# State observers for invariant dynamics on a Lie group

*C. Lageman*[*], *R. Mahony*[†], *J. Trumpf*[‡]

## 1 Introduction

This paper concerns the design of full state observers for state space systems where the state is evolving on a finite dimensional, connected Lie group.

Traditional full state observer and filter designs for systems evolving in vector spaces employ the following design principle, going back to the work of Kalman [6] and Luenberger [7] on linear systems. The observer system is designed as a combination of a copy of the system, i.e. a part that can in principle replicate the observed system's trajectory, plus an innovation term which serves to drive the observer trajectory towards the correct system trajectory in the presence of initialization or measurement errors.

We propose an observer design based on a split of the observer dynamics into a synchronous term producing constant error estimates and an innovation term that reduces the error. We show that there is a canonical way of defining error measures in this context. Under mild assumptions, we prove almost global exponential convergence of the resulting observer and demonstrate its utility for a practical pose estimation scenario, a problem that has received strong interest in recent years [1, 2, 8, 9, 10, 11].

Previous general theoretical work in the area has concentrated on invariant observer design [3, 5]. While it is natural to require the same invariance properties of an observer as those obeyed by the observed system, some of the most successful existing observer designs do not share these properties [4]. In this paper, we

[*]Centre for Mathematics and its Applications, The Australian National University, lageman@maths.anu.edu.au. Work supported by the Australian Research Council Centre of Excellence for Mathematics and Statistics of Complex Systems.

[†]Department of Engineering, The Australian National University, Robert.Mahony@anu.edu.au.

[‡]Department of Information Engineering, The Australian National University, Jochen.Trumpf@anu.edu.au.

systematically study the reason for this discrepancy, namely the (in general) non-abelian nature of the underlying Lie group, and how this influences the observer error dynamics.

## 2 Notation and problem formulation

Let $G$ be a finite dimensional, connected Lie group with Lie algebra $\mathfrak{g}$. We denote the identity element in $G$ by $e$, and the left and right multiplication with an element $X \in G$ by $L_X$ and $R_X$, respectively. We use the representation of the tangent bundle of $G$ by left or right translations of the Lie algebra, i.e. $T_eL_X\mathfrak{g}$ or $T_eR_X\mathfrak{g}$ for the tangent space $T_XG$ of $G$ at $X$. We use the simplified notation $Xv$ for vectors $T_eL_Xv \in T_XG$ and $vX$ for vectors $T_eR_Xv \in T_XG$ with $v \in \mathfrak{g}$. Furthermore, we assume that there is a Riemannian metric $\langle\cdot,\cdot\rangle$ on $G$. We do not make any invariance assumptions on this metric yet.

Consider a left invariant system on $G$ of the form

$$\dot{X} = Xu, \tag{1}$$

where $u\colon \mathbb{R} \to \mathfrak{g}$ is an input function. Here, we call an input $u$ of system (1) *admissible* if solutions of the system are unique, exist for all time and are sufficiently smooth.

This paper discusses observer design for a situation where we have (potentially noisy) measurements of $X$ and $u$ and want to build an observer that estimates $X$.

**Example 1.** *As an example we consider the special orthogonal group* $\mathrm{SO}(3)$. *We choose the representation of* $\mathrm{SO}(3)$ *by real, orthogonal* $3\times 3$ *matrices, denoted by* $R$ *or* $S$. *Its Lie algebra* $\mathfrak{so}(3)$ *is given by the real, skew-symmetric* $3\times 3$ *matrices, denoted by* $\Omega$. *The tangent spaces* $T_R\,\mathrm{SO}(3)$ *are identified with*

$$\{R\Omega \mid \Omega \in \mathfrak{so}(3)\} \subset \mathbb{R}^{3\times 3}.$$

*We equip* $\mathrm{SO}(3)$ *with the Riemannian metric induced by the Euclidean one on* $\mathbb{R}^{3\times 3}$, *i.e.* $\langle R, S\rangle = \mathrm{tr}(R^\top S)$. *On* $\mathrm{SO}(3)$ *we have the system*

$$\dot{R} = R\Omega,$$

*with* $R\Omega$ *denoting the matrix product, and* $\Omega\colon \mathbb{R} \to \mathfrak{so}(3)$ *admissible. This system models the kinematics of the attitude* $R$ *of a coordinate frame fixed to a rigid body in 3D-space relative to an inertial frame. Here,* $\Omega$ *encodes the angular velocity measured in the body-fixed frame.*

*In applications, measurements of* $R$ *are for example provided by a vision system, while measurements of* $\Omega$ *are obtained from on board gyrometers.*

The proposed observer design will consist of a *synchronous* term and an *innovation* term. These notions are introduced in the next two sections.

