On Invariant Observers

Silvère Bonnabel and Pierre Rouchon

Ecole des Mines de Paris, Centre Automatique et Systèmes, 60 Bd Saint-Michel, 75272 Paris cedex 06, France, {silvere.bonnabel,pierre.rouchon}@ensmp.fr

Summary. A definition of invariant observer and compatible output function is proposed and motivated. For systems admitting a Lie symmetry-group G of dimension less or equal to the state dimension and with a G-compatible output, an explicit procedure based on the moving frame method is proposed to construct such invariant observers. It relies on an invariant frame and a complete set of invariant estimation errors. Two examples of engineering interest are considered: an exothermic chemical reactor and an inertial navigation problem. For both examples we show how invariance and the proposed construction can be a useful guide to design non-linear convergent observers, although the part of the design procedure which achieves asymptotic stability is not systematic and must take into account the specific nonlinearities of the case under study.

Key words: Observers, symmetries, invariance, moving frame, chemical reactor, inertial navigation.

1 Introduction

Symmetries are important in physics and mathematics, see, e.g., [13, 15]. In control theory, they also play a fundamental role, especially in feedback design and optimal control, see, e.g., [7, 6, 11, 9, 10, 16, 18, 14]. To our knowledge only very few results exploiting symmetries are available for observer design. Most of the results are based on special structure once a proper set of state coordinates has been chosen [12, 5, 17, 19, 8]. Thus most design techniques are not coordinates free. For mechanical system with position measures, an intrinsic design is proposed in [3]: it relies on the metric derived from the kinetics energy; it is invariant with respect to change of configuration coordinates. In [2, 1], it is shown how to exploit symmetry for the design of asymptotic observer via the notion of invariant estimation errors. The observer dynamics remains unchanged up to any transformation of state coordinates belonging to a Lie-group of symmetries. This paper prolonges and completes theses results and can be seen also the counter-part on the observer side of invariant tracking [14].

For clarity's sake, we consider here only the local and regular case. Global and singular cases are much more difficult and require additional assumptions. However, as shown by the analytical examples of the exothermic reactor and inertial navigation, such local analysis are in fact sufficient to get invariant observer and formulae that are also valid on the entire state manifold.

In section 2 we propose a natural definition of an invariant observer. In section 3, we characterize in terms of invariant frame and invariant estimation errors such invariant observers. In section 4, we show that the existence of invariant observers imposes some strong property on the output map: it has to be compatible with the group action on the state space. In section 5, we adapt the moving frame method and propose a procedure to construct explicitly (i.e., up to inversion of non-linear map) the general form of invariant observers. Sections 6 and 7 are devoted to two engineering examples for which we propose nonlinear and convergent invariant observers.

The authors thank Philippe Martin for interesting discussion and comments relative to the inertial navigation problem.

2 Invariance

Consider the following dynamics

$$\frac{d}{dt}x = f(x,t) \tag{1}$$

where the state x lies in an open subset $\mathcal{X}$ of $\mathbb{R}^n$ and f is a smooth function of its arguments. Let G be a r-dimensional Lie group acting locally on $\mathcal{X}$: φ_g is a diffeomorphism on $\mathcal{X}$ close to identity when g is close to e, the identity element of G. We assume that the action of G is locally free: the dimension of the orbit passing through $x \in \mathcal{X}$, i.e., the set of all $\varphi_g(x)$ for g close to e is r, the dimension of G; for any $x \in \mathcal{X}$ and any $y \in \mathcal{X}$ close to x and belonging to the orbit of x, exists a unique g close to e such that $\varphi_g(x) = y$ and such a g depends smoothly of y on the orbit of x. All along the paper, when we consider $g \in G$, we always implicitly assume that g is in a small neighborhood of e.

Definition 1. *Dynamics* (1) *is called G-invariant if for all $g \in G$, $x \in \mathcal{X}$*

$$f(\varphi_g(x), t) = \frac{\partial \varphi_g}{\partial x}(x) \cdot f(x,t). \tag{2}$$

G is thus a symmetry group: for each $g \in G$, the change of variables $X = \varphi_g(x)$ leaves the dynamics unchanged: $\frac{d}{dt}X = f(X,t)$.

Let $h : x \mapsto y = h(x)$ be a regular output map from $\mathcal{X}$ to $\mathcal{Y}$, an open subset of $\mathbb{R}^p$ ($p \leq n$).

