A Discrete-time Observer Based on the Polynomial Approximation of the Inverse Observability Map

Alfredo Germani\textsuperscript{1,2,*} and Costanzo Manes\textsuperscript{1,3,**}

\textsuperscript{1}Department of Electrical and Information Engineering, University of L'Aquila, Poggio di Rivo, 67040 L'Aquila, Italy;\textsuperscript{2}Universit\'a Campus Bio-Medico of Rome, Via Alvaro del Portillo, 21, 00128 Roma Italy;\textsuperscript{3}Istituto di Analisi dei Sistemi ed Informatica del CNR "A Ruberti", Viale Manzoni 30, 00185 Roma, Italy

This paper investigates the problem of asymptotic state observation for analytic nonlinear discrete-time systems. A new technique for the observer construction, based on the Taylor polynomial approximation of the inverse of the observability map, is presented and discussed. The degree chosen for the approximation is the most important design parameter of the observer, in that it can be chosen such to ensure the convergence of the observation error at any desired exponential rate, in any prescribed convergence region (semiglobal exponential convergence).

Keywords: Observer design, discrete-time nonlinear systems, polynomial approximation.

1. Introduction

The importance of the problem of asymptotic state observation of nonlinear dynamical systems is well known, and a large amount of scientific literature on this subject is available. Although the case of continuous time systems has been more intensively studied in the past, an increasing number of authors is devoting the attention to the state observation problem in the discrete-time framework. Many papers discuss the extension to discrete-time systems of observer design techniques devised and developed for continuous time systems. An approach widely investigated is to find a nonlinear change of coordinates and, if necessary, an output transformation, that transform the system into some canonical form suitable for the observer design using linear methodologies. First papers on the subject are \cite{10}, \cite{23}, \cite{24}, where autonomous systems are considered, and quite restrictive conditions are given for the existence of the coordinate transformation that allows the observer design with linearizable error dynamics. Other papers concerned with the problem of finding conditions for the existence of the change of coordinates are \cite{30} and \cite{31}, for autonomous systems, and \cite{3} and \cite{7} for nonautonomous systems. The drawback of this approach is that in general the computation of the coordinate transformation, when existing, is a very difficult task. Also the technique proposed in \cite{22} requires to solve functional equations that in general do not admit a closed form solution. An interesting approach for the construction of observers with linear error dynamics for systems admitting a differential/ difference representation is reported in \cite{26}. Another approach consists in designing observers in the original coordinates, finding iterative algorithms, typically inspired by the Newton method, that asymptotically solve suitably defined observability maps (see e.g. \cite{12}, \cite{13} and \cite{27}). Sufficient conditions of local convergence are provided, in general, under...
the assumption of Lipschitz nonlinearities. The use of the Extended Kalman Filter as a local observer for noise-free systems has been investigated in [29], [4], [5] and [28]. In [28], sufficient conditions of convergence are given in term of existence of a solution for some Linear Matrix Inequalities (LMI). Some authors propose observers for particular classes of systems. In [32] an observer with standard Luenberger form and constant gain is proposed for autonomous systems with the state transition function characterized by Lipschitz nonlinearities and with linear output transformation. Sufficient convergence conditions are given in term of LMI. Nonlinear systems with linear measurements are also considered in [I] and [6]. The problem of asymptotic state reconstruction for systems described by bilinear state-transition functions and rational output functions is considered in [17], where, using the tool of the Kronecker algebra, the problem has been solved through the immersion of the nonlinear system into a linear one of larger dimension. In [19] the use of polynomial approximations of nonlinear discrete-time systems for solving the state observation problem has been investigated, and conditions for the exponential convergence of the polynomial extended Kalman filter [18], when used as an observer, are studied.

This paper presents a new method for the construction of an observer for nonlinear discrete-time systems of the type

\[
\begin{align*}
x(t+1) &= f(x(t)), \\
y(t) &= h(x(t)),
\end{align*}
\]

where \( x(t) \in \mathbb{R}^n \) is the system state, \( y(t) \in \mathbb{R} \) is the measured output, \( f : \mathbb{R}^n \rightarrow \mathbb{R}^n \) is the one-step state transition map, and \( h : \mathbb{R}^n \rightarrow \mathbb{R} \) is the output function. System (1) is assumed to be globally analytic, that means that both \( f \) and \( h \) are analytic functions in all compact sets of \( \mathbb{R}^n \). The approach employed here can be considered the discrete time equivalent of the approach followed in [11] and [14]. The method is based on the Taylor polynomial approximation of the inverse of the observability map (the function that maps states into output sequences) up to a given degree \( \nu \geq 1 \), and constitute a nontrivial extension of the one presented in [12], where only the first degree Taylor approximation has been considered, and only local convergence is guaranteed. The main novelty of the contribution here presented is that, thanks to the analyticity assumption, the degree of the approximation can be chosen such to guarantee the exponential convergence of the observation error to zero when the initial observation error is inside a ball of prescribed radius, centered at the origin. It follows that the proposed observer has the important property to be a semiglobal exponential observer, in that both the convergence rate and the radius of the convergence region can be arbitrarily assigned.

