Convergence theory for observers: Necessary, and Sufficient conditions

(Preliminary version of a summary of the lectures)

Laurent Praly
Centre Automatique et Systèmes
Mines-ParisTech
35 Rue Saint Honoré
77305 Fontainebleau
France
email: laurent.praly@mines-paristech.fr

July 15, 2014
Contents

1 Observation problem and its basic solutions ................................................................. 1
  1.1 The context .................................................................................................................. 1
  1.2 The observation problem ........................................................................................... 2
  1.3 Observers with no loss of information ....................................................................... 3
  1.4 An optimization approach ......................................................................................... 4

2 Convergent Observers and Necessary conditions ......................................................... 6
  2.1 Convergent Observers ............................................................................................... 6
  2.2 Necessary conditions for observer convergence ....................................................... 6
    2.2.1 No restriction on $\tau$ ......................................................................................... 6
    2.2.2 $\tau$ is injective given $h$ .................................................................................. 7

3 Sufficient conditions for observers with $\tau$ bijective given $h$ ................................... 11
  3.1 Case where $\tau$ is the identity function .................................................................... 11
    3.1.1 $f$ is strictly monotonic with respect to $P$ tangentially to the output function level sets ................................................................. 11
    3.1.2 Systems linear up to output non linearity ......................................................... 12
    3.1.3 $f$ tangentially monotonic and directionally Lipschitz ....................................... 13
    3.1.4 $P\partial f + \partial f^TP$ upper bounded ................................................................... 14
  3.2 Case where $(x,t) \mapsto (\tau^{-1}(x,t), h(x,t), t)$ is a diffeomorphism .................... 14

4 Sufficient condition via differential observability and High-gain observers ............... 15
  4.1 Differential observability of order $m$ ....................................................................... 15
  4.2 High gain observer .................................................................................................... 15

5 Nonlinear Luenberger Observer ..................................................................................... 17

References ......................................................................................................................... 18
1 Observation problem and its basic solutions

1.1 The context

Observers are answers to the question of estimating, from observed/measured empirical variables, denoted \( y \), and delivered by sensors equipping a real world system, some “theoretical” variables, called hidden variables in this text, denoted \( z \), which are involved in a mathematical model related to this system. The measured variables make what is called the a posteriori information on the hidden variables, whereas the model is part of the a priori information. Because a model cannot fit exactly a system, introduction of uncertainties is mandatory.

Typically this model describing the link between hidden and measured variables is made of three components:

- a **dynamic model** describes the dynamics/evolution:\(^1\)
  \[
  \delta^d(t) = \dot{x}(t) - f(x(t),t) \quad \text{resp.} \quad \delta^d_k = x_{k+1} = f_k(x_k),
  \]
  where \( t \), in the continuous case, or \( k \), in the discrete case, is an evolution parameter, called time in this text, \( x \) is a state, assumed finite dimensional in this text. \( \delta^d \) represents the uncertainties in the state dynamics, they are by definition what is needed to get an equality. Any possible known inputs are represented here by the time-dependence of \( f \).

- a **sensor model** relates state and measured variables:
  \[
  \delta^m(t) = y(t) - h(x(t),t) \quad \text{resp.} \quad \delta^m_k = y_k - h_k(x_k)
  \]
  with \( \delta^m \) representing the uncertainties in the measurements.

- a model which relates state and hidden variables:
  \[
  z(t) = h_\Psi(x(t),t) \quad \text{resp.} \quad z_k = h_\Psi(x_k)
  \]
  There is no uncertainty here since typically an hidden variable is a theoretical quantity.

In a deterministic setting, the a priori information on the uncertainties \( \delta^d, \delta^m \) may be that the values of \( \delta^d \) and \( \delta^m \) are unknown but belong to known sets \( \Delta^d \) and \( \Delta^m \). Namely we have:

\[
\delta^d(t) \in \Delta^d(t), \quad \delta^m(t) \in \Delta^m(t), \quad \delta^d_k \in \Delta^d_k, \quad \delta^m_k \in \Delta^m_k.
\]

In a stochastic setting and more specifically in a Bayesian approach, it may be that \( \delta^d \) and \( \delta^m \) are unknown realization of stochastic processes for which we know the probability distributions. Similarly we may also know a priori that we have:

\[
x(t) \in \mathcal{X}(t), \quad z(t) \in \mathcal{Z}(t), \quad x_k \in \mathcal{X}_k, \quad z_k \in \mathcal{Z}_k,
\]

where the sets \( \mathcal{X} \) and \( \mathcal{Z} \) are known or we may have a priori probability distribution for \( x \) and \( z \).

In this context, the a priori information is the data of the functions \( f, h \) and \( h_\Psi \), of the sets \( \Delta^d, \Delta^m \) or the corresponding probability distribution and so maybe also of the sets \( \mathcal{X} \) and \( \mathcal{Z} \) or the corresponding a priori probability distribution.

---

\(^1\)\( \dot{x} \) denotes the time derivative \( \frac{dx}{dt} \).
In the next section, we state the observation problem and give the solutions which are
direct consequences of the deterministic and stochastic setting given above. This will allow us
to see that an observer is actually a dynamical system with the measurements as inputs and
the estimate as output.

The main objective of these lectures is the study of the asymptotic behavior of this system
and in particular the study of the convergence of its output to the hidden variables we want
to estimate. The next chapters of this text is devoted to this convergence topic.

To ease the presentation we deal only with the discrete time case in Section 1.3, and the
continuous time case in Section 1.4 and the next chapters.

1.2 The observation problem

Let $X^d(x, t, s)$, respectively $X^d_t(x, k)$, denote a solution of (1) at time $s$, respectively $l$, going
through $x$ at time $t$, respectively $k$, and under the action of $\delta^d$.