Definition 2. *Take a G-invariant dynamics $\frac{d}{dt}x = f(x,t)$ with a smooth output $y = h(x)$. The dynamical system*

$$\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, h(x), t)$$

is called a G-invariant observer if, and only if, for all $g \in G$ and for all x and $\hat{x}$ in $\mathcal{X}$ we have

$$\hat{f}(x, h(x), t) = f(x,t), \quad \frac{\partial \varphi_g}{\partial x}(\hat{x}) \cdot \hat{f}(\hat{x}, h(x), t) = \hat{f}(\varphi_g(\hat{x}), h(\varphi_g(x)), t).$$

This definition means that the observer dynamics is unchanged under transformations of the form $\hat{X} = \varphi_g(\hat{x})$, $X = \varphi_g(x)$, g being an arbitrary element of G:

$$\frac{d}{dt}\hat{X} = \hat{f}(\hat{X}, h(X), t)$$

Notice that this definition does not deal with convergence issues. We separate intentionally invariance from convergence and robustness. We have kept the terminology observer because, with this definition, the system trajectory $t \mapsto x(t)$ is solution of the observer dynamics: $\hat{x}(t) = x(t)$ when $\hat{x}(0) = x(0)$. Clearly such definition must be completed by convergence and robustness properties. Since we do not have general result relative to convergence and robustness we will just complete the definition simply by the following one : the observer $\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, h(x), t)$ is called asymptotic if $\hat{x}(t)$ converges to $x(t)$ when t tends to $+\infty$.

3 Characterization

Assumptions relative to the action of G on $\mathcal{X}$, imply that exist n point-wise linearly independent G-invariant vector fields w_1, ..., w_n, forming a frame on $\mathcal{X}$, i.e. a basis of the tangent space at x (see, theorem 2.84 of [15]).

Take an invariant observer $\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, y, t)$ as described in definition 2. For each y and t, the vector field $\hat{f}$ can be decomposed according to such invariant frame $(w_i)_{i=1,\dots,n}$:

$$\hat{f}(\hat{x}, y, t) = \sum_{i=1}^{n} F_i(\hat{x}, y, t) w_i(\hat{x})$$

where the F_i's are smooth scalar functions of their arguments. Definition 2 means that

$$\forall x \in \mathcal{X}, \quad f(x,t) = \sum_{i=1}^{n} F_i(x, h(x), t) w_i(x)$$

and for any $i \in \{1, ..., n\}$, any $g \in G$ and x, $\hat{x}$ in $\mathcal{X}$,

$$F_i(\varphi_g(\hat{x}), h(\varphi_g(x)), t) = F_i(\hat{x}, h(x), t)$$

Thus we have

$$\hat{f}(\hat{x}, y, t) = f(\hat{x}, t) + \sum_{i=1}^{n} E_i(\hat{x}, y, t) w_i(\hat{x})$$

where each $E_i(\hat{x}, y, t) = F_i(\hat{x}, y, t) - F_i(\hat{x}, h(\hat{x}), t)$ and satisfies for all $x, \hat{x} \in \mathcal{X}$ and $g \in G$:

$$E_i(x, h(x), t) = 0, \quad E_i(\varphi_g(\hat{x}), h(\varphi_g(x)), t) = E_i(\hat{x}, h(x), t).$$

These E_i can be interpreted thus as scalar invariant errors. This motivates the following definition.

Definition 3. *A scalar smooth function $E(\hat{x}, y, t)$ is called an invariant error if, and only if, it satisfies the following equation*

$$E(x, h(x), t) = 0, \quad E(\varphi_g(\hat{x}), h(\varphi_g(x)), t) = E(\hat{x}, h(x), t).$$

for any $g \in G$, $x, \hat{x} \in \mathcal{X}$.

We have thus the following lemma

Lemma 1. *Any G-invariant observer $\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, y, t)$ (see definition 2) reads:*

$$\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, y, t) = f(\hat{x}, t) + \sum_{i=1}^{n} E_i(\hat{x}, y, t) w_i(\hat{x}) \tag{3}$$

where the E_i's are invariant errors and $(w_1, ..., w_n)$ is an invariant frame.

We will see in section 5 how to build such E_i and w_i with the knowledge, in local coordinates, of the transformation φ_g.

4 *G*-Compatible Output

The existence of G-invariant observer in the sense of definition 2 implies a compatibility condition on the output map $y = h(x)$. Under rank conditions, the action of G on state-space $\mathcal{X}$ can be transported via the output map h on the output space $\mathcal{Y}$ in the sense of the following definition:

Definition 4. *Let $\frac{d}{dt}x = f(x, t)$ be a G-invariant dynamics. The smooth output map $\mathcal{X} \ni x \mapsto y = h(x) \in \mathcal{Y}$ is said G-compatible if, and only if, for any $g \in G$, there exists a smooth invertible transformation ρ_g on $\mathcal{Y}$ such that*

$$\forall x \in \mathcal{X}, \quad h(\varphi_g(x)) = \varrho_g(h(x)).$$

This definition implies that G acts also on the output space $\mathcal{Y}$, via the transformations ϱ_g.