The paper is organized as follows. In section II some preliminary definitions and notations are given, and previous work of the authors on the subject is discussed. In section III the new observation algorithm is presented and its convergence properties are discussed. Simulation results and conclusions follows.

An extended Appendix is also present, where the theory of the Taylor approximation of the inverse of a nonlinear vector function is explained in some detail.

2. Preliminaries

In this section some definitions, notations and concepts that will be used throughout the paper are presented, and relations with the authors’ previous work are discussed. The symbol \( f^r(x) \), with \( r \in \mathbb{N} \), is recursively defined as:

\[
\begin{align*}
f^0(x) &= x, \\
f^{r+1}(x) &= (f \circ f^r)(x) = f(f^r(x)), & r \geq 1,
\end{align*}
\]

(the symbol \( \circ \) denotes the composition of functions), and is used to denote the \( r \)-steps state transition, that allows the computation of future system states as a function of past states:

\[
x(t + r) = f^r(x(t)).
\]

The output at time \( t + r \) can be written as a function of \( x(t) \) as follows

\[
y(t + r) = h \circ f^r(x(t)).
\]

Definition 2.1: The observability map for the system (1) is a vector function \( z = \Phi(x) \), with \( \Phi : \mathbb{R}^n \rightarrow \mathbb{R}^n \), defined as:

\[
\Phi(x) = \begin{bmatrix}
h \circ f^{\nu-1}(x) \\
\vdots \\
h \circ f^1(x) \\
h(x)
\end{bmatrix}.
\]

Let \( Y_t \) denote the stacked output sequence \( y(t) \) in the interval \( \{t, t + n\} \)

\[
Y_t = \begin{bmatrix}
y(t + n - 1) \\
\vdots \\
y(t + 1) \\
y(t)
\end{bmatrix}.
\]
By the definition of the observability map, the following holds:

\[ Y_t = \Phi(x(t)). \]  

(7)

**Definition 2.2:** The observability matrix for the system (1) is:

\[ Q(x) = \frac{d}{dx} \Phi(x). \]  

(8)

Following the notations introduced in the Appendix, \( Q(x) \) will be also written as \( \nabla_x \Phi(x) \) or \( \nabla_x \otimes \Phi(x) \) (see (100) in the Appendix).

**Remark 1:** Note that for linear systems the above definition of observability matrix coincides with the classical linear one.

**Definition 2.3:** The nonlinear system (1) is said to be observable in an open subset \( \Omega \subseteq \mathbb{R}^n \) if its observability map (5) is invertible in \( \Omega \), i.e., there exists \( \Phi^{-1}(z) \) such that \( \Phi^{-1}(\Phi(x)) = x \), for all \( x \in \Omega \). If \( \Omega = \mathbb{R}^n \), then the system (1) is said to be globally observable.

The Definition 2.3 of observability states the theoretical possibility of the reconstruction of the state of a nonlinear system, at a given time \( t \), exploiting the knowledge of the output sequence in the interval \( [t, t+n) \).

\[ x(t) = \Phi^{-1}(Y_t). \]  

(9)

Such a possibility in most cases remains theoretical because, in general, closed forms for the inverse of the observability map are not available. Note that the observability property in a set \( \Omega \) implies that the observability matrix \( Q(x) \) is nonsingular in \( \Omega \).

**Definition 2.4:** The nonlinear system (1) is said to be uniformly Lipschitz observable in an open subset \( \Omega \subseteq \mathbb{R}^n \) iff it is observable, and if both maps \( \Phi(x) \) and \( \Phi^{-1}(z) \), are uniformly Lipschitz in \( \Omega \) and in \( \Phi(\Omega) \), respectively. This means that there exist two real numbers \( \gamma_{\Phi^{-1}} \) and \( \gamma_{\Phi} \) such that

\[ \| \Phi(x_1) - \Phi(x_2) \| \leq \gamma_{\Phi} \| x_1 - x_2 \|, \quad \forall x_1, x_2 \in \Omega, \]

\[ \| \Phi^{-1}(z_1) - \Phi^{-1}(z_2) \| \leq \gamma_{\Phi^{-1}} \| z_1 - z_2 \|, \quad \forall z_1, z_2 \in \Phi(\Omega). \]

(10)

If \( \Omega = \mathbb{R}^n \), then the system (1) is said to be globally uniformly Lipschitz observable.

