Observers for Kinematic Systems with Symmetry ⋆

Robert Mahony * **Jochen Trumpf** * **Tarek Hamel** **

* *Research School of Engineering, Australian National University, Canberra ACT 2601, Australia. e-mail:* `Robert.Mahony@anu.edu.au`, `Jochen.Trumpf@anu.edu.au` .
** *T. Hamel is with I3S UNSA-CNRS, Nice-Sophia Antipolis, France. e-mail:* `thamel@i3s.unice.fr`.

Abstract: This paper considers the design of nonlinear state observers for finite-dimensional equivariant kinematics of mechanical systems. The observer design problem is approached by lifting the system kinematics onto the symmetry group and designing an observer for the lifted system. Two particular classes of lifted systems are identified, which we term type I and type II systems, that correspond to common configurations of sensor suites for mobile robotics applications. We consider type I systems in detail and define an error signal on the symmetry group using the group structure. We propose an observer structure with a pre-observer or internal model augmented by an equivariant innovation term that leads to autonomous error evolution. A control Lyapunov function construction is used to design the observer innovation that both ensures the required equivariance, and leads to strong convergence properties of the observer error dynamics.

1. INTRODUCTION

Systems on Lie groups and their homogeneous spaces have been studied extensively since the early 1970s, starting with the work of Brockett [1972, 1973] and Jurdjevic and Sussmann [1972]. Brockett's work was motivated by analytical mechanics and the study of mechanical systems, see Brockett [1977]. The geometric description of mechanical systems naturally leads to system models on differentiable state manifolds that are acted upon by a Lie group, the symmetry group of the mechanical system. Control theory for mechanical systems with symmetry is now a mature subject, and several textbooks on this topic are available, e.g. Jurdjevic [1997], Bloch [2003], Agrachev and Sachkov [2004] and Bullo and Lewis [2004]. Most of the classical literature on system theory for mechanical systems is focused on structure theory and control. Observers that are specifically targeted at systems on Lie groups and their homogeneous spaces, as opposed to more general nonlinear systems, appear to have only been studied in the last ten years, or so. We will survey this literature in more detail below.

The structural question of observability has, however, been studied in a series of papers starting with Brockett [1972], with a more or less complete characterization given in Cheng et al. [1990]. Apostolou and Kazakos [1996] show how the resulting local observability criterion can alternatively be derived using observability codistributions, thus providing a link to the classical observabilty theory for general nonlinear systems, see e.g. Nijmeijer and van der Schaft [1990] or Isidori [1999]. We will not study observability questions in this paper.

We will study the observer problem for mechanical systems with symmetry in a purely deterministic setup. Our work is partly motivated by the need for highly robust and computationally simple state estimation algorithms for robotic vehicles. The classical approach to state estimation for such applications is based on nonlinear filtering techniques such as extended Kalman filters (Anderson and Moore [1979]) unscented filters (Julier and Uhlmann [1997]) or particle filters (Doucet et al. [2001]). Nonlinear observers offer less information than a nonlinear filter, namely state estimates rather than full *posterior* distributions for the state, however, it is often possible to prove strong stability results with large or almost global basins of attraction and to provide computationally simple implementations of the observers.

Another promising approach to state estimation for robotic vehicles is a deterministic optimality based approach known as deterministic filtering or minimum-energy filtering due to Mortensen [1968]. In fact, it has recently been shown by Zamani et al. [2013] that a second order minimum-energy filter for attitude kinematics provides a geometric correction to the Multiplicative Extended Kalman Filter (MEKF), a state-of-the art stochastic filtering algorithm for attitude estimation, cf. Crassidis et al. [2007]. A generalization of minimum-energy filters to arbitrary mechanical systems with symmetry is still work in progress but the resulting filtering algorithms share many of the advantages and disadvantages of stochastic filters.

One of the earliest applied results concerned the design of a nonlinear observer for attitude estimation of a rigid-body using the unit quaternion representation of the special orthogonal group SO(3), Salcudean [1991]. This

⋆ This research was supported by the Australian Research Council through Discovery Grant DP120100316 "Geometric observer theory for mechanical control systems" and by the SCAR ANR-ASTRID fund.

work is seminal to a series of papers undertaken over the last fifteen years that develop nonlinear attitude observers for rigid-body dynamics; Nijmeijer and Fossen [1999], Thienel and Sanner [2003], Mahony et al. [2005], Bonnabel et al. [2006], Campolo et al. [2006], Maithripala et al. [2006], Metni et al. [2006], Kinsey and Whitcomb [2007], Martin and Salaün [2007], Tayebi et al. [2007], Mahony et al. [2008], Vasconcelos et al. [2008], Brás et al. [2011], Grip et al. [2012], Sanyal and Nordkvist [2012], exploiting either the unit quaternion or the matrix Lie group representation of SO(3). Recent observer designs have comparable performance to state-of-the-art nonlinear filtering techniques, Crassidis et al. [2007], generally have much stronger global stability and robustness properties, Mahony et al. [2008], and are simple to implement. The full pose estimation problem has also attracted recent attention, Vik and Fossen [2001], Rehbinder and Ghosh [2003], Baldwin et al. [2009], Vasconcelos et al. [2010], in which case the underlying state space is the Special Euclidean group SE(3) comprising both attitude and translation of a rigid-body. Another promising body of applied work involves development of heading reference systems for UAV systems, Salaün and Martin [2010].

Aghannan and Rouchon [2003] first recognized the importance of invariance properties of observers for mechanical systems with symmetry. More recent work on understanding the generic structure of observers for left invariant systems on Lie groups and their homogeneous spaces, Bonnabel et al. [2008], Mahony et al. [2008], Lageman et al. [2009], has lead to an understanding of the role of invariance properties of observer designs in relation to the resulting observer error dynamics, see Bonnabel et al. [2009], Lageman et al. [2010], Trumpf et al. [2012]. The present paper contains further results in this direction.

In this paper we propose a general full state observer design methodology for a class of kinematic systems with complete symmetry. The focus on kinematic systems is natural for a range of applied problems in mobile robotics that have motivated the authors' interest in this subject. Most mobile robotic vehicles carry an inertial measurement unit (IMU), global positioning system (GPS), tachometers, and other velocity measurement systems as a matter of course. Such systems provide reliable, low noise measurements of the (inertial) velocity of the vehicle. In contrast, measurement of the exogenous force and torque input to a robotic vehicle is generally impossible. Even if force or torque signals are available they are mostly of poor quality and would degrade, rather than improve, estimates of the vehicle's state. In many real world applications, and certainly most applications in mobile robotics, it is best to rely on velocity measurements directly and use these, along with (partial) state measurements to build an observer for the vehicle state. It follows that kinematic models of the physical system are the natural structure on which to base the design of the observer.

Many physical systems, and most mobile robotic vehicles, have physical models with symmetries that encode the invariance of the laws of motion. That is, the behaviour of the system at one point in space is no different from its behaviour at another point in space, at least when viewed through a *symmetric* transformation of space. Such structure is of particular importance in the design of an observer: if the behaviour of a system with symmetry can be modeled and understood at one point in space, and this model can be transported via the symmetry to all points in space, then an observer design made at the reference point can also be transported to all points in space to obtain a global observer design. Observers designed using this principle are known as *equivariant observers* and the approach offers considerable benefits in design methodology and error stability analysis.

The technical material in the paper begins (§2) by developing a modeling framework for the kinematics of systems with complete symmetry. Much of this material is standard in the literature, however, the focus on kinematic system models changes the perspective and it is well worth covering the material again. The modeling process leads us to identify two special classes of kinematic mechanical systems, type I and type II systems, that we consider in more detail. Type I systems model the physical situation where both the velocity sensors and the state sensors are mounted on the body-fixed frame (or possibly both mounted in the inertial frame) of a mobile vehicle. Type II systems are those where the sensors are mounted in mutually opposite frames of reference; for example, velocity sensors are mounted on-board but the state measurement is provided by an external sensor system mounted in the inertial frame of reference. These two classes of systems comprise the majority of mobile robotics applications that the authors have encountered. In this paper we focus on type I systems, the situation of most interest to the authors. Although we make a number of comments about type II systems, we leave a detailed discussion to future work.

The second technical section (§3) in the paper discusses full state observer structure. We propose a structure based on a pre-observer or internal model designed to replicate the system kinematics, coupled with an innovation or error correction term. The symmetry of the system is used to define global error coordinates and we show how to construct a pre-observer that is globally synchronous with the system, that is the error between the pre-observer and the system is constant along matching trajectories for arbitrary velocity inputs. We consider only equivariant innovations, that is innovations that depend on the relative state of the system to the observer, as seen through the symmetry action. There are a number of important consequences of this structure, the most important being the autonomy of the resulting error dynamics. This is a crucial step in the design process as it is now possible to design the observer innovation in error coordinates in a way that is agnostic to the state of the system.

The next technical section (§4) tackles the observer synthesis problem using Lyapunov design principles. We begin with cost functions on the outputs that can be realized from available measurements. By imposing invariance on the costs we can lift these costs to a non-degenerate cost in the error coordinates. For an actual design problem, the simplest approach at this point is to undertake a direct Lyapunov design process and we provide an example to demonstrate how this can be done. In more generality, we show how the cost can be used to define an equivariant gradient innovation once an invariant metric on the Lie group is defined. This construction leads to a gradient flow in the error coordinates that is straightforward to analyze for stability.

A final technical section (§5) is less formal and provides the intuition and the main formulas required to extend the proposed design methodology to an observer that also estimates an unknown constant bias offset in the measured velocity.

The paper is written from a rigourous point of view (at least until §5) and full proofs of the results are provided. This tends to make the development appear more abstract than is truly the case. For all the applications that we have encountered the calculations can be made using standard matrix calculus and observer design for real world systems can be undertaken without requiring the rather daunting differential calculus that we are forced to use to derive the general results. We have provided a running example throughout the paper that demonstrates the methodology in a way that cannot be seen by just reading the theorems. In practice, the approach is simple and easily workable for a wide range of important applications and has already led to a range of highly effective observers in real world systems.

2. EQUIVARIANT KINEMATIC SYSTEMS

In this section, we consider the structure of kinematic systems with complete symmetries. Although this material is closely related to work on the modeling of mechanical systems (Marsden and Ratiu [1999], Bloch [2003], Bullo and Lewis [2004]) and understanding their symmetries, the focus on only the kinematics of the system leads to new perspectives and warrants a careful development.

Definition 1. Let $\mathcal{X}$, and $\mathcal{Y}_i$ for $i = 1, \dots, p$ be finite-dimensional smooth real manifolds that are termed, respectively, the state and output spaces. Let $\mathbb{V}$ denote a finite-dimensional real vector space that is termed the velocity space. A *kinematic* system is defined by state equations

$$\dot{x} = f(x, v), \tag{1a}$$
$$y_i = h_i(x) \tag{1b}$$

for a smooth dynamics function $f\colon \mathcal{X} \times \mathbb{V} \to T\mathcal{X}$, with $f(x, \cdot)\colon \mathbb{V} \to T_x\mathcal{X}$ a *linear* map, and smooth output maps $h_i\colon \mathcal{X} \to \mathcal{Y}_i$. △

For initial conditions $x(0)$ we will denote the solution of (1) by $x(t; x(0))$. We will assume that, given an exogenous input signal $v(t)$, there exist unique solutions on all time intervals considered. The signals that we will use for the observer construction are the (partial) state measurements $y_i(t)$ and the velocity input $v(t)$.

The structure that makes (1) a kinematic system rather than a general model of non-linear dynamics is the vector space structure of the input space $\mathbb{V}$ and the linearity of the system function $f(x, \cdot)$. This linearity in the input models the natural linear structure of velocity.

Example 1.1. A physical direction of an inertial feature (such as the magnetic field of the earth) relative to a body-fixed frame (of a robotic vehicle to which a suite of magnetometers is attached) can be modeled as a direction on the two-sphere S^2 embedded in $\mathbb{R}^3$. As the robotic vehicle rotates the physical direction of the (inertially known) magnetic field moves relative to the body-fixed frame. Such kinematics are important in attitude estimation for mobile robotic vehicles.

Given the state space $\mathcal{X} = \mathrm{S}^2 \subset \mathbb{R}^3$. The kinematics considered are

$$\dot{x} = x \times \Omega. \tag{2}$$

where $x \in \mathcal{X}$, $\Omega \in \mathbb{V} \equiv \mathbb{R}^3$ and $\times$ denotes the vector product. The output is

$$y = x$$

where $y \in \mathcal{Y} \equiv \mathrm{S}^2 \subset \mathbb{R}^3$.

The state $x \in \mathrm{S}^2$ is the direction of the inertial feature relative to the body-fixed frame and as an element of the coordinate space $\mathbb{R}^3$ is expressed in body-fixed coordinates. Note that the actual state of the vehicle is two-dimensional, while the parametrisation that we use is the embedding into $\mathbb{R}^3$, leading to a three-dimensional coordinate representation. The physical velocity of the system $f(x, \Omega) = x \times \Omega$ (an element of $T_x\mathrm{S}^2$), is the motion of the inertial feature relative to the body-fixed frame. However, this two-dimensional velocity can only be globally parameterised via a three-dimensional object Ω. Physically, Ω is the angular velocity of the body-fixed frame relative to the inertial frame. As an element of the coordinate space $\Omega \in \mathbb{V} \equiv \mathbb{R}^3$ it is expressed in body-fixed coordinates. The output is the full state $y = x$.

This example is of interest for two reasons. Firstly, it is necessary to use a three dimensional parametrisation of velocity in order to get a *global* description of the two-dimensional system kinematics in the form (1). The fact that no global two-dimensional velocity parametrisation exists is a consequence of the fact that $T\mathcal{X}$ is a non-trivial vector bundle. [1] The fact that a global three-dimensional parametrisation exists is a function of the embedding [2] $\mathcal{X} \hookrightarrow \mathbb{R}^3$ into Euclidean space. Secondly, the velocity parametrization that we have used leads to an element Ω that represents a velocity measurement physically made relative to the inertial frame, for example, using a strap down inertial measurement unit. In contrast, the output measurement (in this case a full state measurement) is physically made relative to the body-fixed frame. We will see that this final point is a key observation, and will make this system a type I system (Def. 5). △

Example 2.1. A unicycle kinematic system, typically physically realized by two parallel wheels with castors front and back to keep the vehicle from tipping, is one of the most studied non-holonomic systems in the control literature (see, e.g. Bloch [2003]). The kinematic state of the system can be represented by the position and orientation of the vehicle on a planar surface, the ground plane. Its speed and angular velocity are measured using tachometers on each driving wheel individually. In a typical robotics experiment the vehicle position (but not its orientation) is measured using an overhead camera.

This example is perhaps less compelling from an applications point of view than Example 1.1, however, it provides

[1] Brockett [1977] resolves this issue by modeling such systems as fiber bundle maps where the input fiber has the same dimension as the state manifold.

[2] According to the Whitney embedding theorem, any finite-dimensional smooth manifold may be embedded into a Euclidean space of high enough dimension.

a good demonstration of several of the principles of observer design for symmetric systems, and almost everyone in the audience will be familiar with the system.