Theorem 1. *Consider a G-invariant dynamics $\frac{d}{dt}x = f(x,t)$ with output $y = h(x)$ and assume there exists a G-invariant observer $\frac{d}{dt}\hat{x} = \hat{f}(\hat{x}, y, t)$ in the sense of definition 2. Assume that the rank of the jacobian $\partial f(\hat{x}, y, t)/\partial y$ is maximum and equal to the dimension of y. Then, necessarily, the output map is G-compatible in the sense of definition 4.*

Proof. By lemma 1, we have

$$\hat{f}(\hat{x}, y, t) = f(\hat{x}, t) + \sum_{i=1}^{n} E_i(\hat{x}, y, t) w_i(\hat{x})$$

where the E_i are invariant errors and w_i invariant vector fields. Since the rank of $\frac{\partial \hat{f}}{\partial y}$ is equal to $p = \dim(y)$, there exist p invariant errors, $E = (E_{i_1}, ..., E_{i_p})$, $1 \leq i_1 < i_2 ... < i_p \leq n$ such that the Jacobian $\frac{\partial E}{\partial y}$ is invertible. Denote by F the inverse map versus y (locally defined by the implicit function theorem), then we have

$$F(\hat{x}, E(\hat{x}, y, t), t) \equiv y.$$

Moreover we have, for all $\hat{x}, x \in \mathcal{X}$ and $g \in G$,

$$E(\varphi_g(\hat{x}), h(\varphi_g(x)), t) = E(\hat{x}, h(x), t).$$

Thus

$$h(\varphi_g(x), t) = F(\varphi_g(\hat{x}), E(\hat{x}, h(x), t), t).$$

Let us fix $\hat{x}$ and t to some nominal values, said $\bar{x}$ and $\bar{t}$. The above identity means that for any $g \in G$, exists ρ_g (depending also on $\bar{x}$ and $\bar{t}$, but these dependencies are omitted here) a map from $\mathcal{Y}$ to $\mathcal{Y}$ such that, for all $x \in \mathcal{X}$,

$$h(\varphi_g(x), t) = \rho_g(h(x)).$$

Thus the output map is G-compatible and ρ_g is a transformation on $\mathcal{Y}$ with inverse $\rho_{g^{-1}}$.

5 Construction

The above discussion and results justify the following question. Assume that we have explicitly the action of G on the state space $\mathcal{X}$ with a G-compatible output map $y = h(x)$. This means that we have at our disposal r scalar parameters $a = (a_1, ..., a_r)$ corresponding to a particular parametrization of G, n smooth scalar functions $\varphi = (\varphi^1, ..., \varphi^n)$ of a and $x = (x_1, ..., x_n)$ corresponding to the action of G in local coordinates on $\mathcal{X}$,

$$X_i = \varphi^i(a, x), \quad i = 1, ..., n$$

We have also p output scalar maps of x, $h=(h^1,...,h^p)$, and p scalar functions of a and y, $\varrho=(\varrho^1,...,\varrho^p)$, corresponding to action of G on the output-space $\mathcal{Y}$ such that for any group parameter a and state x

$$h(\varphi(a,x))=\varrho(a,h(x)).$$

These transformations φ and ϱ are derived in general from obvious and physical symmetries.

According to lemma 1, invariant observers are built with invariant errors E_i and invariant frame $(w_1,...,w_n)$. More precisely, we will see that, under standard regularity conditions, we have an explicit parametrization of such invariant observers based on a complete set $(E_1,...,E_p)$ of invariant errors.

Let us first explain how to build an invariant frame from the knowledge of φ. This construction is standard (see [15]) and as follows. Since the orbits are of dimension r, the dimension of G, we can decompose the components of φ into two sets: $\varphi=(\bar\varphi,\tilde\varphi)$ with $\dim(\bar\varphi)=r$ and $\dim(\tilde\varphi)=n-r\geq 0$, such that for each x, the map $a\mapsto\bar\varphi(a,x)$ is invertible. This decomposition $x=(\bar x,\tilde x)$ is clearly not unique. Take a normalization of the first set of components, $\bar x$, denoted by $\bar x_0$. We can solve $\bar\varphi(a,x)=\bar x_0$ versus a to obtain $a=\alpha(\bar x_0,x)$ where α is a smooth function. Thus α is characterized by the identity:

$$\bar\varphi(\alpha(\bar x_0,x),x)\equiv\bar x_0.$$

Consider the canonical frame $(\frac{\partial}{\partial x_1},...,\frac{\partial}{\partial x_n})$. For each $i\in\{1,...,n\}$ and x, set

$$w_i(x)=\left(\frac{\partial\varphi}{\partial x}(\alpha(\bar x_0,x),x)\right)^{-1}\cdot\frac{\partial}{\partial x_i}|_{\varphi(\alpha(\bar x_0,x),x)}.$$

