ be a function. Then the Lie derivative of f at x with respect to X is given by

$$\begin{aligned}\mathcal{L}_X f(x) =& (Xf)(x) = X_x[f] \\ =& \lim_{t\to 0}\left(\frac{(\Phi_t^X)^* f - f}{t}\right)(x) = \lim_{t\to 0}\left(\frac{f((\Phi_t^X)^*(x)) - f(x)}{t}\right).\end{aligned}$$

In local coordinates, the Lie derivative of f at x with respect to X is given by

$$\mathcal{L}_X f(x) = \langle \mathrm{d}f(x), X(x)\rangle.$$

Definition A.23. (Lie derivative of vector fields). Let M be a differentiable manifold and $X, Y \in M$. Then the Lie derivative of Y at x with respect to X is given by

$$\mathcal{L}_X Y = \lim_{t\to 0}\left(\frac{(\Phi_t^X)^* Y - Y}{t}\right).$$

Definition A.24. (Commutator of vector fields). Given $X, Y \in \mathfrak{X}(M)$, the vector fields commutator $[\cdot,\cdot] : \mathfrak{X}(M)\times\mathfrak{X}(M) \to \mathfrak{X}(M)$ is defined as $[X, Y](f) = X(Yf) - Y(X(f))$ for all $f \in C^\infty(M)$.

Proposition A.1. *Let $X, Y \in \mathfrak{X}(M)$. Then $\mathcal{L}_X Y = [X, Y]$.*

We have already presented the notion of a submanifold of a manifold M as a subset of M that is also a manifold of a smaller dimension. It is possible to introduce a similar concept for what concern the tangent space considering a linear subspace that we call distribution. The Frobenius theorem allows us to understand under which conditions a distribution represents the tangent space of a submanifold.

Definition A.25. (Distribution). Let M be a differentiable manifold of dimension m. A distribution D of dimension $n < m$, is a map that assigns to each point $x \in M$ a vector subspace D_x of T_xM of dimension n.

Definition A.26. (Integrable submanifold). A submanifold N of M is called an integral manifold of a distribution D, if $T_xN = D_x$ for any point $x \in N$. If an integral manifold of D exists through each point of M, D is said to be completely integrable.

Definition A.27. (Involutive distribution). A distribution D is said to be involutive if for every X, $Y \in D$ we have that $[X, Y] \in D$.

Theorem A.1. *(Frobenius theorem). A distribution D of a manifold M is completely integrable if and only if it is involutive.*

Let's define now an important object called affine connection. An affine connection on a smooth manifold connects nearby tangent spaces, thus it permits tangent vector fields to be differentiated as if they were functions on the manifold with values in a fixed vector space.

Definition A.28. (Affine connection). Let M be a smooth manifold and let $\mathfrak{X}(M)$ be the space of vector fields on M. Then an affine connection ∇ on M is a bilinear map

$$\begin{aligned} \mathfrak{X}(M) \times \mathfrak{X}(M) &\to \mathfrak{X}(M) \\ (X, Y) &\mapsto \nabla_X Y \end{aligned} \tag{A.1}$$

such that

i) $\nabla_{fX} Y = f \nabla_X Y$ for each $X, Y \in \mathfrak{X}(TM)$ and each $f \in C^\infty(M)$;

ii) $\nabla_X fY = f \nabla_X Y + (\mathcal{L}_X f)Y$ for each $X, Y \in \mathfrak{X}(TM)$ and each $f \in C^\infty(M)$.

The vector field $\nabla_X Y$ is called covariant derivative of Y with respect to X.

An affine connection can be described using its projections on the element of the base. These projections are called Christoffel symbols.

Definition A.29. (Christoffel symbols). Let M be a differentiable manifold with connection ∇, and let $(x^1, ..., x^n)$ be coordinates in a chart $(\mathcal{U}, \phi)$. For each $i, j \in \{1, ..., n\}$, $\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}$ is a vector field on U that is a linear combinaton of $\{\frac{\partial}{\partial x^1}, ..., \frac{\partial}{\partial x^n}\}$, thus we can write

$$\nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j} = \Gamma_{ij}^k \frac{\partial}{\partial x^k}$$

where $\Gamma_{ij}^k : U \to \mathbb{R}$, $i, j, k \in \{1, ..., n\}$ are C^r-functions called Christoffel symbols.

Definition A.30. (Affine connection of functions and tensors). Let ∇ be an affine connection and let X be a vector field,

- let $f \in C^\infty$ be a function, then the covariant derivative of f with respect to X is given by
$$\nabla_X f = \mathcal{L}_X f; \tag{A.2}$$

- let $\alpha \in T^*M$ be a 1-form, then its covariant derivative with respect to X is implicitly given by
$$\langle \nabla_X \alpha; Y \rangle = \mathcal{L}_X \alpha(Y) - \langle \alpha; \nabla_X Y \rangle, \tag{A.3}$$
where $Y \in TM$;

- let T be a general tensor of type (r, s), then its covariant derivative with respect to X is given by

$$\begin{aligned}(\nabla_X T)(\alpha^1, ..., \alpha^r, X_1, ..., X_s) =& \mathcal{L}_X(T(\alpha^1, ..., \alpha^r, X_1, ..., X_s)) \\ &- \sum_{i=1}^{r} T(\alpha^1, ..., \nabla_X \alpha^i, ..., \alpha^r, X_1, ..., X_s) \\ &- \sum_{j=1}^{r} T(\alpha^1, ..., \alpha^r, X_1, ..., \nabla_X X_j, ..., X_s),\end{aligned} \tag{A.4}$$

where $\alpha^1, ..., \alpha^r \in T^*M$, $X_1, ..., X_s \in TM$.

Definition A.31. (Riemannian manifold). A Riemannian metric on a smooth manifold M is a (0,2)-tensor $g : T(M) \times T(M) \to \mathbb{R}$ that is symmetric and positive definite. A Riemannian manifold (M, g) is a manifold together with a Riemannian metric g.

The notion of metric lets us define some concepts as the length of a curve or angle between vectors as it's done in the classical Euclidean geometry.

There exist two important invariants of an affine connection on a manifold M: the torsion and the curvature.

Definition A.32. (Torsion). Let M be a smooth manifold and let ∇ be an affine connection. Then a torsion is a $(1, 2)$-tensor defined on vector fields X and Y by

$$T(X, Y) = \nabla_X Y - \nabla_Y X - [X, Y]\,. \tag{A.5}$$

Definition A.33. (Levi-Civita connection). Let (M, g) be a Riemannian manifold. An affine connection ∇ on M is called a Levi-Civita connection if:

- it preserves the metric (i.e. $\nabla g = 0$);
- it is torsion-free (i.e. $T(X, Y) = 0$).

Proposition A.2. *The Levi-Civita connection exists and is unique.*

Theorem A.2. *(Christoffel symbols for Levi-Civita connection). In the case of a Riemannian manifold with metric g, the Christoffel symbols are given by*

$$\Gamma^i_{kl} = \frac{1}{2} g^{im} \left(\frac{\partial g_{mk}}{\partial x^l} + \frac{\partial g_{ml}}{\partial x^k} - \frac{\partial g_{kl}}{\partial x^m} \right)$$

where g^{ij}, i,j=1,..,n, are the entry of the inverse of the metric tensor.

Definition A.34. (Curvature tensor). Let M be a differentiable manifold and ∇ an affine connection. The curvature tensor associated to ∇ is the tensor field of type $(1, 3)$ defined as

$$R(X, Y)Z = \nabla_X \nabla_Y Z - \nabla_Y \nabla_X Z - \nabla_{[X,Y]} Z$$

for all $X, Y, Z \in \mathfrak{X}(M)$.

Definition A.35. (Geodesic). Let M be a differentiable manifold with an affine connection ∇. A curve $\gamma : I \to M$ is a geodesic if $\nabla_{\dot{\gamma}}\dot{\gamma} = 0$.