**Definition 2.5:** Consider a dynamic system of the form

\[ \zeta(t + 1) = g(\zeta(t), y(t), \theta), \quad t \in \mathbb{Z}, \]

\[ \zeta(t) = m(\zeta(t), y(t), \theta), \]  

(11)

where \( \zeta(t) \in \mathbb{R}^q \) is the system state, \( \theta \in \mathbb{R}^p \) is a set of constant parameters, and \( y(t) \in \mathbb{R} \) is the output of system (1).

- System (11), with a given choice of \( \theta \), is said to be a local asymptotic observer for (1) if there exist an open set \( \Omega \subseteq \mathbb{R}^n \) and \( \delta \in \mathbb{R}^+ \) such that, for any pair \( x(t_0), \xi(t_0) \in \Omega \),

\[ \| x(t_0) - \xi(t_0) \| \leq \delta \Rightarrow \lim_{t \to 0} \| x(t) - \xi(t) \| = 0. \]  

(12)

- The system (11) is said to be a semiglobal asymptotic observer if for any \( \Omega \subseteq \mathbb{R}^n \) and \( \delta \in \mathbb{R}^+ \) there exists a choice \( \theta \) such to guarantee the implication (12) for any pair \( x(t_0), \xi(t_0) \in \Omega \).

- The system (11), with a given choice of \( \theta \), is said to be a global asymptotic observer for (1) if for any pair \( x(t_0), \xi(t_0) \in \mathbb{R}^n \) it is \( \lim_{t \to 0} \| x(t) - \xi(t) \| = 0 \).

**Remark 2:** Assume that a system of the type (11) provides a sequence \( \xi(t) \) such that for all pairs \( x(t_0), \xi(t_0) \in \mathbb{R}^n \) it is

\[ \lim_{t \to 0} \| Y_t - \Phi(\xi(t)) \| = 0. \]  

(13)

If the system (1) is uniformly Lipschitz observable, then (11) is a global asymptotic observer. This happens because the Lipschitz property implies the following

\[ \| x(t) - \xi(t) \| = \| \Phi^{-1}(\Phi(x(t))) - \Phi^{-1}(\Phi(\xi(t))) \| \]

\[ = \| \Phi^{-1}(Y_t) - \Phi^{-1}(\Phi(\xi(t))) \| \]

\[ \leq \gamma_{\Phi^{-1}} \| Y_t - \Phi(\xi(t)) \|. \]  

(14)

and therefore \( \| x(t) - \xi(t) \| \to 0 \).

Consider a globally observable system of the type (1). Thanks to the assumption of invertibility of the observability map, this can be used to define a new set of coordinates for system (1):

\[ z(t) = \Phi(x(t - n + 1)). \]  

(15)

Note that \( z(t) \) is used here to represent the system state at time \( t - n + 1 \). Recalling the identity (7), it follows that

\[ z(t) = Y_{t-n+1}. \]  

(16)
Let \((A_b, B_b, C_b)\) be a Brunowsky triple of dimension \(n\), defined as:

\[
A_b = \begin{bmatrix}
0 & 0 & 0 & \cdots & 0 \\
1 & 0 & 0 & \cdots & 0 \\
0 & 1 & 0 & \cdots & 0 \\
\vdots & \vdots & \ddots & \cdots & \vdots \\
0 & 0 & 0 & \cdots & 1
\end{bmatrix} \in \mathbb{R}^{n \times n},
\]

\[
B_b = \begin{bmatrix}
1 \\
0 \\
\vdots \\
0
\end{bmatrix} \in \mathbb{R}^n, 
C_b = \begin{bmatrix}
0 & 0 & 0 & \cdots & 1
\end{bmatrix} \in \mathbb{R}^n.
\]

(17)

Using these matrices the following equations holds

\[
Y_{t-n+2} = A_b Y_{t-n+1} + B_b y(t+1), \\ y(t-n+1) = C_b Y_{t-n+1}.
\]

(18)

With the substitution \(z(t) = Y_{t-n+1}\), these can be rewritten as

\[
z(t+1) = A_b z(t) + B_b y(t+1), \\ y(t-n+1) = C_b z(t).
\]

(19)

If the inverse map \(\Phi^{-1}(\cdot)\) were available, then the original state \(x(t)\) could be easily computed from \(z(t)\) in two steps: first \(x(t-n+1)\) is computed as

\[
x(t-n+1) = \Phi^{-1}(z(t)),
\]

(20)

and then \(x(t)\) is computed as

\[
x(t) = f^{n-1}(x(t-n+1)) = f^{n-1}(\Phi^{-1}(z(t))).
\]

(21)

