State estimation for a class of non-linear systems
Benoît Schwaller, Denis Ensminger, José Ragot

To cite this version:

HAL Id: hal-00857135
https://hal.science/hal-00857135
Submitted on 28 Apr 2021

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L’archive ouverte pluridisciplinaire HAL, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d’enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives| 4.0 International License
STATE ESTIMATION FOR A CLASS OF NONLINEAR SYSTEMS

BENOÎT SCHWALLER *, DENIS ENSMINGNER *, BIRGITTA DRESP-LANGLEY **, JOSÉ RAGOT ***

* Laboratory of Inventive Design (LGECO), EA 3938 INSA Strasbourg
University of Strasbourg, 24 Boulevard de la Victoire, 67000 Strasbourg, France
e-mail: schwaller@convergence.u-strasbg.fr

** Laboratory of Mechanical and Civil Engineering (LMGC)
University of Montpellier 2, UMR 5508 CNRS, 860 route de St. Priest, 34090 Montpellier, France
e-mail: dresp@lmgc.univ-montp2.fr

***Centre for Automatic Control of Nancy
UMR 7039, University of Lorraine, CNRS, 2 avenue de Haye, 54516 Vandœuvre lès Nancy, France
e-mail: Jose.Ragot@ensem.inpl-nancy.fr

We propose a new type of Proportional Integral (PI) state observer for a class of nonlinear systems in continuous time which ensures an asymptotic stable convergence of the state estimates. Approximations of nonlinearity are not necessary to obtain such results, but the functions must be, at least locally, of the Lipschitz type. The obtained state variables are exact and robust against noise. Naslin’s damping criterion permits synthesizing gains in an algebraically simple and efficient way. Both the speed and damping of the observer response are controlled in this way. Model simulations based on a Sprott strange attractor are discussed as an example.

Keywords: nonlinear systems, state observers, continuous time.

1. Introduction

State observers have been intensely exploited since the work of Luenberger (1966) to model, control, or identify linear and nonlinear systems, as in the studies by Krener and Isidori (1983), Zeitz (1987), Bastin and Gevers (1988), Boutat et al. (2009), and Zheng et al. (2009), relating to nonlinear systems transformable into a canonical form. The key idea in such approaches is to produce approximate measures of nonlinearity of the order 1, as in Extended Luenberger Observers (ELOs) (Ciccarella et al., 1993). The approximation of nonlinearities in the canonical form (which results in an ELO) has already been suggested (Bestle and Zeitz, 1983), and this approach can be extended to higher order approximations (Röbenack and Lynch, 2004). An observer using a Partial nonlinear Observer Canonical Form (POCF) (Röbenack and Lynch, 2006) has weaker observability and integrability existence conditions than the well-established nonlinear Observer Canonical Form (OCF).

Nonlinear sliding mode observers use a quasi-Newtonian approach, applied after pseudo-differentiation of the output signal (Hui and Zak, 2005; Veluvolu et al., 2007; Efimov and Fridman, 2011). State observers using Extended Kalman Filters (EKFs) provide another method of transforming nonlinear systems (Boutayeb and Aubry, 1999; 2010; Khémiri et al., 2011). In particular, Besançon et al. (2010) exploit an adaptive observer based on Kalman filters for physical systems to which Liénard transformations can be applied. Finding an appropriate method for parameter synthesis remains one of the major difficulties with state observers for nonlinear systems (cf. the works of Tornambè (1992), Besançon et al. (2004), Boizot et al. (2010) and Farza et al. (2011), where high-gain state observers to deal with this problem were proposed). High-gain state observers reduce observation errors for a range of predetermined amplitudes or fluctuations by making the observations independent from parameters. A weak point of this method is its sensitivity to noise and uncertainty.
In network identification and encryption, observers with delays are used to synchronize chaotic oscillators, as shown in several studies (Feki, 2009; Ibrir, 2009; Ghosh et al., 2010). Noise and uncertainty are not critical factors in such a context. This can be very different in the case of industrial processes, as shown in a recent study by Bodizs et al. (2011), where the performances of observers using ELOs, EKF's or Integrated Kalman Filters (IKF's) are compared. The influence of noise and uncertainty on these observer types was emphasized, with more reliable results produced by ELOs, which permits the exact state reconstruction of highly perturbed systems. For PI and ELO observer classes, Söffker et al. (1995) as well as Morales and Ramirez (2002) demonstrated a compensation effect on measurement errors. Abdessameud and Khelfi (2006) as well as Chen et al. (2011) addressed the problem of uncertainty of nonlinear models.

In the present study here, we deal with a specific class of nonlinear SISO (Single Input Single Output) systems, described by Fließ (1990), called the generalized controller canonical form. In principle, every uniformly observable (Hermann and Krener, 1977; Gauthier and Bernzdorf, 1981) sufficiently smooth SISO system with input u and output y can be transformed into this normal form, 

\[
\begin{align*}
\dot{x}(t) &= A x(t) + f(t), \\
y(t) &= C^T x(t) + \Phi [ w(t) ],
\end{align*}
\]

with the following definitions.

**Definition 1.**

- \(n\) order of differential equation 
- \(u_1(t)\) input of system 
- \(u_i(t)\) \((i - 1)\)-th temporal derivative of \(u_1(t)\) 
- \(u_1^T(t)\) input vector \([u_1(t), \ldots, u_n(t)]\) 
- \(x_1(t)\) \((i - 1)\)-th temporal derivative of \(x_1(t)\) 
- \(x_1^T(t)\) state vector \([x_1(t), \ldots, x_n(t)]\) 
- \(a_i, c_i\) model parameters 
- \(E, \Psi [ g(t), u(t) ]\) a nonlinear function of the type \(C^1\) 
- \(\Phi [ w(t) ]\) function of inputs \(w(t)\) 
- \(y(t)\) output variable

\[
\theta \leq n, \quad \omega_0 = \sqrt{\frac{a_i}{\r_n}}
\]

\(\omega_0\) index of last coefficient \(c_i \neq 0\) 
\(\omega_0\) eigenfrequency of the system.