A geodesic can be seen as the shortest curve that connects two points on the manifold.

Proposition A.3. *(Geodesic equation). Considering the components of the geodesic curve γ as $\gamma(t) = (x^1(t), ..., x^n(t))$, then they satisfy*

$$\frac{\mathrm{d}^2 x^i(t)}{\mathrm{dt}^2} + \Gamma^i_{jk}(x(t))\frac{\mathrm{d}x^j(t)}{\mathrm{dt}}\frac{\mathrm{d}x^k(t)}{\mathrm{dt}} = 0.$$

Appendix B

Review of Lie group theory

One of the most powerful structure that can be used to model mechanical control systems on manifold is the notion of Lie group. A Lie group combines the concepts of differentiable manifold and continuous group and inherits from them important properties related to the differential structure and the symmetries. In this appendix, we recall the basic concepts of Lie groups theory. For an extension review on Lie groups theory we refer to [7], [15], [21], [25], [38], [44].

Lie groups and Lie algebras

In this section, we provide the definition of Lie group and Lie algebra and the relationship between them. We start with some definitions and properties of Lie group and Lie algebra.

Definition B.1. (Group). A set G endowed with a binary operation $\star : G \times G \to G$, $(a, b) \mapsto a \star b$ is a group if:

i) (associativity) $a \star (b \star c) = (a \star b) \star c$ for all $a, b, c \in G$;

ii) (neutral element) there exists $e \in G$ such that $a \star e = e \star a = a$ for all $a \in G$;

iii) (inverse) there exists $a^{-1} \in G$ such that $a \star a^{-1} = a^{-1} \star a = e$ for all $a \in G$.

A group is abelian (or commutative) if the group operation is commutative, i.e. $a \star b = b \star a$ for all $a, b \in G$.

A subset H of G is a subgroup of G if the pair $(H, \star)$ is a group, i.e. if H is a group with respect to the binary operation defined on G.

We are now ready to give the definition of Lie group.

Definition B.2. (Lie group). A Lie group is a group that is also a manifold for which, given x, $y \in G$, the group operation $(x, y) \mapsto xy$ and inverse operation $x \mapsto x^{-1}$ are smooth.

A Lie group mixes together the notions of differentiable manifold and algebraic group, and from them, it inherits the group structure and the differential one.

There exist many types of Lie groups, among these, are matrix Lie groups.

Example B.1. (Matrix Lie groups). The set $GL(n;\mathbb{R})$ of invertible $n \times n$ matrices with real entries is a Lie group with respect to the operation of matrix multiplication.

Definition B.3. (Group homomorphism). Let $(G, \star)$ and $(H, \bullet)$ be two Lie groups. A Lie group homomorphism is a map $\rho : G \to H$ that satisfies $\rho(a \star b) = \rho(a) \bullet \rho(b)$ for all $a, b \in G$. If the Lie group homomorphism ρ is a bijection, then ρ is called an isomorphism, in that case, the groups are said to be isomorphic and one writes $G \simeq H$.

Definition B.4. (Subgroup). A Lie subgroup of a Lie group G is a subgroup $H \subseteq G$ for which the inclusion $i_H : H \to G$ is an injective immersion; If a Lie subgroup H of G is a submanifold of G, then it is a regular Lie subgroup.

An example of a subgroup is the notion of one-parameter subgroup, which is the image of a smooth group homomorphism $\rho : \mathbb{R} \to G$.

A related concept to Lie group is the one of Lie algebra.

Definition B.5. (Lie algebra). A Lie algebra V is a $\mathbb{R}$-vector space endowed with a biliniear operation $[\cdot, \cdot] : V \times V \to V$, called Lie bracket, such that:

i) anti-commutativity: $[\xi, \eta] = -[\eta, \xi]$ for all $\xi, \eta \in V$;

ii) Jacobi identity: $[\xi, [\eta, \zeta]] + [\eta, [\zeta, \xi]] + [\zeta, [\xi, \eta]] = 0$ for all ξ, η, ζ.

Definition B.6. (Structure constants). Consider a Lie algebra V with basis $\{e_1, ..., e_n\}$ and Lie bracket $[\cdot, \cdot]$. The Lie bracket operation has to satisfy

$$[e_i, e_j] = c_{ij}^k e_k, \quad i, j, k \in \{1, ..., n\},$$

where $c_{ij}^k \in \mathbb{R}$ are called structure constants.

Definition B.7. (Lie algebra homomorphism). Let $(V, [\cdot, \cdot]_V)$ and $(U, [\cdot, \cdot]_U)$ be two Lie algebras. A Lie algebra homomorphism is a map $\rho : V \to U$ that satisfies $\rho([\xi, \eta]_V) = [\rho(\xi), \rho(\eta)]_U$ for all $\xi, \eta \in V$. If the Lie homomorphism ρ is a bijection, then ρ is called a Lie algebra isomorphism, in that case, the groups are said to be isomorphic and one writes $V \simeq U$.

Definition B.8. (Subalgebra). Let V be a Lie algebra. A Lie subalgebra U of V is a nonempty subset of V that is a Lie algebra with respect to the bracket operation in V.

Another way to understand Lie bracket is via the adjoint operator.

Definition B.9. (Adjoint operator). Let V be a Lie algebra and $\xi \in V$. Then the adjoint operator corresponding to $\eta \in V$ is the linear map $\mathrm{ad}_\xi : V \to V$ defined by

$$\mathrm{ad}_\xi \eta = [\xi, \eta]\,.$$

The difference between the Lie bracket and the adjoint operator is that the latter fixes the first element while the second is variable.

The importance of the adjoint operator is underlined by the fact that it is a derivation, which means

$$\mathrm{ad}_x[y,z] = [\mathrm{ad}_x y, z] + [y, \mathrm{ad}_x z].$$

Another important property of the adjoint operator is that it is a representation, that is

$$[\mathrm{ad}_x, \mathrm{ad}_y](z) = \mathrm{ad}_{[x,y]}(z).$$

The adjoint operator is defined as a map from a Lie algebra into itself. It is possible to define the dual operator that operates on the dual.

Definition B.10. (Dual adjoint operator). Consider a Lie algebra V and its dual space V^*. Given $\xi \in V$, we define the dual map $\mathrm{ad}^*_\xi : V^* \to V^*$ as $\langle \mathrm{ad}^*_\xi \alpha; \eta \rangle = \langle \alpha; [\xi, \eta] \rangle$ for all $\alpha \in V^*$.

Given a basis of V, the matrix representation of the dual operator $[\mathrm{ad}^*_\xi]$ is the transpose of $[\mathrm{ad}_\xi]$.

Lie groups and Lie algebras are two different mathematical entities: the firsts are manifolds, the seconds are vector spaces. They are related by the fact that the tangent space of the identity of a Lie group is a vector space with a structure of a Lie algebra. To prove this, we have to construct a Lie bracket operation, thus, we need the notion of left-invariant vector field.

Definition B.11. (Left and right translation map). Let $(G, \star)$ be a Lie group. For $g \in G$, the left translation map is the map

$$\begin{aligned} L_g : G &\to G \\ h &\mapsto g \star h. \end{aligned}$$

In a similar way, one can define the right translation map as

$$\begin{aligned} R_g : G &\to G \\ h &\mapsto h \star g. \end{aligned}$$

The left translation map is smooth, its smooth inverse is $L_{g^{-1}}$, therefore L_g is a diffeomorphism of G. For $g \in G$, the tangent map $T_e L_g : T_e G \to T_g G$, assigns to each tangent vector v to $T_e G$ the tangent vector $T_e L_g(v) = L_{g*}(v)$ as done for a generic tangent map on a manifold (see Definition A.11).