The state-space considered is $\mathcal{X} = \mathbb{R}^2 \times \mathrm{S}^1$. The unicycle kinematics are given by

$$\dot{\xi}_1 = \cos(\theta)u, \tag{3a}$$
$$\dot{\xi}_2 = \sin(\theta)u, \tag{3b}$$
$$\dot{\theta} = q \tag{3c}$$

for $x = ((\xi_1, \xi_2)^\top, \theta) \in \mathcal{X}$ and velocity $v = (u, q) \in \mathbb{V} \equiv \mathbb{R}^2$. The output is $y \in \mathcal{Y} = \mathbb{R}^2$

$$y = (\xi_1, \xi_2)^\top.$$

The state $x = ((\xi_1, \xi_2)^\top, \theta) \in \mathbb{R}^2 \times \mathrm{S}^1$ is the position and orientation of the unicycle with respect to an inertial frame, written in inertial coordinates as an element of the coordinate space $\mathbb{R}^2 \times \mathbb{R}$. The physical velocity $f(x, v) = ((\cos(\theta)u, \sin(\theta)u)^\top, q)$ of the system is the motion of the unicycle with respect to the inertial frame, expressed in inertial coordinates. The system is non-holonomic and there is a velocity constraint that enables one to parameterise the physical velocity with two real parameters u, the scalar speed, and q the angular velocity, $(u, q) \in \mathbb{V} \equiv \mathbb{R}^2$. The output is the location of the unicycle relative to the inertial frame, expressed in inertial coordinates.

This example is of interest for two reasons. Firstly, the input space $\mathbb{V} = \mathbb{R}^2$ is lower dimensional than the tangent space $T_x\mathcal{X}$. Secondly, the velocity measurement and the output measurement are both physically made with respect to the inertial frame. We will see that this final point is a key observation, and will make this system a type II system (Def. 4). △

In local coordinates the system map of a kinematic system has the form

$$f(x, v) = \sum_{i=1}^{m} B_i(x) v_i$$

for suitable smooth functions $B_i(x)$ and where the elements $v = (v_1, \ldots, v_m)$ are associated with a basis decomposition of the vector space $\mathbb{V}$.

Let $\mathbf{G}$ be a finite-dimensional real Lie group. For arbitrary $A, B \in \mathbf{G}$, the group multiplication is denoted by AB, the group inverse by A^{-1}, and I denotes the identity element of $\mathbf{G}$. The associated Lie algebra is denoted $\mathfrak{g}$ with Lie bracket $[V, W]$ for $V, W \in \mathfrak{g}$. Define the left translation on the group by $L_A \colon \mathbf{G} \to \mathbf{G}$, $L_A B := AB$. The right translation $R_A B := BA$ is analogous. Although the results presented in this paper hold for general (finite-dimensional) real Lie groups, all the examples that we have considered involve matrix Lie groups $\mathbf{G} \subseteq \mathrm{GL}(n)$; that is closed subgroups of the general linear group of all real invertible $n \times n$ matrices. In this case the group multiplication is given by matrix multiplication, the identity element is the identity matrix and the group inverse is the matrix inverse. The associated matrix Lie algebra is denoted $\mathfrak{g} \subseteq \mathbb{R}^{n \times n}$ with Lie bracket $[V, W] = VW - WV$ given by the matrix commutator. Left and right translation L_A (resp. R_A) now have simple algebraic expressions, and in particular for $X \in \mathbf{G}$ and $U \in \mathfrak{g}$

$$\mathrm{d}R_X(I)U = UX, \quad \mathrm{d}L_X(I)U = XU$$

where $\mathrm{d}R_X(I) \colon T_I\mathbf{G} \to T_X\mathbf{G}$ is the differential of R_X at I and similarly $\mathrm{d}L_X(I) \colon T_I\mathbf{G} \to T_X\mathbf{G}$ is the differential of L_X at I. More generally $\mathrm{d}R_{X_1}(X_2) \colon T_{X_1}\mathbf{G} \to T_{X_2X_1}\mathbf{G}$, and $\mathrm{d}L_{X_1}(X_2) \colon T_{X_1}\mathbf{G} \to T_{X_1X_2}\mathbf{G}$. Where it is clear from context we will omit the base point from differentials and write $\mathrm{d}R_{X_1}$ (resp. $\mathrm{d}L_{X_1}$) rather than the more general $\mathrm{d}R_{X_1}(X_2)$ (resp. $\mathrm{d}L_{X_1}(X_2)$). Identifying $T_I\mathbf{G} \equiv \mathfrak{g}$, the adjoint representation for $X \in \mathbf{G}$ is the map $\mathrm{Ad}_X \colon \mathfrak{g} \to \mathfrak{g}$, $\mathrm{Ad}_X := \mathrm{d}L_X(X^{-1}) \circ \mathrm{d}R_{X^{-1}}(I) = \mathrm{d}R_{X^{-1}}(X) \circ \mathrm{d}L_X(I)$. The map Ad_X is a Lie algebra automorphism obtained as the derivative of the inner Lie group automorphism $L_X \circ R_{X^{-1}} = R_{X^{-1}} \circ L_X \colon \mathbf{G} \to \mathbf{G}$. Note that the left and right translation operations always commute. In a matrix Lie group we have

$$\mathrm{Ad}_X(U) = XUX^{-1}$$

for $X \in \mathbf{G}$ and $U \in \mathfrak{g}$.

A right group action ϕ of $\mathbf{G}$ on a smooth manifold $\mathcal{X}$ is a smooth mapping

$$\phi \colon \mathbf{G} \times \mathcal{X} \to \mathcal{X},$$

with $\phi(A, \phi(B, x)) = \phi(BA, x)$ and $\phi(I, x) = x$. A left group action is analogous with $\phi(A, \phi(B, x)) = \phi(AB, x)$. The symmetry and invariance structure that we will develop requires a choice of either right or left group actions. Physical system models for observer design can be more natural to model with one or the other handedness of the symmetry, depending on the nature of the sensor systems that are used and the way coordinates are chosen. Right-handed invariance is the more natural representation to analyze systems with body-fixed state sensors in the usual coordinates used for physical system modeling. Left-handed invariance is natural for systems with ground-based state sensors. Since the majority of the applications that we have considered involve body-fixed sensor systems we choose to use right invariance (the less sinister option) to develop the structure theory that we will use. Although it may be more natural to model with one type of handedness, the actual symmetry choice is an arbitrary modeling choice and can be changed by re-defining the group multiplication [3] on the symmetry group $\mathbf{G}$ and all results that we state have direct analogues in the opposite handedness.

A group action induces smooth mappings $\phi_A \colon \mathcal{X} \to \mathcal{X}$ for $A \in \mathbf{G}$ by $\phi_A(x) := \phi(A, x)$, and $\phi_x \colon \mathbf{G} \to \mathcal{X}$ for $x \in \mathcal{X}$ by $\phi_x(A) := \phi(A, x)$. The group action ϕ is termed *transitive* if ϕ_x is surjective and in this case the manifold $\mathcal{X}$ is termed a *homogeneous space* of $\mathbf{G}$ (Boothby [1986]). For a group action $\phi \colon \mathbf{G} \times \mathcal{X} \to \mathcal{X}$, the stabilizer of an element $x \in \mathcal{X}$ is given by

$$\mathrm{stab}_\phi(x) = \{A \in \mathbf{G} \mid \phi(A, x) = x\},$$

and is a subgroup of $\mathbf{G}$.

Definition 2. Consider the system (1). Consider right (resp. left) group actions $\phi \colon \mathbf{G} \times \mathcal{X} \to \mathcal{X}$, $\psi \colon \mathbf{G} \times \mathbb{V} \to \mathbb{V}$ and $\rho^i \colon \mathbf{G} \times \mathcal{Y}_i \to \mathcal{Y}_i$. The structure $(\mathbf{G}, \phi, \psi, \rho^i)$ is termed a *symmetry* of the system (1) if for all $A \in \mathbf{G}$, $x \in \mathcal{X}$ and $v \in \mathbb{V}$ one has

[3] More precisely, the operation $\odot \colon \bar{\mathbf{G}} \times \bar{\mathbf{G}} \to \bar{\mathbf{G}}$, $A \odot B := BA$ turns a copy $\bar{\mathbf{G}}$ of the *set* $\mathbf{G}$ into a group $\bar{\mathbf{G}}$ that is isomorphic to the *group* $\mathbf{G}$ via $A \mapsto A^{-1}$. Given an action ϕ of $\mathbf{G}$ on $\mathcal{X}$ of either handedness, $\bar{\phi} \colon \bar{\mathbf{G}} \times \mathcal{X} \to \mathcal{X}$, $\bar{\phi}(A, x) := \phi(A, x)$ defines an action of $\bar{\mathbf{G}}$ on $\mathcal{X}$ of the opposite handedness.

$$\mathrm{d}\phi_A(x)[f(x,v)] = f(\phi(A,x),\psi(A,v)), \tag{4a}$$
$$\rho^i(A,h_i(x)) = h_i(\phi(A,x)). \tag{4b}$$

The symmetry is termed a *complete symmetry* if $\mathcal{X}$ is a homogeneous space with respect to ϕ. A system with a complete symmetry is said to be *equivariant.* [4] △

Example 1.2. Recall the scenario described in Example 1.1 and note that the velocity structure used is naturally associated with rotation of a frame of reference attached to the robotic vehicle $\{B\}$ relative to an inertial frame $\{A\}$. Express the orthonormal frame vectors of $\{B\}$ in coordinates of $\{A\}$ to obtain an orthogonal matrix B. Let $\{C\}$ denote a rotated frame and express the orthonormal frame vectors of $\{C\}$ in coordinates of $\{A\}$ to obtain another orthogonal matrix C. The physical rotation Q of $\{B\}$ to $\{C\}$ can be written in coordinates as $Q = C^\top B$. A vector $w \in \{A\}$ is rotated by the physical rotation Q by $w \mapsto Qw$. This is the standard orthogonal matrix representations of the special orthogonal group SO(3). Note that, unlike the matrices B and C, the columns of $Q \in \mathrm{SO}(3)$ as a matrix do not carry the interpretation of coordinates of a frame of reference with respect to $\{A\}$.

We claim that SO(3) is a symmetry group for Example 1.1 with actions

$$\phi(Q,x) := Q^\top x,$$
$$\psi(Q,\Omega) := Q^\top \Omega,$$
$$\rho(Q,y) := Q^\top y,$$

for $Q \in \mathrm{SO}(3)$ as described above. It is straightforward to verify that these are right group actions;

$$\phi(Q_1,\phi(Q_2,x)) = Q_1^\top Q_2^\top x = (Q_2 Q_2)^\top x = \phi(Q_2 Q_1, x),$$

etc. It is trivial to verify that ϕ is transitive on S^2. Clearly, $\rho(Q,h(x)) = Q^\top x = h(\phi(Q,x))$ and it remains to show that the kinematics are equivariant. One has

$$\begin{aligned}\mathrm{d}\phi_Q[\dot{x}] &= Q^\top(x \times \Omega)\\ &= Q^\top x \times Q^\top \Omega = \phi(Q,x) \times \psi(Q,\Omega).\end{aligned}$$

The SO(3) symmetry expresses the physical fact that the laws of motion, in this case just the kinematics, do not depend on the orientation of the vehicle. △

Example 2.2. Recall the unicycle from Example 2.1. The special Euclidean group SE(2) is the set of rigid-body transformations of two-dimensional Euclidean space. An element of $Q \in \mathrm{SE}(2)$ is parameterized by a rotation

$$R(\alpha) = \begin{pmatrix} \cos(\alpha) & -\sin(\alpha) \\ \sin(\alpha) & \cos(\alpha) \end{pmatrix}$$

and a translation $z \in \mathbb{R}^2$. The classical homogeneous coordinates of Q are given by

$$Q = \begin{pmatrix} R(\alpha) & z \\ 0 & 1 \end{pmatrix}. \tag{5}$$

We claim that SE(2) is a symmetry group for (3) with actions

$$\phi(Q,(\xi,\theta)) := (R(\alpha)\xi + z, \alpha + \theta),$$
$$\psi(Q,v) := v,$$
$$\rho(Q,y) := R(\alpha)y + z,$$

for $Q \in \mathrm{SE}(2)$ as described above. It is straightforward to show that ϕ is a left group action by representing the state $x = (\xi,\theta)$ in homogeneous coordinates as

$$x^h = \begin{pmatrix} R(\theta) & \xi \\ 0 & 1 \end{pmatrix}$$

and noting that $\phi(Q,x)^h = Qx^h$, i.e. the action ϕ corresponds to left matrix multiplication in homogeneous coordinates. This also implies that the action ϕ is transitive on $\mathcal{X}$. The trivial group action ψ is both right and left handed, while ρ is a left action since it is just the first component of ϕ.

To show that the kinematics are equivariant we need to compute the differential $\mathrm{d}\phi_Q(x)$. For a tangent vector $(\dot{\xi},\dot{\theta}) \in T_x\mathcal{X}$ one has

$$\mathrm{d}\phi_Q(x)[(\dot{\xi},\dot{\theta})] = (R(\alpha)\dot{\xi},\dot{\theta}).$$

Thus, for $f(x,v) = ((\cos(\theta)u, \sin(\theta)u)^\top, q)$ one has

$$\begin{aligned}\mathrm{d}\phi_Q[f(x,v)] &= \left(R(\alpha)\begin{pmatrix}\cos(\theta)\\ \sin(\theta)\end{pmatrix}u, q\right)\\ &= \left(\begin{pmatrix}\cos(\alpha+\theta)\\ \sin(\alpha+\theta)\end{pmatrix}u, q\right)\\ &= f(\phi_Q(x),\psi(Q,v)).\end{aligned}$$

Note that $f(x,v)$ is independent of ξ in $x = (\xi,\theta)$, so only the angle component $\alpha+\theta$ of $\phi_Q(x)$ matters for the last step in this calculation. The output equivariance is trivial since

$$\rho(Q,h(x)) = \rho(Q,\xi) = h(\phi_Q(x))$$

by definition. The SE(2) symmetry expresses the physical fact that the kinematics of the unicycle do not depend on its pose. The left handedness of the symmetry is due to the (natural) choice of coordinates and could be turned into a right handedness by re-defining the group multiplication on $SE(2)$. The only place where this matters in the subsequent calculus is where concatenations of action maps occur. Swapping the handedness of the group multiplication will then turn left translations L_X into right translations R_X and vice versa. △

To simplify the exposition, we will from now on concentrate on the case of right handed symmetries in the theoretical development. We will occasionally point to the necessary modifications for the left handed case.

Consider the kinematics (1) and fix a point $x_0 \in \mathcal{X}$. The approach that we take to the observer design problem is to use the symmetry of a kinematic system (1) to lift to a new system on the symmetry group. To do this we will choose a *velocity lift* $F_{x_0} \colon \mathbb{V} \to \mathfrak{g}$ such that

$$d\phi_{x_0}(I)[F_{x_0}(v)] = f(x_0,v). \tag{6}$$

To see that such a map always exists, at least for the case of equivariant kinematics with a complete symmetry, let $\mathfrak{k} = \ker d\phi_{x_0}(I)$, the Lie algebra of the stabilizer $\mathrm{stab}(x_0)$. Choose a complementary subspace $\mathfrak{h} \subset \mathfrak{g}$ of $\mathfrak{k}$ in $\mathfrak{g}$ such that $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{h}$. Since $\mathcal{X}$ is a homogeneous space of $\mathbf{G}$, note that $\dim(\mathfrak{h}) = \dim(T_x\mathcal{X})$. Define a map $F^{\mathfrak{h}}_{x_0} \colon \mathbb{V} \to \mathfrak{h}$ implicitly by

$$F^{\mathfrak{h}}_{x_0}(v) \in \mathfrak{h}, \tag{7a}$$
$$d\phi_{x_0}(I)[F^{\mathfrak{h}}_{x_0}(v)] = f(x_0,v) \tag{7b}$$

for $v \in \mathbb{V}$. This construction is well defined since the linear map $d\phi_{x_0}(I)$ restricted to $\mathfrak{h}$, $d\phi_{x_0}(I)|_{\mathfrak{h}} \colon \mathfrak{h} \to T_{x_0}\mathcal{X}$, is bijective due to the transitivity of ϕ. Although this construction will always yield *some* velocity lift F_{x_0},

[4] Note that some authors use the term equivariant without requiring that $\mathbf{G}$ acts transitively on $\mathcal{X}$.

it is generally best to choose F_{x_0} carefully with some consideration of the physics of the problem.