**Observation problem**: At each time $t$, respectively $k$, given the function $s \in [t - T, t] \mapsto y(s)$,
respectively the sequence $l \in \{k - K, \ldots, k\} \mapsto y_l$, find an estimation $\hat{z}(t)$, respectively $\hat{z}_k$, of
$z(t)$, respectively $z_k$, satisfying:

$$\hat{z}(t) = h_\delta(\hat{x}(t), t) \quad \text{resp.} \quad \hat{z}_k = h_\delta_k(\hat{x}_k) .$$

where $\hat{x}(t)$, respectively $\hat{x}_k$, is to be found as a solution of:

$$\hat{x}(t) \in X(t) ,$$

$$y(s) = h(X^d(\hat{x}(t), t, s, \delta^m(s)) \quad \forall s \in [t - T, t] ,$$

respectively

$$\hat{x}_k \in X_k ,$$

$$y_l = h_l(\hat{x}_k, k, \delta^m_l) \quad \forall l \in \{k - K, \ldots, k\}$$

and where the time functions $\delta^d$ and $\delta^m$ must agree with the a priori (deterministic/stochastic)
information or minimized in some way.

In this statement $T$, respectively $K$, quantify the time window length or memory length
during which we record the measurement. The accumulation with time of measurements, to-
gether with the model equations (1) to (3) and the assumptions on $(\delta^d, \delta^m)$, give a redundancy
of data compared with the number of unknowns that the hidden variables are. This is why it
may be possible to solve this observation problem.

To simplify the following presentation, we restrict our attention to the case$^2$ where the
hidden variables are actually the full model state, i.e.

$$z = h_\delta(x) = x .$$

$^2$When $z$ differs from $x$, observers are called functional observers.
1.3 Observers with no loss of information

Conceptually the answer to this observation problem is easy at least when the memory increases with time ($\hat{T}(t) = 1$ resp. $K_{k+1} = K_k + 1$ ) leading to an infinite non fading memory. It consists in starting from all what the a priori information makes possible and to eliminate what is not consistent with the a posteriori information. In the set valued observer setting, in the discrete time case, this gives the following observer. To ease its reading, we underline the data given by the a priori information. It requires the introduction of two sets $\xi_k$, and $\xi_{k|k-1}$ which are updated at each time $k$ when a new measurement $y_k$ is made available. $\xi_k$ is the set to which $x_k$ is guaranteed to belong at time $k$, knowing all the measurements up to time $k$, and $\xi_{k|k-1}$ is the same but with measurements known up to time $k - 1$.

Set valued observer:

Initialization:

$$\xi_0 = X_0$$

At each time $k$:

Prediction (flowing)

$$\xi_{k|k-1} = f_{k-1}(\xi_{k-1}, \Delta_{k-1}^d)$$

Restriction (consistency)

$$\xi_k = \{ x \in \left( \xi_{k|k-1} \cap X_k \right) : y_k \in h_k(x, \Delta_{k}^m) \}$$

Estimation

$$\hat{x}_k \in \xi_k$$

A key feature here is that this observer has a state $\xi_k$ – a set – and can be seen as a dynamical system in the form:

$$\xi_{k+1} = \varphi_k(\xi_k, y_k), \quad \hat{x}_k \in \xi_k$$

with $y$ as input and $\hat{x}$ as output which is not single valued.

In the stochastic setting, following the Bayesian paradigm, the observer has the same structure but with the state $\xi_k$ being a conditional probability. See [16, Theorem 6.4] or [10, Table 2.1]. In that setting too the observer is not a single state; it is the (a posteriori) conditional probability of the random variable $x_k$ given the a priori information and the sequence of measurements $l \in \{k - K, \ldots, k\} \mapsto y_l$.

Comments

Lossless data compression. The observers introduced above realize a lossless data compression with extracting and preserving all what concerns the hidden variables in the redundant data given by a priori and a posteriori information.

Not single valued estimate. Because it is a “lossless compression” this answer must be all the estimates which are compatible with the a priori and a posteriori information. With the presence of the uncertainties, these estimates cannot be unique. Namely the answer is not single valued. Indeed it is set valued or conditional probability valued.

To extract a particular estimates from all these possible values, the observation problem must be complemented by making precise for what the estimation is made. For instance we may want to select the most likely or the average or more generally some cost-minimizing estimate $\hat{x}$ among all the possible ones given by $\xi$. In this way we obtain an observer giving
a single valued estimate:
\[ \xi_{k+1} = \varphi_k(\xi_k, y_k), \quad \hat{x}_k = \tau_k(\xi_k) \]
respectively
\[ \dot{\xi}(t) = \varphi(\xi(t), y(t), t), \quad \hat{x}(t) = \tau(\xi(t), t) \] (6)

But then, in general, we loose information and in particular we have no idea on the confidence level this estimate has. Also, since the function \( \tau \), at least, encodes for what the estimate \( \hat{x} \) is used, for different uses, different functions \( \tau \) may be needed.

**Implementation:** For the time being, except for very specific cases (Kalman filter, ...) the set valued and the conditional probability valued observers remain conceptual since we do not know how to manipulate numerically sets and probability laws. Their implementation requires approximations. For instance, see [21, 27] for the set case and [4, 9, 10, 16] for the conditional probability case.