Definition B.12. (Left-invariant function). Let G be a Lie group. A function $f : G \to \mathbb{R}$ is left-invariant if $L^*_g f = f$ for all $g \in G$, or equivalently if $f(h) = f(g \star h)$ for all $g, h \in G$.

Left-invariant functions satisfy $f(g) = f(g^{-1} \star g) = f(e)$ for all $g \in G$. Therefore they are constant functions identified by their value at the identity.

Definition B.13. (Left-invariant vector field). A vector field X on G is left-invariant if $L_g^* X = X$ for all $g \in G$, or equivalently if $X(g \star h) = T_h L_g X(h)$ for all $g, h \in G$. The set of left-invariant vector fields of G will be indicated with $\mathcal{L}(G)$.

As done for left-invariant functions, left-invariant vector fields are identified by their value at the identity, indeed $X(g) = X(g \star e) = T_e L_g(X(e))$. We will denote with ξ_L the left-invariant vector field on G defined by $\xi_L(e) = \xi \in T_e G$.

Theorem B.1. *(Left-invariant vector fields properties). Left-invariant vector fields enjoy the following properties:*

i) left-invariant vector fields are smooth;

ii) the set $\mathcal{L}(G)$ of left-invariant vector fields is a Lie subalgebra of $\mathfrak{X}(G)$;

iii) $T_e G$ and $\mathcal{L}(G)$ are isomorphic as vector spaces by the isomorphism

$$\begin{aligned} \rho : T_e G &\to \mathcal{L}(G) \\ X_e &\mapsto \rho(X_e) = (X_e)_L. \end{aligned}$$

Proof. See e.g. [7], [45]. □

By the previous results, one can show that there exists a relation between Lie groups and Lie algebras, in particular, the tangent space at the identity of a Lie is a Lie algebra.

Definition B.14. (Lie algebra of a Lie group). The Lie algebra $\mathfrak{g}$ of a Lie group G is the tangent space at the identity $T_e G$ with the bracket $[\xi, \eta] = [\xi_L, \eta_L](e)$.

The bracket just defined satisfies the Lie algebra bracket properties thanks to the anti-commutativity and Jacobi properties of the vector fields commutator.

Another important connection between a Lie group and its Lie algebra is given by the so-called "exponential map".

Definition B.15. (Exponential map). Let G be a Lie group, and $\mathfrak{g}$ its Lie algebra. The exponential map is provided by

$$\begin{aligned} \exp : \mathfrak{g} &\to G \\ \xi &\mapsto \exp(\xi) := \Phi_1^{\xi_L}(e). \end{aligned}$$

The following theorem describes some properties of the exponential map.

Theorem B.2. *(Exponential map properties). Let G be a Lie group and $\xi \in \mathfrak{g}$. Then:*

i) the integral curve $\mathbb{R} \ni t \mapsto \Phi_t^{\xi_L}(e) = \exp t\xi$ is a one-parameter subgroup of G;

ii) for all $g \in G$ and $t \in \mathbb{R}$, $\Phi_t^{\xi_L}(g) = L_g \circ \exp \xi t$ and thus, left-invariant vector fields are complete;

iii) $\exp : \mathfrak{g} \to G$ *is a local* C^∞*-diffeomorphism and* $T_0 \exp = \mathrm{id}_\mathfrak{g}$.

Example B.2. (Matrix exponential). For a matrix Lie group G the exponential map $\exp : \mathfrak{g} \to G$ coincides with the restriction to $\mathfrak{g}$ of the matrix exponential map.

Lie group action

The idea behind a group can be used to describe the transformation of a manifold. The structure of this latter can give rise to special symmetries that can be formalized through the notion of action.

Definition B.16. (Left actions). A left action of a Lie group G on a manifold M is a smooth mapping $\Phi : G \times M \to M$ such that:

(i) $\Phi(e, x) = x$ for all $x \in M$;

(ii) $\Phi(g, \Phi(h, x)) = \Phi(gh, x)$ for all $g, h \in G$ and $x \in M$;

(iii) for every $g \in G$, the map $\Phi_g : M \to M$, defined by $\Phi_g(x) := \Phi(g, x)$ is a diffeomorphism.

Similarly, for right action we have the following definition.

Definition B.17. (Right actions). A (smooth) right action of a Lie group G on a manifold M is a smooth mapping $\Phi : G \times M \to M$ satisfying the same conditions as for a left action, except that condition (ii) is replaced by:

(ii') $\Phi(g, \Phi(h, x)) = \Phi(hg, x)$ for all $g, h \in G$ and $x \in M$.

To any left-action $\Phi(g, x) = gx$ corresponds the right action $\Phi(g, x) = g^{-1}x$.

Definition B.18. (Orbits). Let G act on M. For a given point $x \in M$, we consider the equivalence condition

$$x \sim y \Leftrightarrow \exists g \in G : gx = y.$$

The equivalence class

$$Orb(x) := [x] = \{y : y \sim x\} = \{gx : g \in G\}.$$

is called the group orbit through x.

The set of all orbits is called the orbit space and is indicated by M/G.

Definition B.19. (Types of group actions). The action $\Phi : G \times M \to M$ of a group G on a manifold M is said to be:

(i) transitive if for every $x, y \in M$ there exists a $g \in G$ such that $gx = y$;

(ii) free if $gx = x$ implies $g = e$;

(iii) faithful (or effective) if for all $g \in G$ such that $g \neq e$, there exists $x \in M$ such that $gx \neq x$;

(iv) proper if, whenever the sequences $\{x_n\}$ and $\{g_n x_n\}$ converge in M, the sequence $\{g_n\}$ has a convergent subsequence in G.

Any action of G on a manifold M induces the corresponding "lifted" actions on TM and T^*M.

Definition B.20. (Tangent and cotangent lifts). Let $\Phi : G \times M \to M$ be a (left or right) action, so $\Phi_g : M \to M$ for every $g \in G$. The tangent lift of Φ is the action

$$\begin{aligned} T\Phi : G \times TM &\to TM \\ (g, (x, v)) &\mapsto T\Phi_g(x, v) = (\Phi_g(x), T_x\Phi_g(v)), \end{aligned}$$

where $x \in M$ and $v \in T_xM$.

The cotangent lift of Φ is the action

$$\begin{aligned} T^*\Phi : G \times T^*M &\to T^*M \\ (g, (x, \alpha)) &\mapsto T^*\Phi_{g^{-1}}(x, \alpha) = (\Phi_g(x), T^*_{\Phi_g(x)}\Phi_{g^{-1}}(\alpha)) \end{aligned}$$

where $\alpha \in T^*_xM$.

Definition B.21. (Action of a Lie group on its tangent and cotangent bundle). The tangent-lifted left translation of the tangent bundle of a Lie group G under its action is given by

$$\begin{aligned} &G \times TG \to TG \\ &(g, (h, v)) \mapsto (gh, gv) := (gh, T_hL_g(v)) = \left(gh, \frac{\mathrm{d}}{\mathrm{d}t}(gc(t))|_{t=0}\right) \end{aligned}$$

where $c(t)$ is any path in G with $c(0) = h$ and $c'(0) = v$, while the cotangent-lifted left translation of the tangent bundle of a Lie group G under its action is given by

$$\begin{aligned} &G \times T^*G \to T^*G \\ &(g, (h, \alpha)) \mapsto (gh, g\alpha) := (gh, T^*_{gh}L_{g^{-1}}(\alpha)) \end{aligned}$$

where

$$\langle T^*_{gh}L_{g^{-1}}(\alpha), w\rangle = \langle \alpha, T_{gh}L_{g^{-1}}(w)\rangle$$

for all $\omega \in T_{gh}G$.