We introduce a PI observer with a modified high gain observer, which includes nonlinear functions and linear parameters. This approach has two advantages:

- it permits an asymptotically stable reconstruction of state representations without having to increase gains to infinity;
- it obtains the limiting conditions of the observer stability in a systematic way, without making approximations on the nonlinearities, and without having to employ optimization algorithms (e.g., by solving Riccati equations).

The proposed approach is completely deterministic and allows controlling the speed and damping of the dynamic response of the observer, according to its requirements, without having to search for a minimum to guarantee its stability. The only requirement is that the nonlinear functions be at least locally of the Lipschitz type, which can be satisfied for many practical applications (Düffing, Van der Pol, Bernoulli equations, inverted penduli, nonlinear friction models for DC motors or valve actuators (Shuang et al., 2010)).

To achieve this, the mathematical representation of the physical system is normalized with regard to the eigenfrequency \(\tilde{\omega}\), with \(\tilde{\tau} = \frac{\tilde{\omega}}{\omega_0}\). This has the considerable advantage of giving a normalized space that is independent of the temporal constants of the system, and fulfills an important function in the observer gain synthesis. Such a normalized representation is possible in time (Gille et al., 1988) as well as in the frequency domain (Gißler and Schmid, 1990) for linear systems.

**Definition 2.** We define the following scaled state representation:

\[
\begin{align*}
\tilde{x}(\tau) &= \tilde{A} \tilde{x}(\tau) + \tilde{f}(\tau), \\
y(\tau) &= \tilde{C}^T \tilde{x}(\tau) + \tilde{\Phi} [ \tilde{w}(\tau) ],
\end{align*}
\]

\[
\tilde{A} = \begin{bmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \end{bmatrix},
\]

\[
\tilde{f}(\tau) = \begin{bmatrix} \tilde{c}_1 \cdot \tilde{\Psi} [ \tilde{g}(\tau), \tilde{w}(\tau) ] \\ \vdots \\ \tilde{c}_n \cdot \tilde{\Psi} [ \tilde{g}(\tau), \tilde{w}(\tau) ] \end{bmatrix},
\]

\[
\tau = \tilde{\omega} t, \quad \tilde{\omega} = \frac{\tilde{\omega}}{\omega_0}
\]

\[
\begin{align*}
u_1(t) &= u_1(\tau) \tilde{\omega}^{i-1}, & x_1(t) &= x_1(\tau) \tilde{\omega}^{i-1}, \\
\tilde{a}_i &= a_i/\tilde{\omega}^{n-i+1}, & \tilde{c}_i &= c_i \tilde{\omega}^{i-1}, \\
\tilde{x}_n(\tau) &= \tilde{x}_n(\tau)/\tilde{\omega}^{i}, & i &= 1, \ldots, n,
\end{align*}
\]
\( \hat{\theta}(\tau) \) being a vector with dimension \( n \).

Equations \( 21-24 \) define time dilation of the state representation and its new parameters, allowing for dilation or retraction of the temporal scale and the amplitude of derivatives from the order from 1 to \( n \) without changing the pattern of the signal \( x_i(\tau) \). In the function \( \Psi(15) \), the terms \( u_i(t) \) and \( x_i(t) \) are replaced by \( u_i(\tau) \overset{\Delta}{=} \hat{u}_i-1 \) and \( x_i(\tau) \overset{\Delta}{=} \hat{x}_i-1 \), and each expression is divided by \( \tilde{\omega}^n \) to obtain the normalized function \( \Psi \).

In the following paragraphs, we will define the observer structure (Section 2.1), characterize the observation error (Section 2.2), and prove that the observer state vector uniformly converges toward the state vector of the physical system (Section 2.3). Then the problem of parameters synthesis will be discussed (Section 2.4) and model simulations for a strange attractor system (Sprott, 1994) will be shown (Section 3).

2. Structure and synthesis of the observer

2.1. System and observer definitions. To obtain the variable \( x_1(\tau) \), if \( \theta = 1 \), we use

\[
x_1(\tau) = \frac{y(\tau) - \Phi [u(\tau)]}{\tilde{c}_1}.
\]

When \( \theta > 1 \), the filtered using

\[
\dot{\hat{y}}(\tau) = K \dot{y}(\tau) + \dot{h}_i[y(\tau) - \Phi [u(\tau)]] \tag{4a}
\]

\[
K = \begin{bmatrix}
0 & 1 & 0 & \ldots \\
0 & 0 & 1 & \ldots \\
\vdots & \ddots & \ddots & \ddots \\
-\tilde{c}_1 & \ldots & -\tilde{c}_0-1 & \tilde{c}_0
\end{bmatrix},
\]

\[
y(\tau)^T = [y_1(\tau) \ldots y_{\theta-1}(\tau)],
\]

\[
\dot{y}(\tau)^T = [0 \ldots 1/\tilde{c}_0].
\]

To analyze the effect, we rewrite (2b) in scalar form, ignoring \( \tilde{c}_{\theta+1}, \ldots, \tilde{c}_n \), which are all zero:

\[
\dot{y}(\tau) - \Phi [u(\tau)] = \dot{x}_0(\tau) + \sum_{i=1}^{\theta-1} \tilde{c}_i x_i(\tau). \tag{5}
\]

In (4d) we insert (5) in the place of \( \dot{y}_{\theta-1}(\tau) \) and after reducing it to the same denominator it becomes

\[
\dot{c}_0 \dot{y}_{\theta-1}(\tau) + \sum_{i=1}^{\theta-1} \tilde{c}_i y_i(\tau) = \sum_{i=1}^{\theta} \tilde{c}_i x_i(\tau) \tag{6a}
\]

\[
y_1(s) [\tilde{c}_1 + \ldots + \tilde{c}_0 s^{\theta-1}] = x_1(s) [\tilde{c}_1 + \ldots + \tilde{c}_0 s^{\theta-1}] \tag{6b}
\]

with (6a), the Laplace transformation of (6a). The transfer function \( y_1(s)/x_1(s) = 1 \). and \( x_1(\tau) = y_1(\tau), \theta = 1, \ldots, \theta \). In the rest of the study, and without limiting the generality, we will consider solely the case where \( \theta = 1 \).