## 3 Synchrony and error functions

In this section we introduce the general concept of *synchrony* between pairs of systems that evolve on a given Lie group $G$ and have a common input. Synchrony refers to an *error function* $E$, quantifying the instantaneous difference of trajectories between the two systems. The pair of systems is called *$E$-synchronous* if the error $E$ is constant along corresponding trajectories. This concept generalizes the concept of copies of the same system. We show that synchronous systems are necessarily copies of each other if certain invariance conditions hold for one of the systems and for the error function.

**Definition 2.** *Consider a pair of systems on $G$*

$$\dot{X} = F_X(X, u, t), \tag{2}$$

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, u, t) \tag{3}$$

*with $F_X, F_{\hat{X}} \colon G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_X(X, u, t) \in T_X G$ and $F_{\hat{X}}(\hat{X}, u, t) \in T_{\hat{X}} G$. Let $E \colon G \times G \to M$ be a smooth* error function*, $M$ a smooth manifold. We call* (2) *and* (3) $E$-synchronous *if for all admissible $u \colon \mathbb{R} \to \mathfrak{g}$, all initial values $X_0, \hat{X}_0 \in G$ of* (2) *and* (3) *and all $t \in \mathbb{R}$ we have*

$$\frac{\mathrm{d}}{\mathrm{d}t} E(\hat{X}(t, \hat{X}_0, u), X(t, X_0, u)) = 0,$$

*where $X(t, X_0, u)$, $\hat{X}(t, \hat{X}_0, u)$ denote the solutions of* (2) *and* (3)*.*

Two particularly simple error functions on a Lie group $G$ are the canonical *right invariant* error

$$E_r(\hat{X}, X) = \hat{X} X^{-1}$$

and the canonical *left invariant* error

$$E_l(\hat{X}, X) = X^{-1} \hat{X}.$$

Here, the label “invariant” refers to simultaneous state space transformations of both systems, e.g. we have $E_r(\hat{X}S, XS) = E_r(\hat{X}, X)$ for all $X, \hat{X}, S \in G$. We now specialize our notion of synchrony to the canonical errors defined above.

**Definition 3.** *We call the pair of systems* (2) *and* (3) right synchronous *if they are $E_r$-synchronous and* left synchronous *if they are $E_l$-synchronous.*

We show that left invariant systems admit precisely one right synchronous “partner”, allowing a canonical choice for the synchronous part of our observer design.

**Theorem 4.** *Consider the left invariant system* (1) *and let a second system be given by*

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, u, t). \tag{4}$$

*The systems* (1) *and* (4) *are right synchronous if and only if*

$$F_{\hat{X}}(\hat{X}, u, t) = \hat{X}u.$$

# 4 Internal models and innovation terms

In this section we axiomatize the notion of an *innovation term* which will form the second part of our observer design. Intuitively, an innovation term "punishes" an observer for deviating from the "true" system trajectory. To make sense of this concept, we need to first formalize how an observer system may "replicate" the observed system's trajectories. The relevant notion is that of an *internal model.*

We start with a system

$$\dot{X} = F_X(X, u, t) \tag{5}$$

with $F_X : G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_X(X, u, t) \in T_X G$ and an observer

$$\dot{\hat{X}} = F_{\hat{X}}(\hat{X}, Y, w, t) \tag{6}$$

with $F_{\hat{X}} : G \times G \times \mathfrak{g} \times \mathbb{R} \to TG$, $F_{\hat{X}}(\hat{X}, Y, w, t) \in T_{\hat{X}} G$. Note that the observer (6) has two inputs $Y$ and $w$, with $Y$ to be fed with measurements of $X$ and $w$ to be fed with measurements of $u$, respectively. The idea is that the observer should produce the correct trajectory if it is fed with exact measurements and started with the correct initial value. This is formalized in the following definition.