Example B.3. (Matrix tangent and cotangent lifted left translations). Let G be a matrix Lie group that acts by left translation into itself. Let $R \in G$ and $(A, \dot{A}) \in TG$ and consider the curve $C(t)$ such that $C(0) = A$ and $C'(0) = \dot{A}$, then the matrix tangent lifted left translation is given by

$$\begin{aligned} G \times TG &\to TG \\ (R, (A, \dot{A})) &\mapsto (RA, R\dot{A}), \end{aligned}$$

while the matrix cotangent lifted left translations is

$$\begin{aligned} G \times T^*G &\to T^*G \\ (R,(A,P)) &\mapsto (RA, R^{-T}P) \end{aligned}$$

for every $P \in T^*_A G$.

Thanks to the tangent lifted left translation action, it is possible to construct a diffeomorphism that relates the tangent bundle with the cartesian product $G \times \mathfrak{g}$.

Definition B.22. (Left trivialization map). Let G be a Lie group acting on TG by tangent lifted left translation. Then, the left trivialization map is defined as

$$\begin{aligned} \lambda : TG &\to G \times \mathfrak{g} \\ (h,\dot{h}) &\mapsto (h, h^{-1}\dot{h}) = (h, T_h L_{h^{-1}}\dot{h}). \end{aligned}$$

The inverse of the left trivialization map produces $\lambda^{-1}(h,\xi) = (h, h\xi)$ and since are both smooth, λ is a diffeomorphism.

The left trivialization map lets to recognize some important diffeomorphisms that characterized the tangent bundle of a Lie group.

Proposition B.1. *Let G act on TG by left lifted left multiplication, then we have the following diffeomorphisms*

$$TG/G \simeq (G \times \mathfrak{g})/G \simeq \mathfrak{g} \tag{B.1}$$

$$[(h,\dot{h})] \mapsto [(h, h^{-1}\dot{h})] \mapsto h^{-1}\dot{h} \tag{B.2}$$

where $[(\cdot,\cdot)]$ denotes the orbit.

Proof. (see [15])

□

Appendix C

Proof of the second-order minimum energy filter on Lie groups

In this appendix, we recall the proof of Theorem 2.1 as done in [35]. The proofs of the Lemmas used are not reported and can be found in [35].

The optimal estimation problem

The optimal control problem defined by the deterministic system (2.27) measurement equation (2.28) with the energy cost functional (2.29)-(2.31) is built in such a way that the external input $u(\tau)$ and the measure $\mathrm{y}(\tau)$ known for $\tau \in [t_0, t]$. Substituting $\varepsilon(\tau) = D^{-1}(y(\tau) - h(g(\tau), \tau))$ into (2.29) we can rewrite the cost function without the dependence on measurement errors:

$$\min_{(g(\cdot),\delta(\cdot))} m(g(t_0), t, t_0) + \int_{t_0}^{t} \ell(\delta(\tau), D^{-1}(\mathrm{y}(\tau) - h(g(\tau), \tau)), t, \tau) d\tau. \tag{C.1}$$

The "control" input is the model error δ.

We denote by $V(g, t)$ the minimum energy value among all trajectories of (2.27) within the interval $[t_0, t]$ that reach the state $g \in G$ at time t. The optimal estimate $\widehat{g}(t)$ is therefore equal to

$$\begin{aligned} \widehat{g}(t) &= g^*_{[t_0,t]}(t) = \arg\min_{g \in G} V(g, t) \\ V(g, t_0) &= m(g, t_0, t_0) \end{aligned} \tag{C.2}$$

As stated in [26], the key observation is that if we assume $V(g, t)$ to be differentiable in a neighborhood of the optimal estimate $\widehat{g}(t)$ then, as $V(g, t)$ attains its minimum at $\widehat{g}(t)$, we must have

$$d_1 V(\widehat{g}(t), t) \equiv 0 \tag{C.3}$$

(compare with (2.15)).

The optimal Hamiltonian

In order to write the Hamilton-Jacobi-Bellman equation associated to the optimal filter, it is necessary to derive the optimal Hamiltonian. Given the incremental cost (2.31) and the dynamics

$$\dot{g}(t) = g(t)\left[\lambda(g(t), u(t), t) + B\delta(t)\right], \tag{C.4}$$

the (time-varying) Hamiltonian: $\widetilde{H} : T^*G \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ is given by

$$\widetilde{H}(p, \delta, t) := \frac{1}{2} e^{-\alpha(t-t_0)}(\mathcal{R}(\delta) + \mathcal{Q}(D^{-1}(\mathrm{y}(t) - h(g,t)))) - \langle p, g(\lambda(g, u(t), t) + B\delta)\rangle \tag{C.5}$$

where $g \in G$ is the base point of $p \in T^*_g G$. Since the optimal control problem aims to minimize the initial starting point incorporated by (2.30), the function m can be thought of as a terminal cost and the minimum energy $V(g,t)$ as a cost-to-go. This justified the presence of the minus sign in (C.5).

The advantage of working with Lie groups is that it is possible to identify the cotangent vector $p \in T^*_g G$, via left translation, with an element $\mu \in \mathfrak{g}$, defined as $\mu = T_e L^*_g(p)$. Thus it is possible to use $(g, \mu) \in G \times \mathfrak{g}^*$ in place of $p \in T^*G$ and consider the left-trivialized Hamiltonian $\widetilde{H} : G \times \mathfrak{g}^* \times \mathbb{R}^d \times \mathbb{R} \to \mathbb{R}$ defined as

$$\widetilde{H}(g, \mu, \delta, t) := \frac{1}{2} e^{-\alpha(t-t_0)}(\mathcal{R}(\delta) + \mathcal{Q}(D^{-1}(\mathrm{y}(t) - h(g,t)))) - \langle \mu, \lambda(g, u(t), t) + B\delta\rangle. \tag{C.6}$$

We are now ready to compute the left-trivialized optimal Hamiltonian.

Proposition C.1. *The left-trivialized optimal Hamiltonian associated to the optimal control problem* (C.4)*,* (C.5) *is:*

$$\begin{aligned} H^-(g, \mu, t) :=& \frac{1}{2} e^{-\alpha(t-t_0)} \langle \mu, B \circ R^{-1} \circ B^*(\mu)\rangle \\ &+ \frac{1}{2} e^{-\alpha(t-t_0)} \mathcal{Q}(D^{-1}(\mathrm{y}(t) - h(g,t))) - \langle \mu, \lambda(g, u(t), t)\rangle \end{aligned} \tag{C.7}$$

Proof. The vector field $g(\lambda(g, u(t), t) + B\delta)$ is linear in δ, while the incremental cost $\ell(\delta, \varepsilon, t, \tau)$ is quadratic in δ, thus:

$$\delta^{opt} = \arg\min_\delta \widetilde{H}^-(g, \mu, \delta, t) = e^{\alpha(t-t_0)} R^{-1} \circ B^*(\mu).$$

Substituting in (C.6) the result follows. □

The left-trivialized HJB equation and the structure of the optimal filter

The Hamilton-Jacobi-Bellman equation associated to optimal control problem (C.1), (C.4) is:

$$\begin{aligned} &\frac{\partial}{\partial t} V(g,t) - H(d_1 V(g,t), t) = 0 \\ &V(g, t_0) = m(g, t_0, t_0) \end{aligned} \tag{C.8}$$

where the optimal Hamiltonian H is defined on $T^*G \times \mathbb{R}$. Thanks to the left-trivialized formulation (C.7), one can obtain the left-trivialized Hamilton-Jacobi-Bellman equation where the left-trivialized optimal Hamiltonian is defined on $G \times \mathfrak{g}^* \times \mathbb{R}$:

$$\frac{\partial}{\partial t} V(g,t) - H^-(g, \mu(g,t), t) = 0 \tag{C.9}$$

where $\mu : G \times \mathbb{R} \to \mathfrak{g}^*$ is the "left-trivialized differential":

$$\mu(g,t) := T_e L_g^*(d_1 V(g,t)). \tag{C.10}$$

The minimum energy estimator $\widehat{g}(t)$ minimizes the value function $V(g,t)$:

$$\widehat{g}(t) = \arg\min_{g \in G} V(g,t). \tag{C.11}$$

The key observation (compare with (2.15)) is that, assuming the differentiability of V, we obtain the necessary condition

$$d_1 V(\widehat{g}(t), t) = 0, \tag{C.12}$$

or equivalently

$$\mu(\widehat{g}(t), t) = 0. \tag{C.13}$$

Lemma C.1. *Given $f : G \times \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to G$ then $d_1 f(g(t), t) = 0$ for all t implies*

$$\mathrm{Hess}_1 f(g(t),t)\,(\dot{g}(t)) + d_1 \left(\frac{\partial}{\partial t} f\right)(g(t),t) = 0.$$