**Definition 3.** To generate state estimates for the system (2), the following modified high gain observer is defined:

\[
\dot{\hat{z}}(\tau) = \tilde{A} \hat{\theta}(\tau) + \dot{h}_i \Delta y_1(\tau) + \dot{\hat{y}}(\tau) + \dot{\tilde{g}}(\tau). \tag{7a}
\]

\[
\hat{\theta}(\tau) = [\dot{x}_1(\tau) \ldots \dot{x}_n(\tau)],
\]

\[
\dot{h}_i = [h_0 \ldots h_1],
\]

\[
\Delta y_1(\tau) = x_1(\tau) - \dot{x}_1(\tau),
\]

\[
\dot{\tilde{g}}(\tau) = [0 \ldots I_0(\tau) + I_1(\tau)],
\]

\[
I_0(\tau) = h_0 \Delta y_1(\tau). \tag{7g}
\]

Figure 1 shows the functional diagram of an observer of the order \( n = 3 \). The error estimate \( \Delta y_1(\tau) \) (7a) is used with gains \( h_1, \ldots, h_n \) in (7a) to compensate for distances between system and observer states. The matrix \( \tilde{A} \) (2d) used in (7a) retains the generalized controller canonical form in the observer. The vectors \( \dot{\tilde{g}}(\tau) \) and \( g(\tau) \), of dimension \( n \), serve to compensate the effects of \( f(\tau) \) and of possible perturbations or external noise to reduce the state distance. The nonlinear functions \( I_1(\tau) \) and \( I_2(\tau) \) used in \( j(\tau) \) and \( g(\tau) \) are defined later in (18d)–(18h). The vector \( \dot{\tilde{g}}(\tau) + g(\tau) \) used in feedback by parameters \( a_i \) in the equation \( \dot{x}_n(\tau) \) is also the vector for estimating \( \dot{\tilde{g}}(\tau) \). Apart from controlling the synthesis of gain parameters \( h_i \) (Section 2.4), the use of (7) requires setting the state vector \( \dot{x}_n(\tau) \) at initial conditions \( 0 \). This structure can be readily applied to the reference examples given in (4). For the parameters \( a_i = 0, i = 1, \ldots, n, h_0 = 0 \), and the functions \( I_1(\tau) = I_2(\tau) = 0 \), (7) becomes an observer with a high gain for SISO systems. The functional difference between the two approaches is that an observer with a high gain works continually to reduce the error \( \Delta y_1(\tau) \) to produce state estimations, whilst (7) no longer takes into account \( \Delta y_1(\tau) \) after convergence of the error to zero.

2.2. Characterization of the observer error dynamics. In a first step, we rewrite the system given in (7) in terms of a single relation depending only on \( \dot{x}_1(\tau) \) and \( \Delta y_1(\tau) \) and their successive derivatives. In a second step, the differential equation for state distances between system and observer is determined. To achieve this, we deduce from (7a) the recursive relation used to generate estimates for successive state characteristics as functions of the output errors:

\[
\dot{x}_{i+1}(\tau) = \hat{x}_i(\tau) - h_{n-i+1} \Delta y_1(\tau), \tag{8a}
\]

