1

An Overview on Observer Tools for Nonlinear Systems

Gildas Besançon

Control Systems Department (former *Laboratoire d'Automatique de Grenoble*), GIPSA-lab. ENSIEG BP 46, 38402 Saint-Martin d'Hères, France
Gildas.Besancon@inpg.fr

1.1 Introduction and Problem Statement

1.1.1 Context and Motivations

The problem of *observer* design naturally arises in a *system* approach, as soon as one needs some *internal* information from *external* (directly available) measurements. In general indeed, it is clear that one cannot use as many sensors as signals of interest characterizing the system behavior (for cost reasons, technological constraints, etc.), especially since such signals can come in a quite large number, and they can be of various types: they typically include time-varying signals characterizing the system (*state variables*), constant ones (*parameters*), and unmeasured external ones (*disturbances*).

This need for internal information can be motivated by various purposes: modeling (*identification*), monitoring (*fault detection*), or driving (*control*) the system. All those purposes are actually jointly required when aiming at keeping a system under control, as summarized by figure 1.1 hereafter. This makes the reconstruction - or observer - problem the heart of a general control problem.

The purpose here will thus be to give an overview on some possible tools for *observer design* (the related problems of control, fault detection or parameter identification being considered in subsequent chapters): in short, an observer relies on a model, with on-line adaptation based on available measurements, and aiming at information reconstruction, i.e. it can be characterized as a *model-based, measurement-based, closed-loop, information reconstructor*.

Usually the model is a state-space representation, and it will be assumed here that all pieces of information to be reconstructed are born by state variables. In front of this, one can try to design an explicit dynamical system whose state should give an estimate of the actual state of the considered model, or just settle the problem as an optimization one. The present chapter will focus on the first case, the second one being considered in a subsequent one.

About the considered model, it can in general be either continuous-time or discrete-time, deterministic or stochastic, finite-dimensional or infinite-dimensional, smooth or "with singularities". But in order to give a quite consistent presentation, the present chapter will be restricted to the case of smooth,

G. Besançon (Ed.): Nonlinear Observers and Applications, LNCIS 363, pp. 1–33, 2007.
springerlink.com

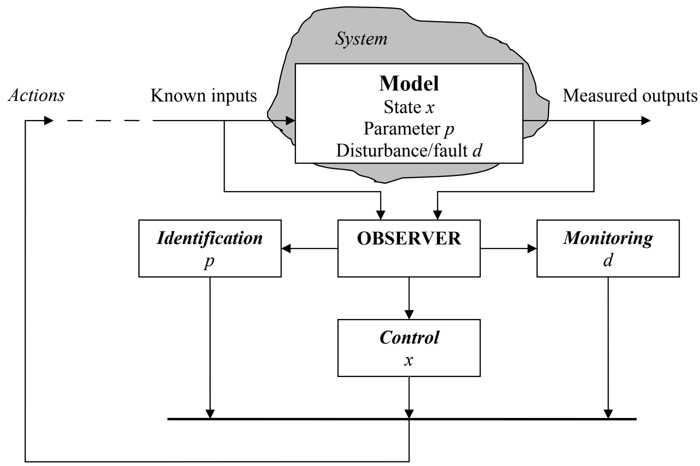


Fig. 1.1. Observer as the heart of control systems

finite-dimensional, deterministic, continuous-time state-space descriptions (even though some elements of observer design can be found for other cases in the literature).

In this framework, the subsequent subsection specifies the problem formulation, while section 1.2 presents the main related observability notions. Section 1.3 then discusses some possible techniques for observer design from a viewpoint which can be summarized by figure 1.2 below:

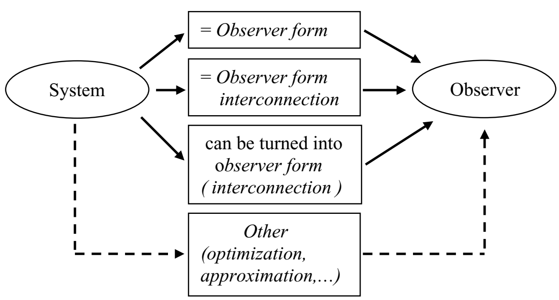


Fig. 1.2. A methodology for observer design

In short, the idea is to rely on specific structures for which observers are available (*observer forms*, in the figure), and try to bring the system under consideration to such cases. Hence, some *basic structures* for observer design are first presented, classified in two categories: those with a constant (input-independent) correction gain, and those with a time-varying (possibly input-dependent) correction gain. Then, methods for possible extensions of those designs to more general classes of systems are proposed, either by means of interconnections, or by transformations. Some conclusions and further remarks are finally given in section 1.4.

Notice that the material here presented roughly updates [7], summarizing some own research and viewpoint on the problem (in the continuity of [6] for instance) as well as various results borrowed from the quite large available ones (including overviews of [15] or [26] for instance).

In all the subsequent sections, the following notations/terminology will be used:

- Standard notations x, u, y and t respectively for state vector, input vector, output vector, and time variable (which might be omitted as an argument when not indispensable),
- I for the identity matrix of appropriate dimensions,
- v_i for the ith component of a vector v,
- $M = M^T > 0$ for a symmetric positive definite matrix M,
- *Stable matrix* for a matrix with all eigenvalues having strictly negative real parts.

while the abbreviation 'w.r.t.' will stand as usual for 'with respect to'.

1.1.2 Observer Problem Statement

Model Under Consideration

All over the chapter, the system under consideration will be considered to be described by a state-space representation generally of the following form:

$$\begin{array}{l} \dot{x}(t) = f(x(t), u(t)) \\ y(t) = h(x(t)) \end{array} \tag{1.1}$$

where x denotes the state vector, taking values in X a connected manifold of dimension n, u denotes the vector of known external inputs, taking values in some open subset U of $I\!R^m$, and y denotes the vector of measured outputs taking values in some open subset Y of $I\!R^p$.

Functions f and h will in general be assumed to be $\mathcal{C}^\infty$ w.r.t. their arguments, and input functions $u(.)$ to be locally essentially bounded and measurable functions in a set $\mathcal{U}$.

The system will be assumed to be forward complete.

More generally, the dynamics might explicitly depend on time via $f(x(t), u(t), t)$, while y might further directly depend on u and even t, via $h(x(t), u(t), t)$. Such an explicitly time-dependent system is usually called 'time-varying' and generalizes (1.1) into:

$$\begin{array}{l} \dot{x}(t) = f(x(t), u(t), t) \\ y(t) = h(x(t), u(t), t) \end{array} \tag{1.2}$$

However, the observer techniques discussed in this chapter are based on more specific forms of state-space representations, among which the following ones can be mentioned:

- Control-affine systems:

$$f(x, u) = f_0(x) + g(x)u$$

- State-affine systems[1]:

$$f(x,u) = A(u)x + B(u), \quad h(x) = Cx \ (or\ C(u)x + D(u))$$

- Linear Time-Varying (LTV) systems:

$$f(x,u,t) = A(t)x + B(t)u, \quad h(x,u,t) = C(t)x + D(t)u$$

- Linear Time-Invariant (LTI) systems:

$$f(x,u) = Ax + Bu, \quad h(x,u) = Cx + Du$$

Finally, the system will be said to be 'uncontrolled' whenever f and h do not depend on u.

In general, $\chi_u(t, x_{t_0})$ will denote the solution of the state equation in (1.1) under the application of input u on $[t_0, t]$ and satisfying $\chi_u(t_0, x_{t_0}) = x_{t_0}$, while u will be omitted for uncontrolled cases.

Observer Problem

Given a model (1.1), the purpose of acting on the system, or monitoring it, will in general need to know $x(t)$, while in practice one has only access to u and y. The observation problem can then be formulated as follows:

> Given a system described by a representation (1.1), find an estimate $\hat{x}(t)$ for $x(t)$ from the knowledge of $u(\tau), y(\tau)$ for $0 \leq \tau \leq t$.

Clearly this problem makes sense when one cannot invert h w.r.t. x at any time.

In front of this, one can look for a solution in terms of optimization, by looking for the best estimate $\hat{x}(0)$ of $x(0)$ which can explain the evolution $y(\tau)$ over $[0, t]$, and from this, get an estimate $\hat{x}(t)$ by integrating (1.1) from $\hat{x}(0)$ and under $u(\tau)$. In order to cope with disturbances, one should rather optimize the estimate of some initial state over a moving horizon, namely minimize some criterion of the form:

$$\int_{t-T}^{t} \|h(\chi_u(\tau, z_{t-T})) - y(\tau)\|^2 d\tau$$

w.r.t. z_{t-T} for any $t > T$, and $y(\tau)$ corresponding to the measured output over $[t-T, t]$ under the effect of the considered input u.

This is a general formulation for a solution to the problem, relying on available optimization tools and results for practical use and guarantees (see e.g. [1, 43, 47]): so it takes advantage of its systematic formulation, but suffers from usual drawbacks of nonlinear optimization (computational burden, local minima...).

Alternatively, one can use the idea of an explicit "feedback" in estimating $x(t)$, as this is done for control purposes: more precisely, noting that if one knows the initial value $x(0)$, one can get an estimate for $x(t)$ by simply integrating (1.1) from $x(0)$, the feedback-based idea is that if $x(0)$ is unknown, one can try to

[1] Including *bilinear systems* as a particular case, for which A, B, C, D are linear w.r.t. u.

correct on-line the integration $\hat{x}(t)$ of (1.1) from some erroneous $\hat{x}(0)$, according to the measurable error $h(\hat{x}(t)) - y(t)$, namely to look for an estimate $\hat{x}$ of x as the solution of a system:

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) + k(t, h(\hat{x}(t)) - y(t)), \ \textit{with } k(t, 0) = 0. \tag{1.3}$$

Such an auxiliary system is what will be defined as an *observer*, and the above equation is the most common form of an observer for a system (1.1) (as in the case of linear systems [36, 42]).

More generally, an observer can be defined as follows:

Definition 1. *Observer*
Considering a system (1.1), an observer is given by an auxiliary system:

$$\begin{aligned} \dot{X}(t) &= F(X(t), u(t), y(t), t) \\ \hat{x}(t) &= H(X(t), u(t), y(t), t) \end{aligned} \tag{1.4}$$

such that:

(i) $\hat{x}(0) = x(0) \Rightarrow \hat{x}(t) = x(t), \quad \forall t \geq 0$*;*
(ii) $\|\hat{x}(t) - x(t)\| \to 0$ *as* $t \to \infty$*;*

If (ii) holds for any $x(0), \hat{x}(0)$*, the observer is* global*.*
If (ii) holds with exponential convergence, the observer is exponential*.*
If (ii) holds with a convergence rate which can be tuned, the observer is tunable*.*

Notice that the overview on observer design presented in the sequel will mainly be dedicated to global exponential tunable observers.