From Lemma C.1 the time differentiation of (C.12) vanishes:

$$\mathrm{Hess}_1 V(\widehat{g}(t),t)\left(\dot{\widehat{g}}(t)\right) + d_1 \left(\frac{\partial}{\partial t} V\right)(\widehat{g}(t),t) = 0, \tag{C.14}$$

here $\mathrm{Hess}_1 V(\widehat{g}(t),t) : T_{\widehat{g}(t)} G \to T^*_{\widehat{g}(t)} G$ is the Hessian operator. By applying the chain rule to the second term:

$$\begin{aligned} d_1 \left(\frac{\partial}{\partial t} V\right)(g,t) &= d_1 \Big(H^-(g, \mu(g,t), t) \Big) \\ &= d_1 H^-(g, \mu(g,t), t) + d_2 H^-(g, \mu(g,t), t) \circ d_1 \mu(g,t). \end{aligned} \tag{C.15}$$

Evaluating in $g = \widehat{g}(t)$, from (C.13) we obtain

$$d_1 \left(\frac{\partial}{\partial t} V\right)(\widehat{g}(t),t) = d_1 H^-(\widehat{g}(t), 0, t) + d_2 H^-(\widehat{g}(t), 0, t), t) \circ d_1 \mu(\widehat{g}(t), t). \tag{C.16}$$

The term $d_1 \mu(\widehat{g}(t), t)$ can be expanded through the following Lemma.

Lemma C.2. *Given $f : G \to \mathbb{R}$ then the derivative of left trivialized differential $T_e L_g^*(df(g))$ is given by*

$$d(T_e L_g^*(df(g))) = T_e L_g^* \circ \mathrm{Hess} f(g) + \omega^{*\leftrightharpoons}_{T_e L_g^*(df(g))} \circ T_g L_{g^{-1}}$$

which equals $T_e L_g^* \circ \mathrm{Hess} f(g)$ whenever df(g)=0.

By Lemma C.2, differentiating (C.10), we get

$$d_1\mu(g,t) = T_eL_g^* \circ \mathrm{Hess}_1 V(g,t) + \omega^{*\leftrightharpoons}_{T_eL_g^*(d_1V(g,t))} \circ T_gL_{g^{-1}} \tag{C.17}$$

$$= T_eL^*_{\widehat{g}(t)} \circ \mathrm{Hess}_1 V(\widehat{g}(t),t), \tag{C.18}$$

the second addend on the right-hand side vanishes thanks to (C.12).

For the first term of the right-hand side of (C.15) we have

$$d_1H^-(\widehat{g}(t),0,t) \circ T_eL_{\widehat{g}(t)} = T_eL^*_{\widehat{g}(t)}(d_1H^-(\widehat{g}(t),0,t)). \tag{C.19}$$

We define the left-trivialized Hessian operator as $Z(g,t) : \mathfrak{g} \to \mathfrak{g}^*$

$$Z(g,t) := T_eL_g^* \circ \mathrm{Hess}_1 V(g,t) \circ T_eL_g. \tag{C.20}$$

With this new formulation of the Hessian operator, we can rewrite the second term of the right-hand side of equation (C.16) as

$$\begin{aligned} d_2H^-(\widehat{g}(t),0,t) &\circ d_1\mu(\widehat{g}(t),t) \circ T_eL_{\widehat{g}(t)} \\ &= d_2H^-(\widehat{g}(t),0,t) \circ T_eL^*_{\widehat{g}(t)} \circ \mathrm{Hess}_1 V \circ T_eL_{\widehat{g}(t)} \\ &= d_2H^-(\widehat{g}(t),0,t) \circ Z(\widehat{g}(t),t) \\ &= Z(\widehat{g}(t),t)^*(d_2H^-(\widehat{g}(t),0,t)) \end{aligned} \tag{C.21}$$

and the first term of (C.14) as

$$\begin{aligned} \mathrm{Hess}_1 V(\widehat{g}(t),t) &\left(\dot{\widehat{g}}(t)\right) \circ T_eL_{\widehat{g}(t)} \\ &= T_eL^*_{\widehat{g}(t)} \circ \mathrm{Hess}_1 V(\widehat{g}(t),t) \left(\dot{\widehat{g}}(t)\right) \\ &= T_eL^*_{\widehat{g}(t)} \circ \mathrm{Hess}_1 V(\widehat{g}(t),t) \circ T_eL_g \circ T_eL_{g^{-1}} \left(\dot{\widehat{g}}(t)\right) \\ &= Z(\widehat{g}(t),t) \left(\widehat{g}(t)^{-1}\dot{\widehat{g}}(t)\right). \end{aligned} \tag{C.22}$$

Substituting (C.21) and (C.19) into (C.22) we get

$$\begin{aligned} Z(\widehat{g}(t),t) &\left(\widehat{g}(t)^{-1}\dot{\widehat{g}}(t)\right) = \mathrm{Hess}_1 V(\widehat{g}(t),t) \left(\dot{\widehat{g}}(t)\right) \circ T_eL_{\widehat{g}(t)} \\ &= -d_1 \left(\frac{\partial}{\partial t} V\right)(\widehat{g}(t),t) \circ T_eL_{\widehat{g}(t)} \\ &= -T_eL^*_{\widehat{g}(t)}(d_1H^-(\widehat{g}(t),0,t)) - Z(\widehat{g}(t),t)^*(d_2H^-(\widehat{g}(t),0,t)) \end{aligned} \tag{C.23}$$

and, since $Z(\widehat{g}(t),t) = Z(\widehat{g}(t),t)^*$ (Z is symmetric), multiplying by $Z(\widehat{g}(t),t)^{-1}$ we obtain

$$\widehat{g}(t)^{-1}\dot{\widehat{g}}(t) = -d_2H^-(\widehat{g}(t),0,t) - Z(\widehat{g}(t),t)^{-1} \circ T_eL^*_{\widehat{g}(t)}(d_1H^-(\widehat{g}(t),0,t)). \tag{C.24}$$

This equation produces a differential equation for $\dot{\widehat{g}}(t)$ as a function of H^- given in (C.7).

In the following, we adopt the shorthand notation $h_t(g)$ and $\lambda_t(g,u)$ for $h(g(t),t)$ and $\lambda(g(t),u(t),t)$, respectively, and drop the explicit dependence on time of signals from our notation where convenient.

The differential with respect to μ of the optimal Hamiltonian is given by

$$d_2H^-(g,\mu,t) = -e^{\alpha(t-t_0)}B\circ R^{-1}\circ B^*(\mu) - \lambda_t(g,u). \tag{C.25}$$

To calculate the differential of the optimal Hamiltonian with respect to g, we need the following Lemma.