**Definition 5.** *Consider the pair of systems* (5) *and* (6)*. We say that* (6) *has an* internal model *of* (5) *if for all admissible* $u : \mathbb{R} \to G$*,* $X_0 \in G$ *and all* $t \in \mathbb{R}$

$$\hat{X}(t; X_0, X(t; X_0, u), u) = X(t; X_0, u), \tag{7}$$

*where* $X(t; X_0, u)$ *and* $\hat{X}(t; \hat{X}_0, Y, w)$ *denote the solutions of* (5) *and* (6)*, respectively.*

Note that by Theorem 4 a right synchronous "observer" for the left invariant system (1) has the form

$$\dot{\hat{X}} = \hat{X}u \tag{8}$$

and hence obviously has an internal model of the observed system. Intuitively, an innovation term is now a term that, added to the right hand side of system (8), tries to drive the observer away from its current trajectory and towards the correct trajectory whenever it deviates from a merely replicated system trajectory, while not destroying the internal model structure. Conceptually, we want to go the other way and identify which terms on the right hand side of a general observer (6) that has an internal model of the observed system act as an innovation in that sense. We choose a more general definition that only requires the term not to disturb the observer's evolution whenever it is already following the correct trajectory.

**Definition 6.** *Assume that* (6) *has an internal model of* (5). *We call a map* $\alpha\colon G \times G \times \mathfrak{g} \times \mathbb{R} \to TG$ *an* innovation term *for* (5) *and* (6) *if*

1. *for all* $\hat{X}, Y \in G, w \in \mathfrak{g}, t \in \mathbb{R}$: $\alpha(\hat{X}, Y, w, t) \in T_{\hat{X}}G$ *and*
2. *for all admissible* $u\colon \mathbb{R} \to \mathfrak{g}$, $X_0 \in G$ *and* $t \in \mathbb{R}$
$$\alpha\left(\hat{X}(t; X_0, X(t; X_0, u), u), X(t; X_0, u), u, t\right) = 0.$$

Note that condition 2 in the above definition is in particular implied by the stronger condition $\alpha(X, X, w, t) = 0$ for all $X \in G$, $w \in \mathfrak{g}$ and $t \in \mathbb{R}$.

We now focus on left invariant observed systems. The next theorem shows that any observer that has an internal model of such a system naturally splits into a right synchronous term (which is unique according to Theorem 4) and an innovation term. This result provides the final justification for our approach to observer design.

**Theorem 7.** *Consider the left invariant system* (1) *and the observer* (6). *Assume that the observer has an internal model of the system. Then the right hand side of the observer*
$$F_{\hat{X}}(\hat{X}, Y, w, t) = \hat{X}w + \alpha(\hat{X}, Y, w, t),$$
*where* $\alpha(\hat{X}, Y, w, t)$ *is an innovation term.*

In summary, we propose the following observer structure
$$\dot{\hat{X}} = \hat{X}w + \alpha(\hat{X}, Y, w, t)$$

# 5 Gradient observers

We still have to choose a *good* innovation term $\alpha$ for our observer. The ultimate goal is to prove almost global convergence results for the observer error. We hence aim at gradient dynamics for the error. For this purpose we assume that we are given a smooth, non-negative cost function $f\colon G \times G \to \mathbb{R}$. Furthermore, let the diagonal $\Delta = \{(X, X) \mid X \in G\}$ consist of global minima of $f$. Then we propose the use of a gradient term as the innovation term for the observer. To make this more precise, recall that the Riemannian gradient of $f$ with respect to the product metric $\langle \cdot, \cdot \rangle_p$ on $G \times G$ is defined by
$$\langle \operatorname{grad} f(\hat{X}, Y), (\eta, \zeta) \rangle_p = df(\hat{X}, Y)(\eta, \zeta)$$
for all $Y, \hat{X} \in G$, $\eta \in T_{\hat{X}}G$, $\zeta \in T_Y G$. Since we use the product metric, the gradient splits into the gradients with respect to the first and second parameter, i.e.
$$\langle \operatorname{grad} f(\hat{X}, Y), (\eta, \zeta) \rangle_p = \langle \operatorname{grad}_1 f(\hat{X}, Y), \eta \rangle + \langle \operatorname{grad}_2 f(\hat{X}, Y), \zeta \rangle.$$
Here, we propose the use of the gradient of $f$ with respect to the first parameter as an innovation term. This yields the following observer
$$\dot{\hat{X}} = \hat{X}w - \operatorname{grad}_1 f(\hat{X}, Y) \tag{9}$$

## 5.1 Error dynamics

To compute the error dynamics of our gradient observers we focus on the case where we have exact measurements of the input $u$ of system (1), as well as exact measurements of the state $X$. In terms of variables we have $Y = X$ and $w = u$ for the left observer (9). Under suitable invariance conditions on the cost function and the Riemannian metric, we get an autonomous error dynamics in this "noise-free" case.