Given a choice of F_{x_0}, define a (system) function $F\colon \mathbf{G} \times \mathbb{V} \to T\mathbf{G}$ by

$$F(X, v) := dR_X(I)[F_{x_0}(\psi(X^{-1}, v))] \in T_X\mathbf{G} \qquad (8)$$

where we associate the tangent space $T_X\mathbf{G} = dR_X(I)[\mathfrak{g}]$ for all $X \in \mathbf{G}$.[5]

Lemma 1. Consider a system (1) that is equivariant with respect to a right handed complete symmetry $(\mathbf{G}, \phi, \psi, \rho^i)$. Choose any point $x_0 \in \mathcal{X}$, a function F_{x_0} that satisfies (6) and define F by (8). Then for any $X \in \mathbf{G}$ and $x = \phi(X, x_0)$ one has

$$\mathrm{d}\phi_{x_0}(X)[F(X, v)] = f(x, v). \qquad (9)$$

△

Proof: Compute

$$\begin{aligned}
\mathrm{d}\phi_{x_0}(X)[F(X, v)] =& \mathrm{d}\phi_{x_0}(X)\left[dR_X(I)[F_{x_0}(\psi(X^{-1}, v))]\right] \\
=& \mathrm{d}\left(\phi_{x_0} \circ R_X\right)(I)[F_{x_0}(\psi(X^{-1}, v))] \\
=& \mathrm{d}\left(\phi_X \circ \phi_{x_0}\right)(I)[F_{x_0}(\psi(X^{-1}, v))] \\
=& \mathrm{d}\phi_X(x_0)\left[\mathrm{d}\phi_{x_0}(I)[F_{x_0}(\psi(X^{-1}, v))]\right] \\
=& \mathrm{d}\phi_X(x_0)[f(x_0, \psi(X^{-1}, v))] \\
=& f(\phi(X, x_0), \psi(X, \psi(X^{-1}, v)))] \\
=& f(\phi(X, x_0), v).
\end{aligned}$$

The step from line 2 to 3 depends on the right group action structure of ϕ. △

Note that $\mathrm{d}\phi_x(A)[F(A, v)] \neq f(\phi(A, x), v)$ in general and (9) does not generalize to shifting the base point. This is because construction of F_{x_0}, and hence of $F(A, v)$, depends explicitly on the choice of base point x_0, effectively the choice of a base point for a frame of reference in which to write down a coordinate expression of the system kinematics.

Lemma 1 is important because it allows us to lift the system (1) to an equivariant system on $\mathbf{G}$. Before defining the lifted system, we need to consider the structure of the outputs in more detail. Fix a reference $x_0 \in \mathcal{X}$ and set $x = \phi(X, x_0)$. For each output y_i we define a reference output $\mathring{y}_i \in \mathcal{Y}_i$ by

$$\mathring{y}_i := h_i(x_0), \qquad i = 1, \ldots, p. \qquad (10)$$

The output model for the lifted equivariant system can now be rewritten as the group action ρ^i acting on the reference point $\mathring{y}_i$,

$$\begin{aligned}
y_i &= h_i(x) = h_i(\phi(X, x_0)) \\
&= \rho^i(X, h_i(x_0)) = \rho^i(X, \mathring{y}_i). \qquad (11)
\end{aligned}$$

The reference $\mathring{y}_i$ is defined using the reference x_0, however, it is really the reference $\mathring{y}_i \in \mathcal{Y}_i$ itself that is fundamental to the observer design problem. In many applications the output h_i comes naturally as a group action acting on an *a-priori* known element $\mathring{y}_i$. The group action ρ^i can then be thought of as encoding the relationship between two separate measurements, generally in different frames of reference, of the same physical variable, $y_i = \rho^i(X, \mathring{y}_i)$.

In the definition of a lifted system we will distinguish between the symmetry group $\mathbf{G}$, and the space in which the lifted kinematics live; that looks like $\mathbf{G}$ but only the smooth manifold structure is required for the system kinematics. A $\mathbf{G}$-torsor, $\mathcal{G}$, is defined as the set of elements of $\mathbf{G}$ equipped with its manifold structure, but without the group structure. Right translation defines a group action of $\mathbf{G}$ on $\mathcal{G}$ via $R\colon \mathbf{G} \times \mathcal{G} \to \mathcal{G}$, $R(A, X) := XA$ where the identification between elements of $\mathbf{G}$ and $\mathcal{G}$ is used to apply the multiplication, in the case of matrix groups it is simply matrix multiplication.

Definition 3. Consider a kinematic system (Def. 1) that is equivariant with respect to a right handed complete symmetry $(\mathbf{G}, \phi, \psi, \rho^i)$ (Def. 2) for the Lie group $\mathbf{G}$. Fix a point $x_0 \in \mathcal{X}$ and a velocity lift $F_{x_0}\colon \mathbb{V} \to \mathfrak{g}$ that satisfies (6). Define the lifted system function $F(X, v)$ by (8). Define reference outputs $\mathring{y}_i$ by (10) for $i = 1, \ldots, p$.

A *lifted equivariant system* is defined to be the system on the $\mathbf{G}$-torsor $\mathcal{G}$

$$\dot{X} := F(X, v), \qquad (12a)$$

$$y_i := \rho^i(X, \mathring{y}_i) =: H_i(X) \qquad (12b)$$

for $v \in \mathbb{V}$ and initial condition $X(0) \in \mathcal{G}$ such that $\phi_{x_0}(X(0)) = x(0)$ projects to the initial condition of (1).

This definition is justified by the following lemma.[6]

Lemma 2. Consider equivariant kinematics (1) with a lifted system (12) (Def. 3). Then the lifted system (12) is equivariant with respect to the complete symmetry $(\mathbf{G}, R, \psi, \rho^i)$. Moreover, the solutions $X(t; X_0)$ of (12) project to solutions $x(t; x_0)$ of (1) via

$$\phi(X(t; X_0), x_0) = x(t; x_0).$$

△

Proof: To prove the equivariance note that R is transitive by definition and compute

$$\begin{aligned}
&\mathrm{d}R_A(X)[F(X, v)] \\
&\quad= \mathrm{d}R_A(X)\mathrm{d}R_X(I)[F_{x_0}(\psi(X^{-1}, v))] \\
&\quad= \mathrm{d}(R_A \circ R_X)(I)[F_{x_0}(\psi(X^{-1}, \psi(A^{-1}A, v)))] \\
&\quad= \mathrm{d}(R_{XA})(I)[F_{x_0}(\psi(A^{-1}X^{-1}, \psi(A, v)))] \\
&\quad= \mathrm{d}(R_{R_A X})(I)[F_{x_0}(\psi((R_A X)^{-1}, \psi(A, v)))] \\
&\quad= F(R_A X, \psi(A, v))
\end{aligned}$$

This proves (4a). To see (4b) note that

$$\begin{aligned}
H_i(R_A X) &= \rho^i((R_A X), \mathring{y}_i) \\
&= \rho^i(A, \rho^i(X, \mathring{y}_i)) \\
&= \rho^i(A, \rho^i(X, \mathring{y}_i)) \\
&= \rho^i(A, H_i(X))
\end{aligned}$$

Consider a solution $X = X(t; X(0))$ of (12) and compute the time derivative of the projection $x(t) = \phi_{x_0}(X(t; X(0)))$:

[5] This construction is adapted to a right handed symmetry; given a left handed symmetry it is necessary to replace R_X by L_X.

[6] If we started out with a left handed symmetry and defined the lifted system function accordingly, the lifted system would turn out to be equivariant with respect to the complete symmetry $(\mathbf{G}, L, \psi, \rho^i)$ instead.

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi_{x_0}(X(t;X(0))) &= \mathrm{d}\phi_{x_0}(X)F(X,v)\\ &= \mathrm{d}\phi_{x_0}(X)dR_X(I)[F_{x_0}(\psi(X^{-1},v))]\\ &= \mathrm{d}(\phi_{x_0}\circ R_X)(I)[F_{x_0}(\psi(X^{-1},v))]\\ &= \mathrm{d}\phi_X(x_0)d\phi_{x_0}(I)[F_{x_0}(\psi(X^{-1},v))]\\ &= \mathrm{d}\phi_X(x_0)f(x_0,(\psi(X^{-1},v))\\ &= f(\phi_X(x_0),\psi(X,\psi(X^{-1},v))) = f(x,v).\end{aligned}$$

Since $x(0) = \mathrm{d}\phi_{x_0}(X(0))$ by definition, then from uniqueness of solutions it is clear that $x(t) = \phi_{x_0}(X(t;X(0))) = x(t;x(0))$ is the solution of (1). △

Example 1.3. Recall the scenario described in Example 1.1. The Lie algebra of SO(3) is the set of skew symmetric 3×3 matrices

$$\mathfrak{so}(3) = \{W \in \mathbb{R}^{3\times 3} \,|\, W = -W^\top\}.$$

Fix $x_0 = \mathbf{e}_3 \in \mathrm{S}^2$ where $\mathbf{e}_3 = (0,0,1)^\top$ is the unit vector in third axis of $\mathbb{R}^3$. We choose a velocity lift

$$F_{\mathbf{e}_3}(\Omega) = \begin{pmatrix} 0 & -\Omega_3 & \Omega_2\\ \Omega_3 & 0 & -\Omega_1\\ -\Omega_2 & \Omega_1 & 0\end{pmatrix} =: \Omega_\times.$$

To verify (6) note that the Fréchet derivative

$$\mathrm{D}\phi_{\mathbf{e}_3}(Q)[WQ] = W^\top Q^\top \mathbf{e}_3 = -WQ^\top \mathbf{e}_3$$

since $W \in \mathfrak{so}(3)$. Evaluating at $Q = I$ and applying to $F_{\mathbf{e}_3}(\Omega)$ yields

$$\begin{aligned}\mathrm{d}\phi_{\mathbf{e}_3}(I)[F_{\mathbf{e}_3}(\Omega)] &= -\Omega_\times \mathbf{e}_3 = -\Omega\times\mathbf{e}_3\\ &= \mathbf{e}_3\times\Omega = f(\mathbf{e}_3,\Omega)\end{aligned}$$

and hence the choice of $F_{\mathbf{e}_3}$ satisfies (6). Note that only the elements Ω_2 and Ω_3 actually contribute to the image $\mathrm{d}\phi_{\mathbf{e}_3}(I)[F_{\mathbf{e}_3}]$ since $\ker \mathrm{d}\phi_{\mathbf{e}_3}(I) = \mathrm{span}\{\mathbf{e}_1\}$.

The system function (8) is

$$\begin{aligned}F(X,\Omega) &:= \mathrm{d}R_X(I)F_{\mathbf{e}_3}(\psi(X^{-1},\Omega)\\ &= ((X^{-1})^\top\Omega)_\times X = (X\Omega)_\times X\\ &= X\Omega_\times X^\top X = X\Omega_\times,\end{aligned}$$

since $X^{-1} = X^\top$ because $X \in \mathrm{SO}(3)$ and using the easily verified result that $(X\Omega)_\times = X\Omega_\times X^\top$. That is the lifted kinematics are

$$\dot{X} = X\Omega_\times = \mathrm{d}R_X\,\mathrm{Ad}_X(\Omega_\times). \tag{13}$$

The reference output is $\mathring{y} = h(\mathbf{e}_3) = \mathbf{e}_3$. Then

$$H(X) := \rho(X,\mathbf{e}_3) = X^\top\mathbf{e}_3 = y.$$

In the context of the attitude estimation application this models an inertial direction $\mathbf{e}_3$ corresponding to a physical direction such as the gravitational field, measured in body fixed coordinates $y = X^\top\mathbf{e}_3$.

It is also interesting to verify that the lifted dynamics do project to the kinematic system (2). This can be seen by computing

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi_{\mathbf{e}_3}(X) &= \mathrm{D}\phi_{\mathbf{e}_3}[X\Omega_\times]\\ &= -\Omega_\times X^\top\mathbf{e}_3 = -\Omega_\times x\\ &= -\Omega\times x = x\times\Omega.\end{aligned}$$

△

Example 2.3. Recall the scenario described in Example 2.1. The Lie algebra of SE(2) is the set

$$\mathfrak{se}(2) = \left\{\begin{pmatrix} 0 & a & w_1\\ -a & 0 & w_2\\ 0 & 0 & 0\end{pmatrix} \in \mathbb{R}^{3\times 3} \,\middle|\, a, w_1, w_2 \in \mathbb{R}\right\}. \tag{14}$$

We write $w = (w_1, w_2)$ and

$$a_\times = \begin{pmatrix} 0 & a\\ -a & 0\end{pmatrix}$$

then $W = W(a, w_1, w_2) \in \mathfrak{se}(2)$ can be written in block form

$$W = \begin{pmatrix} a_\times & w\\ 0 & 0\end{pmatrix}.$$

Fix $x_0 = (0,0,0) \in \mathcal{X}$ corresponding to the origin of the inertial frame with zero orientation. We choose a velocity lift

$$F_0(u,q) = \begin{pmatrix} 0 & q & u\\ -q & 0 & 0\\ 0 & 0 & 0\end{pmatrix}.$$

Note that

$$\phi_0(Q) = (R(\alpha)0 + z, \alpha + 0) = (z,\alpha)$$

for an element $Q \in \mathrm{SE}(2)$ parameterized by (5). For $W \in \mathfrak{se}(2)$ parameterized by (14) and A as in (5) then it is straightforward to verify that

$$QW = \begin{pmatrix} R(\alpha)a_\times & R(\alpha)w\\ 0 & 0\end{pmatrix}. \tag{15}$$

Consider the derivative

$$\mathrm{D}\phi_0(Q)[QW] = (R(\alpha)w, a). \tag{16}$$

This formula follows from the fact that the rate of change of α in $R(\alpha)a_\times$ is a. Evaluating at $Q = I$, i.e. $R(\alpha) = I$ and $z = 0$, and applying to $F_0(v)$ yields

$$\begin{aligned}\mathrm{d}\phi_0(I)[F_0(u,q)] &= ((u,0)^\top, q) = ((\cos(0)u, \sin(0)u)^\top, q)\\ &= f(0,v).\end{aligned}$$

It follows that the choice of F_0 satisfies (6).

The system function (8) is

$$\begin{aligned}F(X,v) &:= \mathrm{d}L_X F_0(\psi(X^{-1},v)) = XF_0(v)\\ &= \begin{pmatrix} R(\theta) & \xi\\ 0 & 1\end{pmatrix}\begin{pmatrix} q_\times & \binom{u}{0}\\ 0 & 0\end{pmatrix} = \begin{pmatrix} R(\theta)q_\times & R(\theta)\binom{u}{0}\\ 0 & 0\end{pmatrix}.\end{aligned}$$

That is the lifted kinematics are

$$\dot{X} = XF_0(v) = \mathrm{d}\bar{R}_X F_0(v), \tag{17}$$

where $\bar{R}_X = L_X$ is the *right* translation with respect to the re-defined group multiplication on SE(2) that is turning our left handed equivariance into a right handed one.