Notice also that with notations of (1.1) and (1.4), the difference $\hat{x} - x$ will be called *observer error*.

Notice finally that with the above point of view, the observation problem turns to be a problem of observer design.

1.2 Nonlinear Observability

The purpose of this section is to discuss some conditions required on the system for possible solutions to the above mentioned observer problem. Such conditions above all correspond to what are usually called *observability* conditions. In short, they must express that there indeed is a possibility that the purpose of the observer can be achieved, namely that it might be possible to recover $x(t)$ from the only knowledge of u and y up to time t: at a first glance, this will be possible only if $y(t)$ bears the information on the full state vector when considered over some time interval: this roughly corresponds to the notion of "*observability*".

However, when restricting the definition of an observer strictly to items (i)-(ii), one can find observers yielding solutions to the observation problem even in cases when y does not bear the full information on the state vector:

Consider for instance the simple system:

$$\dot{x} = -x + u, \ y = 0$$

Clearly one cannot get any information on x from y, and yet the system:

$$\dot{\hat{x}} = -\hat{x} + u$$

satisfies (i)-(ii) and yields an estimate of x, since:

$$\overbrace{\hat{x} - x}^{\cdot} = -(\hat{x} - x).$$

This corresponds to a notion of "*detectability*". Notice that in that case, however, the rate of convergence cannot be tuned. Additional remarks in that respect can be found e.g. in [6].

If we restrict ourselves to the case of observers in the sense of *tunable* observers, then observability becomes a necessary condition. Such a condition can be specified in a geometric way as shown hereafter, while analytical additional conditions are discussed afterwards.

1.2.1 Geometric Conditions of Observability

For a possible design of a (tunable) observer, one must be able to recover the information on the state via the output measured from the initial time, and more particularly to recover the corresponding initial value of the state. This means that observability is characterized by the fact that from an output measurement, one must be able to distinguish between various initial states, or equivalently, one cannot admit *indistinguishable* states (following [33]):

Definition 2. *Indistinguishability*
A paire $(x_0, x_0') \in I\!R^n \times I\!R^n$ *is indistinguishable for a system (1.1) if:*

$$\forall u \in \mathcal{U},\ \forall t \geq 0,\ h(\chi_u(t, x_0)) = h(\chi_u(t, x_0')).$$

A state x *is indistinguishable from* x_0 *if the pair* (x, x_0) *is indistinguishable.*

From this, observability can be defined:

Definition 3. *Observability [resp. at* x_0*]*
A system (1.1) is observable [resp. at x_0*] if it does not admit any indistinguishable paire [resp. any state indistinguishable from* x_0*].*

This definition is quite general (global), and even too general for practical use, since one might be mainly interested in distinguishing states from their neighbors:

Consider for instance the case of the following system:

$$\dot{x} = u, \quad y = sin(x). \tag{1.5}$$

Clearly, y cannot help distinguishing between x_0 and $x_0 + 2k\pi$, and thus the system is not observable. It is yet clear that y allows to distinguish states of $]-\frac{\pi}{2}, \frac{\pi}{2}[$.

This brings to consider a weaker notion of observability:

Definition 4. *Weak observability [resp. at x_0]*
A system (1.1) is weakly observable [resp. at x_0] if there exists a neighborhood U of any x [resp. of x_0] such that there is no indistinguishable state from x [resp. x_0] in U.

Notice that this does not prevent from cases where the trajectories have to go far from U before one can distinguish between two states of U.
Consider for instance the case of a system:

$$\dot{x} = u; \quad y = h(x)$$

with h a $\mathcal{C}^\infty$ function as in figure 1.3 below: clearly the system is weakly observable since any state is distinguishable from any other one by applying some nonzero input u, but distinguishing two points of $[-1, 1]$ needs to wait for y to move away from 0.

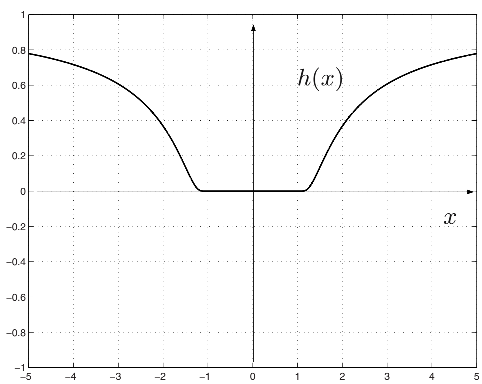


Fig. 1.3. Output function of a weakly but not locally observable system

Hence, to prevent from this situation, an even more local definition of observability can be given:

Definition 5. *Local weak observability [resp. at x_0]*
A system (1.1) is locally weakly observable [resp. at x_0] if there exists a neighborhood U of any x [resp. of x_0] such that for any neighborhood V of x [resp. x_0] contained in U, there is no indistinguishable state from x [resp. x_0] in V when considering time intervals for which trajectories remain in V.

This roughly means that one can distinguish every state from its neighbors without "going too far". This notion is of more interest in practice, and also presents the advantage of admitting some 'rank condition' characterization.

Such a condition relies on the notion of *observation space* roughly corresponding to the space of all observable states:

Definition 6. *Observation space*
The observation space for a system (1.1) is defined as the smallest real vector space (denoted by $\mathcal{O}(h)$) of $\mathcal{C}^\infty$ functions containing the components of h and closed under Lie derivation along $f_u := f(.,u)$ for any constant $u \in I\!R^m$ (namely such that for any $\varphi \in \mathcal{O}(h)$, $L_{f_u}\varphi \in \mathcal{O}(h)$, where $L_{f_u}\varphi(x) = \frac{\partial\varphi}{\partial x} f(x,u)$).

Definition 7. *Observability rank condition [resp. at x_0]*
A system (1.1) is said to satisfy the observability rank condition [resp. at x_0] if:

$$\forall x, \quad dimd\mathcal{O}(h)\mid_x = n \quad \text{[resp. } dimd\mathcal{O}(h)\mid_{x_0} = n]$$

where $d\mathcal{O}(h)\mid_x$ is the set of $d\varphi(x)$ with $\varphi \in \mathcal{O}(h)$.

From this we have [33]:

Theorem 1. *A system (1.1) satisfying the observability rank condition at x_0 is locally weakly observable at x_0.*
More generally a system (1.1) satisfying the observability rank condition is locally weakly observable.
Conversely, a system (1.1) locally weakly observable satisfies the observability rank condition in an open dense subset of X.

In short this follows from the facts that:

(i) the observability rank condition at some x_0 means the existence of n elements of the observation space defining a diffeomorphism around x_0;
(ii) for any indistinguishable pair (x_0, x_0') and any element $\varphi \in \mathcal{O}(h)$, $\varphi(x_0) = \varphi(x_0')$.

As an example of application, consider again system (1.5): for this system one clearly has $d\mathcal{O}(h) = span\{cos(x)dx, sin(x)dx\}$ and thus $dimd\mathcal{O}(h)\mid_{x_0} = 1$ for any x_0, namely the system satisfies the observability rank condition.
As a second example, consider a system of the following form:

$$\begin{array}{l} \dot{x} = Ax \\ y = Cx \quad \text{with } x \in I\!R^n. \end{array} \tag{1.6}$$

For this system, the observability rank condition is equivalent to local weak observability (which is itself equivalent to observability) and is characterized by the so-called *Kalman rank condition*:

Theorem 2. *For a system of the form (1.6)*

- *The observability rank condition is equivalent to $rank\mathcal{O}_m = n$ with $\mathcal{O}_m$ the so-called* observability matrix *defined by $\mathcal{O}_m = \begin{pmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{pmatrix}$ the ;*
- *The observability rank condition is equivalent to the observability of the system.*

The first point results from straightforward computations (e.g. as in [34]) since here the *kth* Lie derivation $L_f^k h(x) = CA^k x$, while the second one results from the definition of observability (see e.g. [40]).

Notice that if system (1.6) satisfies the above observability rank condition, the pair (A, C) is usually called *observable*.

Notice also that the above result also holds for controlled systems with $\dot{x} = Ax + Bu$.

Notice finally that the above observability rank condition is also sufficient for a possible observer design for (1.6) (even necessary and sufficient for a *tunable* observer design - see later).

However, in general, the observability rank condition is not enough for a possible observer design: this is due to the fact that in general, observability depends on the inputs, namely it does not prevent from the existence of inputs for which observability vanishes.

As a simple example, consider the following system:

$$\begin{aligned} \dot{x} &= \begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix} x \\ y &= \begin{pmatrix} 1 & 0 \end{pmatrix} x \end{aligned} \tag{1.7}$$

it is clearly observable for any constant input $u \neq 0$, but not observable for $u = 0$.

This means that the purpose of observer design requires a look at the inputs.

1.2.2 Analytic Conditions for Observability

In view of example (1.7) additional conditions to those previously presented might be required for possible observer designs, related to inputs. The purpose here is to discuss such conditions, while effective designs will be proposed later on.

More precisely, notions of *universal inputs* and *uniform observability* for systems (1.1) are first introduced (as in [15] for instance), and the stronger notions of *persistency* and *regularity* more usually defined for state affine systems [15] are then presented for the more general case of systems (1.1).

Definition 8. *Universal inputs [resp. on* $[0,t]$*]*
An input u is universal (resp. on $[0,t]$*) for system (1.1) if* $\forall x_0 \neq x_0'$, $\exists \tau \geq 0$ *(resp.* $\exists \tau \in [0,t]$*) s.t.* $h(\chi_u(\tau, x_0)) \neq h(\chi_u(\tau, x_0'))$.

An input u is a singular input if it is not universal.

As an example, for system (1.7), $u(t) = 0$ is a singular input.

It can be underlined here that for $\mathcal{C}^w$ systems, universal $\mathcal{C}^w$ inputs are dense in the set of $\mathcal{C}^w$ functions for the topology induced by $\mathcal{C}^\infty$ [46].