**Need of finite or infinite but fading memory and of contraction:** In these observers, model states \( x \) which are consistent with the a priori information but do not agree with the a posteriori information are eliminated (set intersection or probability product). But once a point is eliminated, this is for ever. As a consequence if there is, at some time, a misfit between a priori and a posteriori information, it is mistakenly propagated in future times. And what we claimed as “guaranteed” above may not be true. A way to round this problem is to keep the information memory finite or infinite but fading. In particular, with fixed length memory, consistent points which were disregarded due to measurements which are no more in the memory are reintroduced. This says in particular that observers should not be sensitive to their initial condition. In the continuous time case, for the dynamics of an observer given by
\[ \dot{\xi}(t) = \varphi(\xi(t), y(t), t) \] (7)
whose solution at time \( s \), going through \( \xi \) at time \( t \) and under the action of \( y \), is denoted by \( \Xi^y(\xi, t, s) \), for example, it is well known that the initial condition \( \xi \) is forgotten if, given \( t \) and \( \{ \sigma \mapsto y(\sigma) \} \), the flow \( \xi \mapsto \Xi^y(\xi, t, s) \) generated by (7) is a strict contraction\(^3\) for each \( s > t \) or, at least, if a distance between any two solutions \( \Xi^y(\xi_1, t, s) \) and \( \Xi^y(\xi_2, t, s) \), with the same input \( y \), converges to 0.

### 1.4 An optimization approach

A short-cut to obtain directly an observer giving a single valued estimate is to design it by trading off among a priori and a posteriori information (see [12, pages 7-10], [1], ...). For example, in the continuous time case, we can select the estimate \( \hat{x}(t) \) among the minimizers (in \( x \)) of:
\[
C(\{ s \mapsto \delta^d(s) \}, x, t) = \int_{-\infty}^{t} C\left( \delta^d(s), y(s), X^\delta^d(x, t, s), s \right) ds
\]
where \( X^\delta^d(x, t, s) \) is still the notation for a solution to (1) and \( \{ s \mapsto \delta^d(s) \} \), representing the unmodelled effect on the dynamics, is among the arguments for the minimization. The

\(^3\)See [17] for a bibliography on contraction.
infinitesimal cost $C$ is chosen to take non negative values and be such that $C(0, h(x, s), x, s)$ is zero. For instance, it can be:

$$C(\delta^d, y, x, s) = \|\delta^d\|^2_x + d_y(y, h(x, s))^2$$

where $\|.\|_x$ is a norm at the point $x$ and $d_y$ is a distance in the measurement space. In the same spirit, instead of optimization, a minimax approach can be followed. See [6], [5, Chapter 7].

With $x$ fixed, the minimization of $C$ is an infinite horizon optimal control problem in reverse time. Solving on line this problem is extremely difficult and again approximations are needed. We do not go on with this approach, but we remark that, under extra assumptions, the observer we obtain following this approach is also in the form of a dynamical system (6) but with the specificity that the estimate $\hat{x}$ is part of the observer state $\xi$ and its dynamics are a copy of the undisturbed model with a correction term which is zero when the estimated state reproduce the measurement. Namely we get:

$$\dot{\hat{x}}(t) = f(\hat{x}(t), t, 0) + E\left(\{\sigma \mapsto y(\sigma)\}, \hat{x}(t), y(t), t\right)$$

where $E$ is zero when $h(\hat{x}(t), t) = y(t)$. 

5
2 Convergent Observers and Necessary conditions

2.1 Convergent Observers

Once we have imposed the observer to give a single valued estimate, because of data loss and/or approximation in the implementation, the main problem we are facing is the convergence of this estimate to the “true” value, at least when there is no uncertainties. We concentrate now our attention on the study of this convergence, but, to simplify, in the continuous time case only.

Let the model and observer dynamics be:

\[
\dot{x}(t) = f(x(t), t), \quad y(t) = h(x(t), t) \tag{8}
\]

\[
\dot{\xi}(t) = \varphi(\xi(t), y(t), t), \quad \hat{x}(t) = \tau(\xi(t), y(t), t) \tag{9}
\]

with \( x \), called from now on system state, of dimension \( n \), measurement \( y \) of dimension \( p \leq n \) and observer state \( \xi \) of dimension \( m \). As before, we denote by \((X(x, t, s), \Xi((x, \xi), t, s))\) a solution of (8)-(9).

Since we are dealing with convergence, the focus is on what is going on when the time becomes very large and in particular on the set \( \Omega \) of model states which are accumulation points of some solution. Specifically we are interested in the stability properties of the set

\[
\mathcal{Z}(t) = \left\{ (x, \xi) : x \in \Omega \text{ } & \text{ } x = \tau(\xi, h(x(t), t)) \right\}
\]

which is contained in the zero estimation error set associated with the given model-observer pair.

Definition 1 (Convergent observer) We say the observer (9) is convergent if, for each \( t \), there exists a set \( \mathcal{Z}_\omega(t) \subset \mathcal{Z}(t) \), such that, on the domain of existence of the solution, a distance between the point \((X(x, t, s), \Xi((x, \xi), t, s))\) and the set \( \mathcal{Z}_\omega(s) \) is upperbounded by a real function \( s \mapsto \beta^c_{x,\xi,t}(s) \), maybe dependent on \((x, \xi, t)\), with non negative values, strictly decreasing and going to zero as \( s \) goes to infinity.

In this definition and in the following the index \( \omega \) is used to denote that the object is related to a behavior for the time going to +\( \infty \).

2.2 Necessary conditions for observer convergence

2.2.1 No restriction on \( \tau \)

It is possible to prove that, if the observer is convergent then,

Necessity of detectability: When \( h \) and \( \tau \) are uniformly continuous in \( x \) and \( \xi \) respectively, the estimate \( \hat{x} \) does converge to the model state \( x \). In this case, two solutions of the model (8) which produce the same measurement must converge to each other. This is an asymptotic distinguishability property called detectability.