To see that this projects to the kinematics (3) then we compute

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\phi_0(X) &= \mathrm{D}\phi_0[XF_0(v)] = (R(\theta)\tbinom{u}{0}, q)\\ &= ((\cos(\theta)u, \sin(\theta)u)^\top, q) = f(x,v),\end{aligned}$$

where the second equality on the first line follows from (16) with suitable substitution of variables. The reference $\mathring{y} = h(0) = 0$. Then

$$H(X) := \rho(X,0) = \xi = y.$$

That is, the measurement is the observation of the position of the unicycle expressed in inertial coordinates. △

There are two special cases of equivariant kinematics (1) that we identify corresponding to particular properties of the velocity group action ψ and the resulting properties of the lifted system on the symmetry group torsor. We distinguish between general (right) equivariant kinematics on a torsor, such as (12a), and those that can be written as $\dot{X} = \mathrm{d}R_X \operatorname{Ad}_X(F_{x_0}(v))$ or $\dot{X} = \mathrm{d}R_X F_{x_0}(v)$. [7] We refer to these two special classes of systems as type I and type II, respectively. Type II systems have the simpler physical intuition and have been studied in the control literature (Brockett [1972, 1973], Jurdjevic and Sussmann [1972], Cheng et al. [1990], Apostolou and Kazakos [1996]). It turns out that it is type I systems that model the typical applications in mobile robotics that motivated the authors' work and the authors are unaware of work that considers such systems explicitly other than their own (Lageman et al. [2010], Trumpf et al. [2012]). We will describe type II systems first as they have the simpler physical intuition, and then describe type I systems.

Definition 4. Consider a right equivariant kinematic system (1) with complete symmetry $(\mathbf{G}, \phi, \psi, \rho^i)$. Then this system is said to be a *type II system* if $\psi(A, v) = v$ for all $A \in \mathbf{G}$.

Consider the lifted equivariant system associated with a type II system,

$$\begin{aligned}\dot{X} = F(X, v) &= \mathrm{d}R_X F_{x_0}(\psi(X^{-1}, v)) \\ &= \mathrm{d}R_X F_{x_0}(v),\end{aligned}$$

yielding a type II system on the $\mathbf{G}$-torsor $\mathcal{G}$. For a matrix Lie group $\mathbf{G}$, this system has the standard form $\dot{X} = UX$, $U \in \mathfrak{g}$ on $\mathcal{G}$. This corresponds to a family of right invariant vector fields on the Lie group $\mathbf{G}$ parametrized by the input $U \in \mathfrak{g}$.

The underlying structure that leads to type II systems (and later to type I systems) comes from the physics of the system, in particular the way in which the velocity measurement interacts with the system symmetry. In the case of type II systems, the velocity parametrization of (1) is independent of the frame of reference in which the system is expressed. That is,

$$\mathrm{d}\phi_A(x)[f(x, v)] = f(\phi_A(x), v). \tag{18}$$

For the usual choice of coordinates and the right-handed equivariance, this corresponds to the case where velocity is measured with respect to the body-fixed frame. For the same coordinates but left-handed equivariance, this analogously corresponds to the case where velocity is measured with respect to the inertial frame.

Example 2.4. The non-holonomic unicycle from Example 2.1 is left equivariant in the usual coordinates and the velocity $v = (u, q)$ is measured with respect to the inertial frame by the onboard tachometers. △

Definition 5. Consider an equivariant kinematic system (1) with right-handed symmetry $(\mathbf{G}, \phi, \psi, \rho^i)$. Then this system is said to be a *type I system* if there exists a velocity lift F_{x_0} satisfying (7) such that

$$\operatorname{Ad}_X(F_{x_0}(v)) = F_{x_0}(\psi(X^{-1}, v)). \tag{19}$$

The intuition in this definition is again seen at the level of the lifted system on the $\mathbf{G}$-torsor $\mathcal{G}$. In this case one has

$$\begin{aligned}\dot{X} = F(X, v) &= \mathrm{d}R_X F_{x_0}(\psi(X^{-1}, v)) \\ &= \mathrm{d}R_X \operatorname{Ad}_X(F_{x_0}(v)) = \mathrm{d}R_X \mathrm{d}L_X \mathrm{d}R_{X^{-1}} F_{x_0}(v) \\ &= \mathrm{d}L_X F_{x_0}(v),\end{aligned}$$

since left and right translation commute. For a matrix Lie group $\mathbf{G}$, these kinematics have direct velocity parametrization $\dot{X} = XU$, $U \in \mathfrak{g}$ on $\mathcal{G}$, corresponding to a family of *left* invariant vector fields on the Lie group $\mathbf{G}$ parametrized by the input $U \in \mathfrak{g}$. For the usual choice of coordinates and right-handed equivariance, this structure corresponds to the case where velocity measurements are made with respect to the inertial frame.

Example 1.4. Recall the scenario described in Example 1.1. A typical velocity measurement for an attitude observation problem comes from a strap down inertial measurement unit (IMU). The gyroscopes on such a device use inertial effects to measure the angular velocity of the body-fixed frame with respect to the inertial frame. Note that the measured velocity is expressed in the body-fixed frame, but this should not be confused with the reference frame with respect to which the measurements are made. The group coordinates $X \in \mathrm{SO}(3)$ are chosen to represent the orientation of the body-fixed frame with respect to the inertial frame. The key driver for the choice of right-handed symmetry in the model is the output equation

$$y = X^\top \mathring{y}.$$

This physically models a process where the inertial direction $\mathring{y}$ is measured in the body-fixed frame. However, the inverse $X^{-1} = X^\top$ in the group action imposes a right-handed equivariance on the output measurement in these group coordinates. Given that the output equation has imposed a right-handed symmetry it is necessary to model the kinematic symmetry as a set of right-handed actions. Such a symmetry is provided by the definitions of ϕ and ψ given in Example 1.2. The structure of the ψ action is a consequence of the relationship of the physical velocity measurement to the various choices of symmetry imposed by the choice of coordinates and the natural symmetry of the output measurement.

The type I nature of the system is easily verified

$$\begin{aligned}F_{x_0}(\psi(X^{-1}, v)) &= ((X^{-1})^\top \Omega)_\times = (X\Omega)_\times \\ &= \operatorname{Ad}_X(\Omega_\times) = \operatorname{Ad}_X(F_{x_0}(v)).\end{aligned}$$

△

It appears to be much more challenging to verify the algebraic conditions of Definition 5 than Definition 4 since the former requires a choice of F_{x_0}. However, if the physical modeling is undertaken with care, our experience shows that this leads naturally to the required structure and makes the identification of type I or type II systems straightforward.

The difference between type I, type II and general equivariant systems on the symmetry group torsor is really a consequence of how velocity is parameterized in the physical system. If one allows the freedom to change the way in which the velocity is parameterized in the model in a state-dependent way then one may choose a new velocity

$$v_{\text{new}} = \psi(X^{-1}, v_{\text{old}})$$

to transform any equivariant system into a type II system $\dot{X} = \mathrm{d}R_X(F_{x_0}(v_{\text{new}}))$. Such a transformation is state-dependent of course, and the new velocity v_{new} cannot be 'measured' using the old velocity sensors because they will be physically placed incorrectly in the system. If the full system state were known, then the state dependent transformation $\psi(X^{-1}, \cdot)$ would be exactly the operation required to infer the new velocity from the old physical

[7] In the left equivariant calculus we would need to replace R_X by L_X and Ad_X by $\operatorname{Ad}_{X^{-1}}$.

velocity sensor, a transformation that is clearly of no practical use in designing observers.

The terms type I and type II are deliberately chosen to be agnostic to the handedness of equivariance of the underlying kinematic system, since the handedness is a matter of modeling choice, while the type of system depends on the physical relationship between the sensors.

3. OBSERVER STRUCTURE THEORY

In this section, we propose a structure for the design of equivariant observers and derive invariance properties of the associated error kinematics.

The proposed approach is to design an observer for the lifted equivariant system (12) and then project this down to $\mathcal{X}$ to obtain an observer on the original system state space. We use $\hat{X} \in \mathcal{G}$ (where $\mathcal{G}$ is the $\mathbf{G}$-torsor) to denote an estimate for the lifted system state $X(t; X(0))$ for unknown $X(0)$. The fundamental structure for the observer that we consider is that of a *pre-observer* (or internal model) (Bonnabel et al. [2009, 2008], Lageman et al. [2010]) with innovation. The pre-observer is a copy of the lifted system kinematics (12) with $\hat{X}$ replacing X

$$\dot{\hat{X}} = F_{x_0}(\psi(\hat{X}^{-1}, v))\hat{X} \tag{20}$$

and depends on $v \in \mathbb{V}$ and the observer state $\hat{X}$. The innovation $\Delta \in \mathfrak{g}$ is an error correction term. It takes outputs $\{y_i\}$ and the observer state $\hat{X}$ and generates a correction term for the observer dynamics with the goal that $\hat{X} \to X(t, X(0))$, or at least that $\hat{x} = \phi(\hat{X}, x_0)$ converges to $x(t; x(0))$. Thus, the proposed observer for a lifted equivariant systems has the form

$$\dot{\hat{X}} = F_{x_0}(\psi(\hat{X}^{-1}, v))\hat{X} - \mathrm{d}R_{\hat{X}}\Delta(\hat{X}, y_1, \dots, y_p), \tag{21a}$$

$$\hat{x}(t) = \phi_{x_0}(\hat{X}(t; \hat{X}(0)) \tag{21b}$$

for $\hat{X}(0) \in \mathcal{G}$ some initial condition, typically $\hat{X}(0) = I$, and x_0 chosen as the best *a priori* guess of $x(0)$.

The observer (21) described above could be analyzed locally by exploiting the equivariance of the underlying kinematics, see Bonnabel et al. [2008, 2009]. Indeed, if a local approach based on equivariance is taken, then there is no real need to lift the system on to the symmetry group torsor, a moving frame approach can be taken directly on the state space, cf. Bonnabel et al. [2008]. By lifting onto the symmetry group torsor, however, it is possible to provide global analyses and design methodologies for the cases of type I and type II systems, see Mahony et al. [2008], Lageman et al. [2010], Trumpf et al. [2012]. Since there is a large class of applications, indeed most applications the authors are aware of, that can be modeled as type I and type II systems, this warrants a careful development of this approach.

3.1 Pre-observers and error functions

In order to study the relationship between two trajectories on $\mathcal{G}$, we will introduce a (smooth) *error function*

$$E \colon \mathcal{G} \times \mathcal{G} \to M, \tag{22}$$

where M is a smooth manifold. The role of the error E is analogous to the vector error $\tilde{x} = \hat{x} - x$ in providing a *global* comparison between trajectories in classical linear observer theory. Note that in linear observer theory the scalar norm of the error $||\tilde{x}||^2$ is also used as a quantitative measure of observer performance. We make a distinction between the error function E ($\tilde{x}$ in the linear theory), a multi-dimensional map that allows comparison of trajectories of the pre-observer; and a cost function, that we will define in §4 ($||\tilde{x}||^2$ in the linear theory), that is used as a Lyapunov function during the design of the innovation term.

Two particularly simple error functions on a Lie group torsor $\mathcal{G}$ are the canonical type I error $E_{\mathrm{I}} \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}$,

$$E_{\mathrm{I}}(\hat{X}, X) := \hat{X}X^{-1}, \tag{23}$$

and the canonical type II error $E_{\mathrm{II}} \colon \mathcal{G} \times \mathcal{G} \to \mathcal{G}$

$$E_{\mathrm{II}}(\hat{X}, X) := X^{-1}\hat{X}, \tag{24}$$

both defined for an underlying right equivariant model. [8] Where the arguments $\hat{X}$ and X are clear from the context we simply write E_{I} and E_{II}. Observe that both E_{I} and E_{II} are non-degenerate in the sense that the partial maps $E(\hat{X}, \cdot) \colon \mathcal{G} \to \mathcal{G}$ and $E(\cdot, X) \colon \mathcal{G} \to \mathcal{G}$ from either error are global diffeomorphisms.

Both type I and type II errors have natural invariance properties. The type I error has a symmetry that matches that of the underlying kinematic model. That is, given right-handed symmetry of the kinematics then the matching symmetry on the Lie-group is right translation R_S. For all $\hat{X}, X \in \mathcal{G}$, $S \in \mathbf{G}$, one has

$$E_{\mathrm{I}}(R_S\hat{X}, R_S X) = \hat{X}SS^{-1}X^{-1} = E_{\mathrm{I}}(\hat{X}, X).$$

The type II error has the opposite symmetry to that of the underlying kinematic model

$$E_{\mathrm{II}}(L_S\hat{X}, L_S X) = X^{-1}S^{-1}S\hat{X} = E_{\mathrm{II}}(\hat{X}, X).$$

The structure of type I and type II errors is coupled to the structure of type I and type II systems, respectively. Consider a type I lifted equivariant system (Def. 5) and the pre-observer (20). For any initial conditions $X(0), \hat{X}(0) \in \mathcal{G}$ one has

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}} &= \frac{\mathrm{d}}{\mathrm{d}t}[\hat{X}]X^{-1} + \hat{X}\frac{\mathrm{d}}{\mathrm{d}t}[X^{-1}] \\
&= \left(\mathrm{Ad}_{\hat{X}} F_{x_0}(v)\right)\hat{X}X^{-1} - \hat{X}X^{-1}\frac{\mathrm{d}}{\mathrm{d}t}[X]X^{-1} \\
&= \hat{X}F_{x_0}(v)X^{-1} - \hat{X}X^{-1}\left(\mathrm{Ad}_X F_{x_0}(v)\right)XX^{-1} \\
&= \hat{X}(F_{x_0}(v) - F_{x_0}(v))X^{-1} \\
&= 0.
\end{aligned}$$

It is worth observing that the same error properties do not hold for the type II error and type I systems

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{II}} &= \frac{\mathrm{d}}{\mathrm{d}t}[X^{-1}]\hat{X} + X^{-1}\frac{\mathrm{d}}{\mathrm{d}t}[\hat{X}] \\
&= -X^{-1}(\mathrm{Ad}_X F_{x_0}(v)))XX^{-1}\hat{X} \\
&\qquad + X^{-1}(\mathrm{Ad}_{\hat{X}} F_{x_0}(v)))\hat{X} \\
&= -F_{x_0}(v)E_{\mathrm{II}} + E_{\mathrm{II}}F_{x_0}(v) \\
&= \left(-F_{x_0}(v) + \mathrm{Ad}_{E_{\mathrm{II}}} F_{x_0}(v)\right)E_{\mathrm{II}}.
\end{aligned}$$

A similar computation to the above shows that type II lifted equivariant systems are compatible with the E_{II} error and not compatible with the E_{I} error.

[8] Should one choose to model the underlying kinematic system with the opposite equivariance then the definitions of type I and type II errors would switch.

The analyses for type I and type II systems have many similarities but vary substantially in detail at the point where the outputs are considered. In the name of conceptual simplicity, and in order to keep the length of the paper within reason, we will concentrate on type I systems from this point on in this paper. Type I systems provide a model for a wide range of applications in mobile robotics that have motivated much of the authors' work in this area.

3.2 Equivariant innovations

Since we will need to deal extensively with elements of all output spaces at the same time in the sequel we introduce notation. Define the product output space $\mathcal{Y}$ by

$$\mathcal{Y} = \mathcal{Y}_1 \times \cdots \times \mathcal{Y}_p.$$

We will write $y, \mathring{y}, e^{\mathrm{I}} \in \mathcal{Y}$ with

$$y = (y_1, \ldots, y_p), \quad \mathring{y} = (\mathring{y}_1, \ldots, \mathring{y}_p), \quad e^{\mathrm{I}} = (e^{\mathrm{I}}_1, \ldots, e^{\mathrm{I}}_p)$$

where e^{I} is the output error that we will introduce in Definition 7 below.