Then w_i is invariant because for any group parameter b and x we have

- $w_i(\varphi(b,x))=\left(\frac{\partial\varphi}{\partial x}(\alpha(\bar x_0,\varphi(b,x)),\varphi(b,x))\right)^{-1}\cdot\frac{\partial}{\partial x_i}$ and thus

$$\left(\frac{\partial\varphi}{\partial x}(b,x)\right)^{-1}\cdot w_i(\varphi(b,x))=\left(\frac{\partial\varphi}{\partial x}(\alpha(\bar x_0,\varphi(b,x)),\varphi(b,x))\frac{\partial\varphi}{\partial x}(b,x)\right)^{-1}\cdot\frac{\partial}{\partial x_i}$$

- the group structure implies that, for any group parameters c, d, we have $\varphi(d,\varphi(c,x))=\varphi(d\cdot c,x)$; thus

$$\varphi(\alpha(\bar x_0,\varphi(b,x)),\varphi(b,x))=\varphi(\alpha(\bar x_0,\varphi(b,x))\cdot b,x)$$

  and also

$$\frac{\partial\varphi}{\partial x}(\alpha(\bar x_0,\varphi(b,x)),\varphi(b,x))\frac{\partial\varphi}{\partial x}(b,x)=\frac{\partial\varphi}{\partial x}(\alpha(\bar x_0,\varphi(b,x))\cdot b,x);$$

  where $\cdot$ corresponds to the composition law on G.

- since $\alpha(\bar{x}_0, \varphi(b,x)) \cdot b = \alpha(\bar{x}_0, x)$, we have thus

$$\left(\frac{\partial \varphi}{\partial x}(b,x)\right)^{-1} \cdot w_i(\varphi(b,x)) = \left(\frac{\partial \varphi}{\partial x}(\alpha(\bar{x}_0,x),x)\right)^{-1} \cdot \frac{\partial}{\partial x_i} = w_i(x).$$

The computation of a complete set of invariants E_i is obtained via similar manipulations based on the normalization function α. We recall here the procedure presented in [2] based on the moving frame method (see also [14] for closely related computations of invariant tracking errors). We consider here the action of G on the product space $\mathcal{X} \times \mathcal{Y}$ defined by

$$\mathcal{X} \times \mathcal{Y} \ni (\hat{x}, y) \mapsto (\varphi(a,\hat{x}), \varrho(a,y)) \in \mathcal{X} \times \mathcal{Y}.$$

Consider once again the normalization function $\alpha(\bar{x}_0, \hat{x})$. Then each component of $\tilde{\varphi}(\alpha(\bar{x}_0,\hat{x}),\hat{x})$ and of $\varrho(\alpha(\bar{x}_0,\hat{x}),y)$ are invariant scalar function of $(\hat{x}, y)$. Moreover, they form a complete set of invariant relative to the action of G on $\mathcal{X} \times \mathcal{Y}$: this means that any invariant function I of $\hat{x}$ and y is a function of these $p+n-r$ fundamental invariants. Thus every invariant errors $E_i(\hat{x}, y, t)$ admits the following expression:

$$E_i(\hat{x}, y, t) = F_i(\tilde{\varphi}(\alpha(\bar{x}_0,\hat{x}),\hat{x}), \varrho(\alpha(\bar{x}_0,\hat{x}),y), t)$$

where F_i is a smooth function of its arguments. To be a little more explicit, denote by $I_1(\hat{x})$, ..., $I_{n-r}(\hat{x})$ the components of $\tilde{\varphi}(\alpha(\bar{x}_0,\hat{x}),\hat{x})$ and by $\varepsilon_1(\hat{x},y)$, ..., $\varepsilon_p(\hat{x},y)$ the components of $\varrho(\alpha(\bar{x}_0,\hat{x}),y) - \varrho(\alpha(\bar{x}_0,\hat{x}),h(\hat{x}))$. Then the invariant errors E_i have the form

$$E_i(\hat{x}, y, t) = \mathcal{E}_i(I, \varepsilon, t)$$

where $\mathcal{E}_i$ is a smooth function of its arguments I, ε and t, with $\mathcal{E}_i(I, 0, t) \equiv 0$.