If we are interested, not only in the asymptotic behavior, but also in the transient (as for output feedback) a property stronger than detectability is needed. In particular instantaneous distinguishability (see Section 4) is necessary if we want to be able to impose the decay rate of the function \( \beta^c_{x,\xi,t} \).
Necessity of $m \geq n - p$: For each $t$, there exists a subset $O_\omega(t)$ of $\Omega$, supposed to collect the model states which can be asymptotically estimated, and such that we can associate, to each of its point $x$, a set $\tau^{inv}(x, t)$ allowing us to redefine the set $Z_\omega(t)$ as:

$$Z_\omega(t) = \{(x, \xi) : x \in O_\omega(t) \& \xi \in \tau^{inv}(x, t)\}.$$  

This implies that, for each $t$ and each $x$ in $O_\omega(t)$, there is a point $\xi$ satisfying:

$$x = \tau(\xi, h(x, t), t).$$ (10)

This is a surjectivity property of the function $\tau$ but of a special kind since $h(x, t)$ is an argument of $\tau$. We say that, for each $t$, the function $\tau$ is surjective to $O_\omega(t)$ given $h$. In a “generic” situation this property requires the dimension $m$ of the observer state $\xi$ to be larger or equal to the dimension $n$ of the model state $x$ minus the dimension $p$ of the measurement $y$.

2.2.2 $\tau$ is injective given $h$

We consider now the case where the observer has been designed with a function $\tau$ which is injective given $h$, namely we have the following implication, when $x$ is in $O_\omega(t)$,

$$\left[\tau(\xi_1, h(x, t), t) = \tau(\xi_2, h(x, t), t) \& \xi_1 \in \tau^{inv}(x, t)\right] \implies \xi_1 = \xi_2.$$

In a “generic” situation, this property together with the surjectivity given $h$, implies that the dimension $m$ of the observer state $\xi$ should be between $n - p$ and $n$.

To ease the presentation of the following statement, we assume:

- the system is time independent, i.e., it is:

$$\dot{x}(t) = f(x(t)) \quad y(t) = h(x(t)).$$

- An asymptotic state observer associated to an open set $O$ is a system

$$\dot{\xi}(t) = \varphi(\xi(t), y(t)) \quad ,$$

with state $\xi$ in $\bar{U} \subset \mathbb{R}^m$.

Regularity Assumption :
For all $(x, \xi) \in O \times O$, the solution $t \mapsto \Xi(x, \xi, t)$ is uniquely defined in $(\sigma_{\bar{U}}(x), \sigma_{\bar{U}}^+(x))$.

Convergence Assumption :
The zero error set is

$$Z = \{(x, \xi) \in O \times \bar{U} : x = \tau(\xi, h(x))\}$$

contains a non empty subset $Z_\omega$ which is locally asymptotically stable.
Proposition 1  Assume the Regularity Assumption and Convergence Assumption hold. Let $O_\omega$ be the projection of $Z_\omega$ on $\mathcal{O}$. If there exists a class $\mathcal{K}^\infty$ function $\alpha_\tau$ such that 

$$d_\xi(\xi_a, \xi_b) \leq \alpha_\tau(d_a(x_a, \tau(\xi_b, h(x_a)))) \quad \forall (x_a, \xi_b) \in O_\omega \times \tilde{U} : \xi_a \in \tau^{inv}(x_a),$$

then

- $\tau^{inv}$ is singled valued and there exists a function $\ell : \tilde{U} \times \mathbb{R}^p \rightarrow \mathbb{R}^n$ such that the $\varphi$ function of the observer can be decomposed as

$$\varphi(\xi, y) = L_f \tau^{inv}(\tau(\xi, y)) + \ell(\xi, y)$$

with:

$$\ell(\xi, y) = 0 \quad \forall (\xi, y) \in \tilde{U} \times h(O_\omega) : y = h(\tau(\xi, y)), \tau(\xi, y) \in O_\omega.$$

- There exist a class $\mathcal{K}L$ function $\beta$ and a continuous function $\omega : O_\omega \times \tilde{U} \rightarrow \mathbb{R}_+$ such that:

$$d_\xi(\Xi((x, \xi_a), t), \Xi((x, \xi_b), t)) \leq \omega(x, \xi_a, \xi_b) \beta(d_\xi(\xi_a, \xi_b), t) \quad \forall (x, \xi_a, \xi_b, t) \in O_\omega \times \tilde{U} \times \mathbb{R}_+.$$

Together with the necessary surjectivity condition we have seen, the assumption in this Proposition implies that $\tau$ is a bijection given $h(x)$, i.e. when

$$x = \tau(\tau^{inv}(x), h(x)) \quad \forall x \in O_\omega,$$

$$\xi = \tau^{inv}(\tau(\xi, y)) \quad \forall (\xi, y) \in \tilde{U} \times h(O) : y = h(\tau(\xi, y)), \tau(\xi, y) \in O_\omega.$$

As a consequence $\tau^{inv}$ can be considered formally as part of of a possible set of coordinates for $x$. This view point allows to get a better understanding of the first property. In the above context, it says nothing but the dynamic of the partial coordinates $\tau^{inv}$ is, of course, the image of the vector field $f$, under the change of coordinates $(x, t) \mapsto (\tau^{inv}(x, t), t)$ but again all this given $h$, namely we have:

$$L_f \tau^{inv}(x, t) = \varphi(\tau^{inv}(x, t), h(x, t), t) \quad \forall x \in O_\omega(t)$$

Hence the observer dynamic is a copy of the image by $\tau^{inv}$ of the system dynamic plus a correction term which is zero when the estimated measurement $h(\tau(\xi, y))$ is equal to the actual measurement $y$. We have seen this already in the optimization approach.