The innovation is a function of the observer state $\hat{X} \in \mathcal{G}$ and measurements $y \in \mathcal{Y}$ with parameters $\mathring{y} \in \mathcal{Y}$

$$\begin{aligned} \Delta \colon \mathcal{G} \times \mathcal{Y} \times \mathcal{Y} &\to \mathfrak{g}, \\ (\hat{X}, y; \mathring{y}) &\mapsto \Delta_{\mathring{y}}(\hat{X}, y). \end{aligned} \tag{25}$$

The reference outputs $\mathring{y}_i = h_i(x_0)$ are important constant parameters in the observer design and it is useful to keep them explicitly in the notation. The driving term in (21) is actually $-\mathrm{d}R_{\hat{X}}\Delta = -\Delta\hat{X}$, however, since the goal of our work is to design equivariant observers then it is appropriate to left trivialize the innovation to the Lie algebra. In general, one may allow Δ to also depend on velocity and time, however, neither of these generalizations lead to any advantage in observer design (Lageman et al. [2010]) and the development is simpler and more direct without the added complexity. Where the arguments $(\hat{X}, y; \mathring{y})$ are clear from context we will simply write Δ.

Definition 6. An innovation (25) is termed *equivariant* if for all $S \in \mathbf{G}$, $y_i, \mathring{y}_i \in \mathcal{Y}_i$, then

$$\Delta_{\mathring{y}}(R_S\hat{X}, \rho^1(S, y_1), \ldots, \rho^p(S, y_p)) = \Delta_{\mathring{y}}(\hat{X}, y_1, \ldots, y_p).$$

The key advantage of working with an equivariant innovation is that the observer design problem can be reduced to a Lyapunov argument in suitable error coordinates. The structure developed so far leads to the following theorem.

Theorem 1. Consider a type I lifted system (Def. 5) with the observer (21). Then the dynamics of the canonical error E_{I} is autonomous if and only if the innovation term Δ is equivariant. The autonomous error dynamics for an equivariant innovation has the form

$$\frac{d}{dt}E_{\mathrm{I}} = -\mathrm{d}R_{E_{\mathrm{I}}}(I)\Delta_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}). \tag{26}$$

$\triangle$

Proof: The error dynamics is given by

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}} &= \frac{\mathrm{d}}{\mathrm{d}t}[\hat{X}]X^{-1} + \hat{X}\frac{\mathrm{d}}{\mathrm{d}t}[X^{-1}] &(27)\\ &= \left(\mathrm{Ad}_{\hat{X}} F_{x_0}(v)\right)\hat{X}X^{-1} - \Delta_{\mathring{y}}(\hat{X}, y)\hat{X}X^{-1} \\ &\qquad - \hat{X}X^{-1}\left(\mathrm{Ad}_X F_{x_0}(v)\right)XX^{-1} \\ &= -\Delta_{\mathring{y}}(\hat{X}, y)E_{\mathrm{I}}. &(28) \end{aligned}$$

If $\Delta_{\mathring{y}}(\hat{X}, y)$ is equivariant then

$$\Delta_{\mathring{y}}(\hat{X}, y) = \Delta_{\mathring{y}}(R_{X^{-1}}\hat{X}, \mathring{y}) = \Delta_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$$

and the error dynamics are autonomous and of the form (26).

On the other hand, if the error dynamics are autonomous then from (28)

$$\Delta_{\mathring{y}}(\hat{X}, y) = -\mathrm{d}R_{E_{\mathrm{I}}^{-1}}\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}}$$

and the right hand side is a function only of E_{I}; that is, it cannot depend independently on the signals $\hat{X}$, X or y. Since the error is equivariant then for any $S \in \mathbf{G}$ one has $E_{\mathrm{I}}(R_S\hat{X}, R_SX) = E_{\mathrm{I}}(\hat{X}, X)$ and autonomy of the above expression ensures that

$$\mathrm{d}R_{E_{\mathrm{I}}(R_S\hat{X}, R_SX)^{-1}}\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}}(R_S\hat{X}, R_SX) = \mathrm{d}R_{E_{\mathrm{I}}^{-1}}\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}}.$$

Write $\Delta_{\mathring{y}}(\hat{X}, y) = \Delta_{\mathring{y}}(\hat{X}, \rho^i(X, \mathring{y}))$ and compute

$$\begin{aligned} \Delta_{\mathring{y}}(R_S\hat{X}, \rho^i(S, y)) &= \Delta_{\mathring{y}}(R_S\hat{X}, \rho^i(R_SX, \mathring{y})) \\ &= -\mathrm{d}R_{E_{\mathrm{I}}(R_S\hat{X}, R_SX)^{-1}}\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}}(R_S\hat{X}, R_SX) \\ &= -\mathrm{d}R_{E_{\mathrm{I}}^{-1}}\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}} = \Delta_{\mathring{y}}(\hat{X}, \rho^i(X, \mathring{y})) \\ &= \Delta_{\mathring{y}}(\hat{X}, y). \end{aligned}$$

It follows that the innovation is equivariant. $\triangle$

3.3 Output errors

Since we will ultimately construct an invariant innovation based on invariant cost functions on the output spaces, we introduce output error coordinates that are adapted to the given equivariance structure.

Definition 7. Consider a lifted equivariant system (Def. 3). Given an estimate $\hat{X} \in \mathcal{G}$ then define the *type I output error* $e^{\mathrm{I}}_i \colon \mathcal{G} \times \mathcal{Y}_i \to \mathcal{Y}_i$ by

$$e^{\mathrm{I}}_i(\hat{X}, y_i) := \rho^i(\hat{X}^{-1}, y_i). \tag{29}$$

Where the arguments are clear from context we will simply write e^{I}_i.

Note that e^{I}_i is implementable, since y_i is measured and $\hat{X}$ is known from the observer state. The error $e^{\mathrm{I}}_i(\hat{X}, y_i)$ has a natural equivariance;

$$\begin{aligned} e^{\mathrm{I}}_i(R_S\hat{X}, \rho^i(S, y_i)) &= \rho^i((R_S\hat{X})^{-1}, \rho^i(S, y_i)) \\ &= \rho^i(S^{-1}\hat{X}^{-1}, \rho^i(S, y_i)) \\ &= \rho^i(\hat{X}^{-1}, \rho^i(S^{-1}, \rho^i(S, y_i))) \\ &= \rho^i(\hat{X}^{-1}, y_i) = e^{\mathrm{I}}_i(\hat{X}, y_i) \end{aligned}$$

associated with the equivariance of the underlying system kinematics. Indeed, the error e^{I}_i inherits this equivariance from the symmetry of the output y_i encoded in the ρ^i group action. The above construction is a specific example of an *invariant output error* (Bonnabel et al. [2008]).

This invariance leads to the following structure

$$\begin{aligned} e^{\mathrm{I}}_i(\hat{X}, y_i) &= \rho^i(\hat{X}^{-1}, \rho^i(X, \mathring{y}_i)) = \rho^i(X\hat{X}^{-1}, \mathring{y}_i) \\ &= \rho^i(E_{\mathrm{I}}^{-1}, \mathring{y}_i) = e^{\mathrm{I}}_i(E_{\mathrm{I}}^{-1}, \mathring{y}_i) \end{aligned}$$

It follows that for $\hat{X} = X$, then $e^{\mathrm{I}}_i = \mathring{y}_i$, that is e^{I}_i is 'centred' on the reference $\mathring{y}_i$. The above computation also demonstrates that e^{I}_i is adapted to the E_{I} error.

Example 1.5. Recall the scenario described in Example 1.1. The reference output was $\mathring{y} = \mathbf{e}_3$ with measurement $y = \rho(X, \mathbf{e}_3) = X^\top \mathbf{e}_3$.

The output error is

$$e^{\mathrm{I}} = \rho(\hat{X}^{-1}, y) = \hat{X} y.$$

This can be written

$$e^{\mathrm{I}} = \hat{X} X^\top \mathbf{e}_3 = \hat{X} X^\top \mathbf{e}_3 = E_{\mathrm{I}} \mathbf{e}_3 = \rho(E_{\mathrm{I}}^{-1}, \mathbf{e}_3).$$

$\triangle$

Remark 1. In prior work (Trumpf et al. [2012], Lageman et al. [2010]) an output y_i was termed complementary to an invariant system on the Lie-group if it had the opposite equivariance to that of the group kinematics. In those papers, complementary outputs to a left equivariant system would have right symmetry, and vice versa. The equivalent situation in the present development has the underlying symmetry of the kinematics as the primary equivariance. However, type I systems are defined in such a way that they have the opposite equivariance on the group when written without group action on the Lie algebra. The measurement y_i has the underlying symmetry of the kinematics, and hence, in the old language, is a complementary measurement for a type I system.

Type II systems have a matching equivariance on the group when written in their direct velocity parameterization. The measurement y_i also has the same handed equivariance. Such measurements were termed compatible measurements in prior work. $\triangle$

4. OBSERVER DESIGN METHODOLOGY

The approach taken in this paper is to define invariant cost functions on the output spaces. Properly chosen, these functions can be lifted and aggregated to define a Lyapunov function in error coordinates on the **G**-torsor $\mathcal{G}$ that has a global minimum at the identity with positive definite Hessian. Since the error dynamics of the system are autonomous, a straightforward Lyapunov design leads to desirable observer behaviour.

Definition 8. An *invariant cost function at* $\mathring{y}_i$ on the output space $\mathcal{Y}_i$ is a function $\ell^i_{\mathring{y}_i} : \mathcal{G} \times \mathcal{Y}_i \to \mathbb{R}^+$, $(\hat{X}, y_i) \mapsto \ell^i_{\mathring{y}_i}(\hat{X}, y_i)$ such that for all $S \in \mathbf{G}$

$$\ell^i_{\mathring{y}_i}(R_S \hat{X}, \rho^i(S, y_i)) := \ell^i_{\mathring{y}_i}(\hat{X}, y_i)$$

and $(I, \mathring{y}_i)$ is a global minimum of the cost.

The cost is termed *non-degenerate* if the Hessian in the second variable, $\mathrm{Hess}_2 \ell^i_{\mathring{y}_i}(I, \mathring{y}_i) > 0$ is positive definite at $(I, \mathring{y}_i)$. $\triangle$

It is straightforward to verify that for an invariant cost function

$$\begin{aligned} \ell^i_{\mathring{y}_i}(\hat{X}, y_i) &= \ell^i_{\mathring{y}_i}(R_{\hat{X}^{-1}} \hat{X}, \rho^i(\hat{X}^{-1}, y_i)) = \ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}}_i) \\ &= \ell^i_{\mathring{y}_i}(I, \rho^i(E_{\mathrm{I}}^{-1}, \mathring{y}_i)) = \ell^i_{\mathring{y}_i}(E_{\mathrm{I}}, \mathring{y}_i) \end{aligned} \tag{30}$$

and hence $\ell^i_{\mathring{y}_i}$ can be directly written as a cost in terms of the error e^{I}_i and in turn as a function of the group error E_{I}. In fact, if one can find a single variable cost function on an output space $\mathcal{Y}_i$, then it is straightforward to build an invariant cost function using a related construction. Let $f : \mathcal{Y}_i \to \mathbb{R}^+$ be a function with a global minimum at a point $\mathring{y}_i$. Define

$$\ell^i_{\mathring{y}_i}(\hat{X}, y_i) := f(\rho^i(\hat{X}^{-1}, y_i)) = f(e^{\mathrm{I}}).$$

This function is invariant by construction since

$$\begin{aligned} \ell^i_{\mathring{y}_i}(R_S \hat{X}, \rho^i(S, y_i)) &= f(\rho^i(S^{-1} \hat{X}^{-1}, \rho(S, y_i))) \\ &= f(\rho^i(S S^{-1} \hat{X}^{-1}, y_i)) \\ &= f(\rho^i(\hat{X}^{-1}, y_i)) \\ &= \ell^i_{\mathring{y}_i}(\hat{X}, y_i). \end{aligned}$$

The global minimum is a direct consequence of the global minimum of f. Similarly, if f is non-degenerate at $\mathring{y}_i$ then $\ell^i_{\mathring{y}_i}$ is non-degenerate.

The aggregate cost that we consider is written as a function

$$\begin{aligned} &\ell_{\mathring{y}} : \mathcal{G} \times \mathcal{Y} \to \mathbb{R}^+, \\ &\ell_{\mathring{y}}(\hat{X}, y) := \sum_{i=1}^{n} \ell^i_{\mathring{y}_i}(\hat{X}, y_i). \end{aligned} \tag{31}$$

This cost is a function over the product of the observer state space and all output spaces that depends on a set of parameters $\mathring{y} \in \mathcal{Y}$.

If all the cost functions are invariant then, recalling (30), the aggregate cost function can be written

$$\ell_{\mathring{y}}(\hat{X}, y) = \sum_{i=1}^{p} \ell^i_{\mathring{y}_i}(E_{\mathrm{I}}, \mathring{y}_i) = \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$$

Thus, the aggregate cost can be rewritten as an error cost $\ell_{\mathring{y}}(\cdot, \mathring{y}) : \mathcal{G} \to \mathbb{R}^+$ in the coordinate E_{I} on the **G**-torsor $\mathcal{G}$ that depends on constant parameters $\mathring{y}$. Although they have the same algebraic form, we will make a distinction between the *aggregate cost* $\ell_{\mathring{y}}$ with domain $\mathcal{G} \times \mathcal{Y}$ and the *error cost* $\ell_{\mathring{y}}(\cdot, \mathring{y})$ with domain $\mathcal{G}$. The fact that the error cost is a map from $\mathcal{G}$ to $\mathbb{R}^+$ makes it suitable to use in the Lyapunov design procedure that we propose later in the section.

Lemma 3. Assume that $\ell^i_{\mathring{y}_i}(\hat{X}, y_i)$ are non-degenerate invariant cost functions at $\mathring{y}_i$ for $i = 1, \ldots, p$. Then the aggregate cost $\ell_{\mathring{y}} : \mathcal{G} \times \mathcal{Y} \to \mathbb{R}^+$, $(\hat{X}, y) \mapsto \ell_{\mathring{y}}(\hat{X}, y)$ is invariant. The error cost $\ell_{\mathring{y}}(\cdot, \mathring{y}) : \mathcal{G} \to \mathbb{R}^+$, $E_{\mathrm{I}} \mapsto \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$ has a global minimum at $E_{\mathrm{I}} = I$. Moreover, if

$$\bigcap_{i=1}^{p} \mathrm{stab}_{\rho^i}(\mathring{y}_i) = \{I\}$$

then $\ell_{\mathring{y}}(\cdot, \mathring{y})$ is non-degenerate at I. $\triangle$

Proof: Invariance of the aggregate cost is straightforward consequence of invariance of the individual output costs. The global minimum of the error cost follows from the fact that all the output costs in output error coordinates have global minima at $(I, \mathring{y}_i)$ (30).

Let $\mathfrak{s}_i = \ker \mathrm{d}\rho^i_{\mathring{y}_i}(I)$ denote the Lie-algebra associated with $\mathrm{stab}_{\rho^i}(\mathring{y}_i)$. It is straightforward to verify that $\bigcap \mathrm{stab}_{\rho^i}(\mathring{y}_i) = \{I\}$ is equivalent to $\bigcap \mathfrak{s}_i = \{0\}$.