6 An Exothermic Reactor

Let us consider the classical exothermic reactor of [4]. With slightly different notations, the dynamics reads

$$\begin{aligned} \frac{d}{dt}X &= D(t)(X^{in} - X) - k \exp\left(-\frac{E}{RT}\right) X \\ \frac{d}{dt}T &= D(t)(T^{in}(t) - T) + c \exp\left(-\frac{E}{RT}\right) X + u(t) \end{aligned}$$

where (E, R, k, c) are positive and known constant parameters, $D(t)$, $T^{in}(t)$ and $u(t)$ are known time functions. The parameter $X^{in} > 0$, the inlet composition is unknown. The available online measure is T, the temperature inside the reactor. The reactor composition X is not measured.

These two differential equations correspond to material and energy balances. Their structure is independent of the units. Let us formalize such independence in terms of invariance. We just consider change of material unit corresponding to the following scaling $X \mapsto aX$ and $X^{in} \mapsto aX^{in}$ with $a > 0$. The group G will be the multiplicative group $\mathbb{R}_+^*$.Take $x = (X^{in}, X, T, c)$ as state and the action of G is defined for each $a > 0$ via the transformation

$$\begin{pmatrix} X^{in} \\ X \\ T \\ c \end{pmatrix} \mapsto \varphi(a, x) = \begin{pmatrix} aX^{in} \\ aX \\ T \\ c/a \end{pmatrix}.$$

Then the dynamics

$$\begin{aligned}
\frac{d}{dt}X^{in} &= 0 \\
\frac{d}{dt}X &= D(t)(X^{in} - X) - k\exp\left(-\frac{E}{RT}\right)X \\
\frac{d}{dt}T &= D(t)(T^{in}(t) - T) + c\exp\left(-\frac{E}{RT}\right)X + u(t) \\
\frac{d}{dt}c &= 0
\end{aligned}$$

is invariant and $y = (T, c)$ is a G-compatible output. Notice that we have added to the original state (X, T), the inlet composition X^{in} and the parameter c. Since c it is known, it has been added to the output function.

An obvious invariant frame is $(X^{in}\frac{\partial}{\partial X^{in}}, X\frac{\partial}{\partial X}, \frac{\partial}{\partial T}, c\frac{\partial}{\partial c})$. Invariant errors are built with the following complete set of invariants

$$\frac{\hat{X}}{\hat{X}^{in}}, \quad \hat{T}, \quad \hat{c}\hat{X}, \quad \hat{T} - T, \quad (\hat{c} - c)\hat{X}^{in}.$$

Since c is known ($\hat{c} = c$), we do not care for the estimation of c and we thus consider invariant observers for (X^{in}, X, T). X^{in} and X must be estimated since they are unknown. Since X^{in} can only be constructed as an integral of the estimation error $\hat{T} - T$, $\hat{T}$ must be calculated anyway although measured. Moreover the temperature measure can be noisy and thus noiseless estimation $\hat{T}$ of T could be useful for feedback.

According to lemma 1, invariant observers have the following form

$$\begin{aligned}
\frac{d}{dt}\hat{X}^{in} &= A\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right)\hat{X}^{in} \\
\frac{d}{dt}\hat{X} &= D(t)(\hat{X}^{in} - \hat{X}) - k\exp\left(-\frac{E}{R\hat{T}}\right)\hat{X} + B\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right)\hat{X} \\
\frac{d}{dt}\hat{T} &= D(t)(T^{in}(t) - \hat{T}) + c\exp\left(-\frac{E}{R\hat{T}}\right)\hat{X} + u(t) + C\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right)
\end{aligned}$$

where A, B and C are smooth functions such that

$$A\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, 0\right) = B\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, 0\right) = C\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, 0\right) = 0.$$

Such observers preserve the fact that $\hat{X}$ and $\hat{X}^{in}$ are positive quantities.

Although we have no general method for choosing A, B and C to ensure convergence, we propose here a stabilizing design.

First, up to a change of A, B and C, we can replace the Arrhenius term $\exp\left(-\frac{E}{R\hat{T}}\right)$ by $\exp\left(-\frac{E}{RT(t)}\right)$ where $T(t)$ is the measure (kind of output injection). Thus the invariant observer reads (without changing the notations for A, B and C):

$$\frac{d}{dt}\hat{X}^{in} = A\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right)\hat{X}^{in}$$

$$\frac{d}{dt}\hat{X} = D(t)(\hat{X}^{in} - \hat{X}) - k\exp\left(-\frac{E}{RT(t)}\right)\hat{X} + B\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right)\hat{X}$$

$$\frac{d}{dt}\hat{T} = D(t)(T^{in}(t) - T(t)) + c\exp\left(-\frac{E}{RT(t)}\right)\hat{X} + u(t) + C\left(\frac{\hat{X}}{\hat{X}^{in}}, c\hat{X}, \hat{T}, \hat{T} - T\right).$$