The second property says that, if moreover the functions $h$ and $\tau$ are uniformly continuous in $x$ and $\xi$ respectively, then, given $\xi_1$ and $\xi_2$ a distance between $\Xi((x, \xi_1), t, s)$ and $\Xi((x, \xi_2), t, s)$ goes to zero as $s$ goes to infinity. This property is related to what was called extreme stability (see [28]) in the 50’s and 60’s and is called incremental stability today (see [3]). It holds when the flow generated by the observer is contracting, property we have mentioned already as desirable to reduce the sensitivity to modelling errors. We develop a little bit more this property of contraction

Definition : In general, the system

$$\dot{\xi}(t) = \varphi(\xi(t), t)$$

8
defined on an open set $\mathcal{U}$ of $\mathbb{R}^m$ (equipped with the distance $d_\xi$) is said to generate a (strictly) contracting flow if, for any $t_0$ in $\mathbb{R}^m$, the function

$$t \in [t_0, T[ \mapsto d_\xi(\Xi(\xi_a, t_0, t), \Xi(\xi_b, t_0, t))$$

is (strictly when not zero) decreasing for any pair $(\xi_a, \xi_b)$ in $\mathcal{U}^2$.

In the following, to simplify the presentation we restrict our attention to the case of an Euclidean metric but all this is available for a complete Riemannian metric.

**Definition :** The system $\dot{\xi} = \varphi(\xi, t)$ defined on an open set $\mathcal{U}$ of $\mathbb{R}^m$ (equipped with the distance $d_\xi$) is said to generate a (strictly) contracting flow for an Euclidean distance if there exist coordinates for $\xi$ and a symmetric positive definite matrix such that for any $t_0$ in $\mathbb{R}^m$, the function

$$t \in [t_0, T[ \mapsto \sqrt{\Xi(\xi_a, t_0, t) - \Xi(\xi_b, t_0, t)}$$

is (strictly when not zero) decreasing for any pair $(\xi_a, \xi_b)$ in $\mathcal{U}^2$.

Let us consider now the case of an observer for which the function $\tau$ is simply the identity. We know it must have the following dynamic :

$$\varphi(\xi, y, t) = f(\xi, t) + \ell(\xi, y, t) ,$$

with :

$$\ell(\xi, y, t) = 0 \quad \forall (\xi, y, t) \in \mathcal{U} \times \mathbb{R}^p \times \mathbb{R} : y = h(\xi, t) .$$

**Proposition 2** Assume the observer above generates a contracting flow for the Euclidean distance and $(x, \xi, y) \mapsto (f(x, t), h(x, t), \varphi(\xi, y, t))$ is a $C^1$ function for almost all $t$ in $\mathbb{R}$. Then we have

$$v^T P \frac{\partial f}{\partial x}(x, t) v \leq 0 \quad \forall (x, v) \in \mathcal{O} \times \mathbb{R}^n : \frac{\partial h}{\partial x}(x, t) v = 0 ,$$

for almost all $t$ in $\mathbb{R}$.

The property expressed in this Proposition is actually a local detectability property being related to the standard detectability property of the linear time varying system given by the first order approximation along a solution.

**Definition :** Let $P$ be a positive definite symmetric matrix and $f : \mathbb{R} \times \mathcal{O} \rightarrow \mathbb{R}^n$ and $h : \mathbb{R} \times \mathcal{O} \rightarrow \mathbb{R}^m$ be two functions and with $x \mapsto (f(x, t), h(x, t))$ $C^1$ for almost all $t$. We say $f$ is monotonic with respect to $P$ tangentially to the output function level sets if there exists a non negative real number $p$ such that we have, for almost all $t$,

$$v^T P \frac{\partial f}{\partial x}(x, t) v \leq -p v^T P v \quad \forall (x, v) \in \mathcal{O} \times \mathbb{R}^n : \frac{\partial h}{\partial x}(x, t) v = 0 . \quad (12)$$

It is strictly monotonic if $p$ is strictly positive.

Due to the fact that we have restricted our study to the case of an Euclidean distance, the monotonicity property mentioned above is coordinate dependent. Hence the importance of choosing the coordinates in such a proper way that we can find the matrix $P$. A case where we are guaranteed of the existence of such coordinates is...
**Proposition 3** When we have one output $y$ only, assume that the function $\Phi$ defined as:

$$\Phi(x) = \begin{pmatrix} h(x) \\ L_f h(x) \\ \vdots \\ L_f^{n-1}(x) \end{pmatrix}$$

is a diffeomorphism on some open set $\Omega$. Then there exists a matrix $P$ such that, when expressed in the coordinates $\Phi(x)$, $f$ is strictly monotonic with respect to $P$ tangentially to the output function level sets at least on any compact subset of $\Omega$.

This can be proved by exploiting the fact that in the $\Phi(x)$ coordinates, the system is:

$$\begin{align*}
\dot{x}_1 &= x_2, \\
\dot{x}_2 &= x_3, \\
\vdots &\\
\dot{x}_{n-1} &= x_n, \\
\dot{x}_n &= \psi(x_1, \ldots, x_n),
\end{align*}$$

with a function $\psi$ which is defined and Lipschitz on the image by $\Phi$ of any compact subset of $\Omega$.

The monotonicity property is by itself a sufficient condition for the existence of a convergent observer.