But one has to also notice that in general characterizing singular inputs is not easy. Things are easier for systems which do not admit such singular inputs:

Definition 9. *Uniformly observable systems (resp. locally)*
A system is uniformly observable (UO) if every input is universal (resp. on $[0,t]$*).*

Example 1. The system (1.8) below is uniformly observable [23]:

$$\dot{x} = \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ & & \ddots & \ddots & \\ & & & & 0 \\ \vdots & & & & 1 \\ 0 & \cdots & & & 0 \end{pmatrix} x + \begin{pmatrix} \varphi_1(x_1) \\ \varphi_2(x_1, x_2) \\ \vdots \\ \varphi_n(x_1, \ldots, x_{n-1}) \\ \varphi_{n-1}(x_1, \ldots, x_n) \end{pmatrix} u \tag{1.8}$$
$$y = x_1; \quad x = (x_1 \ldots, x_n)^T$$

This can be checked by considering any pair of distinct states $x \neq x'$: assuming indeed that their respective components x_k and x'_k coincide up to order i and that $x_{i+1} = x'_{i+1}$, then it is clear from (1.8) that $\dot{x}_{i-1} - \dot{x}'_{i-1} \neq 0$ and thus there exists t_0 such that $x_i(t) \neq x'_i(t)$ for $0 < t < t_0$. By induction, we easily end up with the existence of some time for which $x_1(t) \neq x'_1(t)$, which is true for any u.

This property actually means that observability is independent of the inputs and thus can allow an observer design also independent of the inputs, as in the case of LTI systems (see later).

For systems which are not uniformly observable, in general possible observers will depend on the inputs, and not all inputs will be admissible. Restricting the set of inputs to universal ones, as in the case of uniformly observable systems - for which *all* inputs are universal, is actually not enough:

Consider for instance the following system:

$$\dot{x} = \begin{pmatrix} 0 & u \\ -u & 0 \end{pmatrix} x; \quad y = (1 \;\; 0)x$$

For this system, the input defined by $u(t) = 1$ for $t < t_1$ and $u(t) = 0$ for $t \geq t_1$ is clearly universal, but if a disturbance appears after t_1, it is also clear that x cannot be correctly reconstructed.

This shows that universality must be guaranteed over the time, namely must be *persistent*. In order to characterize this persistency, notice first that we have the following property:

Proposition 1. *An input u is a universal input on $[0, t]$ for system (1.1) if and only if $\int_0^t ||h(\chi_u(\tau, x_0)) - h(\chi_u(\tau, x'_0))||^2 d\tau > 0$ for all $x_0 \neq x'_0$.*

This can be easily checked from definition 8.

Then one can define persistency as follows:

Definition 10. *Persistent inputs*
An input u is a persistent input for a system (1.1) if

$$\exists t_0, T : \forall t \geq t_0, \, \forall x_t \neq x'_t, \quad \int_t^{t+T} ||h(\chi_u(\tau, x_t)) - h(\chi_u(\tau, x'_t))||^2 d\tau > 0$$

Equivalently, this can be expressed as:

$$\int_{t-T}^{t} ||h(\chi_u(\tau, x_{t-T})) - h(\chi_u(\tau, x'_{t-T}))||^2 d\tau > 0, \quad \forall x_{t-T} \neq x'_{t-T} \tag{1.9}$$

which might be more suitable thinking of t as a current time and the inequality as a property on the past measurements.

This basically guarantees observability over a given time interval.

However this does not prevent observability from possibly vanishing as time goes to infinity. If this happens, effective observers would in general have to compensate this by a correction gain going to infinity:

Consider for instance the system defined by:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= x_2 + u \\ y &= \begin{pmatrix} \frac{x_1}{1+x_2^2} \\ x_2 \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} \end{aligned}$$

For this system any (bounded) input can be checked to be persistent in the sense of definition 10, since whenever x_{t-T} and x'_{t-T} differ from one another in their second component, (1.9) clearly holds, while if they only differ in their first component, it can be shown that the left-hand side roughly behaves as $(e^{-4(t-T)} - e^{-4t})$ which is indeed positive. But from this it is also clear that the system is *less and less observable* as $t \to \infty$.

It can be noticed that the state variables x_1, x_2 could here be reconstructed by an auxiliary system of the form (1.3), for instance given as follows:

$$\begin{aligned} \dot{\hat{x}}_1 &= y_2 - k_1(t)(\tfrac{\hat{x}_1}{1+y_2^2} - y_1) \\ \dot{\hat{x}}_2 &= \hat{x}_2 + u - k_2(\hat{x}_2 - y_2) \end{aligned}$$

for any $k_2 > 0$ and some k_1 growing as y_2 (namely of the form $\kappa_1(1 + y_2^2)$ for $\kappa_1 > 0$). This system indeed clearly guarantees that $\hat{x}_2 - x_2 \to 0$ (since $\frac{d}{dt}(\hat{x}_2 - x_2) = -k_2(\hat{x}_2 - x_2)$), and also that $\hat{x}_1 - x_1 \to 0$ (from $\frac{d}{dt}(\hat{x}_1 - x_1) = -\kappa_1(\hat{x}_1 - x_1)$), but with a correction gain $k_1(t)$ growing to infinity.

In order to avoid this, one needs a *guarantee* of observability, namely some *regular persistency*:

Definition 11. *Regularly persistent inputs*
An input u is a regularly persistent input for a system (1.1) if:

$$\begin{aligned} &\exists t_0, T : \forall x_{t-T}, x'_{t-T},\ \forall t \geq t_0, \\ &\int_{t-T}^{t} ||h(\chi_u(\tau, x_{t-T})) - h(\chi_u(\tau, x'_{t-T}))||^2 d\tau \geq \beta(\|x_{t-T} - x'_{t-T}\|) \end{aligned}$$

for some class $\mathcal{K}$ function β.

From the above proposed definitions of persistency and regular persistency, we recover the usual definitions already available for state affine systems (of [15] for instance):

Proposition 2. *For state affine systems, regularly persistent inputs are inputs u such that:*

$$\exists t_0, T, \alpha : \int_{t-T}^{t} \Phi_u^T(\tau, t-T) C^T C \Phi_u(\tau, t-T) d\tau \geq \alpha I > 0 \quad \forall t \geq t_0, \tag{1.10}$$

with $\Phi_u(\tau, t)$ the transition matrix classically defined by:

$$\frac{d\Phi_u(\tau, t)}{d\tau} = A(u(\tau))\Phi_u(\tau, t),\ \Phi_u(t, t) = I.$$

This is a straight consequence of the application of definition 11 to the case of state affine systems, with $\beta(\|z\|) = \alpha\|z\|^2$.

The left-hand side quantity in (1.10) corresponds to the so-called *observability Grammian*, classically defined for LTV systems, for any $t_1 < t_2 \in I\!R$, as:

$$\Gamma(t_1, t_2) = \int_{t_1}^{t_2} \Phi^T(\tau, t_1) C^T(\tau) C(\tau) \Phi(\tau, t_1) d\tau \tag{1.11}$$

where Φ as above denotes the transition matrix for the autonomous part of the system.

Remark 1

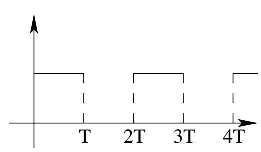
- Regularly persistent inputs for state affine systems are those making the system an LTV system *Uniformly Completely Observable* in the sense of Kalman [36] (since uniform complete observability for LTV systems is typically defined by (1.10);
- For general nonlinear systems, the definition is not of easy use, while for state affine or LTV systems, it is independent of initial states.

As an example of input properties, consider the following system:

$$\dot{x} = \begin{pmatrix} 0 & u \\ 0 & 0 \end{pmatrix} x;\ y = (1\ 0)x$$

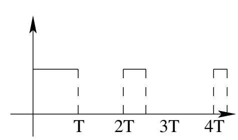
For this system, the input for instance defined by:

$$u(t) = 1 \text{ on } t \in [2kT, (2k+1)T[,\ k \geq 0$$
$$u(t) = 0 \text{ on } t \in [(2k+1)T, (2k+2)T[,\ k \geq 0$$



is regularly persistent, while that defined by:

$$u(t) = 1 \text{ on } t \in [2kT, (2k + \tfrac{1}{k+1})T[,\ k \geq 0$$
$$u(t) = 0 \text{ on } t \in [(2k + \tfrac{1}{k+1})T, (2k+2)T[,\ k \geq 0$$



is not [15].

Notice that for the reasons previously mentioned, regular persistency appears to be the property actually needed for effective state reconstruction.

However, it can be noticed that it depends on some time T roughly required to get enough information. If one is interested by an estimation 'in short time', he will need some kind of stronger observability property, corresponding to the application of what was originally called *locally regular inputs* on the basis of state affine systems [15]. In a more general context, this property can be formulated as follows:

Definition 12. *Locally regular inputs*
An input u is a locally regular input for a system (1.1) if:

$$\begin{array}{l}\exists T_0, \alpha : \forall x_{t-T}, x'_{t-T},\ \forall T \leq T_0,\ \forall t \geq T, \\ \displaystyle\int_{t-T}^{t} ||h(\chi_u(\tau, x_{t-T})) - h(\chi_u(\tau, x'_{t-T}))||^2 d\tau \geq \beta(\|x_{t-T} - x'_{t-T}\|, \frac{1}{T})\end{array}$$

for some class $\mathcal{KL}$ function β.

This property characterizes in some sense observability for arbitrarily short times. Obviously when T decays to zero, the observability cannot be kept guaranteed, which explains the decaying characteristic of β. When again considering state affine systems, we can roughly recover the definition previously used in [15, 4, 11] for instance, by considering some appropriate $\beta(\|x_t - x'_t\|, \frac{1}{T})$:

Proposition 3. *For state affine systems, locally regular inputs are inputs u such that:*

$$\exists T_0, \alpha : \forall T \leq T_0,\ \forall t \geq T,$$

$$\int_{t-T}^{t} \Phi_u^T(\tau, t-T) C^T C \Phi_u(\tau, t-T) d\tau \geq \alpha \frac{1}{T} \begin{pmatrix} T & & & 0 \\ & T & & \\ & & \ddots & \\ 0 & & & T^n \end{pmatrix}^2 \tag{1.12}$$

with $\Phi_u(\tau, t)$ the transition matrix as in proposition 2.

Here β is given by the right-hand side multiplied by $\|x_{t-T} - x'_{t-T}\|$: this is in particular motivated by the form of the Grammian for the linear part of a uniformly observable system (1.8) [15]:

$$\Gamma(t-T, t) = T \begin{pmatrix} 1 & \frac{T}{2} & \frac{T^2}{6} & \cdots \\ \frac{T}{2} & \frac{T^2}{3} & \frac{T^3}{8} & \cdots \\ \frac{T^2}{6} & \frac{T^3}{8} & \frac{T^4}{20} & \\ \vdots & \vdots & & \ddots \end{pmatrix},$$

which can indeed be lower bounded as in (1.12) for α small enough.