Let us choose

$$A = B = -\beta c\hat{X}\exp\left(-\frac{E}{RT(t)}\right)(\hat{T} - T(t))$$

and

$$C = -\gamma c\hat{X}\exp\left(-\frac{E}{RT(t)}\right)(\hat{T} - T(t))$$

with β and γ strictly positive design parameters. Take the variables $\hat{\xi} = \log(\hat{X}/\hat{X}^{in})$ and $\hat{Z} = \log(\hat{X})$ instead of $\hat{X}^{in}$ and $\hat{X}$, since we have homogeneous equations in $\hat{X}^{in}$ and $\hat{X}$. Then we have the following triangular structure

$$\frac{d}{dt}\hat{\xi} = D(t)\exp(-\hat{\xi}) - D(t) - k\exp\left(-\frac{E}{RT(t)}\right)$$

$$\frac{d}{dt}\hat{Z} = D(t)\exp(-\hat{\xi}) - \beta c\exp\hat{Z}\exp\left(-\frac{E}{RT(t)}\right)(\hat{T} - T(t)) - D(t) - k\exp\left(-\frac{E}{RT(t)}\right)$$

$$\frac{d}{dt}\hat{T} = D(t)(T^{in}(t) - T(t)) + c\exp\left(-\frac{E}{RT(t)}\right)\exp\hat{Z}\ (1 - \gamma(\hat{T} - T(t))) + u(t).$$

Take $\tilde{Z} = \hat{Z} - \log(X(t))$ and $\tilde{T} = \hat{T} - T(t)$ instead of $\hat{Z}$ and $\hat{T}$. We get

$$\frac{d}{dt}\hat{\xi} = D(t)\exp(-\hat{\xi}) - D(t) - k\exp\left(-\frac{E}{RT(t)}\right)$$

$$\frac{d}{dt}\tilde{Z} = D(t)(\exp(-\hat{\xi}) - \exp(-\xi(t))) - \beta c\exp\left(-\frac{E}{RT(t)} + Z(t)\right)\exp\tilde{Z}\ \tilde{T}$$

$$\frac{d}{dt}\tilde{T} = c\exp\left(-\frac{E}{RT(t)} + Z(t)\right)(\exp\tilde{Z} - 1) - \gamma c\exp\left(-\frac{E}{RT(t)} + Z(t)\right)\exp\tilde{Z}\ \tilde{T}$$

where $\xi = \log(X/X^{in})$ and $Z = \log X$. Assume that exist M and $\eta > 0$ such that for all $t \geq 0$, $M \geq X^{in}, D(t), X(t), T(t) \geq \eta$. Then $(\hat{\xi}, \tilde{Z}, \tilde{T})$ remain bounded for all $t \geq 0$ and $\lim_{t \mapsto +\infty}(\hat{\xi}(t) - \xi(t)) = 0$. It is thus enough to analyze the convergence to 0 of the following reduced system in $(\tilde{Z}, \tilde{T})$:

$$
\begin{aligned}
\frac{d}{dt}\tilde{Z} &= -\beta c \exp\left(-\frac{E}{RT(t)} + Z(t)\right) \exp \tilde{Z}\, \tilde{T} \\
\frac{d}{dt}\tilde{T} &= c \exp\left(-\frac{E}{RT(t)} + Z(t)\right)(\exp \tilde{Z} - 1) - \gamma c \exp\left(-\frac{E}{RT(t)} + Z(t)\right) \exp \tilde{Z}\, \tilde{T}.
\end{aligned}
$$

Consider the regular change of time scale $\tau = \int_0^t c \exp\left(-\frac{E}{RT(s)} + Z(s)\right)\, ds$. Then:

$$
\begin{aligned}
\frac{d\tilde{Z}}{d\tau} &= -\beta \exp \tilde{Z}\, \tilde{T} \\
\frac{d\tilde{T}}{d\tau} &= (\exp \tilde{Z} - 1) - \gamma \exp \tilde{Z}\, \tilde{T}
\end{aligned}
$$

This system admits $\tilde{Z} + \exp(-\tilde{Z}) + \frac{\beta}{2}\tilde{T}^2$ as Lyapounov function. A standard application of Lasalle invariance principle shows that 0 is globally asymptotically stable.