**Proposition 4** Assume the system can be decomposed as:

$$
\begin{pmatrix}
\dot{x}_1(t) \\
\dot{x}_2(t)
\end{pmatrix}
= 
\begin{pmatrix}
f_1(x_1(t), x_2(t), t) \\
f_2(x_1(t), x_2(t), t)
\end{pmatrix},
\quad
y(t) = h(x(t), t) = x_1(t)
$$

and $f$ is monotonic tangentially to the output function level sets with respect to $P = \begin{pmatrix} P_1 & P_3^T \\ P_3 & P_2 \end{pmatrix}$.

Then:

$$
\begin{align*}
\dot{\xi}(t) &= f_2\left(\xi(t) - P_2^{-1}P_3y(t), y(t), t\right) + P_2^{-1}P_3 f_1\left(\xi(t) - P_2^{-1}P_3y(t), y(t), t\right) \\
\dot{x}(t) &= \tau(\xi(t), y(t)) = \begin{pmatrix} y(t) \\ \xi(t) - P_2^{-1}P_3y(t) \end{pmatrix}
\end{align*}
$$

is an observer which solves the observer convergence problem.

See [7, Proposition 3.2]

This observer is called a reduced order observer since $m < n$. Its functions $\tau$ is a bijection given $h(x)$ with $\tau^{inv}(x) = x_2 + P_2^{-1}P_3x_1$ as “inverse given $h(x)$”

Knowing now how a convergent observer should look like, we move to a quick description of some such observers.
3 Sufficient conditions for observers with $\tau$ bijective given $h$

3.1 Case where $\tau$ is the identity function

A convergent observer whose function $\tau$ is the identity has the following form:

$$\dot{\xi}(t) = f(\xi(t), t) + E(\{\sigma \mapsto y(\sigma)\}, \xi(t), y(t), t), \quad \hat{x}(t) = \xi(t). \quad (13)$$

The only piece remaining to be designed is the correction term $E$. It has to ensure convergence\(^4\). For such a design, the typical first step is to look for appropriate coordinates to exhibit some specific properties of the vector field $f$ – monotonicity, convexity, . . . . .

We list here some results along these lines.

3.1.1 $f$ is strictly monotonic with respect to $P$ tangentially to the output function level sets

We first observe that we can always rewrite $f$ as

$$f(x, t) = \bar{f}(x h(x, t), t)$$

Assume we have found coordinates for $x$ and a positive definite matrix $P$ such that $\bar{f}$ is strictly monotonic with respect to $P$ tangentially to the output function level sets, in the following way:

$$v^T P \frac{\partial \bar{f}}{\partial x}(x, y, t) v \leq -p v^T P v \tag{14}$$

$$\forall (x, y, v, t) \in \mathcal{O} \times \mathbb{R}^p \times \mathbb{R}^n \times \mathbb{R} : \quad \frac{\partial h}{\partial x}(x, t) v = 0 \quad \& \quad y = h(x, t).$$

with $p$ a strictly positive real number. Then we can get a locally convergent observer. A nonlocal one can be obtained in the case where $h$ is linear

$$h(x, t) = H(t) x, \tag{15}$$

when expressed in the above particular coordinates for $x$. Indeed we have.

**Proposition 5** Assume $\mathcal{O} = \mathbb{R}^n$, and (14) and (15) hold. Then, for each compact subset $E$ of $\mathbb{R}^n$, we can find a function $\ell : \mathbb{R}^n \to \mathbb{R}_+$ such that, for all locally Lipschitz function $\ell$ satisfying

$$\ell(\xi, t) \geq \ell_E(\xi) \quad \forall (\xi, t) \in \mathbb{R}^n \times \mathbb{R},$$

the observer

$$\dot{\xi}(t) = \bar{f}(\xi(t), y(t), t) + \ell(\xi(t)) P^{-1} H(t)^T [y(t) - H(t)\xi(t)], \quad \hat{x}(t) = \xi(t)$$

solves the observer convergence problem for all initial condition $(x, \xi)$ in $\mathbb{R}^n \times \mathbb{R}^n$ satisfying $x - \xi \in E$.

\(^4\)And may be also other properties like symmetry preserving (see [8]).
See [22, 24].

If the monotonicity property holds for all direction and not only tangentially to the output function level sets, then the correction term is not needed. This leads to

**Proposition :** Let the system be

\[
\dot{x}(t) = f(x(t), h(x(t), t), t) \quad , \quad y(t) = h(x(t), t) .
\]

Assume there exist a positive definite symmetric matrix \( P \) and a strictly positive real number \( p \) such that we have

\[
P \frac{\partial f}{\partial x}(x, y, t) + \frac{\partial f}{\partial x}(x, y, t)^T P \leq -2pP \quad \forall (x, y, t) \in \mathcal{O} \times \mathbb{R}^p \times \mathbb{R} : y = h(x, t) .
\]

Then the observer convergence problem is solved by :

\[
\dot{\xi}(t) = f(\xi(t), y(t), t) \quad , \quad \dot{x}(t) = \xi(t) .
\]

### 3.1.2 Systems linear up to output non-linearities

Consider the system

\[
\dot{x} = f(x) \quad , \quad y = h(x)
\]

with \( x \) in \( \mathbb{R}^n \), \( y \) in \( \mathbb{R} \) and functions \( f \) and \( h \) sufficiently many times differentiable. We are interested in the case where there exist coordinates \( x \) for \( x \) such that the expression of the system dynamics become simply linear may be up to the addition of terms depending only on the output. Our motivation for this comes from the following result

**Proposition 6** Assume :

1. we have found coordinates \( x \) such that the system dynamic is

\[
\dot{x}(t) = F(y(t), t) x(t) + g(y(t), t) \quad , \quad y(t) = H(y(t), t) x(t) + k(t) ,
\]

where \( g \) and \( k \) are some arbitrary functions;

2. there exist a strictly positive real number \( c \) and functions \( \gamma_f : \mathbb{R} \times \mathbb{R}^p \rightarrow [c, +\infty[ \) and \( \gamma_h : \mathbb{R} \times \mathbb{R}^p \rightarrow \mathbb{R}^+ \), such that the pair \( (\frac{F}{\gamma_f}, \frac{H}{\gamma_h}) \) is constant and detectable.