Obviously for a linear observable system, every input is locally regular.

Notice that the characterization (1.12) actually slightly differs from that previously considered in [15, 4, 11] ($\Phi_u(\tau, t)$ was considered instead of $\Phi_u(\tau, t-T)$ in the inequality), but they become equivalent whenever Φ_u (i.e. A) is bounded.

All this will tell us on some possible observer designs for classes of systems, as discusses in next section. Notice that more specific notions of observability, which have been introduced in connection with more specific designs not presented in details here will be omitted (such as 'infinitesimal observability' or 'differential observability', related to 'high gain techniques' as in [26], or 'generic observability' used in algebraic approaches as in [19] for instance). Additionally, some final remarks can be given as follows:

Remark 2

- If a system, e.g. control affine, is not observable in the sense of rank condition, it can be decomposed into observable and non observable subsystems as follows [34]:
$$\begin{aligned}\dot{\zeta}_1 &= f_1(\zeta_1,\zeta_2) + g_1(\zeta_1,\zeta_2)u\\ \dot{\zeta}_2 &= f_2(\zeta_2) + g_2(\zeta_2)u\\ y &= h_2(\zeta_2)\end{aligned}$$
where the subsystem in ζ_2 satisfies the observability rank condition. In that case one has to work on ζ_2.
- If the considered system is not observable, but satisfies the following:
$\forall u$ *such that* x_0 *and* x_0' are indistinguishable by u :
$$\chi_u(t,x_0) - \chi_u(t,x_0') \to 0 \, as \, t \to \infty$$
it satisfies a property of *detectability*, and in that case one may have the opportunity to design an observer in the sense of (i) and (ii).
- The analytic observability conditions which have here been presented (persistency, regular persistency, local regularity of definitions 10, 11 or 12 respectively) have been defined in terms of inputs, for controlled nonlinear systems of the form (1.1). But those definitions clearly still hold for time-varying systems (1.2). They can even be considered for uncontrolled systems (time-varying or not) since they are basically defined by output evolutions w.r.t. initial conditions. In other words, those notions could have been defined as various observability properties, parameterized by the input in the controlled case.

1.3 Nonlinear Observer Design

In view of the previously presented notions of observability, and in particular the problem of inputs which has been highlighted, it can easily be understood that observer designs will in general depend on the inputs of the system (or the analytic observability conditions previously highlighted).

However, in some cases, the observer design might be independent of the input, as in the case of uniformly observable systems for instance. Hence, following the viewpoint of [6], observer designs can be classified into 'uniform observers' and 'non uniform observers' w.r.t. inputs (or time), and we can roughly consider the following cases:

- For uniformly observable systems, one might design uniform observers;
- For non-uniformly observable systems, one might design non uniform observers.

The first ones correspond to the so-called *Luenberger observer* for LTI systems [42], while the second ones typically correspond to the case of *Kalman observers* for LTV systems [36]. Those observers will be first recalled in the next subsection, and then extended to nonlinear systems as *Luenberger-like* and *Kalman-like* designs. Subsection 1.3.2 will then discuss possible extensions of such 'basic designs'.

Notice that one might also design *non uniform* observers for uniformly observable systems (for instance using a Kalman approach), while in some cases *uniform* designs might be achieved for non uniformly observable systems (when the system satisfies some detectability property for instance, as discussed in [12]).

1.3.1 Basic Structures

Some observers are presented here for particular structures of systems. In the whole section, an observer is to be understood as a *global, exponential, tunable* observer.

Remember that the observer approach we consider is that of designing an auxiliary system intended to give an estimate $\hat{x}$ of the actual state vector x in the sense that $\hat{x}(t) - x(t) \to 0$ as $t \to \infty$. Hence the main problem turns to be an observer design so as to make the origin asymptotically stable for the corresponding observer error system. In all the presented results hereafter, this can be mostly studied by classical Lyapunov tools, as they are recalled in the appendix section 1.5.

Observer Designs for Linear Structures

The cases of LTI and LTV systems are here first considered.

Luenberger observer (for LTI systems)

Let us consider here LTI systems of the following form:

$$\begin{aligned} \dot{x}(t) &= Ax(t) + Bu(t) \\ y(t) &= Cx(t) \end{aligned} \tag{1.13}$$

For those systems we have the following classical (Luenberger) result [42]:

Theorem 3. *If system (1.13) satisfies the observability rank condition then there exists an observer of the form:*

$$\dot{\hat{x}}(t) = A\hat{x}(t) + Bu(t) - K(C\hat{x}(t) - y(t))$$

with K such that $A - KC$ is stable.

Remark 3. The rate of convergence can be arbitrarily chosen by appropriate design of K.

This can be established by showing that observability guarantees the existence of a transformation into a so-called observability canonical form, for which the design of an appropriate observer gain is straightforward (see e.g. [40]).

Kalman observer (for LTV systems)

Let us consider here LTV systems of the following form:

$$\begin{aligned} \dot{x}(t) &= A(t)x(t) + Bu(t) \\ y(t) &= C(t)x(t) \end{aligned} \tag{1.14}$$

with $A(t), C(t)$ uniformly bounded.

For those systems we have the following (Kalman-related) result [36, 16, 32, 10, 26]:

Theorem 4. *If system (1.14) is uniformly completely observable, then there exists an observer of the form:*

$$\dot{\hat{x}}(t) = A(t)\hat{x}(t) + B(t)u(t) - K(t)(C(t)\hat{x}(t) - y(t))$$

with $K(t)$ given by:

$$\begin{aligned} \dot{M}(t) &= A(t)M(t) + M(t)A^T(t) - M(t)C^T(t)W^{-1}C(t)M(t) + V + \delta M(t) \\ M(0) &= M_0 = M_0^T > 0, \; W = W^T > 0 \\ K(t) &= M(t)C^T(t)W^{-1} \end{aligned} \tag{1.15}$$

with either $\delta > 2\|A(t)\|$ for all t, or $V = V^T > 0$.

Remark 4

- The rate of convergence can be tuned by δ or V.
- For $\delta = 0$, we get the classical Kalman observer, the usual related condition for convergence being that (A, V) be *uniformly completely controllable* (dual of uniform complete observability).
- For $\delta = 0$, the observer is optimal in the sense of minimizing w.r.t. z:

$$\begin{aligned} &\int_0^t [(C(\tau)z(\tau) - y(\tau))^T W^{-1}(C(\tau)z(\tau) - y(\tau)) + v^T(\tau)V^{-1}v(\tau)]d\tau \\ &+(z_0 - \hat{x}_0)^T M_0^{-1}(z_0 - \hat{x}_0) \end{aligned}$$

  under $\dot{z}(t) = A(t)z(t) + v(t)$ $y(t) = C(t)z(t)$.

  Namely, it provides an explicit solution to the optimization-based approach mentioned in the introduction.

  It is also optimal in the sense of minimizing the mean of the square estimation error for a system affected by state white noises and measurement white noises, uncorrelated to each other, with V and W as respective variance matrices [40].

- The observer gain can also be computed as $K(t) = S^{-1}(t)C^T W^{-1}$ where S is the solution of:

$$\begin{aligned} &\dot{S}(t) = -A^T(t)S(t) - S(t)A(t) + C^T(t)W^{-1}C(t) - \delta S(t) - S(t)VS(t) \\ &S(0) = S^T(0) > 0 \end{aligned}$$

  which makes it a linear equation in S whenever V is chosen equal to 0. This is also true for all subsequent *Kalman-like* designs, even if they will be expressed in terms of (1.15).

The result of theorem 4 can be established by showing that:

(i) $\exists \alpha_1, \alpha_2, t_0$ such that $\forall t \geq t_0 :\ 0 < \alpha_1 I \leq M^{-1}(t) \leq \alpha_2 I$ basically from the condition of uniform complete observability;
(ii) $V(e,t) = e^T(t)M^{-1}(t)e(t)$ where $e := \hat{x} - x$ is a Lyapunov function for the observer error equation, which is exponentially decaying with a rate of decay tunable via δ or the minimal eigenvalue of V.

This can be shown either when $V = 0$ and $\delta > 2\|A(t)\|$ [10, 32], or when $V = V^T > 0$ and $\delta = 0$ [26].

On the basis of theorem 4, an extension can be intuitively derived for *nonlinear* systems relying on its first order approximation along the estimated trajectories, and known as *Extended Kalman Filter* (see e.g. [27]):

Definition 13. *Extended Kalman Filter (EKF)*
Given a nonlinear system of the form:

$$\begin{aligned} \dot{x}(t) &= f(x(t), u(t)) \\ y(t) &= h(x(t)) \end{aligned}$$

the corresponding Extended Kalman Filter is given by:

$$\dot{\hat{x}}(t) = f(\hat{x}(t), u(t)) - K(t)(h(\hat{x}(t)) - y(t))$$

where $K(t)$ is given as in the Kalman observer (1.15) with:

$$A(t) := \frac{\partial f}{\partial x}(\hat{x}(t), u(t)), \quad C(t) := \frac{\partial h}{\partial x}(\hat{x}(t))$$

This yields a candidate for a systematic observer design in front of a nonlinear system, but in general the convergence is not guaranteed, except under specific structure conditions (or domain of validity). This motivates the inspection of more specific nonlinear structures.

Observer Designs for Nonlinear Structures

Some observer designs are here presented for specific structures of nonlinear systems, extending the Luenberger and Kalman observers above recalled for linear systems.

Luenberger-like design (for UO systems)

Let us first consider classes of systems for which observability does not depend on the input, namely Uniformly Observable systems.

The idea is basically to rely on a linear time-invariant part in order to design a gain as in Luenberger observers, and either compensate exactly all nonlinear elements when possible (by output injection for instance), or dominate them via the linear part.

Additive output nonlinearity

Consider here a system of the form:

$$\begin{aligned} \dot{x} &= Ax + \varphi(Cx, u) \\ y &= Cx \end{aligned} \tag{1.16}$$

Here the nonlinearity can be constructed from direct measurements and thus compensated in the observer design (as originally proposed in [38, 39] for instance):

Theorem 5. *If (A, C) is observable, system (1.16) admits an observer of the form:*

$$\dot{\hat{x}} = A\hat{x} + \varphi(y, u) - K(C\hat{x} - y)$$

with K such that $A - KC$ is stable.

Remark 5

Clearly here, the observer error is exactly linear, and thus the convergence rate can be arbitrarily tuned by appropriate choice of K as in the case of linear systems.