Guided by invariance considerations, we have obtained the following globally converging non-linear observer:

$$
\begin{aligned}
\frac{d}{dt}\hat{X}^{in} &= -\beta \exp\left(-\frac{E}{RT(t)}\right)\ (\hat{T} - T(t))\ \ c\hat{X}\ \hat{X}^{in} \\
\frac{d}{dt}\hat{X} &= D(t)(\hat{X}^{in} - \hat{X}) - \exp\left(-\frac{E}{RT(t)}\right)\left(k + \beta(\hat{T} - T(t))c\hat{X}\right)\ \hat{X} \\
\frac{d}{dt}\hat{T} &= \exp\left(-\frac{E}{RT(t)}\right)\left(1 - \gamma(\hat{T} - T(t))\right)\ c\hat{X} + D(t)(T^{in}(t) - T(t)) + u(t)
\end{aligned}
$$

where the design parameters β and γ have to be chosen strictly positive.

7 Inertial Navigation

Consider a flying body carrying an Inertial Measurement Unit which measures the earth magnetic field (denoted $\mathbf{B}$ in the earth frame) and the instantaneous rotation vector $\boldsymbol{\omega}$. Let $\mathbf{b}$ denote the magnetic field in the body frame (measured by the IMU). Let $R \in SO(3)$ denote the rotation matrix which maps the mobile body frame to the earth frame. The kinematic relations between R and $\boldsymbol{\omega}$ reads,

$$
\frac{d}{dt} R\mathbf{a} = R(\boldsymbol{\omega}(t) \wedge \mathbf{a})
$$

for any vector $\mathbf{a}$. Omitting $\mathbf{a}$, we have the following dynamics for R in $SO(3)$:

$$\frac{d}{dt}R = R(\boldsymbol{\omega}(t) \wedge \cdot) \tag{4}$$

with output

$$y = \mathbf{b} = R^{-1}\mathbf{B} \tag{5}$$

If the earth frame rotates via $g \in SO(3)$, then R becomes gR, $\mathbf{B}$ becomes $g\mathbf{B}$ and the above equations remain unchanged. This means that the dynamics

$$\frac{d}{dt}R = R(\boldsymbol{\omega}(t) \wedge \cdot), \quad \frac{d}{dt}\mathbf{B} = 0$$

is invariant under the action of the group $G = SO(3)$ via left multiplication:

$$\varphi_g(R, \mathbf{B}) = (gR, g\mathbf{B}), \quad g \in SO(3)$$

and the output map $y = R^{-1}\mathbf{B}$ is clearly G-compatible since invariant ($\varrho_g = I_d$).

Let $\mathbf{e}_1$,$\mathbf{e}_2$,$\mathbf{e}_3$ be a basis of $\mathbb{R}^3$. Then, an invariant frame on $SO(3)$, the state-space for R, is as follows:

$$SO(3) \ni R \mapsto w_i(R) = R(\mathbf{e}_i \wedge \cdot) \quad i = 1, 2, 3$$

Notice that $w_i(R)$ belongs to the tangent space at R.

Since each component of $\mathbf{b} = R^{-1}\mathbf{B}$ is invariant, they form a complete set of scalar invariants relative to the action of $G = SO(3)$ on $SO(3) \times \mathbb{R}^3$, the space of $(R, \mathbf{B})$. Every invariant observer reads

$$\frac{d}{dt}\hat{R} = \hat{R}(\boldsymbol{\omega}(t) \wedge \cdot) + \sum_{i=1}^{3} E_i \left(\hat{R}^{-1}\mathbf{B}, \hat{R}^{-1}\mathbf{B} - \mathbf{b}\right) \; \hat{R}(\mathbf{e}_i \wedge \cdot) \tag{6}$$

where the scalar functions E_i are such that $E_i(\hat{R}^{-1}\mathbf{B}, 0) \equiv 0$.

We can adjust the function E_i in order to have the convergence of the estimation error $\hat{R}^{-1}\mathbf{B} - R^{-1}\mathbf{B}$ to 0 when t tends to $+\infty$. Assume that the E_i correspond to the coordinates of the vector $\mathbf{b} \wedge (\hat{R}^{-1}\mathbf{B})$. More precisely we set

$$\sum_{i=1}^{3} E_i \mathbf{e}_i = K\mathbf{b} \wedge (\hat{R}^{-1}\mathbf{B})$$

where $K > 0$ is some design parameter. Then we get the following invariant observer

$$\frac{d}{dt}\hat{R} = \hat{R}\left([\boldsymbol{\omega}(t) + K\mathbf{e}] \wedge \cdot\right)$$

with $\mathbf{e} = \mathbf{b}(t) \wedge (\hat{R}^{-1}\mathbf{B})$. Standard computations (using $R^{-1} = R^T$) show that

$$\frac{d}{dt}(\hat{R}^{-1}\mathbf{B} - \mathbf{b}) = -\boldsymbol{\omega} \wedge (\hat{R}^{-1}\mathbf{B} - \mathbf{b}) - K\mathbf{e} \wedge (\hat{R}^{-1}\mathbf{B})$$

and

$$\frac{1}{2}\frac{d}{dt}\left\|\hat{R}^{-1}\mathbf{B}-\mathbf{b}\right\|^2 = -K\|\mathbf{e}\|^2$$

Thus $\mathbf{e}$ tends to zeros when t tends to $+\infty$. This implies that $\mathbf{b}$ and $\hat{R}^{-1}\mathbf{B}$ tend to be co-linear. Since their modules are equal, this means the convergence of the estimated output to the measured output (a direct analysis shows that the situation when $\mathbf{b} = -\hat{R}^{-1}\mathbf{B}$ is unstable and thus will not be obtained in practice).