\[
\dot{x}_{n}(\tau) = \hat{x}_{n-1}(\tau) - h_2 \Delta y_1(\tau) + I_2(\tau). \tag{8b}
\]
From (10), (11) and (12) we obtain with as many derivatives of \( \hat{x}_1(\tau) \) as needed. The vector of successive derivatives for \( \hat{x}_1(\tau) \) is thus
\[
\tilde{y}(\tau) = [\tilde{x}_1(\tau) \ldots \tilde{x}_n(\tau)],
\]
which leads to defining the output errors given in (7d) in terms of
\[
\Delta y_1(\tau) = x_1(\tau) - \tilde{x}_1(\tau).
\]
Successive derivatives of \( \Delta y_1(\tau) \) are written as
\[
\Delta y_{i+1}(\tau) = \frac{d^{(i)}}{d\tau^i} \Delta y_1(\tau), \quad i = 1, \ldots, n-1.
\]
From (12) we deduce
\[
\Delta y_{i+1}(\tau) = \Delta y_i(\tau), \quad i = 1, \ldots, n-1.
\]
From (10), (11) and (12) we obtain
\[
\Delta y_i(\tau) = x_i(\tau) - \tilde{x}_i(\tau),
\]
\[
\Delta y_i(\tau) = [\Delta y_1(\tau) \ldots \Delta y_n(\tau)].
\]
Repeatedly applying (8) with this new notation leads to a new form of the vector \( \tilde{y}(\tau) \):
\[
\dot{\tilde{y}}(\tau) = \tilde{y}(\tau) + \tilde{G} \Delta y_1(\tau) - g(\tau),
\]
\[
\tilde{G} = \begin{bmatrix}
0 & \ldots & 0 \\
0 & \ldots & 0 \\
-1 & \ldots & 0 \\
-2 & \ldots & -n_{\text{max}}
\end{bmatrix},
\]
Then, we re-introduce (15a) into (7a) to obtain
\[
\dot{\tilde{y}}(\tau) = \tilde{A} \tilde{y}(\tau) + \tilde{A} \tilde{G} \Delta y_1(\tau) + \tilde{A} \Delta y_1(\tau) + \dot{\tilde{y}}(\tau).
\]
The temporal derivative of (15a) can be written as follows:
\[
\dot{\tilde{y}}(\tau) = \tilde{A} \tilde{y}(\tau) + \tilde{A} \tilde{G} \Delta y_1(\tau) - \dot{\tilde{y}}(\tau) + \dot{\tilde{y}}(\tau).
\]
\[
\tilde{A} \tilde{G} = \begin{bmatrix}
0 & 0 & \ldots & 0 \\
0 & -h_{\text{max}} & 0 & \ldots \\
0 & \ldots & 0 & 0 \\
0 & \ldots & -h_{\text{max}} & 0 \\
-2 & \ldots & -h_{\text{max}} & 0
\end{bmatrix},
\]
Then, reducing the right-hand side of (16) to (17a), we obtain a new state system, which only depends on \( \hat{x}_1(\tau), \Delta y_1(\tau) \) and their successive derivatives:
\[
\dot{\tilde{y}}(\tau) = \tilde{A} \tilde{y}(\tau) + \tilde{H} \Delta y_1(\tau) + \tilde{f}(\tau),
\]
\[
\tilde{A} \tilde{H} = \begin{bmatrix}
0 & \ldots & 0 & 0 \\
0 & \ldots & 0 & 0 \\
\ldots & \ldots & \ldots & \ldots \\
0 & \ldots & 0 & h_{n-1} \ldots h_{\text{max}}
\end{bmatrix},
\]
\[
g_i = \sum_{j=1}^{n_{\text{max}}} h_{n_{\text{max}}} \tilde{a}_{n_{\text{max}}},
\]
The vector \( \tilde{f}(\tau) \), composed of functions \( I_0(\tau), I_1(\tau) \) and \( I_2(\tau) \), is of the type \( C^1 \). The variables \( \tilde{x}_1(\tau), \tilde{x}_2(\tau) = d\tilde{x}_1(\tau)/d\tau \) used in functions \( f_1, f_2 \) are directly obtainable. The derivative of \( I_2(\tau) \) allows constructing functions by involving \( \tilde{x}_1(\tau), \tilde{x}_2(\tau), \tilde{x}_3(\tau) \). This allows a large number of choices of \( I_2(\tau) + I_1(\tau) \). Second order of choices of \( I_2(\tau) + I_1(\tau) \) and third order nonlinear systems are often composed of such a, particularly polynomial, function (41d). We will now calculate the
difference between (2a) and (18a). This gives
\[
\Delta \tilde{y}(\tau) = \Delta \hat{y}(\tau) = \left[ \Delta \hat{A} - H \right] \Delta \tilde{y}(\tau) + \Delta \tilde{f}(\tau) - \tilde{f}(\tau),
\]
(19a)
\[
\hat{A} - H = \begin{bmatrix}
0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{bmatrix},
\]
(19b)
\[
\hat{A} \tau_i = \begin{cases} 
\tilde{a}_i + h_i, & i = n, \\
\tilde{a}_i + h_i + g_i, & i < n,
\end{cases}
\]
(19c)
\[
\tilde{f}(\tau)^T = \begin{bmatrix} 0 & \ldots & \tilde{\Psi}(x(\tau), u(\tau)) \end{bmatrix},
\]
(19d)
which enables us to obtain the state representation of the distances between the physical system (vector \(f(\tau)\)) and vector \(\hat{f}(\tau)\) (2b) and (10) and the observer. Here \(\hat{f}(\tau)\) is composed of the function \(\tilde{\Psi}\) of the physical system, which is of the type \(C^1\). The distance \(\hat{f}(\tau)\) only takes part in the equation \(n\) of (19a). The matrix \(\hat{A} - H\) retains the canonical form of regulation.

If the distances \(\Delta \tilde{y}(\tau)\) tend towards \(\tilde{0}\) given (15), the variables \(\tilde{x}(\tau) + \tilde{y}(\tau)\) tend towards \(\tilde{x}(\tau)\), which tend towards \(\hat{x}(\tau)\). As a consequence, the distance due to nonlinearities
\[
\Delta \tilde{\Psi}(\tau) = \tilde{\Psi}(x(\tau), u(\tau)) - \tilde{I}_1(\tau) - \tilde{I}_2(\tau)
\]
(20)
tends towards 0. For the rest of the derivation, it is useful to define the augmented vectors \(\Delta \hat{y}_n(\tau)\), \(\tilde{x}_n(\tau)\), \(\tilde{\tilde{x}}_n(\tau)\):
\[
\Delta \hat{y}_n(\tau) = \begin{bmatrix} \Delta \hat{y}(\tau)^T, \Delta \hat{y}_{n+1}(\tau) \end{bmatrix}^T, \quad (21a)
\]
\[
\tilde{x}_n(\tau) = \begin{bmatrix} x(\tau)^T, x_{n+1}(\tau) \end{bmatrix}^T, \quad (21b)
\]
\[
\tilde{\tilde{x}}_n(\tau) = \begin{bmatrix} \tilde{x}(\tau)^T, \tilde{x}_{n+1}(\tau) \end{bmatrix}^T, \quad (21c)
\]
\[
\Delta y_{n+1}(\tau) = x_{n+1}(\tau) - \tilde{x}_{n+1}(\tau) = \Delta \hat{y}_n(\tau) \quad (21d)
\]
\[
\dot{x}_{n+1}(\tau) = \frac{d^n}{dt^n} x_{n+1}(\tau) = \dot{x}_{n}(\tau), \quad (21e)
\]
\[
\dot{\tilde{x}}_{n+1}(\tau) = \frac{d^n}{dt^n} \tilde{x}_{n+1}(\tau) = \dot{\tilde{x}}_{n}(\tau). \quad (21f)
\]

For analyzing the integrator effects (2b), we temporarily differentiate (19a) through (13) and (21), and reintroduce (21d) into \(d^n/df\) for \(\tilde{I}_n(\tau)\). The function \(\tilde{\Psi}(t)\) (13d) in the vector \(\tilde{f}(\tau)\) is also differentiated, and terms \(u_i(t)\) and \(x_i(t)\) are replaced by \(u_i(\tau) - \hat{u}^i\) and \(x_i(\tau) - \hat{u}^i\), with each expression divided by \(\hat{u}^i\) to obtain the normed functions \(\tilde{\Psi}(\tau)\), \(\tilde{I}_1(\tau)\) and \(\tilde{I}_2(\tau)\) (see (41d), (48) as an example). Hence, we obtain the following \((n + 1) \times (n + 1)\) state representation:
\[
\Delta \hat{y}(\tau) = \hat{A} \Delta \hat{y}(\tau) + \Delta \tilde{\Psi}(\tau),
\]
(22a)
\[
\hat{A} = \begin{bmatrix} 0 & 1 & 0 & \cdots & 0 \\
0 & 0 & 1 & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & \cdots & 0 & 1
\end{bmatrix},
\]
(22b)
\[
\tilde{A}_0 = h_0, \quad (22c)
\]
\[
\Delta \tilde{\Psi}(\tau)^T = \begin{bmatrix} 0 & \ldots & \Delta \tilde{\Psi}(\tau) \end{bmatrix}, \quad (22d)
\]
\[
\Delta \tilde{\Psi}(\tau) = \tilde{\Psi}(x(\tau), u(\tau)) - I_1(\tau) - I_2(\tau). \quad (22e)
\]

Given (22b) and (22c), the observer gain has become unitary.