Lemma C.3. *Let V be a vector space, let $A : V \to V$ be linear and let $\mathcal{Q} : V \to \mathbb{R}$ be a quadratic form with associated symmetric positive definite linear map $Q : V \to V^*$. Given $f : G \to V$, then*

$$d\left(\frac{1}{2}\mathcal{Q}(A(f(g)))\right) = (A^*\circ Q\circ A(f(g)))\circ df(g) \tag{C.26}$$

$$\begin{aligned}\mathrm{Hess}\left(\frac{1}{2}Q(A(f(g)))\right) =& (df(g))^*\circ A^*\circ Q\circ A\circ df(g)\\ &+ (A^*\circ Q\circ A(f(g)))^{T_gG}\circ \mathrm{Hess}f(g).\end{aligned} \tag{C.27}$$

Applying Lemma C.3 to (C.7) (with $A = D^{-1}$, $(\mathrm{y}(t) - h(g,t)) = f(g)$) we obtain

$$d_1H^-(g,\mu,t) = -e^{\alpha(t-t_0)}\cdot((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(g)))\circ dh_t(g) - \mu\circ d_1\lambda_t(g,u). \tag{C.28}$$

Substituting $g = \widehat{g}$ and $\mu = 0$, the two expression above become

$$d_2H^-(\widehat{g},0,t) = -\lambda_t(\widehat{g},u) \tag{C.29}$$

$$d_1H^-(\widehat{g},0,t) = -e^{-\alpha(t-t_0)}\cdot\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(\widehat{g}))\big)\circ dh_t(\widehat{g}) \tag{C.30}$$

Defining $r_t(\widehat{g})\in\mathfrak{g}^*$ by

$$r_t(\widehat{g}) := T_eL^*_{\widehat{g}}\left[\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(\widehat{g}))\big)\circ dh_t(\widehat{g})\right] \tag{C.31}$$

we can rewrite (C.24) as

$$\widehat{g}(t)^{-1}\dot{\widehat{g}}(t) = \lambda_t(\widehat{g},u) + e^{-\alpha(t-t_0)}\cdot Z(\widehat{g},t)^{-1}r_t(\widehat{g}). \tag{C.32}$$

Since the integral part of (2.29) at t_0 is equal to zero, the cost functional satisfy

$$J(\delta,\varepsilon,g_0;t_0,t_0) = m(g(t_0),t_0,t_0) \tag{C.33}$$

and thus, the initial condition for the optimal filter is as in (2.33).

Approximate time evolution of Z

The solution of (C.32), together with a differential equation for $Z(\widehat{g}(t),t)$, constitutes the optimal filter. Unfortunately, such an approach is going to fail as Z satisfies an infinite-dimensional differential equation (linear dynamics with quadratic cost represents an exception). For this reason, in the following, we compute an approximation of the time evolution of $Z(\widehat{g},t)$ along the optimal solution $\widehat{g}(t)$ by neglecting the third covariant derivative of the value function V. Such an approximation is denoted by $X(g,t)$, and in order to produce a differential equation for it, we need the following Lemmas.

Lemma C.4. *Given $f : G \times \mathbb{R} \to \mathbb{R}$ and $g : \mathbb{R} \to G$ then*

$$\begin{aligned}\frac{d}{dt}\big(T_eL^*_{g(t)}\circ\mathrm{Hess}_1 f(g(t),t)\circ T_eL_{g(t)}\big) =& \\ \nabla^*_{g(t)^{-1}\dot g(t)}\circ T_eL^*_{g(t)}\circ\mathrm{Hess}_1 f(g(t),t)\circ T_eL_{g(t)}& \\ + T_eL^*_{g(t)}\circ\mathrm{Hess}_1 f(g(t),t)\circ T_eL_{g(t)}\circ\nabla_{g(t)^{-1}\dot g(t)}& \\ + T_eL^*_{g(t)}\circ\frac{\partial}{\partial t}(\mathrm{Hess}_1 f)(g(t),t)\circ T_eL_{g(t)} + h.o.t.&\end{aligned} \tag{C.34}$$

Lemma C.5. *Let $X \in \mathfrak{g}$ then*

$$d(g\mapsto T_eL_g(X)) = T_eL_g\circ\omega^{\leftrightharpoons}_X\circ T_gL_{g^{-1}}. \tag{C.35}$$

We are now ready to provide a differential equation for $X(g,t)$.

Proposition C.2. *$X(t) := X(\widehat{g}(t),t) \in \mathcal{L}(\mathfrak{g},\mathfrak{g}^*)$ fulfills the operator Riccati equation*

$$\dot X = e^{-\alpha(t-t_0)}\cdot S - F^*\circ X - X\circ F - e^{\alpha(t-t_0)}\cdot X\circ B\circ R^{-1}\circ B^*\circ X \tag{C.36}$$

with

$$X(t_0) = X_0 = T_eL^*_{\widehat{g}_0}\circ\mathrm{Hess}\, m_0(\widehat{g}_0)\circ T_eL_{\widehat{g}_0} \tag{C.37}$$

$$F(t) = -\omega_{\widehat{g}^{-1}\dot{\widehat{g}}} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} + d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} \tag{C.38}$$

$$\begin{aligned}S(t) =& -T_eL^*_{\widehat{g}}\circ[\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(\widehat{g}))\big)^{T_{\widehat{g}}G}\circ\mathrm{Hess}\, h_t(\widehat{g})& \\ & -(dh_t(\widehat{g}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ dh_t(\widehat{g})]\circ T_eL_{\widehat{g}}&\end{aligned} \tag{C.39}$$

$$\widehat{g}_0 = \arg\min_{g\in G} m_0(g). \tag{C.40}$$

Proof. Using Lemma C.4 we get

$$\begin{aligned}\frac{d}{dt}Z(\widehat{g}(t),t) =& \frac{d}{dt}\left(T_eL^*_{\widehat{g}(t)}\circ\mathrm{Hess}_1 V(\widehat{g}(t),t)\circ T_eL_{\widehat{g}(t)}\right)& \\ =& \omega^*_{\widehat{g}^{-1}\dot{\widehat{g}}}\circ Z(\widehat{g}(t),t) + Z(\widehat{g}(t),t)\circ\omega_{\widehat{g}^{-1}\dot{\widehat{g}}}& \\ & + T_eL^*_{\widehat{g}(t)}\circ\frac{\partial}{\partial t}(\mathrm{Hess}_1 V)(\widehat{g}(t),t)\circ T_eL_{\widehat{g}(t)} + h.o.t.&\end{aligned} \tag{C.41}$$

To compute the partial derivative of $\mathrm{Hess}_1 V$ we will exploit the fact that it commutes with covariant differentiation on G and then apply equation (C.15). We rewrite equation (C.15) as

$$d_1\left(\frac{\partial}{\partial t}V\right)(g,t) = d_1H^-(g,\mu(g,t),t) + (d_1\mu(g,t))^*(d_2H^-(g,\mu(g,t),t)) \quad \text{(C.42)}$$

and consider the dual of equation (C.17)

$$(d_1\mu(g,t))^*(W) = (\mathrm{Hess}_1 V(g,t))^* \circ T_eL_g(W) + T_eL^*_{g^{-1}} \circ \omega^{\rightleftharpoons *}_W(T_eL^*_g(d_1V(g,t))) \quad \text{(C.43)}$$

Combining this with (C.42) we conclude that

$$\begin{aligned} &d_1\left(\frac{\partial}{\partial t}V\right)(g,t) = \\ &\quad d_1H^-(g,\mu(g,t),t) + (\mathrm{Hess}_1 V(g,t))^* \circ T_eL_g(d_2H^-(g,\mu(g,t),t)) \\ &\quad T_gL^*_{g^{-1}} \circ \omega^{\rightleftharpoons *}_{d_2H^-(g,\mu(g,t),t),t)}(T_eL_g(d_1V(g,t))). \end{aligned} \quad \text{(C.44)}$$