The convergence of $\hat{R}$ to R is only partial since, if $t \mapsto \hat{R}(t)$ is solution of the invariant observer (6), then $t \mapsto O_\theta \hat{R}(t)$ is also solution of (6) where O_θ is the rotation around $\mathbf{B}$ by an arbitrary angle θ. In a certain sense, we cannot improve the convergence since with the output $\mathbf{b} = R^{-1}\mathbf{B}$, the dynamics on R is not observable: the trajectories $t \mapsto R(t)$ and $t \mapsto O_\theta R(t)$ lead to exactly the same output trajectory $t \mapsto \mathbf{b}(t)$.

8 Conclusion

We have proposed a systematic method to design observers preserving the symmetries of the original system. We do not have, up to now, a similar systematic procedure to tackle convergence. Nevertheless, the two previous examples indicate that invariance could be a useful guide for the design of convergent non-linear observers.

References

1. N. Aghannan. *Contrôle de Réacteurs de Polymérisation, observateur et invariance.* PhD thesis, Ecole des Mines de Paris, November 2003.
2. N. Aghannan and P. Rouchon. On invariant asymptotic observers. In *Proceedings of the 41st IEEE Conference on Decision and Control*, volume 2, pages 1479– 1484, 2002.
3. N. Aghannan and P. Rouchon. An intrinsic observer for a class of lagrangian systems. *IEEE AC*, 48(6):936–945, 2003.
4. R. Aris and N.R. Amundson. An analysis of chemical reactor stability and control- i,ii,iii. *Chem. Engng. Sci.*, 7:121–155, 1958.
5. J. Birk and M. Zeitz. Extended Luenberger observers for nonlinear multivariable systems. *Int. J. Control*, 47:1823–1836, 1988.
6. A. M. Bloch, P. S. Krishnaprasad, J. E. Marsden, and R. Murray. Nonholonomic mechanical systems with symmetry. *Arch. Rational Mech. Anal.*, 136:21–99, 1996.
7. F. Fagnani and J. Willems. Representations of symmetric linear dynamical systems. *SIAM J. Control and Optim.*, 31:1267–1293, 1993.
8. J.P. Gauthier and I. Kupka. *Deterministic Observation Theory and Applications.* Cambridge University Press, 2001.

9. J.W. Grizzle and S.I. Marcus. The structure of nonlinear systems possessing symmetries. *IEEE Trans. Automat. Control*, 30:248–258, 1985.
10. B. Jakubczyk. Symmetries of nonlinear control systems and their symbols. In *Canadian Math. Conf. Proceed.*, volume 25, pages 183–198, 1998.
11. W. S. Koon and J. E. Marsden. Optimal control for holonomic and nonholonomic mechanical systems with symmetry and lagrangian reduction. *SIAM J. Control and Optim.*, 35:901–929, 1997.
12. A.J. Krener and W. Respondek. Nonlinear observers with linearizable error dynamics. *SIAM J. Control Optim.*, 23:197–216, 1985.
13. J.E. Marsden and T.S. Ratiu. *Introduction to Mechanics and Symmetry*. Springer-Verlag, New York, 1994.
14. Ph. Martin, P. Rouchon, and J. Rudolph. Invariant tracking. *ESAIM: Control, Optimisation and Calculus of Variations*, 10:1–13, 2004.
15. P. J. Olver. *Equivalence, Invariants and Symmetry*. Cambridge University Press, 1995.
16. W. Respondek and I.A. Tall. Nonlinearizable single-input control systems do not admit stationary symmetries. *Systems and Control Letters*, 46:1–16, 2002.
17. J. Rudolph and M. Zeitz. A block triangular nonlinear observer normal form. *Systems and Control Letters*, 23:1–8, 1994.
18. A.J. van der Schaft. Symmetries in optimal control. *SIAM J. Control Optim.*, 25:245259, 1987.
19. M. Zeitz and X. Xia. On nonlinear continuous observers. *Int. J. Control*, 66(6):943–954, 1997.