Given (22b) and (22c), the observer gain has become unitary. Considering that \(\tilde{\Psi}(x(\tau), u(\tau))\) is a nonlinear function Lipschitz in \(x(\tau)\) and uniformly bounded in \(u(\tau)\) in an invariant set, with a Lipschitz constant \(L\), i.e.,
\[
\|\Delta \tilde{\Psi}(\tau)\| \leq L \|\Delta \hat{y}(\tau)\| \quad (23)
\]
For many systems, if functions \(\Delta \tilde{\Psi}(\tau)\) are not globally of the Lipschitz type, they may be locally so.

2.3. Convergence of state observations. The observer convergence analysis consists in proving the globally asymptotic evolution of the error estimate for state reconstruction. In other words, regardless of the initial conditions, the observer state is to converge toward the state of the physical system. This leads to the following theorem.

**Theorem 1.** Let us consider a physical SISO system described by (2), for which the observer structure (7) may be used and the distance (22d) exists. If the latter is locally of the Lipschitz type (23), then the observer (7) will be locally stable if the gains \(h_i\) controlling the constants \(\tilde{A}_i\) (15) can be adjusted so that they satisfy the following conditions:
\[
\tilde{A}_i \geq \hat{u}^i + \frac{2}{\omega^{n+1}}, \quad i = 0, \ldots, n - 1, \quad (24a)
\]
\[
\tilde{A}_n \geq \hat{u}^n + \frac{1}{2} \frac{n + 32}{32} + \frac{1}{2} \frac{1}{\omega^{n+1}}. \quad (24b)
\]

If the distance (22e) is globally of the Lipschitz type and if the gains \(h_i\) satisfy (24), then the observer (7) will be globally asymptotically stable.

**Proof** It can be achieved by proving the stability of (22e) using an appropriate Lyapunov function, like the
following quadratic function:

$$V_0(\tau) = \Delta \underline{y}_1(\tau)^T P \Delta \underline{y}_1(\tau).$$  \hfill (25a)

where $P$ is an $(n + 1) \times (n + 1)$ lower triangular matrix defined as positive and satisfying the Sylvester criteria. To prove the uniform convergence of state distances, (25) needs to be derived:

$$\dot{V}_0(\tau) = \Delta \underline{y}_1(\tau)^T P \Delta \underline{y}_1(\tau) + \Delta \underline{y}_1(\tau)^T P \Delta \underline{y}_1(\tau).$$  \hfill (26)

We reintroduce (22a) replacing $\Delta \underline{y}_1(\tau)$ and obtain

$$\dot{V}_0(\tau) = \Delta \underline{y}_1(\tau)^T \hat{Q} \Delta \underline{y}_1(\tau) + N(\tau),$$  \hfill (27a)

where $N(\tau)$ describes the influence of the nonlinear functions on state distances. By factorizing the first term of the right-hand side of (27a) through the products $\Delta \underline{y}_1(\tau) \Delta \underline{y}_1(\tau)$, we obtain a lower triangular matrix $Q$ of dimension $(n + 1) \times (n + 1)$. The coefficients of the principal diagonal element are written as

$$q_{ii} = \begin{cases} -\tilde{A}_{i-1}/2, & i = 1, \ldots, n, \\ (1/2) - 2 \tilde{A}_{ii}, & i = n, \end{cases}$$  \hfill (28a)

$$\tilde{A}_1 > 0, \quad i = 0, \ldots, n - 1, \quad \tilde{A}_n > 1.$$  \hfill (28b)

If the conditions (28b) hold, the coefficients $q_{ii}$ (28a) satisfy the Sylvester criteria to obtain the semi-negativity of $Q$: all successive minors have opposite signs.

Conditions applying to coefficients $\tilde{A}_i$ determine the observer gains $h_i$ and follow from (19a) and (22c). To determine the sign of $\dot{V}_0(\tau)$ requires expanding $N(\tau)$ through (22a) and (25b):

$$N(\tau) = \Delta \underline{y}_1(\tau)^T \hat{\Delta} \hat{\Psi}(\tau),$$  \hfill (29a)

$$\hat{\Delta} \hat{\Psi}(\tau) = \hat{Q} + \hat{P}. $$  \hfill (29b)

To determine the sign of $\dot{V}_0(\tau)$, we need to majorize the term $N(\tau)$ in (27a) and (29a). This is possible on the basis of the following inequality (Raghavan and Hedrick, 1994):

$$\Delta \underline{y}_1(\tau)^T \hat{Q} \Delta \hat{\Psi}(\tau) \leq \Delta \underline{y}_1(\tau)^T \hat{R} \Delta \underline{y}_1(\tau),$$  \hfill (30a)

$$\hat{R} = \frac{\sigma^{n+1}}{4} \frac{Q + \hat{P}}{\sigma^{n+1}}.$$  \hfill (30b)