For each output cost $\ell^i_{\mathring{y}_i}(E_{\mathrm{I}}, \mathring{y}_i)$ in error coordinates, compute the differential

$$
\begin{aligned}
\mathrm{d}_1\ell^i_{\mathring{y}_i}(E_\mathrm{I},\mathring{y}_i) &= \mathrm{D}_{E_\mathrm{I}}\ell^i_{\mathring{y}_i}(I,\rho^i(E_\mathrm{I}^{-1},\mathring{y}_i)) \\
&= \mathrm{d}_2\ell^i_{\mathring{y}_i}(I,e^\mathrm{I})\mathrm{D}_{E_\mathrm{I}}\rho^i_{\mathring{y}_i}(E_\mathrm{I}^{-1}) \qquad (32)\\
&= \mathrm{d}_2\ell^i_{\mathring{y}_i}(I,e^\mathrm{I})\mathrm{d}\rho^i_{\mathring{y}_i}(E_\mathrm{I}^{-1})\mathrm{D}_{E_\mathrm{I}}\,\mathrm{inv}(E_\mathrm{I})\\
&= -\mathrm{d}_2\ell^i_{\mathring{y}_i}(I,e^\mathrm{I})\mathrm{d}\rho^i_{\mathring{y}_i}(E_\mathrm{I}^{-1})\mathrm{d}L_{E_\mathrm{I}^{-1}}(I)\mathrm{d}R_{E_\mathrm{I}^{-1}}(E_\mathrm{I}) \qquad (33)
\end{aligned}
$$

where d_1 and d_2 indicate differential with respect to the first and second argument respectively, $\mathrm{D}_{E_\mathrm{I}}$ is the Fréchet derivative with respect to the argument E_I, $\mathrm{inv}(E) = E_\mathrm{I}^{-1}$ and we use the well known formula $\mathrm{d}\ \mathrm{inv}(X) = -\mathrm{d}L_{X^{-1}}(I)\circ\mathrm{d}R_{X^{-1}}(X)$. Evaluating this expression at $E_\mathrm{I} = I$ one obtains

$$\mathrm{d}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i) = -\mathrm{d}_2\ell^i_{\mathring{y}_i}(I,\mathring{y}_0)\mathrm{d}\rho^i_{\mathring{y}_i}(I).$$

Since $(I,\mathring{y}_i)$ is a global minimum of $\ell^i_{\mathring{y}_i}$ then $\mathrm{d}_2\ell^i_{\mathring{y}_i}(I,\mathring{y}_i) = 0$ and $\mathrm{d}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i) = 0$ as expected at a global minimum.

The Hessian operator tensors

$$
\begin{aligned}
&\mathrm{Hess}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i)\colon T_I\mathbf{G}\to T_I^*\mathbf{G}\\
&\mathrm{Hess}_2\ell^i_{\mathring{y}_i}(I,\mathring{y}_i)\colon T_{\mathring{y}_i}\mathcal{X}\to T_{\mathring{y}_i}^*\mathcal{X}
\end{aligned}
$$

that map tangent vectors to co-vectors are intrinsically defined at a critical point of the cost (Absil et al. [2008]). From the above derivation, standard computation shows that the relationship between the Hessian operators is given by

$$\mathrm{Hess}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i) = \mathrm{d}\rho^i_{\mathring{y}_i}(I)^\star\mathrm{Hess}_2\ell^i_{\mathring{y}_i}(I,\mathring{y}_0)\mathrm{d}\rho^i_{\mathring{y}_i}(I)$$

where $\mathrm{d}\rho^i_{\mathring{y}_i}(I)^\star$ denotes the pull back $\mathrm{d}\rho^i_{\mathring{y}_i}(I)^\star\colon T^\star_{\mathring{y}_i}\mathcal{X}\to T^\star_I\mathbf{G}$ induced by $\mathrm{d}\rho^i_{\mathring{y}_i}$. Since $\mathrm{Hess}_2\ell^i_{\mathring{y}_i}(I,\mathring{y}_0)$ is positive definite then it follows that $\mathrm{Hess}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i)$ is positive semi-definite with kernel

$$\ker\mathrm{Hess}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i) = \ker\mathrm{d}\rho^i_{\mathring{y}_i}(I) = \mathfrak{s}_i.$$

One has that

$$\mathrm{Hess}_1\ell_{\mathring{y}}(I,\mathring{y}) = \sum_{i=1}^n \mathrm{Hess}_1\ell^i_{\mathring{y}_i}(I,\mathring{y}_i).$$

Since $\bigcap\mathfrak{s}_i = \{0\}$, it follows that $\mathrm{Hess}_1\ell_{\mathring{y}}(I,\mathring{y}) > 0$ is positive definite. $\triangle$

For type I systems (with constant parameters $\mathring{y}$) the error cost $\ell_{\mathring{y}}(\cdot,\mathring{y})$ on the $\mathbf{G}$-torsor $\mathcal{G}$ can be thought of directly as a candidate Lyapunov function for the design of the observer. The design process can be undertaken explicitly working with the algebraic structure of the problem or tackled using the general theory presented later in the section. It is illustrative to consider an example first, and show how a practical problem can be approached using classical Lyapunov design methodology and simple algebraic manipulations.

Example 1.6. Recall the scenario described in Example 1.1. The reference output is $\mathring{y} = \mathbf{e}_3$ with measurement $y = \rho(X,\mathbf{e}_3) = X^\top\mathbf{e}_3$ and output error $e^\mathrm{I} = \hat{X}y = E_\mathrm{I}\mathbf{e}_3$.

We propose an output cost based on the chordal distance on the sphere $\mathrm{S}^2\subset\mathbb{R}^3$,

$$\ell_{\mathbf{e}_3}(\hat{X},y) = \frac{1}{2}\|\hat{X}y-\mathbf{e}_3\|^2.$$

The invariance of the cost is a consequence of the invariance of the Euclidean norm operator with respect to rotation

$$\frac{1}{2}\|Q\hat{X}y - Q\mathbf{e}_3\|^2 = \frac{1}{2}\|Q(\hat{X}y-\mathbf{e}_3)\|^2 = \frac{1}{2}\|\hat{X}y-\mathbf{e}_3\|^2$$

for any $Q\in\mathrm{SO}(3)$. Applying the rotation $\hat{X}^{-1}$, it follows that

$$\ell_{\mathbf{e}_3}(\hat{X},y) = \frac{1}{2}\|e^\mathrm{I}-\mathbf{e}_3\|^2 = \frac{1}{2}\|E_\mathrm{I}\mathbf{e}_3-\mathbf{e}_3\|^2 = \ell_{\mathbf{e}_3}(E_\mathrm{I},\mathbf{e}_3)$$

as expected. Clearly $\ell_{\mathbf{e}_3}$ has a unique minimum at the point $e^\mathrm{I} = \mathbf{e}_3$ corresponding to $E_\mathrm{I} = I$. It is straightforward to see that the cost is non-degenerate since it is the restriction of a non-degenerate quadratic function on $\mathbb{R}^{3\times3}$ to $\mathrm{SO}(3)$.

Taking the time-differential of $\ell_{\mathbf{e}_3}(E_\mathrm{I},\mathbf{e}_3)$ along error dynamics $\dot{E}_\mathrm{I} = -\Delta E_\mathrm{I}$ with $\Delta\in\mathfrak{so}(3)$ yields

$$
\begin{aligned}
\frac{\mathrm{d}}{dt}\ell_{\mathbf{e}_3}(E_\mathrm{I},\mathbf{e}_3) =&(E_\mathrm{I}\mathbf{e}_3-\mathbf{e}_3)^\top\dot{E}_\mathrm{I}\mathbf{e}_3\\
=&-(E_\mathrm{I}\mathbf{e}_3-\mathbf{e}_3)^\top\Delta E_\mathrm{I}\mathbf{e}_3\\
=&-(e^\mathrm{I}-\mathbf{e}_3)^\top\Delta e^\mathrm{I}\\
=&-\mathrm{tr}\left(e^\mathrm{I}(e^\mathrm{I}-\mathbf{e}_3)^\top\Delta\right).
\end{aligned}
$$

Define the orthogonal projection onto the skew-symmetric matrices with respect to the trace inner product $\mathrm{tr}(Z_1^\top Z_2)$ on $\mathbb{R}^{n\times n}$, $\mathbb{P}_{\mathfrak{so}(3)}\colon\mathbb{R}^{3\times3}\to\mathfrak{so}(3)$, by

$$\mathbb{P}_{\mathfrak{so}(3)}(Z) := \frac{1}{2}(Z-Z^\top).$$

Since $\Delta\in\mathfrak{so}(3)$ then

$$\mathrm{tr}\left(e^\mathrm{I}(e^\mathrm{I}-\mathbf{e}_3)^\top\Delta\right) = \mathrm{tr}\left(\mathbb{P}_{\mathfrak{so}(3)}\left(e^\mathrm{I}(e^\mathrm{I}-\mathbf{e}_3)^\top\right)\Delta\right).$$

This leads us to propose the innovation

$$
\begin{aligned}
\Delta :=& k\left[\mathbb{P}_{\mathfrak{so}(3)}\left(e^\mathrm{I}(e^\mathrm{I}-\mathbf{e}_3)^\top\right)\right]^\top\\
=& k\frac{1}{2}\left(-e^\mathrm{I}\mathbf{e}_3^\top + \mathbf{e}_3{e^\mathrm{I}}^\top\right)^\top\\
=& \frac{k}{2}\left(-\mathbf{e}_3{e^\mathrm{I}}^\top + e^\mathrm{I}\mathbf{e}_3^\top\right) = \frac{k}{2}(\mathbf{e}_3\times e^\mathrm{I})_\times
\end{aligned}
$$

for $k>0$ a scalar gain and where the last line follows from the identity $uw^\top - wu^\top = (w\times u)_\times$ for $u,w\in\mathbb{R}^3$. Note that the innovation Δ is implementable since it is a function of known variables e^I and $\mathbf{e}_3$. Written in terms of the observer state $\hat{X}$ and measurement y one has

$$\Delta_{\mathbf{e}_3}(\hat{X},y) = \frac{k}{2}(\mathbf{e}_3\times\hat{X}y)_\times. \qquad (34)$$

To see that the innovation is equivariant compute

$$
\begin{aligned}
\Delta_{\mathbf{e}_3}(R_Q\hat{X},\rho(Q,y)) &= \Delta_{\mathbf{e}_3}(\hat{X}Q,Q^\top y))\\
&= \frac{k}{2}(\mathbf{e}_3\times\hat{X}QQ^\top y)_\times\\
&= \frac{k}{2}(\mathbf{e}_3\times\hat{X}y)_\times = \Delta_{\mathbf{e}_3}(\hat{X},y).
\end{aligned}
$$

The observer that we consider is then, using (21), (13) and (34),

$$\dot{\hat{X}} = \hat{X}\Omega_\times - \frac{k}{2}(\mathbf{e}_3\times\hat{X}y)_\times\hat{X}, \qquad (35a)$$

$$\hat{x}(t) = \hat{X}(t)\mathbf{e}_3. \qquad (35b)$$

It follows from Theorem 1 that the error dynamics are

$$\dot{E} = -\frac{k}{2}(\mathbf{e}_3\times E_\mathrm{I}\mathbf{e}_3)_\times E_\mathrm{I}.$$

With this choice it is easily verified that

$$\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathbf{e}_3}(E_\mathrm{I},\mathbf{e}_3) = -\frac{k}{2}\|(\mathbf{e}_3\times e^\mathrm{I})_\times\|_F^2 = -k\|\mathbf{e}_3\times e^\mathrm{I}\|^2.$$

From Lyapunov theory, see e.g. Khalil [1996], it follows that E_I converges to the largest forward invariant set contained in $\{E_\mathrm{I}\in\mathrm{SO}(3)\mid\mathbf{e}_3\times E_\mathrm{I}\mathbf{e}_3 = 0\}$. $\triangle$

A key aspect of the proposed design methodology is that the explicit Lyapunov design process is relatively straightforward to undertake and leads to effective observer construction. We go on to show that this approach to observer design will always yield an implementable observer construction by considering the general case.

Consider the time differential of $\ell_{\mathring{y}}(\cdot, \mathring{y})$ along solutions of a type I system with observer (21) and for constant reference outputs $\mathring{y}$. One has

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) &= \mathrm{d}_1\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) \left[\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}}\right] \\
&= -\sum_{i=1}^{n} \mathrm{D}_{E_{\mathrm{I}}}\ell^i_{\mathring{y}_i}(I, \rho^i_{\mathring{y}_i}(E_{\mathrm{I}}^{-1})) \left[\mathrm{d}R_{E_{\mathrm{I}}}(I)\Delta\right] \\
&= -\sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{D}_{E_{\mathrm{I}}}\rho^i_{\mathring{y}_i}(E_{\mathrm{I}}^{-1}) \left[\mathrm{d}R_{E_{\mathrm{I}}}(I)\Delta\right] \\
&= -\sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{\mathring{y}_i}(E_{\mathrm{I}}^{-1})\mathrm{D}_{E_{\mathrm{I}}}\,\mathrm{inv}(E_{\mathrm{I}}) \left[\mathrm{d}R_{E_{\mathrm{I}}}(I)\Delta\right] \\
&= \sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{\mathring{y}_i}(E_{\mathrm{I}}^{-1})\mathrm{d}L_{E_{\mathrm{I}}^{-1}}(I)\left[\Delta\right] \\
&= \sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{e^{\mathrm{I}}_i}(I)\left[\Delta\right]. \qquad (36)
\end{aligned}
$$

The transition from line 4 to 5 is based on the derivation in (32) to (33) along with the cancelation $\mathrm{d}R_{E_{\mathrm{I}}^{-1}}\mathrm{d}R_{E_{\mathrm{I}}} = \mathrm{id}$. For the final line, observe that

$$
\begin{aligned}
\mathrm{D}_Y[\rho(XY, z)] &= \mathrm{D}_Y[\rho(Y, \rho(X, z))] \\
&= \mathrm{d}_1\rho(Y, \rho(X, z)) \\
&= \mathrm{d}\rho_{\rho(X,z)}(Y)
\end{aligned}
$$

and

$$
\begin{aligned}
\mathrm{D}_Y[\rho(XY, z)] &= \mathrm{D}_Y[\rho(L_X(Y), z)] \\
&= \mathrm{d}_1\rho(L_X(Y), z) \circ \mathrm{d}L_X(Y) \\
&= \mathrm{d}\rho_z(XY) \circ \mathrm{d}L_X(Y).
\end{aligned}
$$

It follows that

$$
\mathrm{d}\rho_z(XY) \circ \mathrm{d}L_X(Y) = \mathrm{d}\rho_{\rho(X,z)}(Y).
$$

Setting $X = E_{\mathrm{I}}^{-1}$, $Y = I$ and $z = \mathring{y}_i$ verifies (36).

Equation (36) is of critical importance in the design of observers since the resulting relationship for $\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathring{y}}$ is expressed entirely in terms of known variables. The individual cost functions $\ell^i_{\mathring{y}_i}$ are known, and their differentials with respect to the second argument $\mathrm{d}_2\ell^i_{\mathring{y}_i}(\mathring{y}_i, e^{\mathrm{I}})$ can be computed. The error e^{I} is implementable, $\mathring{y}_i$ is known and the first term in the expansion can be computed. The group action ρ^i is known and the output y_i is measured so that the second term is also implementable.

To provide a general framework for the design of the innovation Δ based on (36) we need to provide a methodology to take the differential information $\sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{e^{\mathrm{I}}_i}(I)$, that can be thought of as an element of the dual $\mathfrak{g}^\star$ to the Lie-algebra $\mathfrak{g}$, and map this to an element $\Delta \in \mathfrak{g}$. In Example 1.6 this correspondence was undertaken algebraically using the structure of the trace operator and the projection $\mathbb{P}_{\mathfrak{so}}(3)$. In general, an elegant and well motivated approach to lifting a differential to a tangent vector is done by defining a Riemannian metric and using the gradient construction.