Then, using the chain rule and Lemma C.5

$$\begin{aligned} \frac{\partial}{\partial t}(\mathrm{Hess}_1 V)(\widehat{g}(t),t) =& \mathrm{Hess}_1\left(\frac{\partial}{\partial t}V\right)(\widehat{g}(t),t) \\ =& \mathrm{Hess}_1 H^-(\widehat{g},0,t) + d_2(d_1H^-)(\widehat{g},0,t) \circ d_1\mu(\widehat{g},t) \\ &+ \mathrm{Hess}_1 V(\widehat{g},t) \circ T_eL_{\widehat{g}} \circ \omega^{\rightleftharpoons}_{d_2H^-(\widehat{g},0,t)} \circ T_{\widehat{g}}L_{\widehat{g}^{-1}} \\ &+ \mathrm{Hess}_1 V(\widehat{g},t) \circ T_eL_{\widehat{g}} \circ d_1(d_2H^-)(\widehat{g},0,t) \\ &+ \mathrm{Hess}_1 V(\widehat{g},t) \circ T_eL_{\widehat{g}} \circ \mathrm{Hess}_2 H^-(\widehat{g},0,t) \circ d_1\mu(\widehat{g},t) \\ &+ T_{\widehat{g}}L^*_{\widehat{g}^{-1}} \circ \omega^{\rightleftharpoons *}_{d_2H^-(\widehat{g},0,t)} \circ T_eL^*_{\widehat{g}} \circ \mathrm{Hess}_1 V(\widehat{g},t) + h.o.t. \end{aligned} \quad \text{(C.45)}$$

Here we have used (C.13) and the fact that the Hessian operator at a critical point is symmetric. From (C.41), (C.45) and (C.18) we obtain

$$\begin{aligned} &\frac{d}{dt}Z(\widehat{g}(t),t) \\ &\approx \omega^*_{\widehat{g}^{-1}\dot{\widehat{g}}} \circ Z(\widehat{g},t) + Z(\widehat{g},t) \circ \omega_{\widehat{g}^{-1}\dot{\widehat{g}}} \\ &\quad + T_eL^*_{\widehat{g}} \circ \mathrm{Hess}_1 H^-(\widehat{g},0,t) \circ T_eL_{\widehat{g}} + T_eL^*_{\widehat{g}} \circ d_2(d_1H^-)(\widehat{g},0,t) \circ Z(\widehat{g},t) \\ &\quad + \omega^{\rightleftharpoons *}_{d_2H^-(g,0,t)} \circ Z(\widehat{g},t) + Z(\widehat{g},t) \circ \omega^{\rightleftharpoons}_{d_2H^-(g,0,t)} \\ &\quad + Z(\widehat{g},t) \circ d_1(d_2H^-)(\widehat{g},0,t) \circ T_eL_{\widehat{g}} + Z(\widehat{g},t) \circ \mathrm{Hess}_2 H^-(\widehat{g},0,t) \circ Z(\widehat{g},t). \end{aligned} \quad \text{(C.46)}$$

Differentiating (C.28) and using Lemma C.3 the differential of the left-trivialized optimal Hamiltonian becomes

$$\begin{aligned} &\mathrm{Hess}_1 H^-(\widehat{g}(t),0,t) \\ &\quad = -e^{-\alpha(t-t_0)} \cdot \left((D^{-1})^* \circ Q \circ D^{-1}(\mathrm{y} - h_t(\widehat{g}))\right)^{T_{\widehat{g}}G} \circ \mathrm{Hess}\, h_t(\widehat{g}) \\ &\qquad + e^{-\alpha(t-t_0)} \cdot (dh_t(\widehat{g}))^* \circ (D^{-1})^* \circ Q \circ D^{-1} \circ dh_t(\widehat{g}) \end{aligned} \quad \text{(C.47)}$$

From (C.28) and (C.25)

$$\begin{aligned} d_2(d_1H^-)(\widehat{g}(t),0,t) =&(d_1(d_2H^-))(\widehat{g}(t),0,t)^* \\ =&-(d_1\lambda(\widehat{g}(t),u(t),t))^* \end{aligned} \tag{C.48}$$

and differentiating (C.25) yields

$$\mathrm{Hess}_2H^-(\widehat{g}(t),0,t) = -e^{\alpha(t-t_0)}\cdot B\circ R^{-1}\circ B^*. \tag{C.49}$$

Substituting (C.47), (C.48) and (C.49) into (C.46) we obtain

$$\begin{aligned} \frac{d}{dt}Z(\widehat{g}(t),t) \approx& e^{-\alpha(t-t_0)}\cdot S - F^*\circ Z(\widehat{g},t) - Z(\widehat{g},t)\circ F \\ &- e^{\alpha(t-t_0)}\cdot Z(\widehat{g},t)\circ B\circ R^{-1}\circ B^*\circ Z(\widehat{g},t) \end{aligned} \tag{C.50}$$

with

$$F(t) = -\omega_{\widehat{g}^{-1}\dot{\widehat{g}}} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} + d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} \tag{C.51}$$

$$\begin{aligned} S(t) =& -T_eL^*_{\widehat{g}}\circ[\big((D^{-1})^*\circ Q\circ D^{-1}(\mathrm{y}-h_t(\widehat{g}))\big)^{T_{\widehat{g}}G}\circ \mathrm{Hess}\, h_t(\widehat{g}) \\ &-(dh_t(\widehat{g}))^*\circ(D^{-1})^*\circ Q\circ D^{-1}\circ dh_t(\widehat{g})]\circ T_eL_{\widehat{g}} \end{aligned} \tag{C.52}$$

□

To conclude we define $K(t) := e^{-\alpha(t-t_0)}X^{-1}(t)$ where X satisfies (C.36). The time derivative of K satisfies

$$\dot{K}(t) = -\alpha K(t) + K(t)\circ S\circ K(t) + K(t)\circ F^* + F\circ K(t) + B\circ R^{-1}\circ B^* \tag{C.53}$$

From the definition of torsion

$$\begin{aligned} T_{\lambda_t(\widehat{g},u)}(X) =&\omega_{\lambda_t(\widehat{g},u)}X - \omega_X\lambda_t(\widehat{g},u) - \mathrm{ad}_{\lambda_t(\widehat{g},u)}X \\ =&\omega_{\lambda_t(\widehat{g},u)}X - \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)}X - \mathrm{ad}_{\lambda_t(\widehat{g},u)}X \end{aligned} \tag{C.54}$$

We observe that we can rewrite the operator F as

$$\begin{aligned} F(t) =& -\omega_{\widehat{g}^{-1}\dot{\widehat{g}}} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} + d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} \\ =& -\omega_{\lambda_t(\widehat{g},u)+Kr} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} + d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} \\ =&d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} - \omega_{\lambda_t(\widehat{g},u)} - \omega_{Kr} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} \\ =&d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} - \omega_{Kr} - \mathrm{ad}_{\lambda_t(\widehat{g},u)} - \omega_{\lambda_t(\widehat{g},u)} + \omega^{\leftrightharpoons}_{\lambda_t(\widehat{g},u)} + \mathrm{ad}_{\lambda_t(\widehat{g},u)} \\ =&d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} - \mathrm{ad}_{\lambda_t(\widehat{g},u)} - T_{\lambda_t(\widehat{g},u)} - \omega_{Kr} \end{aligned} \tag{C.55}$$

Defining

$$\begin{aligned} A(t) :=&d_1\lambda_t(\widehat{g},u)\circ T_eL_{\widehat{g}} - \mathrm{ad}_{\lambda_t(\widehat{g},u)} - T_{\lambda_t(\widehat{g},u)} \\ E(t) :=&S(t) \end{aligned} \tag{C.56}$$

we finally obtain the operators for the second-order optimal filter.

Bibliography

[1] B. D. O. Anderson and J. B. Moore. *Optimal filtering*. Courier Corporation, 2012.