In (30a), factorizing the products $\Delta \underline{y}_1(\tau) \Delta \underline{y}_1(\tau)$ yields a positive lower triangular matrix $R$ of dimension $(n + 1) \times (n + 1)$, which permits rewriting (29a) as an inequality:

$$N(\tau) \leq \Delta \underline{y}_1(\tau)^T \hat{R} \Delta \underline{y}_1(\tau),$$  \hfill (31a)

$$r_{ii} = \begin{cases} -\sigma^{n+1} (\frac{17}{16} + \frac{1}{\sigma^{n+1}}), & i \neq n, \\ -\sigma^{n+1} (\frac{n}{16} + \frac{1}{\sigma^{n+1}}), & i = n, \end{cases}$$  \hfill (31b)

with $r_{ii}$ (31b) as positive coefficients of the principal diagonal element. The inequality (31a) and the Lipschitz conditions (23) permit majorizing $N(\tau)$ in (27a):

$$\dot{V}_0(\tau) \leq \Delta \underline{y}_1(\tau)^T (Q + \hat{R}) \Delta \underline{y}_1(\tau).$$  \hfill (32)

To determine the sign of $\dot{V}_0(\tau)$, we need

$$Q + \hat{R} \leq 0.$$  \hfill (33)

The sum $Q + \hat{R}$ yields a lower triangular matrix that satisfies Sylvester criteria of semi-negativity if the $n$ inequalities (33) are satisfied. Then, if $I_z(\tau)$ (18a) and $I_z(\tau)$ (18b) exist, if $\Delta \hat{\Psi}$ (22a) is Lipschitz (23) $V_0(\tau)$ is semi-negative and (22a) is globally and asymptotically stable. The observer is locally stable if (23) is locally Lipschitz.

### 2.4. Synthesis of observer parameters.

The foregoing proof of uniform convergence and overall asymptotic stability of the observer make it possible to conceive a synthesis of parameters $h_i$ by applying Lyapunov’s first-order approximation (Lyapunov, 1892; Gille et al., 1988; Fuller, 1992) to (22a). This implies neglecting nonlinear terms in functions $\Delta \hat{\Psi}(\tau)$, which then takes the form of a classical linear differential equation. More precisely, this approximation deals with singular points of a nonlinear system. Lyapunov, in particular, made an important theoretical claim (first-order method) according to which, except in certain so-called “critical” cases, the nature and especially the stability of a singular point can be determined neglecting terms with degrees greater than 2. This would consist in linearizing a nonlinear system by limiting it to the smallest domains of variation in variables (“small movements”); in other words, by assimilating the nonlinear functions of variables to the first term of their serial development.

The proof of globally asymptotic stability given in Section 2.3 limits the problem of parameter synthesis near the point of equilibrium $\Delta \hat{y}(\tau)^T \rightarrow 0$. In our case here, this leads to a synthesis of gains $h_i$ and keeps the roots of the characteristic equation of the linearized observer in conditions where such an approximation is
valid. Lyapunov’s “critical case” is avoided. The “small movements” in our case are the differences in state trajectories between the physical system and its adapted observer. Since the role of an observer is to constantly reduce or minimize these differences, the theorem proposed here appears ideally suited. Otherwise, very complex equations, specific to each model of an observer’s input-output behaviour, would be required. Such an unnecessary complexity seriously complicates attempts to synthesize the coefficients \( h_i \), while the stability of a linearized system conforms to that of a real system. On the basis of these considerations, we can re-write (22a) as follows:

\[
\Delta y_{n+2}(\tau) + \sum_{i=1}^{n} \tilde{A}_i \Delta y_{i+1}(\tau) + h_0 \Delta y_1(\tau) = 0. \tag{34}
\]

Observer stability and convergence, as proven in the previous section here, do not account for error damping. Naslin’s normal damping polynomials (Naslin, 1960; 1963, Humbert and Ragot, 1970; Gissler and Schmid, 1990; Kim, 2002) define algebraic criteria for a simple and rapid analysis of the damping associated with the parameter space in (33), allowing a parameter synthesis of gains \( h_i \) which ensures the convergence of vectors \( \tilde{x}(\tau) \) (10, 11, 12, 13) and allows controlling the transitory states. A normal polynomial \( D(s) = a_0 + a_1 s + \ldots + a_{n+1} s^{n+1} + s^{n+2} \) of degree \( n + 2 \) is defined in terms of a characteristic pulsation \( \omega_0 \) and an associated interval \( \gamma \). Characteristic pulsations of \( D(s) \) as defined by Naslin are written in terms of

\[
\omega_i = \frac{a_i}{a_{i+1}}, \quad i = 0, \ldots, n. \tag{35}
\]

The relationship between characteristic pulsations and interval \( \gamma \) is

\[
\omega_i = \omega_0 \gamma^i, \quad i = 0, \ldots, n. \tag{36}
\]

The definition of the first pulsation \( \omega_0 \) determines that of all others. The coefficient \( \gamma \) is assimilated with the general damping of \( D(s) \). Given (23), and by normalizing the parameters of \( D(s) \), the relation between characteristic pulsations and intervals is

\[
\tilde{\omega}_i = \frac{\tilde{A}_i}{A_{i+1}} = \frac{\gamma^i}{\tau}. \tag{37}
\]

This condition applies to the characteristic polynomial associated with (34). The Naslin technique of gain adjustment satisfies the Routh–Hurwitz criteria within an adimensional space. For the observer gain synthesis, coefficient \( h_n \) is determined for the highest order taking into account (22d):

\[
\tilde{A}_n = \frac{\gamma^n}{\tau}, \quad h_n = \frac{\gamma^n}{\tau} - \tilde{a}_n \tag{38}
\]

For practical reasons, the coefficient \( \gamma \) is often chosen arbitrarily between 2 and 5/2, which produces a specific response and damping in the system. The coefficients \( h_{n-1}, \ldots, h_1, h_0 \) are obtained by using

\[
h_i = \prod_{j=1}^{n} \left( \frac{\gamma^j}{\tau} \right) - \tilde{a}_i - g_i, \quad i = 1, \ldots, n-1, \tag{39a}
\]

\[
h_0 = \prod_{j=0}^{n} \left( \frac{\gamma^j}{\tau} \right). \tag{39b}
\]

Naslin damping polynomials provide a simple and efficient means of observer parameter synthesis fully respecting the conditions for convergence (24), which ensure stability. The speed of convergence is controlled by \( r \) through parameter \( \tilde{r} \) and the damping by \( \gamma \). Given that \( \tilde{a}_i, g_i \) are normalized, the gains obtained satisfy the conditions (24) independently of the time scale of the system.