Additive triangular nonlinearity

Consider here a system of the form:

$$\begin{aligned} \dot{x} &= A_0 x + \varphi(x, u) \\ y &= C_0 x \\ \textit{with} \quad A_0 &= \begin{pmatrix} 0 & 1 & & 0 \\ & & \ddots & \\ & & & 1 \\ 0 & & & 0 \end{pmatrix}, \quad C_0 = (1 \; 0 \; \cdots \; 0). \end{aligned} \tag{1.17}$$

Here the idea will be to use the uniform observability, and thus a structure as in (1.8), to weight a gain based on the linear part, so as to make the linear dynamics of the observer error to dominate the nonlinear one [24, 15, 26]:

Theorem 6. *If φ is globally Lipschitz w.r.t. x, uniformly w.r.t. u and such that:*

$$\frac{\partial \varphi_i}{\partial x_j}(x, u) = 0 \; for \; j \geq i + 1, \quad 1 \leq i, j \leq n,$$

system (1.17) admits an observer of the form:

$$\dot{\hat{x}} = A_0\hat{x} + \varphi(\hat{x}, u) - \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix} K_0(C_0\hat{x} - y)$$

with K_0 *such that* $A_0 - K_0C_0$ *is stable, and* λ *large enough.*

Remark 6

- This design is known as *high gain observer* since it relies on the choice of some sufficiently large tuning parameter λ;
- The larger λ is, the faster the convergence is.
- Output injection can also be used as in theorem 6.
- This design can be extended to systems of the following form [20, 25, 26]:

  $$\dot{x}(t) = f(x(t), u(t)),\ y(t) = C_0x(t)$$

  where $\frac{\partial f_i}{\partial x_j} = 0$ for $j > i + 1$ and $\frac{\partial f_i}{\partial x_{i+1}} \geq \alpha_i > 0$ for all x, u;
- The design can also be extended to multi-output uniformly observable systems [17, 18];
- This design has been shown to be very useful for observer-based control.

The result of theorem 6 can be established by showing that $V(e) = e^TP(\lambda)e$ is a Lyapunov function for the observer error equation, exponentially decaying with a rate of decay being tunable via λ, where:

$$e = \hat{x} - x \quad \text{and} \quad P(\lambda) = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix}^{-1} P_0 \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix}^{-1},$$

with P_0 such that:

$$P_0(A_0 - K_0C_0) + (A_0 - K_0C_0)^T = -I.$$

Kalman-like design (for non-UO systems)

In the case when observability depends on the inputs (systems which are not uniformly observable), the design will be restricted to some appropriate classes of inputs. Then the two possible cases of compensable or non compensable nonlinearities can again be considered.

<u>State affine systems</u>
Consider here a system of the form:

$$\begin{aligned} \dot{x}(t) &= A(u(t))x(t) + B(u(t)) \\ y(t) &= Cx(t) \end{aligned} \tag{1.18}$$

with $A(u(t))$ uniformly bounded.

Here the idea is that imposing the input function yields a linear time-varying system. Hence the following Kalman-like result holds [32, 15, 10]:

Theorem 7. *If u is regularly persistent for (1.18), then the system admits an observer of the form:*

$$\dot{\hat{x}}(t) = A(u(t))\hat{x}(t) + B(u(t)) - K(t)(C\hat{x}(t) - y(t))$$

with $K(t)$ given by:

$$\begin{aligned} \dot{M}(t) &= M(t)A^T(u(t)) + A(u(t))M(t) - M(t)C^TW^{-1}CM(t) + V + \delta M(t) \\ M(0) &= M^T(0) > 0,\ W = W^T > 0 \\ K(t) &= M(t)C^TW^{-1} \end{aligned}$$

with $\delta > 2\|A(u(t))\|$ or $V = V^T > 0$ as in LTV systems.

Remark 7

The convergence rate can be tuned by appropriate choice of δ or V.

This design can clearly be extended to systems which are affine in the unmeasured states, up to additive output nonlinearity, of the following form [29, 10]:

$$\begin{aligned} \dot{x}(t) &= A(u(t), Cx(t))x(t) + B(u(t), Cx(t)) \\ y(t) &= Cx(t) \end{aligned} \tag{1.19}$$

with $A(u(t), C\chi_u(t, x_0))$ bounded for any x_0.

Theorem 8. *If u is regularly persistent for (1.19), in the sense that it makes $v(t) := \begin{pmatrix} u(t) \\ C\chi_u(t, x_0) \end{pmatrix}$ regularly persistent for $\dot{x}(t) = A(v(t))x(t)$, $y(t) = Cx(t)$ for any x_0, then the system admits an observer of the form:*

$$\dot{\hat{x}}(t) = A(u(t), y(t))\hat{x}(t) + B(u(t), y(t)) - K(t)(C\hat{x}(t) - y(t))$$

with $K(t)$ given by:

$$\begin{aligned} \dot{M}(t) &= M(t)A^T(u(t), y(t)) + A(u(t), y(t))M(t) - M(t)C^TW^{-1}CM(t) \\ &\quad + V + \delta M(t) \\ M(0) &= M^T(0) > 0,\ W = W^T > 0 \\ K(t) &= M(t)C^TW^{-1} \end{aligned}$$

with $\delta > 2\|A(u(t), y(t))\|$ or $V = V^T > 0$.

State affine systems and additive triangular nonlinearity

Combining structure of system (1.18) or more generally (1.19) with that of system (1.17) leads to consider systems of the following form:

$$\begin{aligned} \dot{x} &= A_0(u,y)x + \varphi(x,u) \\ y &= C_0 x \quad \text{with} \\ A_0(u,y) &= \begin{pmatrix} 0 & a_{12}(u,y) & & 0 \\ & & \ddots & \\ & & & a_{n-1n}(u,y) \\ 0 & & & 0 \end{pmatrix} \textit{bounded}, \quad C_0 = (1\ 0 \cdots\ 0), \end{aligned} \tag{1.20}$$

and with φ as in theorem 6.

This means that the observer will need to rely on high gain, but for a non uniformly observable system. As a consequence, the observability property corresponding to observability for short times of proposition 3 will be here required, but parameterized by y as above:

Theorem 9. *If φ is globally Lipschitz w.r.t. x, uniformly w.r.t. u and such that:*

$$\frac{\partial \varphi_i}{\partial x_j}(x,u) = 0 \, for \, j \geq i+1, \quad 1 \leq i,j \leq n,$$

and u is locally regular for (1.17), in the sense that it makes $v(t) := \begin{pmatrix} u(t) \\ C\chi_u(t,x_0) \end{pmatrix}$ locally regular for $\dot{x}(t) = A(v(t))x(t)$, $y(t) = Cx(t)$ for any x_0, then the system admits an observer of the form:

$$\dot{\hat{x}} = A_0(u,y)\hat{x} + \varphi(\hat{x},u) - \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix} K_0(t)(C_0\hat{x} - y)$$

with $K_0(t)$ given by:

$$\begin{aligned} \dot{M}(t) &= \lambda(M(t)A^T(u(t),y(t)) + A(u(t),y(t))M(t) - M(t)C^T W^{-1} C M(t) + \delta M(t)) \\ M(0) &= M^T(0) > 0,\, W = W^T > 0 \\ K(t) &= M(t)C^T W^{-1} \end{aligned}$$

$\delta > 2\|A(u,y)\|$ and $\lambda = \frac{1}{T}$ large enough.

This can be established by showing that [4]:

(i) From local regularity assumption:

$$\exists \lambda > 0,\ \forall \lambda \geq \lambda_0,\ \forall t \geq \frac{1}{\lambda}, \quad 0 < \alpha_1 I \leq M^{-1}(t) \leq \alpha_2 I$$

for α_1, α_2 independent of λ.

(ii) $V(e,t) = e^T P(\lambda,t)e$ is a Lyapunov function for the observer error equation, exponentially decaying, with a rate of decay tunable by λ, where $e = \hat{x} - x$ and

$$P(\lambda,t) = \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix}^{-1} M^{-1}(t) \begin{pmatrix} \lambda & & 0 \\ & \ddots & \\ 0 & & \lambda^n \end{pmatrix}^{-1}$$

Notice that this kind of design also holds for systems of the form (1.20) where the a_{ii+1}'s in A_0 are matrices instead of scalars [11].

Finally, notice again that we have presented observer designs in terms of UO systems (for which observers with constant gains have been given) and non UO systems (for which observers with varying gains have been presented). But obviously one could design an observer with a varying gain for UO systems, since in that case any input will satisfy the appropriate condition for the observer to work. Conversely, in some cases one can design an observer with a *constant* gain even if the system is not UO: this can be done provided the system satisfies some detectability property as mentioned before [12].

1.3.2 Advanced Designs

The presentation of possible observer designs in previous section has been restricted to very specific structures of systems. In this section are presented some ways to deal with nonlinear systems which do not a priori satisfy the structures previously presented.

Interconnection-Based Design

The first way to extend the class of systems for which an observer can be designed is to interconnect observers in order to design an observer for some interconnected system, when possible. If indeed a system is not under a form for which an observer is already available, but can be seen as an interconnection between several subsystems each of which would admit an observer if the states of the other subsystems were known, then a candidate observer for the interconnection of these subsystems is given by interconnecting available sub-observers (e.g. as in [13]). This is sketched by figure 1.4 below for the case of two subsystems.

As a simple example, consider the following system:

$$\begin{aligned} \dot{x}_1 &= x_2 \\ \dot{x}_2 &= u_1 \\ \dot{x}_3 &= x_4 + \varphi(x_2) \\ \dot{x}_4 &= u_2 \\ y &= \begin{pmatrix} x_1 \\ x_3 \end{pmatrix} \end{aligned} \tag{1.21}$$

Clearly here one can consider the system as the interconnection of the following two subsystems:

$$(\Sigma_1) \begin{cases} \dot{x}_1 = x_2 \\ \dot{x}_2 = u_1 \\ y_1 = x_1 \end{cases} \quad \text{and} \quad (\Sigma_2) \begin{cases} \dot{x}_3 = x_4 + \varphi(v) \\ \dot{x}_4 = u_2 \\ y = x_3 \end{cases}$$

where $v = x_2$ defines the interconnection.

It is also clear that (Σ_1) being linear and observable, it admits an observer (say O_1), as well as (Σ_2) whenever v is considered as a known input for (Σ_2) (let (say $O_2(v)$) denote the corresponding observer).