Then there exists a matrix \( L \) such that the observer convergence problem is solved by :

\[
\dot{\xi}(t) = F(y(t), t) \xi(t) + g(y(t), t) + \frac{\gamma_f(y(t), t)}{\gamma_h(y(t), t)} L \left[ y(t) - H(y(t), t) \xi(t) \right]
\]

\[
\dot{x}(t) = \xi(t) .
\]

A case where such specific coordinates exist is identified in the following result.
Proposition 7 Assume there exist \( x_0 \) in \( \mathbb{R}^n \) and a neighborhood \( \mathcal{V} \) of this point so that:

\[
\text{Rank}(dh(x_0), dL_f h(x_0), \ldots, dL_f^{n-1} h(x_0)) = n .
\]

Assume also there exist three functions \( a, b \) and \( c \) satisfying, for all \( x \) in \( \mathcal{V} \),

\[
L_{adj} g(x) = 0 \quad \text{if} \quad 0 \leq j \leq n - 2 , \quad = (-1)^{n-1} a(x) \quad \text{if} \quad j = n - 1 ;
\]

\[
\begin{bmatrix}
ad_f^i \bar{a}, ad_f^j \bar{g} \\
\end{bmatrix} a(x) = 0 \quad 0 \leq i < j \leq n - 1 ,
\]

\[
dL_g (L_f^j h)(x) a(x) = \left( \frac{n(n-1)}{2} + 1 \right) a a(x) dL_f h(x) + b(x) dh(x) ,
\]

with the notations

\[
\bar{f} = \exp(-c) f , \quad \bar{g} = \exp(-(n-1)c) g
\]

and where \( g \) is a vector field defined on \( \mathcal{V} \) satisfying

\[
L_g L_f^j h(x) = 0 \quad \text{if} \quad 0 \leq j \leq n - 2 , \quad = 1 \quad \text{if} \quad j = n - 1 .
\]

Then there exists a diffeomorphism \( \phi : \mathcal{V} \to \mathbb{R}^n \) and a function \( d : \mathbb{R} \to \mathbb{R}^n \) such that, with he coordinates:

\[
x = \phi(x)
\]

the system reads:

\[
\dot{x}(t) = \exp(c(y(t))) [F x(t) + d(y(t))] , \quad y(t) = H x(t)
\]

where the pair \((F,H)\) is observable.

See [18, 23]

3.1.3 \( f \) tangentially monotonic and directionally Lipschitz

Proposition 8 For the system

\[
\dot{x}(t) = f(x(t), y(t)) , \quad y(t) = H x(t)
\]

with \( \mathcal{O} = \mathbb{R}^n \), assume :

1. the function \( x \in \mathcal{O} \mapsto f(x,y) \) is \( C^1 \) for all \( y \) in \( \mathbb{R}^p \).
2. \( f \) is strictly monotonic with respect to \( P \) tangentially to the output function level sets (see (12)).
3. \( f \) is Lipschitz in some direction, i.e. there exist an \( n \times p \) matrix \( \ell_0 \) and a non negative real number \( \theta \) such that

\[
H \ell_0 = I
\]

\[
\left| \ell_0^T P \frac{\partial f}{\partial x} (x,y) v \right| \leq \theta |v| \quad \forall (x,y,v) \in \mathcal{O} \times \mathbb{R}^p \times \mathbb{R}^n : Hv = 0 .
\]

Then we can find an \( n \times p \) matrix \( \ell \) such that the observer convergence problem is solved by:

\[
\begin{align*}
\dot{\xi}(t) &= f \left( \xi(t) + \ell_0 [y(t) - H \xi(t)] , y(t) \right) + \ell \left[ y(t) - H(t) \xi(t) \right] , \\
\dot{x}(t) &= \xi(t) .
\end{align*}
\]

See [22]
3.1.4 $P \partial f + \partial f^T P$ upper bounded

Assume we can find coordinates so that the system reads:

$$f(x) = Fx + \sum_{i=1}^{l} G_i \phi_i(K_i x)$$

**Proposition 9** Assume $\text{Rank}(H) = p$, the functions $\phi_i$ are gradients of functions whose Hessian have a spectrum contained in the interval $[a_i, b_i]$, $(b_i \leq +\infty)$. If we can find a positive definite symmetric matrix $P$ satisfying

$$v^T P F v + \sum_{i=1}^{l} \frac{b_i}{4} \left| (PG_i + K_i^T) v \right|^2 - \sum_{i=1}^{l} \frac{a_i}{4} \left| (PG_i - K_i^T) v \right|^2 < 0 \quad \forall v \in \mathbb{R}^n : H v = 0 ,$$

then there exist matrices $L_{0,i}$ and $L$ such that the observer convergence problem is solved by:

$$\dot{\xi}(t) = F \xi(t) + \sum_{i=1}^{l} G_i \phi_i(K_i \xi(t) - L_{0,i}[H \xi(t) - y(t)]) - L[H \xi(t) - y(t)]$$

$$\hat{x}(t) = \xi(t) .$$

See [13].