3. Simulations

To illustrate our developments, we consider an observer of a strange attractor of type D (Sprott, 1994). This corresponds to a system of the order \( n = 3 \) defined in terms of

\[
\dot{z}_1(t) = -z_2(t), \tag{40a}
\]

\[
\dot{z}_2(t) = z_1(t) + z_3(t), \tag{40b}
\]

\[
\dot{z}_3(t) = z_1(t) z_2(t) + 3 z_2(t)^2, \tag{40c}
\]

\[
y(t) = z_1(t). \tag{40d}
\]

It is possible to define the system (40) in terms of its successive derivatives:

\[
\dot{x}_1(t) = x_2(t), \tag{41a}
\]

\[
\dot{x}_2(t) = x_3(t), \tag{41b}
\]

\[
\dot{x}_3(t) = \tilde{\Psi}(x(t)), \tag{41c}
\]

\[
\tilde{\Psi}(x(t)) = x_1(t)^2 + x_1(t) x_3(t) - x_2(t) - 3 x_2(t)^2, \tag{41d}
\]

and a linking system with (41b) and (41d):

\[
z_1(t) = x_1(t), \tag{42a}
\]

\[
z_2(t) = -x_2(t), \tag{42b}
\]

\[
z_3(t) = -x_1(t) - x_3(t), \tag{42c}
\]

\[
y(t) = z_1(t). \tag{42d}
\]

This transformation permits observing the system through (4). Figure (a) visualizes the state trajectory in the two-dimensional plane \( x = y \) coordinates, with the initial parameters \( z_1(0) = -4, 27 z_2(0) = -0, 39 z_3(0) = 0, 3 \). Figure (b) permits observing estimates of the
system’s temporal response, showing a superposition of several eigenfrequencies and random amplitudes. The highest is chosen as the norm. We consider $\omega_0 = 2\pi$ with a ratio $r = 6$. This choice of initial values places the observer in a “pessimistic” regime where the assumed state distance is the largest possible. Normalization of (2) applied to our example here gives

\[
\begin{align*}
\dot{x}_1(\tau) &= x_2(\tau), \\
\dot{x}_2(\tau) &= x_3(\tau), \\
\dot{x}_3(\tau) &= \hat{\Psi}[x(\tau)], \\
\hat{\Psi}[x(\tau)] &= \frac{1}{\omega^3} x_1(\tau)^2 + \frac{1}{\omega} x_1(\tau) x_3(\tau) - \frac{1}{\omega^2} x_2(\tau) - \frac{3}{\omega} x_2(\tau)^2.
\end{align*}
\]

The linking system for (43) and (44) becomes

\[
\begin{align*}
z_1(t) &= x_1(\tau), \\
z_2(t) &= -\hat{\omega} x_2(\tau) = -37.7 x_2(\tau), \\
z_3(t) &= -x_1(\tau) - \hat{\omega}^2 x_3(\tau) \\
y(\tau) &= x_1(\tau).
\end{align*}
\]

The observer of the system (43) illustrated by Fig. 1 is written in terms of

\[
\begin{align*}
\dot{\hat{x}}_1(\tau) &= \dot{x}_2(\tau) + h_3 \Delta y_1(\tau), \\
\dot{\hat{x}}_2(\tau) &= \dot{x}_3(\tau) + I_2(\tau) + h_2 \Delta y_1(\tau), \\
\dot{\hat{x}}_3(\tau) &= I_0(\tau) + I_1(\tau) + h_1 \Delta y_1(\tau), \\
I_0(\tau) &= h_0 \Delta y_1(\tau), \\
\Delta y_1(\tau) &= x_1(\tau) - \hat{x}_1(\tau),
\end{align*}
\]

with the nonlinear functions decomposed into

\[
\begin{align*}
I_2(\tau) &= \frac{1}{\omega} \dot{\hat{x}}_1(\tau) \dot{\hat{x}}_2(\tau), \\
I_2(\tau) &= \frac{1}{\omega} \left[ \dot{\hat{x}}_1(\tau)^2 + \dot{\hat{x}}_1(\tau) \dot{\hat{x}}_2(\tau) \right], \\
I_1(\tau) &= \frac{1}{\omega^3} \dot{\hat{x}}_1(\tau)^2 - \frac{1}{\omega^2} \dot{\hat{x}}_2(\tau) - \frac{4}{\omega} \dot{\hat{x}}_2(\tau)^2.
\end{align*}
\]

A Lyapunov approximation of the first order (38) and (39), generates the observer gain synthesis

\[
\begin{align*}
h_3 &= \frac{\gamma^3}{r}, \\
h_2 &= \frac{\gamma^2 \gamma_3}{r^2}, \\
h_1 &= \frac{\gamma^1 \gamma^2 \gamma_3}{r^3}, \\
h_0 &= \frac{\gamma_0 \gamma^1 \gamma^2 \gamma_3}{r^4}.
\end{align*}
\]

For three different $\gamma$, we obtain the gains given in Table 1.

<table>
<thead>
<tr>
<th>$\gamma$</th>
<th>$h_3$</th>
<th>$h_2$</th>
<th>$h_1$</th>
<th>$h_0$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.5</td>
<td>0.56</td>
<td>0.89</td>
<td>0.053</td>
<td>0.009</td>
</tr>
<tr>
<td>2</td>
<td>1.33</td>
<td>0.89</td>
<td>0.3</td>
<td>0.05</td>
</tr>
</tbody>
</table>

The Lipschitz constant $L$ is determined to prove that the observer converges irrespective of the initial conditions of the attractor. The temporal derivative of (41d) is computed and then normalized:

\[
\begin{align*}
\dot{\hat{\Psi}}[x(\tau)] &= 2 x_1(t) x_2(t) + x_1(t) x_4(t) \\
&\quad - 5 x_2(t) x_3(t) - x_4(t), \\
\dot{\hat{\Psi}}[x(\tau)] &= 2 x_1(t) x_2(t) + x_1(t) x_4(t) \\
&\quad - 5 x_2(t) x_3(t) - x_4(t).}
\]