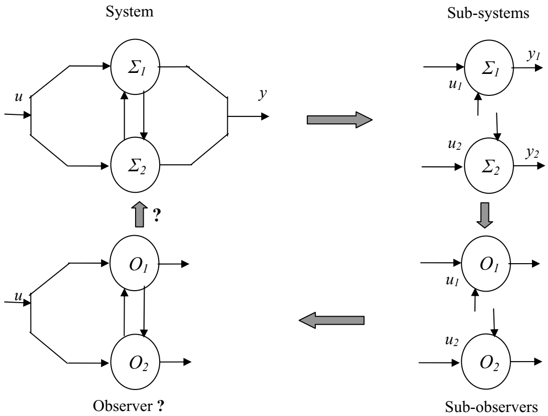


Fig. 1.4. Interconnection-based observer design

The idea is then to get an observer for the whole system from the interconnection $(O_1)+(O_2(\hat{x}_2))$ where $\hat{x}_2$ is provided by (O_1).

It can here be checked that for instance if φ is globally Lipschitz, $(O_1)+(O_2(\hat{x}_2))$ can indeed yield an observer.

Now if (Σ_2) is replaced by:

$$(\Sigma_2') \begin{cases} \dot{x}_3 = \varphi(x_2)x_4 \\ \dot{x}_4 = u_2 \\ y_2 = x_3 \end{cases}$$

it also results from previous section that an observer can be designed for (Σ_2') if x_2 is considered to be a known input, provided that this input is regularly persistent for (Σ_2'). If φ is globally Lipschitz, it can again be checked that this is enough for making it possible to get an observer for the whole system by interconnecting sub-observers (e.g. as in [10]).

This shows that under appropriate conditions separate possible designs can indeed yield some overall observer. But it does not go that well in any case. Consider for instance the following system:

$$\dot{x}_1 = -\frac{1}{2(t+1)}x_1;\ y_1 = 0 \tag{1.22}$$

$$\dot{x}_2 = -\frac{1}{4(t+1)}x_2 + x_1;\ y_2 = 0 \tag{1.23}$$

This system can be seen as an interconnection via x_1 between two subsystems respectively defined by (1.22) and (1.23). Clearly each of them admits an observer (here not tunable) as follows, as long as x_1 is assumed to be known for the second one:

$$\dot{\hat{x}}_1 = -\frac{1}{2(t+1)}\hat{x}_1;\ y_1 = 0 \tag{1.24}$$

$$\dot{\hat{x}}_2 = -\frac{1}{4(t+1)}\hat{x}_2 + x_1;\ y_2 = 0 \tag{1.25}$$

But if we inject $\hat{x}_1$ given by (1.24) into (1.25), one can check that the error equation is not stable.

This just illustrates the fact that in general, the stability of the interconnected observer is not guaranteed by that of each sub-observer, in the same way as separate designs of observer and controller do not in general result in some stable observer-based controller for nonlinear systems (no *separation principle*).

This means that the stability of interconnection of sub-observers requires a specific attention. Conditions can indeed be derived so as to guarantee a possible design by interconnection of separate subdesigns, either in the case of *cascade* interconnection as in the above examples, or even in the case of *full* interconnection [13].

Full interconnection

Let us first consider the general case of *full* interconnection, via the example of systems made of two subsystems for the sake of illustration, and described by the following representation:

$$(\Sigma)\begin{cases} \dot{x}_1 = f_1(x_1, x_2, u),\ u \subset U \subset I\!R^m;\ f_i\ \mathcal{C}^\infty \text{ function},\ i = 1,2; \\ \dot{x}_2 = f_2(x_2, x_1, u),\ x_i \in X_i \subset I\!R^{n_i},\ i = 1,2; \\ y\ = (h_1(x_1),\, h_2(x_2))^T = (y_1,\, y_2)^T,\ y_i \in I\!R^{\eta_i},\ i = 1,2. \end{cases} \tag{1.26}$$

Assume also that $u(.) \in \mathcal{U} \subset \mathcal{L}^\infty(I\!R^+, U)$, and set $\mathcal{X}_i := \mathcal{AC}(I\!R^+, I\!R^{n_i})$ the space of absolutely continuous function from $I\!R^+$ into $I\!R^{n_i}$. Finally, when $i \in \{1,2\}$, let $\bar{\imath}$ denote its complementary index in $\{1,2\}$.

The idea here is that system (1.26) can be seen as the interconnection of two subsystems (Σ_i) for $i = 1,2$ given by:

$$(\Sigma_i) \quad \dot{x}_i = f_i(x_i, v_{\bar{\imath}}, u), \quad y_i = h_i(x_i), \qquad (v_{\bar{\imath}}, u) \in \mathcal{X}_{\bar{\imath}} \times \mathcal{U}. \tag{1.27}$$

Assume that for each system (Σ_i), one can design an observer $(\mathcal{O}_i)$ of the following form:

$$(\mathcal{O}_i) \quad \dot{z}_i = f_i(z_i, v_{\bar{\imath}}, u) + k_i(g_i, z_i)(h_i(z_i) - y_i), \quad \dot{g}_i = G_i(z_i, v_{\bar{\imath}}, u, g_i), \tag{1.28}$$

for smooth k_i, G_i and $(z_i, g_i) \in (I\!R^{n_i} \times G\!\!\!G_i)$, $G\!\!\!G_i$ positively invariant by (1.28).

The point is to look for an observer for (1.26) under the form of the following interconnection:

$$(\mathcal{O})\begin{cases} \dot{\hat{x}}_i = f_i(\hat{x}_i, \hat{x}_{\bar{\imath}}, u) + k_i(\hat{g}_i, \hat{x}_i)(h_i(\hat{x}_i) - y_i);\ i = 1,2; \\ \dot{\hat{g}}_i = G_i(\hat{x}_i, \hat{x}_{\bar{\imath}}, u, \hat{g}_i);\ i = 1,2 \end{cases} \tag{1.29}$$

Set $e_i := z_i - x_i$, and for any $u \in \mathcal{U}, v_{\bar{\imath}} \in \mathcal{X}_i$ consider the following system (where $k_i^{v_{\bar{\imath}}}(t)$ denotes gain $k_i(g_i, z_i)$ defined in (1.28)) :

$$\mathcal{E}_i^{(u,v_{\bar{\imath}})} \begin{cases} \dot{e}_i = f_i(z_i, v_{\bar{\imath}}, u) - f_i(z_i - e_i, v_{\bar{\imath}}, u) + k_i^{v_{\bar{\imath}}}(t)(h_i(z_i) - h_i(z_i - e_i)) \\ \dot{z}_i = f_i(z_i, v_{\bar{\imath}}, u) + k_i^{v_{\bar{\imath}}}(t)(h_i(z_i) - h_i(z_i - e_i)) \\ \dot{g}_i = G_i(z_i, v_{\bar{\imath}}, u, g_i). \end{cases}$$

Then sufficient conditions for (1.29) to be an observer for (1.26) have been expressed in [13] as follows:

Theorem 10. *[13] If for $i = 1, 2$, any signal $u \in \mathcal{U}, v_{\bar{\imath}} \in \mathcal{AC}(\mathbb{R}^+, \mathbb{R}^{n_{\bar{\imath}}})$, and any initial value $(z_i^0, g_i^0) \in \mathbb{R}^{n_i} \times \mathbb{G}_i$, $\exists V_i(t, e_i), W_i(e_i)$ positive definite functions such that:*

(i) $\forall x_i \in X_i; \forall e_i \in \mathbb{R}^{n_i}; \forall t \geq 0,$

$$\frac{\partial V_i}{\partial t}(t, e_i) + \frac{\partial V_i}{\partial e_i}(t, e_i)[f_i(x_i + e_i, v_{\bar{\imath}}(t), u(t)) - f_i(x_i, v_{\bar{\imath}}(t), u(t)) \\ + k_i^{v_{\bar{\imath}}}(t)(h_i(x_i + e_i) - h_i(x_i))] \leq -W_i(e_i)$$

(ii) $\exists \alpha_i > 0; \forall x_i \in X_i; \forall x_{\bar{\imath}} \in \mathbb{R}^{n_{\bar{\imath}}}; \forall e_i \in \mathbb{R}^{n_i}; \forall e_{\bar{\imath}} \in \mathbb{R}^{n_{\bar{\imath}}}; \forall t \geq 0,$

$$\left\| \frac{\partial V_i}{\partial e_i}(t, e_i)[f_i(x_i, x_{\bar{\imath}} + e_{\bar{\imath}}, u(t)) - f_i(x_i, x_{\bar{\imath}}, u(t))] \right\| \leq \alpha_i \sqrt{W_i(e_i)} \sqrt{W_{\bar{\imath}}(e_{\bar{\imath}})},$$

(iii) $\alpha_1 + \alpha_2 < 2,$

then (1.29) is an asymptotic observer for (1.26).

•

This can be established on the basis of Lyapunov arguments by appropriately combining V_1 and V_2. The result can be extended to more than two subsystems by using Lyapunov stability analysis of interconnected systems for instance as in [37].

As an example, this approach can yield observers for systems of the form:

$$\begin{aligned} \dot{x}_1 &= A_1 x_1 + f_1(x_1, x_2, u) \\ \dot{x}_2 &= A_2 x_2 + f_2(x_1, x_2, u) \\ y &= \begin{pmatrix} C_1 x_1 \\ C_2 x_2 \end{pmatrix} \end{aligned}$$

relying on high gain separate designs for x_1 and x_2 for instance, or even Kalman separate designs (in particular if $A_i = A_i(u)$ for some $i \in \{1, 2\}$ for instance) [6, 13].