Remark 2. It is worth noting that gradient innovation construction is not the only option for observer design. A more general structure would be to consider general gain maps

$$
K_{\mathring{y}}(\hat{X}, y) \colon \mathfrak{g}^\star \to \mathfrak{g}.
$$

This is the situation encountered in recent work on deterministic optimal observer design (Zamani et al. [2013]). Further discussion of general gain maps is beyond the scope of the present work. △

Let $\langle\cdot,\cdot\rangle \colon \mathfrak{g}\times\mathfrak{g} \to \mathbb{R}$ be a positive definite inner product on $\mathfrak{g}$. It is straightforward to place a right invariant Riemannian metric on $\mathbf{G}$ induced by the inner product $\langle\cdot,\cdot\rangle$,

$$
\langle \mathrm{d}R_S V, dR_S W\rangle_S = \langle V, W\rangle
$$

where $\mathrm{d}R_S V$ and $\mathrm{d}R_S W \in T_S\mathbf{G}$ are arbitrary elements of the tangent space of $\mathbf{G}$ at S expressed as right translations of elements $V, W \in \mathfrak{g}$ of the Lie-algebra.

Remark 3. For a matrix Lie-group, with matrix Lie-algebra, a simple and effective choice of inner product is the trace inner product

$$
\langle V, W\rangle := \mathrm{tr}(V^\top W).
$$

△

By identification, the above construction induces an invariant Riemannian metric on the $\mathbf{G}$-torsor $\mathcal{G}$ where invariance is to be interpreted with respect to the natural $\mathbf{G}$-action. The *gradient* of a function $f \colon \mathcal{G} \to \mathbb{R}$, denoted $\mathrm{grad}\, f$, is defined implicitly by

$$
\langle \mathrm{grad}\, f(S), \mathrm{d}R_S W\rangle = \mathrm{D}_S f(S)[\mathrm{d}R_S W],
$$

for all $W \in \mathfrak{g} \equiv T_I\mathcal{G}$ (or equivalently $\mathrm{d}R_S W \in T_S\mathcal{G}$), and for $\mathrm{grad}\, f(S) \in T_S\mathcal{G}$.

Consider the derivative $\mathrm{D}_{E_{\mathrm{I}}}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$ of the aggregate cost taken in (the general) direction $\mathrm{d}R_{E_{\mathrm{I}}}W \in T_{E_{\mathrm{I}}}\mathcal{G}$ where $W \in \mathfrak{g}$. An analogous computation to (36) yields

$$
\begin{aligned}
\mathrm{D}_{E_{\mathrm{I}}}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})[\mathrm{d}R_{E_{\mathrm{I}}}W] &= \mathrm{D}_{E_{\mathrm{I}}}\ell_{\mathring{y}}(I, \rho^i_{\mathring{y}_i}(E_{\mathrm{I}}^{-1}, \mathring{y}))[\mathrm{d}R_{E_{\mathrm{I}}}W] \\
&= -\sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{e^{\mathrm{I}}_i}(I)\left[W\right]. \qquad (37)
\end{aligned}
$$

The gradient with respect to the first variable E_{I} of the cost $\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$ is then the solution to the implicit relationship

$$
\begin{aligned}
&\langle \mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}), \mathrm{d}R_{E_{\mathrm{I}}}W\rangle \\
&\qquad = -\sum_{i=1}^{n} \mathrm{d}_2\ell^i_{\mathring{y}_i}(I, e^{\mathrm{I}})\mathrm{d}\rho^i_{e^{\mathrm{I}}_i}(I)\left[W\right] \qquad (38)
\end{aligned}
$$

for arbitrary $W \in \mathfrak{g}$ and $\mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) \in T_{E_{\mathrm{I}}}\mathcal{G}$. Looking at equations (36) and (38), a good gradient innovation fulfils

$$
\Delta(E_{\mathrm{I}}, \mathring{y}) = k\mathrm{d}R_{E_{\mathrm{I}}^{-1}}\, \mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}). \qquad (39)
$$

This choice will lead to a decrease of the cost

$$
\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) &= -\,k\langle \mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}), \mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})\rangle \\
&= -\,k\|\, \mathrm{grad}_1\, \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})\|^2,
\end{aligned}
$$

providing the basis for the stability analysis of the observer given in Theorem 2. However, before this analysis is undertaken we still need to actually define an equivariant innovation according to Definition 6. To this end we simply extend the definition suggested by (39) to

$$
\Delta_{\mathring{y}}(\hat{X}, y) := k\mathrm{d}R_{\hat{X}^{-1}}\, \mathrm{grad}_1\, \ell_{\mathring{y}}(\hat{X}, y) \qquad (40)
$$

for arbitrary $\hat{X} \in \mathcal{G}$ and $y \in \mathcal{Y}$.

Lemma 4. Consider a type I lifted equivariant system (Definition 5) with an invariant aggregate cost $\ell_{\mathring{y}}(\hat{X}, y)$ (Equation (31)). Let $\langle \cdot, \cdot \rangle$ be a right invariant Riemannian metric on $\mathcal{G}$. Then the innovation given by (40) is equivariant (Definition 6) and

$$\Delta_{\mathring{y}}(\hat{X}, y) = \Delta_{\mathring{y}}(I, e^{\mathrm{I}}) = \Delta_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}).$$

Proof: From (30) one has that for all $W \in \mathfrak{g}$

$$\begin{aligned}
&\langle \operatorname{grad}_1 \ell^i_{\mathring{y}_i}(R_S\hat{X}, \rho^i(S, y_i)), \mathrm{d}R_{R_S\hat{X}}W \rangle \\
&\quad = \mathrm{d}_1\ell^i_{\mathring{y}_i}(R_S\hat{X}, \rho^i(S, y_i)) \left[\mathrm{d}R_{R_S\hat{X}}W\right] \\
&\quad = \mathrm{d}_1\ell^i_{\mathring{y}_i}(R_S\hat{X}, \rho^i(S, y_i)) \circ \mathrm{d}R_S \left[\mathrm{d}R_{\hat{X}}W\right] \\
&\quad = \mathrm{D}_{\hat{X}}\ell^i_{\mathring{y}_i}(R_S\hat{X}, \rho^i(S, y_i)) \left[\mathrm{d}R_{\hat{X}}W\right] \\
&\quad = \mathrm{D}_{\hat{X}}\ell^i_{\mathring{y}_i}(\hat{X}, y_i) \left[\mathrm{d}R_{\hat{X}}W\right] \\
&\quad = \mathrm{d}_1\ell^i_{\mathring{y}_i}(\hat{X}, y_i) \left[\mathrm{d}R_{\hat{X}}W\right] \\
&\quad = \langle \operatorname{grad}_1 \ell^i_{\mathring{y}_i}(\hat{X}, y_i), \mathrm{d}R_{\hat{X}}W \rangle \\
&\quad = \langle \mathrm{d}R_S \operatorname{grad}_1 \ell^i_{\mathring{y}_i}(\hat{X}, y_i), \mathrm{d}R_S\mathrm{d}R_{\hat{X}}W \rangle \\
&\quad = \langle \mathrm{d}R_S \operatorname{grad}_1 \ell^i_{\mathring{y}_i}(\hat{X}, y_i), \mathrm{d}R_{R_S\hat{X}}W \rangle.
\end{aligned}$$

This implies

$$\operatorname{grad}_1 \ell^i_{\mathring{y}_i}(R_S\hat{X}, \rho^i(S, y_i)) = \mathrm{d}R_S \operatorname{grad}_1 \ell^i_{\mathring{y}_i}(\hat{X}, y_i).$$

Using the suggestive notation

$$\rho(S, y) := (\rho^1(S, y_1), \ldots, \rho^p(S, y_p)),$$

it then follows from (40) that

$$\begin{aligned}
&\Delta_{\mathring{y}}(R_S\hat{X}, \rho(S, y)) \\
&\quad = k\mathrm{d}R_{(R_S\hat{X})^{-1}} \operatorname{grad}_1 \ell_{\mathring{y}}(R_S\hat{X}, \rho(S, y)) \\
&\quad = k\mathrm{d}R_{\hat{X}^{-1}}\mathrm{d}R_{S^{-1}}\mathrm{d}R_S \operatorname{grad}_1 \ell_{\mathring{y}}(\hat{X}, y) \\
&\quad = k\mathrm{d}R_{\hat{X}^{-1}} \operatorname{grad}_1 \ell_{\mathring{y}}(\hat{X}, y) \\
&\quad = \Delta_{\mathring{y}}(\hat{X}, y)
\end{aligned}$$

showing that the gradient innovation is equivariant.

In particular,

$$\begin{aligned}
\Delta_{\mathring{y}}(\hat{X}, y) &= \Delta_{\mathring{y}}(R_{\hat{X}^{-1}}E_{\mathrm{I}}, \rho(\hat{X}^{-1}, \mathring{y})) = \Delta_{\mathring{y}}(I, e^{\mathrm{I}}) \\
&= \Delta_{\mathring{y}}(R_X E_{\mathrm{I}}, \rho(X, \mathring{y})) = \Delta_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})
\end{aligned}$$

as claimed. △

For a type I system, the observer for a gradient innovation associated with an invariant aggregate cost function is

$$\dot{\hat{X}} = \mathrm{Ad}_{\hat{X}^{-1}}\left(F_{x_0}(v)\right)\hat{X} - k \operatorname{grad}_1 \ell_{\mathring{y}}(\hat{X}, y), \tag{41a}$$

$$\hat{x}(t) = \phi_{x_0}(\hat{X}(t; \hat{X}(0)) \tag{41b}$$

for $\hat{X}(0) \in \mathcal{G}$ some initial condition, typically $\hat{X}(0) = I$ and x_0 chosen as the best *a-priori* guess of $x(0)$. For this observer, then Lemma 4 and Theorem 1 show that the associated error dynamics (26) are of gradient type

$$\frac{\mathrm{d}}{\mathrm{d}t}E_{\mathrm{I}} = -\mathrm{d}R_{E_{\mathrm{I}}}\Delta = -k \operatorname{grad}_1 \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}).$$

Theorem 2. Consider a type I lifted equivariant system (Def. 5). Let $\ell_{\mathring{y}}$ denote an aggregate cost constructed from non-degenerate equivariant output costs according to Equation (31). Assume that

$$\bigcap \operatorname{stab}_{\rho^i}(\mathring{y}_i) = \{I\}.$$

Let $\langle \cdot, \cdot \rangle$ be an inner-product on $\mathfrak{g}$ that induces a right-invariant Riemannian metric on $\mathcal{G}$. Define an innovation by (39) and consider the observer system (21a) resp. (41a). Then there exists a basin of attraction $\mathcal{B} \subseteq \mathcal{G}$ containing I such that for any initial conditions $X(0)$ and $\hat{X}(0) \in \mathcal{G}$ such that $E_{\mathrm{I}}(0) \in \mathcal{B}$ then $E_{\mathrm{I}}(t) \to I$ and $\hat{X}(t) \to X(t)$. Moreover, if $\phi_{x_0}(X(0)) = x(0)$, then $\hat{x}(t) = \phi_{x_0}(\hat{X}) \to x(t; x_0)$. △

Proof: The time differential of $\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})$ along the error flow (26) yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) = -k\| \operatorname{grad}_1 \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y})\|^2 \tag{42}$$

where the norm $\| \cdot \|$ is the norm on $T_{E_{\mathrm{I}}}\mathcal{G}$ induced by the right invariant Riemannian metric.

Since the error cost $\ell_{\mathring{y}}(\cdot, \mathring{y})$ is non-degenerate (Lemma 3), there exists an open neighbourhood of I in $\mathcal{G}$ where the cost has compact connected sub-level sets all containing the global minimum I and no other critical point of the cost. Let $\mathcal{B} \subseteq \mathcal{G}$ be the largest such sub-level set of the cost. Convergence of the error dynamics to the unique minimum follows from classical Lyapunov theory by noting that $\operatorname{grad}_1 \ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) = 0$ on $\mathcal{B}$ implies that $E_{\mathrm{I}} = I$. It follows trivially that $\hat{X}(t) \to X(t)$. The final statement is also straightforward to see given the projection properties of the observer and state flows. △

Example 1.7. Recall the scenario described in Example 1.1. The reference output is $\mathring{y} = \mathbf{e}_3$ with measurement $y = \rho(X, \mathbf{e}_3) = X^\top\mathbf{e}_3$ and output error $e^{\mathrm{I}} = \hat{X}y = E_{\mathrm{I}}\mathbf{e}_3$. The innovation is

$$\Delta_{\mathbf{e}_3}(\hat{X}, y) = \frac{k}{2}(\mathbf{e}_3 \times \hat{X}y)_\times$$

and the observer is given by (35).

In Example 1.6 it was shown that

$$\frac{\mathrm{d}}{\mathrm{d}t}\ell_{\mathbf{e}_3}(E_{\mathrm{I}}, \mathbf{e}_3) = -k\|\mathbf{e}_3 \times e^{\mathrm{I}}\|^2$$

and hence E_{I} converges to the largest forward invariant set contained in $\{E_{\mathrm{I}} \in \mathrm{SO}(3) \mid \mathbf{e}_3 \times E_{\mathrm{I}}\mathbf{e}_3 = 0\}$. The reason that we don't get convergence of $E_{\mathrm{I}} \to I$, as was the case in Theorem 2, is because with only one measurement y the lifted cost function

$$\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) = \frac{1}{2}\|E_{\mathrm{I}}\mathbf{e}_3 - \mathbf{e}_3\|^2$$

is degenerate. Indeed, since there is only a single measurement

$$\ker \mathrm{Hess}\ell_{\mathring{y}}(E_{\mathrm{I}}, \mathring{y}) = \ker \mathrm{d}\rho_{\mathbf{e}_3}(I) = \operatorname{span}\{(\mathbf{e}_3)_\times\}.$$

One has

$$\begin{aligned}
\operatorname{stab}_\rho(\mathbf{e}_3) &= \{E_{\mathrm{I}} \in \mathrm{SO}(3) \mid E_{\mathrm{I}}\mathbf{e}_3 = \mathbf{e}_3\} \\
&= \{E_{\mathrm{I}} \in \mathrm{SO}(3) \mid \mathbf{e}_3 \times E_{\mathrm{I}}\mathbf{e}_3 = 0\}
\end{aligned}$$

and one can see that the unobservable subspace of the error system is the stabilizer of the output group action. The proposed observer converges to $\operatorname{stab}_\rho(\mathbf{e}_3)$ but not necessarily to I.

Observe that

$$\operatorname{stab}_\phi(x_0) = \operatorname{stab}_\phi(\mathbf{e}_3) = \operatorname{span}\{(\mathbf{e}_3)_\times\},$$

the kernel of the group action $\phi_{\mathbf{e}_3} \colon \mathrm{SO}(3) \to \mathrm{S}^3$. Since $\operatorname{stab}_\phi(\mathbf{e}_3) \subset \operatorname{stab}_\rho(\mathbf{e}_3)$ then the unobservable subgroup of the error flow lies in the kernel of the group action projection and it follows that $\hat{x}(t) \to x(t)$.

The only way to ensure convergence of the error estimate on the group (for constant reference $\mathring{y}$) is by adding an additional measurement with reference output $\mathring{y}_2$ that is not co-linear with $\mathbf{e}_3$. However, a similar effect is obtained if the reference output $\mathring{y} := \mathring{y}(t)$ is time varying and observability can be recovered based on a persistence of excitation condition (Trumpf et al. [2012]). The error cost construction $\ell_{\mathring{y}}(\cdot, \mathring{y})$ and the Lyapunov argument given in Theorem 2 cannot be used for this analysis since it depends on the constant $\mathring{y}$ assumption. More details can be found in Trumpf et al. [2012]. $\triangle$

Remark 4. The general condition for the state $\hat{x}(t) \to x(t)$ is that

$$\mathfrak{k} \subseteq \bigcap_{i=1}^{p} \mathfrak{s}_i$$

for $\mathfrak{k} = \ker \mathrm{d}\phi_{x_0}(I)$ and $\mathfrak{s}_i = \ker \mathrm{d}\rho^i_{\mathring{y}_i}(I)$. $\triangle$

5. VELOCITY BIAS

Most of the common velocity sensor systems used by mobile robotic systems suffer from slowly time-varying disturbances caused by temperature or vibration sensitivity of the micro-electronic mechanical (MEMS) architecture used to generate the physical measurement or calibration error in more general sensors. It is of interest to consider how the design methodology proposed in this paper can be extended to deal with such a situation. This section presents the structure of a general nonlinear observer for type I systems with bias estimation. The development is based on work in progress (Khosravian et al. [2013]) and we present only the main formulas and intuition in the present paper.