To obtain the expression of $x_4(\tau)$, we differentiate (42c) and replace (42b) by $x_1(t)$. Then the whole is normalized (26), which yields

\[
x_4(\tau) \hat{\omega}^3 = z_2(t) - z_3(t).
\]

This allows completing (44) to construct the vector $[x_1(\tau) \ldots x_4(\tau)]$:

\[
\begin{align*}
x_1(\tau) &= z_1(t), \\
x_2(\tau) &= -(1/\hat{\omega}) z_2(t), \\
x_3(\tau) &= -(1/\hat{\omega}^2) (z_1(t) + z_3(t)), \\
x_4(\tau) &= (1/\hat{\omega}^3) (z_2(t) - z_3(t)),
\end{align*}
\]

which is used in (48b). It is possible to determine the maximum Lipschitz constant $L$ for vector $\dot{\Psi}(\tau) = 0$ of the system on the basis of simulations and through (23). In our example here, $L = 3 \cdot 10^{-5}$. The gains determined in Table 1 satisfy the conditions (24), yielding local stability since $\Delta \dot{\Psi}(\tau)$ can only be locally Lipschitz.

Figure 2(c) illustrates the convergence of observer state trajectories towards the system state trajectories for the three parameter types. The observer stabilizes asymptotically, with oscillations that are stronger when $\gamma$ is weaker. The choice for $\gamma$ between 2 and 2.5 is experimentally justified. The speed of convergence is not affected by this coefficient.

A second simulation allows comparison of the observer (7) for a $\gamma = 2.5$ with that of a high gain observer ($I_1(\tau) = I_2(\tau) = h_0 = 0$) having a chain of temporarily nonnormalized integrators with gains $h_2 = 39, h_3 = 616, h_0 = 3875, 8$, and with an observer structure (Gauthier et al., 1992) suggesting parameters $\theta = 62.8$ to determine gains $h_i$. Figures 2(e) and
Fig. 2. Observation of the Sprott D strange attractor: state trajectory of the system (a), oscillations of $z_1(t)$ (b), influence of damping $\gamma$ (c), observer output under the influence of white noise (d), high-gain observer (e), fluctuations of the high-gain observer (f).

Paragraph: (2f) demonstrate the convergence of (7) which tends constantly towards zero, while there is still a residual error for the high gain and Gauthier observers, which ends by following fluctuations of $\Psi [x(t), u(t)]$ within the limits of an amplitude factor.

A third simulation is performed with parameters $\bar{r} = 20$, $\omega_0 = 2\pi$, $\gamma = 2$. A limited band-pass white noise with a signal to noise ratio of 10 is added to the
measured output $y(\tau)$. Reconstruction of the latter $\tilde{x}_1(\tau)$ in Fig. 2(d) perfectly matches $x_1(\tau)$. This is ensured by the integration $I_0(\tau)$, which compensates for any static bias due to the noise on the basis of the functions $I_1(\tau)$, $I_2(\tau)$.

4. Conclusions and perspectives

The observer introduced here proposes a new proportional integral strategy applied to continuous nonlinear systems. It allows the exact and stable reconstruction of state trajectories of a physical system. The observer structure allows expressing the convergence dynamics in terms of a nonlinear differential equation with constant coefficients. The stability of the observer was demonstrated as long as the nonlinear functions were at least locally of the Lipschitz type. In this case, it is possible to determine conditions which systematically guarantee this stability, using a weighted quadratic Lyapunov function.

We proceed to an observer parameter synthesis which allows providing its transitory states with temporal specifications. Naslin’s damping polynomials are well-adapted and simple to use here.

The reconstructed state vector has the major advantage of being directly exploitable for state control without any additional transformations.

The approach introduced here should prove useful in the context of state estimates for nonlinear systems using a decoupled multiple model. Further simulations under conditions of parametric uncertainty, minimized by online parameter identification, may help increase the robustness of such estimates, even in the presence of instrumental noise and external system perturbation. It may ultimately be extended to MIMO systems and to the modelling of systems based on Lure decompositions.

References


State estimation for a class of nonlinear systems


**Benoît Schwaller** received his engineer’s degree with specialisation in control from Fachhochschule Offenburg (Germany) in 1990. Then he joined the University of Strasbourg (France), where he obtained his Ph.D. in automatic control and biomechanics in 1993. Since 1999, Benoît Schwaller has been a Maître Assistant at the University Institute of Technology of Schiltigheim and the University of Strasbourg, and conducts his research at INSA Strasbourg, Graduate School of Science and Technology (LGECO). His fields of interest include real time control, nonlinear observer modelling and identification.

**Denis Ensminger** began with an informatics and management technological diploma, then passed a master degree in computer science and finally a Ph.D. in informatics and bio-mechanics at the Louis Pasteur University of Strasbourg (France). He worked as an engineer for companies such as INTERTECHNIQUE, SIEMENS or ATRAITS. Since 1992, he has been teaching computer science at the Louis Pasteur Institute of Technology at the University of Strasbourg. Presently, his research in INSA Strasbourg, Graduate School of Science and Technology (LGECO), is mostly on signal processing, applied artificial intelligence and embedded systems. He still works for industry as a consultant.
Birgitta Dresp-Langley obtained a Ph.D. in cognitive sciences in 1992 at University Paris V. She has been a tenured research scientist with CNRS since 1993. Her fields of interest include cognitive systems, perceptual processes, modelling, and abstract expressionist paintings.

José Ragot received his engineer’s degree with specialisation in control from ECN Nantes, Graduate School of Science and Technology (France), in 1969. Then, he joined the University of Nancy (France), where he received his master’s degree in 1970. In 1973 he obtained a Ph.D. and in 1980 a Diplôme de Doctorat-es-Science. Since 1985, José Ragot has been a full professor at the National Polytechnic Institute of Lorraine and a researcher in the Automatic Control Research Center of Nancy (CRAN). His fields of interest include data validation, process diagnosis, fault detection and isolation, as well as modelling and identification.

Received: 2 April 2012
Revised: 5 August 2012
Re-revised: 23 October 2012