Cascade interconnection

In the *weaker* case of *cascade* interconnection, namely when $f_1(x_1, x_2, u) = f_1(x_1, u)$ in (1.26), various results have been proposed for the stability of the interconnected system. Let us report here the weakened assumptions proposed in [13] in this context of observer design:

Theorem 11. *Assume that:*

I. System $\dot{x}_1 = f_1(x_1, u)$; $y_1 = h_1(x_1)$ *admits an observer* $(\mathcal{O}_1)$ *as in (1.28) (without* v_2*), s.t.* $\forall u \in \mathcal{U}$ *and* $\forall x_1(t)$ *admissible trajectory of the system associated to u:*

$$\lim_{t\to\infty} e_1(t) = 0 \text{ and } \int_0^{+\infty} \|e_1(t)\| dt < +\infty \quad (\text{with } e_1 := z_1 - x_1); \tag{1.30}$$

II. $\exists c > 0$; $\forall u \in U$; $\forall x_2 \in X_2$, $\|f_2(x_2, x_1, u) - f_2(x_2, x_1', u)\| \leq c\|x_1 - x_1'\|$;

III. $\forall u \in \mathcal{U}, \forall v_1 \in \mathcal{AC}(\mathbb{R}^+, \mathbb{R}^{n_1})$, $\forall z_2^0, g_2^0$, $\exists v(t, e_2), w(e_2)$ *positive definite functions s.t for every trajectory of* $\mathcal{E}_2^{(u,v_1)}$ *with* $z_2(0) = z_2^0, g_2(0) = g_2^0$*:*
(i) $\forall x_2 \in X_2, e_2 \in \mathbb{R}^{n_2}, t \geq 0,$

$$\frac{\partial v}{\partial t}(t, e_2) + \frac{\partial v}{\partial e_2}(t, e_2)[f_2(x_2 + e_2, v_1(t), u(t)) - f_2(x_2, v_1(t), u(t)) + k_2^{v_1}(t)(h_2(x_2 + e_2) - h_2(x_2))] \leq -w(e_2)$$

(ii) $\forall e_2 \in \mathbb{R}^{n_2}, t \geq 0;\ v(t, e_2) \geq \bar{w}(e_2)$
(iii) $\forall e_2 \in \mathbb{R}^{n_2} \backslash \mathcal{B}(0, r), t \geq 0;\ \left\| \frac{\partial v}{\partial e_2}(t, e_2(t)) \right\| \leq \lambda(1 + v(t, e_2(t)))$ *for some constants* $\lambda, r > 0$ *and* $\mathcal{B}(0, r) := \{e_2 : \|e_2\| \leq r\}$.

Then:

$$\begin{aligned} \dot{\hat{x}}_1 &= f_1(\hat{x}_1, u) + k_1(\hat{g}_1, \hat{x}_1)(h_1(\hat{x}_1) - h_1(x_1)) \\ \dot{\hat{x}}_2 &= f_2(\hat{x}_1, \hat{x}_2, u) + k_2(\hat{g}_2, \hat{x}_2)(h_2(\hat{x}_1) - h_2(x_1)) \\ \dot{\hat{g}}_1 &= G_1(\hat{x}_1, u, \hat{g}_1); \\ \dot{\hat{g}}_2 &= G_2(\hat{x}_2, \hat{x}_1, u, \hat{g}_2). \end{aligned} \tag{1.31}$$

is an observer for (1.26) where $f_1(x_1, x_2, u) = f_1(x_1, u)$.

•

Once again this result can be established by Lyapunov analysis.

Notice that the here above proposed conditions might be modified by using more specific stability results for cascade systems.

A typical example of cascade observer design can be found in [10], where the Kalman-like design was extended to systems of the following form:

$$\begin{aligned} \dot{x}_1 &= A_1(u, y)x_1 + B_1(u, y) \\ \dot{x}_2 &= A_2(u, y, x_1)x_2 + B2(u, y, x_1) \\ y &= \begin{pmatrix} C_1 x_1 \\ C_2 x_2 \end{pmatrix} \end{aligned}$$

Those examples show how using available observers for systems in some particular forms, one might be able to design observers for further nonlinear systems. Next section proposes another way to do so.

Transformation-Based Design

Principle

The observer designs presented till now are still all based on particular structures of the system (either isolated or interconnected). The subsequent idea is that these designs can also give state observers for systems which can be turned into one of these forms by an appropriate transformation. The most commonapproach in that respect is to consider changes of state coordinates. Such a relationship defines some *system equivalence*:

Definition 14. *System equivalence [resp. at x_0]*
A system described by:

$$\begin{cases} \dot{x} = f(x,u) = f_u(x)\, x \in I\!R^n, u \in I\!R^m \\ y = h(x) \in I\!R^p \end{cases} \tag{1.32}$$

will be said to be **equivalent** *[resp.* **at x_0**] *to the system:*

$$\begin{cases} \dot{z} = F(z,u) = F_u(z) \\ y = H(z) \end{cases} \tag{1.33}$$

if there exists a diffeomorphism $z = \Phi(x)$ defined on $I\!R^n$ [resp. some neighbourhood of x_0] such that:

$$\forall u \in I\!R^m, \quad \frac{\partial \Phi}{\partial x} f_u(x) \,|_{x=\Phi^{-1}(z)} = F_u(z) \qquad \text{and} \qquad h \circ \Phi^{-1} = H.$$

Systems (1.32) and (1.33) are then said to be equivalent by $z = \Phi(x)$.

The interest of such a property for observer design can then be illustrated by the following proposition (e.g. as in [6]):

Proposition 4. *Given two systems (Σ_1) and (Σ_2) respectively defined by:*
$(\Sigma_1) \begin{cases} \dot{x} = X(x,u) \\ y = h(x) \end{cases}$ *and* $(\Sigma_2) \begin{cases} \dot{z} = Z(z,u) \\ y = H(z) \end{cases}$
and equivalent by $z = \Phi(x)$,
If:

$$(\mathcal{O}_2) \begin{cases} \dot{\hat{z}} = Z(\hat{z},u) + k(w, H(\hat{z}) - y)) \\ \dot{w} = F(w,u,y) \end{cases}$$

is an observer for (Σ_2),
Then:

$$(\mathcal{O}_2) \begin{cases} \dot{\hat{x}} = X(\hat{x},u) + \left(\frac{\partial \Phi}{\partial x}\right)^{-1}_{|\hat{x}} k(w, h(\hat{x}) - y) \\ \dot{w} = F(w,u,y) \end{cases}$$

is an observer for (Σ_1).

From this indeed, if a system is not of an appropriate structure for an observer design in view of previous sections, but is equivalent to some other system which does have some appropriate structure, then the observer problem can be solved for the original system.

Examples

The idea of proposition 4 has motivated various works on characterizing systems which can be turned into some appropriate structures for observer design, from the linear one up to output injection [38, 14, 39] to several forms of cascade block state affine systems up to nonlinear injections from block to block as in (1.34) below for instance [30, 45, 10, 31, 8, 9, ...].

$$\begin{cases} \dot{z}_1 = A_1(u, y^1)z_1 + \varphi_1(u, y^1) \\ \dot{z}_2 = A_2(u, y^2, z_1)z_2 + \varphi_2(u, y^2, z_1) \\ \quad \vdots \\ \dot{z}_q = A_q(u, y^q, z_1, \dots z_{q-1})z_q + \varphi_q(u, y^q, \dots z_{q-1}) \\ y = \begin{pmatrix} C_1 z_1 \\ \vdots \\ C_q z_q \end{pmatrix} = \begin{pmatrix} y^1 \\ \vdots \\ y^q \end{pmatrix} \\ \quad u \in I\!\!R^m, z_i \in I\!\!R^{n_i}, y^i \in I\!\!R^{\nu_i}, \end{cases} \tag{1.34}$$

As a simple illustrative example, let us consider here the problem of turning a nonlinear system:

$$\dot{x} = f(x)$$
$$y = h(x), \ x \in I\!\!R^n$$

into a linear observable form up to output injection as follows:

$$\dot{x} = Ax + \varphi(Cx)$$
$$y = Cx$$

Necessary and sufficient conditions for this problem to be solvable have been given in terms of differential geometry in [38].

A constructive algorithm to simultaneously check the possibility of the transformation and construct φ can alternatively be given in the spirit of [28] as follows:

1. Get the representation:
$$y^{(n)} = \Phi(y, \dot{y}, \dots y^{(n-1)})$$
and set $z_1 := y$.
2. For $i \geq 1$, define φ_i by: $\frac{\partial \varphi_i}{\partial y} = \frac{\partial z_i^{(n-i+1)}}{\partial y^{(n-i)}}$;
If φ_i is not only a function of y, the transformation fails and the procedure ends. Else, set: $z_{i+1} := \dot{z}_i - \varphi_i$
3. Continue until $i = n$ or the procedure aborts.

The procedure is clearly sufficient, and it can be checked that it is indeed necessary.

As a second simple example, turning some $n-$dimensional nonlinear control affine system into the appropriate structure for high gain observer design, if possible, is obtained by the following transformation [23, 24]:

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

Finally, it can be underlined that some enlargement of the class of systems admitting an observer on the basis of the particular structures highlighted in the above presentation can also be obtained by further considering output transformations (e.g. as in [39, 28, 5]), or state extension (e.g. using immersion [22, 41, 35, 11]), for instance.

In particular, it has been shown in [5] that any control-affine system satisfying the observability rank condition can be turned into a form (1.20) for appropriate dimensions of the a_{ii+1}'s.

1.4 Conclusion

The purpose in this chapter was to give some overview on techniques of observer design for nonlinear systems. Clearly this presentation follows a particular viewpoint on the problem, and does not claim to be exhaustive. In particular the most important notions of observability (from this viewpoint) have been reviewed, and some observers have been presented according to two types of designs in that respect: uniform and non uniform ones w.r.t. input (or time). Those designs are in particular driven by specific structures of systems, and admit smooth explicit gains. Extensions of such designs to more general structures by interconnections and transformations have also been discussed. More details on some of the mentioned techniques can be found in the subsequent chapters - such as high gain designs, immersion-based results or optimization-based approaches. On the other hand, further comments on *detectability* and related designs have for instance been omitted, as well as various other technical approaches where the design is not necessarily smooth (as in sliding modes [21, 3, ...]), explicit (as in LMI-based designs [2, 44, ...]) or exact (as in many approximate approaches).

1.5 Appendix: Lyapunov Tools

The purpose of observers being asymptotic state reconstruction, an observer for a given system is to be an auxiliary system such that the error between the observer state and the system state asymptotically decays to zero, namely 0 is to be an asymptotically stable equilibrium for the error system.

From this, Lyapunov tools for stability analysis are instrumental in designing observers, and the purpose here is thus to recall the main results in that respect (as they can be found for instance in [37]).

In general the considered stability will be that of some nonautonomous system, namely a system of the form:

$$\dot{x}(t) = f(x(t), t) \tag{1.35}$$

such that $f(0,t) = 0$ for any $t \geq 0$, and where f is regular enough (at least piecewise continuous in t and locally Lipschitz in x on $D \times [0, \infty)$ where D is some state domain of $\mathbb{R}^n$ containing 0).