3.2 Case where $(x, t) \mapsto (\tau^{inv}(x, t), h(x, t), t)$ is a diffeomorphism

At each time $t$ we know already that the model state $x$ we want to estimate satisfy

$$y(t) = h(x, t) .$$

So, as remarked in [19], when $(h(x, t), t)$ can be used as part of coordinates for $(x, t)$, we need to estimate the remaining part only. This can be done if we find a function $\tau^{inv}$, whose values are $n - p$ dimensional, such that $(x, t) \mapsto (y, \eta, t) = (h(x, t), \tau^{inv}(x, t), t)$ is a diffeomorphism and the flow $\eta \mapsto \eta^\theta(\eta, t, s)$ generated by

$$\dot{\eta}(t) = \frac{\partial \tau^{inv}}{\partial x}(x(t), t)f(x(t), t) + \frac{\partial \tau^{inv}}{\partial t}(x(t), t) ,$$

$$= \varphi(\eta(t), y(t), t)$$

is a strict contraction for all $s > t$. Indeed in this case the observer dynamics can be chosen as:

$$\dot{\xi}(t) = \varphi(\xi(t), y(t), t)$$

and the estimate $\hat{x}(t)$ is obtained as solution of:

$$\tau^{inv}(\hat{x}(t), t) = \xi(t) , \quad h(\hat{x}(t), t) = y(t) .$$

This is the reduced order observer paradigm we have encountered already in Proposition 4. See [11], [19, Theorem 4], …
4 Sufficient condition via differential observability and High-gain observers

4.1 Differential observability of order $m$

Instantaneous distinguishability means that we can distinguish as quickly as we want two model states by looking at the paths of the measurements they generate. A sufficient condition to have this property can be obtained by looking at the Taylor expansion in $s$ of $h(X(x, t, s), s)$.

Indeed, we have:

$$h(X(x, t, s), s) = \sum_{i=0}^{m-1} h_i(x, t) \frac{(s - t)^i}{i!} + o((s - t)^{m-1})$$

where $h_i$ is a function obtained recursively as

$$h_0(x, t) = h(x, t)$$

$$h_{i+1}(x, t) = \frac{\partial h_i}{\partial x}(x, t) f(x, t) + \frac{\partial h_i}{\partial t}(x, t).$$

If there exists an integer $m$ such that, in some uniform way with respect to $t$, the function $x \mapsto H_m(x, t) = (h_0(x, t), \ldots, h_{m-1}(x, t))$ is injective then we do have instantaneous distinguishability. We say the system is differentially observable of order $m$ when this injectivity property holds. When a system has such a property, the model state space has a very specific structure as discussed in [15, Section 1.9]. It means that we can reconstruct $x$ from the knowledge of $y$ and its $m - 1$ first time derivatives, i.e. there exists a function $\Phi$ such that we have:

$$x = \Phi(H_m(x, t), t).$$

This way, we are left with estimating the derivatives of $y$.

4.2 High gain observer

The estimation of the derivatives of $y$ can be done as follows. With the notation

$$\eta_i = h_{i-1}(x, t),$$

we obtain:

$$\dot{\eta}(t) = F \eta + G h_m (\Phi(\eta(t), t), t)$$

where

$$F \eta = (\eta_2, \ldots, \eta_m, 0), \quad G = (0, \ldots, 0, 1).$$

When the last term on the right hand side is Lipschitz, we can find a convergent observer in the form:

$$\dot{\xi}(t) = F \xi(t) + G h_m (\hat{x}(t), t) + K(y(t) - \xi_1(t)),$$

$$\hat{x}(t) = \tau(\xi(t), t),$$
with $\xi$ being actually an estimation of $\eta$ and where $K$ is a constant matrix and $\tau$ is a modified version of $\Phi$ keeping the estimated state in some a priori given set $\mathcal{X}(t)$.

This is the high gain observer paradigm. See [14, 26]. The implementation difficulty is in the function $\hat{\Phi}$, not to mention sensitivity to measurement uncertainty.
5 Nonlinear Luenberger Observer

To ease the presentation here, we assume\(^5\) the system is time independent, i.e., it is:

\[
\dot{x} = f(x) , \quad y = h(x) .
\]

We have mentioned that having the flow generated by the observer contracting is a desirable property. If we do not impose any structure to the function \(\tau\) this is easily achieved by starting the observer design with picking the function \(\varphi\) as:

\[
\dot{\xi}(t) = \varphi(\xi(t), y(t)) = A\xi(t) + B(y(t))
\]

where \(A\), not related to \(f\), is a matrix whose eigen values have strictly negative real part. It turns out that, under weak restriction, there exists a function \(\tau^{inv}\) satisfying (11), namely:

\[
L_f\tau^{inv}(x) = A\tau^{inv}(x) + B(h(x)) .
\] (16)

To obtain a convergent observer it is then sufficient to find a (uniformly continuous) function \(\tau\) satisfying:

\[x = \tau(\tau^{inv}(x, t), h(x, t), t)\]

For this to be possible, the function \(\tau^{inv}\) should be injective given \(h\). It can be proved that this injectivity holds when the observer state has dimension \(m \geq 2(n + 1)\), the model is distinguishable and provided the eigen values of \(A\) have a sufficiently negative real part and are not in a set of zero Lebesgue measure. See [2, 19, 25].

Unfortunately, we are facing again a possible difficulty in the implementation since an expression for a function \(\tau^{inv}\) satisfying (16) is needed and the function \(\tau : (\xi, y, t) \mapsto \dot{x}(t)\) is known implicitly only as:

\[\xi = \tau^{inv}(\dot{x}(t), t) .\]

It is to round this difficulty that the function \(B\) should be appropriately chosen and/or that an approximation theory is needed. See [2, 20].

\(^5\)Results without these assumptions are available.
References