**Theorem 8.** *If $f$ and the Riemannian metric on $G$ are right invariant then we have for the right invariant error $E_r$ of the observer (9) that*

$$\dot{E}_r = -\operatorname{grad}_1 f(E_r, e).$$

The gradient dynamics of the error yield the following convergence result in the noise-free case.

**Theorem 9.** *Assume that $Y \mapsto f(Y, e)$ is a Morse-Bott function with a global minimum at $e$ and no other local minima. If $f$ and the Riemannian metric on $G$ are both right invariant, then both errors $E_r$ and $E_l$ of the observer (9) converge to $e$ for generic initial conditions. Furthermore, $f(E_r, e)$ converges monotonically to $f(e, e)$ and this convergence is locally exponential near $e$.*

**Remark 10.** *Consider the left invariant system (1) and noisy measurements $w = u + \delta$ and $Y = N_l X$, with additive driving noise $\delta \in \mathfrak{g}$ and left multiplicative state noise $N_l \in G$. A straightforward calculation yields*

$$\dot{E}_r = \operatorname{Ad}_{\hat{X}} \delta E_r - \operatorname{grad}_1 f(E_r, N_l)$$

*for the canonical right invariant error of the observer (9). A suitably bounded noise will yield at least a practical stability result in these cases.*

## 5.2 Example: Attitude estimation on $\mathrm{SO}(3)$

We revisit Example 1 on the special orthogonal group SO(3),

$$\dot{R} = R\Omega,$$

where $R$ denotes the attitude of a coordinate frame fixed to a rigid body in 3D-space relative to an inertial frame and $\Omega$ encodes the angular velocity measured in the body-fixed frame. The velocity measurements are given by $w = \Omega$ and the state measurements are given by $Y = R$. We define the cost function $f(\hat{R}, Y) = \frac{k}{2}\|\hat{R} - Y\|_F^2$, with $\|\cdot\|_F$ the Frobenius norm and $k$ a positive constant. Our observer construction yields

$$\dot{\hat{R}} = \hat{R}w + k\hat{R}\mathbb{P}_{\mathfrak{so}(3)}(\hat{R}^\top Y)$$

which coincides with the passive filter proposed in [8].

# Bibliography


[1] G. Baldwin, R. Mahony, J. Trumpf, T. Hamel, and T. Cheviron, *Complementary filter design on the Special Euclidean group SE(3)*, in Proceedings of the European Control Conference (ECC), 2007, pp. 3763–3770.

[2] S. Bonnabel, P. Martin, and P. Rouchon, *A non-linear symmetry-preserving observer for velocity-aided inertial navigation*, in American Control Conference, 2006, pp. 2910–2914.

[3] ——, *Invariant asymptotic observers*, IEEE Transactions on Automatic Control, (2008). Accepted for publication.

[4] ——, *Non-linear observer on lie groups for left-invariant dynamics with right-equivariant output*, in 17th IFAC World Congress, 2008. Accepted for publication.

[5] S. Bonnabel and P. Rouchon, *On invariant observers*, in Control and Observer Design for Nonlinear Finite and Infinite Dimensional Systems, vol. 322 of Lecture Notes in Control and Information Sciences, Springer, 2005, pp. 53–67.

[6] R. Kalman, *A new approach in linear filtering and prediction problems*, Transactions of the American Society of Mechanical Engineers [ASME], Journal of Basic Engineering, 82D (1960), pp. 35–45.

[7] D. Luenberger, *Observing the state of a linear system with observers of low dynamic order*, IEEE Trans. Mil. Electr., ME-8 (1964), pp. 74–80.

[8] R. Mahony, T. Hamel, and J.-M. Pflimlin, *Non-linear complementary filters on the special orthogonal group*, IEEE Transactions on Automatic Control, (2008). Accepted for publication.

[9] 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, 2007, pp. 1039–1045.

[10] 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, 2007, pp. 6424–6429.

[11] J. Vasconcelos, R. Cunha, C. Silvestre, and P. Oliveira, *Landmark based nonlinear observer for rigid body attitude and position estimation*, in Proceedings of the 46th IEEE Conference on Decision and Control, 2007, pp. 1033–1038.