For such a system one can consider the following:

Definition 15. *Stability*
The equilibrium $x = 0$ of (1.35) will be said to be uniformly stable *if:*

$$\forall \epsilon > 0,\, \exists \delta > 0,\, \textit{independent of } t_0 : \|x(t_0)\| < \delta \Rightarrow \|x(t)\| < \epsilon,\, \forall t \geq t_0 \geq 0.$$

The equilibrium is uniformly asymptotically stable *if it is uniformly stable and:*

$$\exists c > 0 \textit{ independent of } t_0 : \forall \|x(t_0)\| < c,\, lim_{t \to \infty} \|x(t)\| = 0,\, \textit{uniformly in } t_0,$$

namely:

$$\|x(t)\| \leq \beta(\|x(t_0)\|, t - t_0),\, \forall t \geq t_0 \geq 0, \forall \|x(t_0)\| < c$$

for a constant $c > 0$ independent of t_0, and a continuous function $\beta(r,s)$ vanishing at 0 and strictly increasing w.r.t. its first argument, and decreasing w.r.t. its second argument even going to zero at infinity (class $\mathcal{KL}$ function).

The equilibrium is globally uniformly asymptotically stable *if it satisfies the above inequality for any initial state $x(t_0)$.*

The equilibrium is exponentially stable *if it satisfies the above inequality with $\beta(r,s) = kre^{-\gamma s}$, $k, \gamma > 0$, and* globally exponentially stable *if this condition holds for any initial state.*

Then we can recall the following:

Theorem 12. *Let $V : D \times [0, \infty) \to \mathbb{R}$ be a $\mathcal{C}^1$ function such that:*

1. $W_1(x) \leq V(x,t) \leq W_2(x)$
2. $\dfrac{\partial V}{\partial t} + \dfrac{\partial V}{\partial x} f(t,x) \leq -W_3(x)$

for $t \geq 0, x \in D$ and W_1, W_2, W_3 are continuous positive definite functions on D. Then the equilibrium $x = 0$ is uniformly asymptotically stable.

If the above conditions hold globally and in addition W_1 is radially unbounded *($W_1(x) \to \infty$ if $\|x\| \to \infty$), then $x = 0$ is* globally *uniformly asymptotically stable.*

If in fact:

$$W_i(x) \geq k_i \|x\|^c \textit{ for } i = 1, 3,\; W_2(x) \leq k_2 \|x\|^c$$

for $k_1, k_2, k_3, c > 0$, in the above conditions, then then $x = 0$ is exponentially *stable.*

If those conditions hold globally, then then $x = 0$ is globally exponentially *stable.*

Notice that a function V satisfying the first inequality of item 1 above is called *proper*, and it is called *decrescent* if the second inequality holds.

References

[1] M. Alamir. Optimization-based nonlinear observers revisited. *Int. Journal of Control*, 72(13):1204–1217, 1999.

[2] M. Arcak and P. Kokotović. Nonlinear observers: a circle criterion design and robustness analysis. *Automatica*, 37(12):1923–1930, 2001.

[3] J.P. Barbot, T. Boukhobza, and M. Djemai. Sliding mode observer for triangular input form. In *Proc. 35th IEEE Conf. on Decision and Control, Kobe, Japan*, pages 1489–90, 1996.

[4] G. Besançon. Further results on high gain observers for nonlinear systems. In *Proc. 38th IEEE Conf. on Decision and Control, Phoenix, USA*, 1999.

[5] G. Besançon. On output transformations for state linearization up to output injections of nonlinear systems. *IEEE Trans. on Automatic Control*, 44(11):1975–1981, 1999.

[6] G. Besançon. A viewpoint on observability and observer design for nonlinear systems. In *New Directions in Nonlinear Observer Design - Lecture Notes in Control and Info. Sciences 244*, pages 2–21. H. Nijmeijer and T. Fossen Eds. Springer-Verlag, London, 1999.

[7] G. Besançon. Observer design for nonlinear systems. In *Advanced topics in control systems theory - Lecture Notes in Control and Info. Sciences 328*, pages 61–89. A. Loría, F. Lamnabhi-Lagarrigue and E. Panteley Eds, Springer-Verlag, London, 2006.

[8] G. Besançon and G. Bornard. State equivalence based observer design for nonlinear control systems. In *Proc. of IFAC World Congress, San Francisco, CA, USA*, pages 287–292, 1996.

[9] G. Besançon and G. Bornard. On charactering classes of observer forms for nonlinear systems. In *Proc. 4th European Control Conf., Brussels, Belgium*, 1997.

[10] G. Besançon, G. Bornard, and H. Hammouri. Observer synthesis for a class of nonlinear control systems. *European J. of Control*, 3(1):176–193, 1996.

[11] G. Besançon and A. Ţiclea. Immersion-based observer design for rank-observable nonlinear systems. *IEEE Trans. on Automatic Control*, 52(1):83–88, 2007.

[12] G. Besançon and H. Hammouri. On uniform observation of non-uniformly observable systems. *Systems & Control Letters*, 33(1):1–11, 1996.

[13] G. Besançon and H. Hammouri. On observer design for interconnected systems. *Journal of Math. Systems, Estimation and Control*, 8(3), 1998.

[14] D. Bestle and M. Zeitz. Canonical form observer design for nonlinear time-variable systems. *Int. Journal of Control*, 38(2):419–431, 1983.

[15] G. Bornard, F. Celle-Couenne, and G. Gilles. Observability and observers. In *Nonlinear Systems - T.1, Modeling and Estimation*, pages 173–216. Chapman & Hall, London., 1995.

[16] G. Bornard, N. Couenne, and F. Celle. Regularly persistent observer for bilinear systems. In *New Trends in Nonlinear Control Theory; Lecture Notes in Control and Info. Sci., 122*, pages 130–140. Springer-Verlag, Berlin, 1988.

[17] G. Bornard and H. Hammouri. A high gain observer for a class of uniformly observable systems. In *Proc. 30th IEEE Conf. on Decision and Control, Brighton, England*, pages 1494–1496, Dec 1991.

[18] G. Bornard and H. Hammouri. A graph approach to uniform observability of nonlinear multi output systems. In *40th IEEE Conf. on Decision and Control, Las Vegas, USA*, 2002.

[19] G. Conte, C. Moog, and AM. Perdon. *Nonlinear control systems: an algebraic setting.* Springer, Berlin, 1999.
[20] F. Deza, D. Bossanne, E. Busvelle, J.P. Gauthier, and D. Rakotopara. Exponential observers for nonlinear systems. *IEEE Trans. on Automatic Control*, 38(3):482–484, 1993.
[21] S. Drakunov and V. Utkin. Sliding mode observers. tutorial. In *Proc. 34th IEEE Conf. on Decision and Control, New Orleans, LA, USA*, pages 3376–8, 1995.
[22] M. Fliess and Kupka. A finiteness criterion for nonlinear input-output differential systems. *Siam Journal on Control and Optimization*, 21(5):721–728, 1983.
[23] J.P. Gauthier and G. Bornard. Observability for any $u(t)$ of a class of nonlinear systems. *IEEE Trans. on Automatic Control*, 26(4):922–926, 1981.
[24] J.P. Gauthier, H. Hammouri, and S. Othman. A simple observer for nonlinear systems - applications to bioreactors. *IEEE Trans. on Automatic Control*, 37(6):875–880, 1992.
[25] J.P. Gauthier and A.K. Kupka. Observability and observers for nonlinear systems. *Siam Journal on Control and Optimization*, 32(4):975–994, 1994.
[26] J.P. Gauthier and A.K. Kupka. *Deterministic observation theory and applications.* Cambridge Univ Press, 1997.
[27] A. Gelb. *Applied optimal estimation.* MIT Press, Cambridge, 1992.
[28] A. Glumineau, C. Moog, and F. Plestan. New algebraic-geometric conditions for the linearization by input-output injection. *IEEE Trans. on Automatic Control*, 41(4):598–603, 1996.
[29] H. Hammouri and F. Celle. Some results about nonlinear systems equivalence for the observer synthesis. In *New Trends in Systems Theory*, pages 332–339. Birkhäuser, 1991.
[30] H. Hammouri and J.P. Gauthier. Bilinearization up to output injection. *Systems & Control Letters*, 11:139–149, 1988.
[31] H. Hammouri and M. Kinnaert. A new formulation for time-varying linearization up to output injection. *Systems & Control Letters*, 28:151–157, 1996.
[32] H. Hammouri and J. De Leon Morales. Observer synthesis for state-affine systems. In *Proc. 29th IEEE Conf. on Decision and Control, Honolulu, HA, UA*, pages 784–785, 1990.
[33] R. Hermann and A.J. Krener. Nonlinear controllability and observability. *IEEE Trans. on Automatic Control*, 22(5):728–740, 1977.
[34] A. Isidori. *Nonlinear Control Systems.* Springer-Verlag, Berlin, 3^{rd} edition, 1995.
[35] P. Jouan. Immersion of nonlinear systems into linear systems modulo output injection. *Siam Journal on Control and Optimization*, 41:1756–78, 2003.
[36] R.E. Kalman and R.S. Bucy. New results in linear filtering and prediction theory. *J. Basic Eng.*, 83:95–108, 1961.
[37] H. Khalil. *Nonlinear Systems.* Prentice Hall, 2^{nd} edition, 1996.
[38] A. J. Krener and A. Isidori. Linearization by output injection and nonlinear observers. *Systems & Control Letters*, 3:47–52, 1983.
[39] A. J. Krener and W. Respondek. Nonlinear observers with linearizable error dynamics. *Siam Journal on Control and Optimization*, 23(2):197–216, 1985.
[40] H. Kwakernaak and R. Sivan. *Linear Optimal Control Systems.* Wiley-Interscience, New York, 1972.
[41] J. Lévine and R. marino. Nonlinear systems immersion, observers and finite dimensional filters. *Systems & Control Letters*, 7:137–142, 1986.

[42] D.G. Luenberger. Observers for multivariable systems. *IEEE Trans. on Automatic Control*, 11(2):190–197, 1966.
[43] H. Michalska and D.Q. Mayne. Moving horizon observers and observer-based control. *IEEE Trans. on Automatic Control*, 40(6):995–1006, 1995.
[44] J. Moreno. Observer design for nonlinear systems: a dissipative approach. In *IFAC Symp. Systems Structure and Control, Oaxaca, Mex.*, 2004.
[45] J. Rudolph and M. Zeitz. A block triangular nonlinear observer normal form. *Systems & Control Letters*, 23:1–8, 1994.
[46] H.J. Sussmann. Single input observability of continuous time systems. *Math. Systems Theory*, 12:371–393, 1979.
[47] G. Zimmer. State observation by on-line minimization. *Int. Journal of Control*, 60(4):595–606, 1994.