[2] M. S. Arulampalam et al. "A tutorial on particle filters for online nonlinear/ non-Gaussian Bayesian tracking". In: *IEEE Transactions on signal processing* 50.2 (2002), pp. 174–188.

[3] A. M. Bloch. *Nonholonomic Mechanics and Control*. Vol. 24. Springer, 2015.

[4] A. M. Bloch, J. E. Marsden, and D. V. Zenkov. "Quasivelocities and symmetries in non-holonomic systems". In: *Dynamical systems* 24.2 (2009), pp. 187–222.

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

[6] A. Bravo-Doddoli and L. C. Garcia-Naranjo. "The dynamics of an articulated n-trailer vehicle". In: *Regular and Chaotic Dynamics* 20.5 (2015), pp. 497–517.

[7] F. Bullo and A. D. Lewis. *Geometric control of mechanical systems: modeling, analysis, and design for simple mechanical control systems*. Vol. 49. Springer Science & Business Media, 2004.

[8] W. Chen, S. Chern, and K. S. Lam. *Lectures on differential geometry*. Vol. 1. World Scientific Publishing Company, 1999.

[9] A. Cogliati and P. Mastrolia. "Cartan, Schouten and the search for connection". In: *Historia Mathematica* 45.1 (2018), pp. 39–74.

[10] J. David and P. V. Manivannan. "Control of truck-trailer mobile robots: a survey". In: *Intelligent Service Robotics* 7 (2014), pp. 245–258.

[11] M. P. Do Carmo and J. Flaherty Francis. *Riemannian geometry*. Vol. 6. Springer, 1992.

[12] J. M. Esposito and M. Graves. "An algorithm to identify docking locations for autonomous surface vessels from 3-D LiDAR scans". In: *2014 IEEE International Conference on Technologies for Practical Robot Applications (TePRA)*. IEEE. 2014, pp. 1–6.

[13] J. Q. Gallier and J. Quaintance. *Differential geometry and lie groups*. Vol. 12. Springer, 2020.

[14] J. Grabowski et al. “Nonholonomic constraints: a new viewpoint”. In: *Journal of mathematical physics* 50.1 (2009), p. 013520.

[15] D. D. Holm, T. Schmah, and C. Stoica. *Geometric mechanics and symmetry: from finite to infinite dimensions*. Vol. 12. Oxford University Press, 2009.

[16] A. Isidori. *Nonlinear control systems*. Springer Science & Business Media, 2013.

[17] Andrew H. Jazwinski. *Stochastic processes and filtering theory*. Courier Corporation, 2007.

[18] S. J. Julier and J. K. Uhlmann. “Reduced sigma point filters for the propagation of means and covariances through nonlinear transformations”. In: *Proceedings of the 2002 American Control Conference*. Vol. 2. IEEE. 2002, pp. 887–892.

[19] R. Kalman. “A new approach to linear filtering and prediction problems”. In: *ASME–Journal of Basic Engineering* 82.1 (Mar. 1960), pp. 35–45.

[20] A. Khanpoor, A. K. Khalaji, and S. A. A. Moosavian. “Modeling and control of an underactuated tractor–trailer wheeled mobile robot”. In: *Robotica* 35.12 (2017), pp. 2297–2318.

[21] A. W. Knapp and A. W. Knapp. *Lie groups beyond an introduction*. Vol. 140. Springer, 1996.

[22] S. Kobayashi and K. Nomizu. *Foundations of differential geometry*. Vol. 1. 2. New York, London, 1963.

[23] O. Ljungqvist et al. “A Path Planning and Path-following Control Framework for a General 2-trailer with a Car-like Tractor”. In: *Journal of field robotics* 36.8 (2019), pp. 1345–1377.

[24] R. Mahony, V. Kumar, and P. Corke. “Multirotor aerial vehicles: Modeling, estimation, and control of quadrotor”. In: *IEEE Robotics and Automation magazine* 19.3 (2012), pp. 20–32.

[25] J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry: a basic exposition of classical mechanical systems*. Vol. 17. Springer Science & Business Media, 2013.

[26] R. E. Mortensen. “Maximum-likelihood recursive nonlinear filtering”. In: *Journal of Optimization Theory and Applications* 2.6 (1968), pp. 386–394.

[27] K. Nomizu. “Invariant affine connections on homogeneous spaces”. In: *American Journal of Mathematics* 76.1 (1954), pp. 33–65.

[28] B. Øksendal. *Stochastic differential equations*. Springer, 2003.

[29] J. M. Osborne and D. V. Zenkov. “Steering the Chaplygin sleigh by a moving mass”. In: *Proceedings of the 44th IEEE Conference on Decision and Control*. IEEE. 2005, pp. 1114–1118.

[30] M. M. Postnikov. *Geometry VI: Riemannian Geometry*. Vol. 91. Springer Science & Business Media, 2013.

[31] R. Quilez et al. "Docking autonomous robots in passive docks with Infrared sensors and QR codes". In: *EAI Endorsed Transactions on Self-Adaptive Systems* 1.2 (2015).

[32] D. Rigo et al. "Second-Order-Optimal Filter on Lie Groups for Planar Rigid Bodies". In: *IEEE Transactions on Automatic Control* 67.9 (2022), pp. 4971–4977. DOI: 10.1109/TAC.2022.3179988.

[33] A. Saccon, A. P. Aguiar, and J. Hauser. "Lie group projection operator approach: Optimal control on TSO(3)". In: *2011 50th IEEE Conference on Decision and Control and European Control Conference*. IEEE. 2011, pp. 6973–6978.

[34] A. Saccon, J. Hauser, and A. P. Aguiar. "Optimal control on Lie groups: The projection operator approach". In: *IEEE Transactions on Automatic Control* 58.9 (2013), pp. 2230–2245.

[35] A. Saccon et al. "Second-order-optimal minimum-energy filters on Lie groups". In: *IEEE Transactions on Automatic Control* 61.10 (2015), pp. 2906–2919.

[36] C. Samson. "Control of chained systems application to path following and time-varying point-stabilization of mobile robots". In: *IEEE transactions on Automatic Control* 40.1 (1995), pp. 64–77.

[37] N. Sansonetto and M. Zoppello. "On the trajectory generation of the hydrodynamic Chaplygin sleigh". In: *IEEE Control Systems Letters* 4.4 (2020), pp. 922–927.

[38] D. H. Sattinger and O. L. Weaver. *Lie groups and algebras with applications to physics, geometry, and mechanics*. Vol. 61. Springer Science & Business Media, 2013.

[39] L. Shen and P. R. Stopher. "Review of GPS travel survey and GPS data-processing methods". In: *Transport reviews* 34.3 (2014), pp. 316–334.

[40] C. Spagnol et al. "Trajectory reconstruction by integration of GPS and a swarm of MEMS accelerometers: model and analysis of observability". In: *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems*. IEEE. 2004, pp. 64–69.

[41] S. Ç. Tekkök et al. "A Novel Docking Algorithm Based On The LiDAR And The V-shape Features". In: *Avrupa Bilim ve Teknoloji Dergisi* 26 (2021), pp. 35–40.

[42] S. Thrun, W. Burgard, and D. Fox. "Probalistic robotics". In: *Kybernetes* (2006).

[43] D. Tilbury, R. M. Murray, and S. S. Sastry. "Trajectory generation for the n-trailer problem using Goursat normal form". In: *IEEE Transactions on Automatic Control* 40.5 (1995), pp. 802–819.

[44] V. S. Varadarajan. *Lie groups, Lie algebras, and their representations*. Vol. 102. Springer Science & Business Media, 2013.

[45] F. W. Warner. *Foundations of differentiable manifolds and Lie groups*. Vol. 94. Springer Science & Business Media, 1983.

[46] D. V. Zenkov. "On Hamel's equations". In: *Theoretical and Applied Mechanics* 43.2 (2016), pp. 191–220.