Consider a velocity measurement $v \in \mathbb{V}$, modeled by

$$v = \mathring{v} + b$$

where $b \in \mathbb{V}$ is an unknown constant bias, and $\mathring{v} \in \mathbb{V}$ is the true system velocity. Define $B := F_{x_0}(b)$ to be a *constant* bias on the Lie-algebra for the lifted system that corresponds to the measurement bias b and let $\hat{B}$ denote an estimate for B.

For a system of type I (Def. 5) with innovation (39)

$$\Delta(\hat{X}, y) := k\mathrm{d}R_{\hat{X}^{-1}} \operatorname{grad}_1 \ell_{\mathring{y}}(\hat{X}, y)$$

the proposed observer (41) with bias estimation is

$$\dot{\hat{X}} = \mathrm{Ad}_{\hat{X}} \left(F_{x_0}(v) - \hat{B}\right) \hat{X} - \Delta(\hat{X}, y)\hat{X}, \tag{43a}$$

$$\dot{\hat{B}} = -\gamma \mathrm{Ad}^{\star}_{\hat{X}} \left(\Delta(\hat{X}, y)\right), \tag{43b}$$

$$\hat{x}(t) = \phi_{x_0}(\hat{X}(t; \hat{X}(0)), \tag{43c}$$

where $\gamma > 0$ is a positive constant and $\mathrm{Ad}^{\star}_{\hat{X}} : \mathfrak{g} \to \mathfrak{g}$ denotes the Hermitian adjoint of $\mathrm{Ad}_{\hat{X}}$ with respect to the inner product $\langle \cdot, \cdot \rangle$ on $\mathfrak{g}$ associated with the right-invariant Riemannian metric used. That is $\langle \mathrm{Ad}_{\hat{X}}(U_1), U_2 \rangle = \langle U_1, \mathrm{Ad}^{*}_{\hat{X}} U_2 \rangle$ for all $U_1, U_2 \in \mathfrak{g}$.

Example 1.8. Recall the scenario described in Example 1.1 and the discussion in Example 1.7. The reference output is $\mathring{y} = \mathbf{e}_3$ with measurement $y = \rho(X, \mathbf{e}_3) = X^{\top}\mathbf{e}_3$, output error $e^{\mathrm{I}} = \hat{X}y = E_{\mathrm{I}}\mathbf{e}_3$ and innovation $\Delta_{\mathbf{e}_3}(\hat{X}, y) = \frac{k}{2}(\mathbf{e}_3 \times \hat{X}y)_{\times}$.

The observer (43) has the form

$$\dot{\hat{X}} = \hat{X}\left(\Omega_{\times} - \hat{B}\right) - \frac{k}{2}(\mathbf{e}_3 \times \hat{X}y)_{\times}\hat{X},$$
$$\dot{\hat{B}} = -\gamma \mathrm{Ad}_{\hat{X}^{\top}} \left(\mathbf{e}_3 \times \hat{X}y\right)_{\times},$$

where $\mathrm{Ad}^{\star}_{\hat{X}} = \mathrm{Ad}_{\hat{X}^{\top}}$ since

$$\begin{aligned}
\langle \mathrm{Ad}_{\hat{X}}(U_1), U_2 \rangle &= \operatorname{tr}\left(\left[\mathrm{Ad}_{\hat{X}}(U_1)\right]^{\top} U_2\right) \\
&= \operatorname{tr}\left(\hat{X}U_1^{\top}\hat{X}^{\top}U_2\right) \\
&= \operatorname{tr}\left(U_1^{\top}\hat{X}^{\top}U_2\hat{X}\right) \\
&= \operatorname{tr}\left(U_1^{\top}\,\mathrm{Ad}_{\hat{X}^{\top}}(U_2)\right) \\
&= \langle U_1, \mathrm{Ad}^{*}_{\hat{X}} U_2 \rangle.
\end{aligned}$$

Since the map $F_{\mathbf{e}_3}$ is one-to-one, the bias evolution (43b) can be projected directly onto the velocity space $\mathbb{V}$

$$\dot{\hat{b}} = -\gamma(\hat{X}^{\top}\mathbf{e}_3) \times y \tag{44}$$

where the $\mathrm{Ad}_{\hat{X}^{\top}}$ has also been factored through the vector product. This observer ((43) with the projected bias evolution (44)) is the literature standard nonlinear attitude observer that has been extensively studied over the last ten years; Mahony et al. [2005], Bonnabel et al. [2006], Campolo et al. [2006], Maithripala et al. [2006], Metni et al. [2006], Kinsey and Whitcomb [2007], Martin and Salaün [2007], Tayebi et al. [2007], Mahony et al. [2008], Vasconcelos et al. [2008], Brás et al. [2011], Grip et al. [2012], Sanyal and Nordkvist [2012]. $\triangle$

Detailed stability proofs for the convergence of a nonlinear observer with bias estimate of this form for SO(3) and SE(3) are provided in application papers available in the literature. The proofs of these results are not as straightforward as that given for Theorem 2 since the resulting error dynamics, including the bias error $\tilde{B} = \hat{B} - B$, are no longer autonomous. Typical proofs use a combination of Lyapunov stability analysis and Barbalat's lemma to prove convergence to an invariant set, and then a separate stability analysis of the time-varying linearisation using the work of Morgan and Narendra [1977b,a] and Loría and Panteley [2002], or a uniform observability condition (Thienel and Sanner [2003]), to prove uniform local exponential stability at the global minimum. A fully general development for the case of matrix Lie-groups is given in Khosravian et al. [2013].

6. CONCLUSION

This paper has provided a general development of a full state observer design methodology for kinematic systems with complete symmetry. The approach proposed applies in particular to type I systems, as described in Definition 5 in the paper, a class of systems that includes a large range of applications in mobile robotics that have motivated the authors' work. The approach is simple and practical and has led to a range of highly effective observers for real world applications.

ACKNOWLEDGEMENTS

This research was supported by the Australian Research Council through Discovery Grant DP120100316 "Geometric observer theory for mechanical control systems" and by the SCAR ANR-ASTRID fund.

REFERENCES

P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds.* Princeton University Press, Princeton, NJ, USA, January 2008.

N. Aghannan and P. Rouchon. An intrinsic observer for a class of Lagrangian systems. *IEEE Transactions on Automatic Control*, 48(6):936–945, 2003.

A. Agrachev and Yu. Sachkov. *Control theory from the geometric viewpoint.* Springer, 2004.

B. D. O. Anderson and J. B. Moore. *Optimal filtering.* Prentice Hall, 1979.

N. Apostolou and D. Kazakos. On local observability for invariant systems on Lie groups and coset spaces. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 452 (1955):2801–2805, 1996.

G. Baldwin, R. Mahony, and J. Trumpf. A nonlinear observer for 6 DOF pose estimation from inertial and bearing measurements. In *Proceedings of the IEEE International Conference on Robotics and Automation (ICRA)*, pages 2237–2242, 2009.

A. M. Bloch. *Nonholonomic mechanics and control.* Springer, 2003.

S. Bonnabel, P. Martin, and P. Rouchon. A non-linear symmetry-preserving observer for velocity-aided inertial navigation. In *Proceedings of the American Control Conference*, pages 2910–2914, 2006.

S. Bonnabel, P. Martin, and P. Rouchon. Symmetry-preserving observers. *IEEE Transactions on Automatic Control*, 53(11):2514–2526, 2008.

S. Bonnabel, P. Martin, and P. Rouchon. Non-linear symmetry-preserving observers on Lie groups. *IEEE Transactions on Automatic Control*, 54(7):1709–1713, 2009.

W. M. Boothby. *An introduction to differentiable manifolds and Riemannian geometry.* Academic Press Inc., London, Britain, 1986.

S. Brás, R. Cunha, J. F. Vasconcelos, C. Silvestre, and P. Oliveira. A nonlinear attitude observer based on active vision and inertial measurements. *IEEE Transactions on Robotics*, 27(4):664–677, 2011.

R. W. Brockett. System theory on group manifolds and coset spaces. *SIAM Journal on Control*, 10(2):265–284, 1972.

R. W. Brockett. Lie theory and control systems defined on spheres. *SIAM Journal on Applied Mathematics*, 25(2): 213–225, 1973.

R. W. Brockett. Control theory and analytical mechanics. In C. Martin and R. Hermann, editors, *Geometric Control Theory*, volume VII of *Lie Groups: History, Frontiers and Applications*, pages 1–48. Math. Sci. Press, 1977.

F. Bullo and A. D. Lewis. *Geometric Control of Mechanical Systems.* Springer, 2004.

D Campolo, F. Keller, and E. Guglielmelli. Inertial/magnetic sensors based orientation tracking on the group of rigid body rotations with application to wearable devices. In *Proceedings of the IEEE/RSJ International Conference on Intelligent Robots and Systems*, 2006.

D. Cheng, W. P. Dayawansa, and C. F. Martin. Observability of systems on Lie groups and coset spaces. *SIAM Journal on Control and Optimization*, 28(3):570–581, 1990.

J. L. Crassidis, F. L. Markley, and Y. Cheng. Nonlinear attitude filtering methods. *Journal of Guidance, Control,and Dynamics*, 30(1):12–28, 2007.

A. Doucet, N. de Freitas, and N. Gordon. *An introduction to sequential Monte Carlo methods.* Springer, 2001.

H. F. Grip, T. I. Fossen, T. A. Johansen, and A. Saberi. Attitude estimation using biased gyro and vector measurements with time-varying reference vectors. *IEEE Transactions on Automatic Control*, 57(5):1332–1338, 2012.

A. Isidori. *Nonlinear control systems.* Springer, 2nd edition, 1999.

S. J. Julier and J. K. Uhlmann. New extension of the kalman filter to nonlinear systems. In *AeroSense'97*, pages 182–193. International Society for Optics and Photonics, 1997.

V. Jurdjevic. *Geometric Control Theory.* Cambridge University Press, 1997.

V. Jurdjevic and H. J. Sussmann. Control systems on Lie groups. *Journal of Differential Equations*, 12:313–329, 1972.

H. K. Khalil. *Nonlinear Systems.* Prentice Hall, New Jersey, U.S.A., second edition, 1996.

A. Khosravian, J. Trumpf, R. Mahony, and C. Lageman. Bias estimation for invariant systems on Lie groups with homogeneous outputs. Submitted to the IEEE Conference on Decision and Control, 2013.

J. C. Kinsey and L. L. Whitcomb. Adaptive identification on the group of rigid-body rotations and its application to underwater vehicle navigation. *IEEE Transactions on Robotics*, 23(1):124–136, 2007.

C. Lageman, J. Trumpf, and R. Mahony. Observers for systems with invariant outputs. In *Proceedings of the European Control Conference*, pages 4587–4592, 2009.

C. Lageman, J. Trumpf, and R. Mahony. Gradient-like observers for invariant dynamics on a Lie group. *IEEE Transactions on Automatic Control*, 55(2):367–377, 2010.

A. Loría and E. Panteley. Uniform exponential stability of linear time-varying systems: revisited. *Systems and Control Letters*, 47(1):13–24, 2002.

R. Mahony, T. Hamel, and J.-M. Pflimlin. Complementary filter design on the special orthogonal group SO(3). In *Proceedings of the IEEE Conference on Decision and Control*, pages 1477–1484, 2005.

R. Mahony, T. Hamel, and J.-M. Pflimlin. Non-linear complementary filters on the special orthogonal group. *IEEE Transactions on Automatic Control*, 53(5):1203–1218, 2008.

D. H. S. Maithripala, J. M. Berg, and W. P. Dayawansa. Almost-global tracking of simple mechanical systems on a general class of Lie groups. *IEEE Transactions on Automatic Control*, 51(2):216–225, 2006.

J. E. Marsden and T. S. Ratiu. *Introduction to mechanics and symmetry.* Springer, 2nd edition, 1999.

P. Martin and E. Salaün. Invariant observers for attitude and heading estimation from low-cost inertial and magnetic sensors. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 1039–1045, 2007.

N. Metni, J.-M. Pflimlin, T. Hamel, and P. Soueeres. Attitude and gyro bias estimation for a VTOL UAV. *Control Engineering Practice*, 14(12):1511–1520, 2006.

A.P Morgan and K.S. Narendra. On the stability of nonautonomous differential equations $\dot{x} = A + B(t)x$, with skew symmetric matrix $B(t)^*$. *SIAM Journal on Control and Optimization*, 15(1):163–176, 1977a.

A.P. Morgan and K.S. Narendra. On the uniform asymptotic stability of certain linear nonautonomous differential equations. *SIAM Journal on Control and Optimization*, 15(1):5–24, 1977b.

R. E. Mortensen. Maximum-likelihood recursive nonlinear filtering. *Journal of Optimization Theory and Applications*, 2(6):386–394, 1968.

H. Nijmeijer and T. I. Fossen, editors. *New Directions in Nonlinear Observer Design*, volume 244 of *Lecture Notes in Control and Information Sciences*. Springer, 1999.

H. Nijmeijer and A. van der Schaft. *Nonlinear dynamical control systems*. Springer, 1990.

H. Rehbinder and B. K. Ghosh. Pose estimation using line-based dynamic vision and inertial sensors. *IEEE Transactions on Automatic Control*, 48(2):186–199, 2003.

E. Salaün and P. Martin. Design and implementation of a low-cost observer-based attitude and heading reference system. *Control Engineering Practice*, 18(7):712–722, 2010.

S. Salcudean. A globally convergent angular velocity observer for rigid body motion. *IEEE Transactions on Automatic Control*, 46(12):1493–1497, 1991.

A. K. Sanyal and N. Nordkvist. Attitude state estimation with multi-rate measurements for almost global attitude feedback tracking. *Journal of Guidance, Control, and Dynamics*, 35(3):868–880, 2012.

A. Tayebi, S. McGilvray, A. Roberts, and M. Moallem. Attitude estimation and stabilization of a rigid body using low-cost sensors. In *Proceedings of the 46th IEEE Conference on Decision and Control*, pages 6424–6429, 2007.

J. Thienel and R. M. Sanner. A coupled nonlinear spacecraft attitude controller and observer with an unknow constant gyro bias and gyro noise. *IEEE Transactions on Automatic Control*, 48(11):2011–2015, 2003.

J. Trumpf, R. Mahony, T. Hamel, and C. Lageman. Analysis of non-linear attitude observers for time-varying reference measurements. *IEEE Transactions on Automatic Control*, 57:2789–2800, 2012.

J. F. Vasconcelos, C. Silvestre, and P. Oliveira. A nonlinear observer for rigid body attitude estimation using vector observations. In *Proceedings of the 17th IFAC World Congress*, pages 8599–8604, 2008.

J. F. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira. A nonlinear position and attitude observer on SE(3) using landmark measurements. *Systems and Control Letters*, 59(3–4):155–166, 2010.

B. Vik and T. Fossen. A nonlinear observer for GPS and INS integration. In *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 2956–2961, 2001.

M. Zamani, J. Trumpf, and R. Mahony. Minimum-energy filtering for attitude estimation. *IEEE Transactions on Automatic Control*, 2013. Accepted for publication 04 April 2013.