It follows that, using the \(z\)-coordinates, system (1) can be rewritten as:

\[
z(t+1) = A_b z(t) + B_b h \circ f^n(\Phi^{-1}(z(t))), \\ y(t) = h \circ f^{n-1}(\Phi^{-1}(z(t))).
\]

(22)

The problem is that a closed form for the inverse of the observability map of a given system is not available, in general. Numerical iterative methods can be used for the computation of \(x(t-n+1)\), by finding the (unique) root of

\[
z(t) - \Phi(x) = 0.
\]

(23)

In [12] the following observer structure was proposed

\[
\psi(t+1) = f(\psi(t)) \\ + (Q(f(\psi(t))))^{-1} \left( B_b \left( y(t+1) - h \circ f^n(\psi(t)) \right) \\ + Q \left( y(t-n+1) - h(\psi(t)) \right) \right) \\ \hat{x}(t) = f^{n-1}(\psi(t)), t > t_0 = n-1,
\]

(24)

where \(Q(\cdot)\) is the observability matrix (8) of system (1). For this observer, the following local convergence result holds true (Theorem 2.1 of [12])

**Theorem 1:** For a system of the type (1) there exists a finite gain vector \(K \in \mathbb{R}^n\) and a suitable \(b > 0\) such that the system (24) is a local observer, provided that:

- System (1) is globally Lipschitz observable (see definition 2.4);
- \(h\) and \(f\) are uniformly Lipschitz in \(\mathbb{R}^n\);
- \(\|f^{n-1}(\psi(t_0)) - x(t_0)\| > \delta\).

The proof of this theorem in [12] exploits the representation (22) of system (1). Note that the theorem is stated here using the notations introduced in this paper, and that the Lipschitz observability condition implies the full rank property of the observability matrix \(Q\).

### 3. The Polynomial Observer

The observer proposed in this paper is based on the idea of using the Taylor approximation of a chosen degree of the inverse map as a numerical method for solving (23). This approach has strong connections with the iterative technique presented in [16], which generalizes the Newton-Rapson approach. The observation algorithm derived is characterized by the classical prediction-correction structure, but in this observer the correction term is a polynomial function of the output prediction error, whereas standard algorithms have a linear correction term.

The main assumption needed for the construction of the proposed observer is that both functions \(f\) and \(h\) in (1) are analytic in all \(\mathbb{R}^n\). Well known results of real analysis establish that the composition of analytic functions gives back an analytic function, and that if an analytic function is invertible, then the inverse is analytic too. It follows that the observability map \(z = \Phi(x)\) defined in (5) is analytic in all \(\mathbb{R}^n\), and that globally observable analytic systems admit an analytic inverse \(x = \Phi^{-1}(z)\) in all \(\mathbb{R}^n\).
It is important to stress that the Taylor polynomial approximation of the inverse of the observability map \( z = \Phi (x) \) can be computed without the explicit knowledge of the inverse function \( x = \Phi^{-1}(z) \), by using only the derivatives of \( \Phi (x) \). The mathematical tools needed for the construction of the polynomial Taylor approximation of inverse functions are described in the Appendix. Using the notations stated in the Appendix, the following derivatives can be defined

\[
G_k(z) = \nabla_z^{[k]} \otimes \Phi^{-1}(z), \quad k = 0, 1, \ldots ,
\]

where \( G_0(z) = \Phi^{-1}(z) \) and \( G_1(z) \) is the Jacobian \( \nabla_z \Phi^{-1}(z) \). Let

\[
\Gamma_k = \sup_{z \in \mathbb{R}^n} \| G_k(z) \|.
\]

The Taylor theorem states that

\[
\Phi^{-1}(z) = \sum_{k=0}^{\nu} \frac{1}{k!} G_k(z)(z - \Phi^{-1}(z)) + r_{\nu+1}(z, \Phi^{-1}(z)),
\]

where the remainder of the series is infinitesimal of order \( \nu + 1 \) w.r.t. \( z - \Phi^{-1}(z) \):

\[
\| r_{\nu+1} \| \leq \frac{\Gamma_{\nu+1}}{(\nu + 1)!} \| z - \Phi^{-1}(z) \|^{\nu+1}.
\]

The exponents \( k/j \) in the Taylor formula (27) denote the Kronecker powers of matrices and vectors (see Appendix).

**Proposition 2:** If, \( \forall \rho > 0 \), the sequence \( \Gamma_k \) defined in (26) is such that

\[
\lim_{k \to 0} \frac{\Gamma_k \beta^k}{(k + 1)!} = 0,
\]

then \( \Phi^{-1}(z) \) is analytic in all \( \mathbb{R}^n \).

**Proof:** The proof is a straightforward consequence of the inequality (28).

Let

\[
\tilde{G}_k(x) = G_k \circ \Phi(x) = \left( \nabla_z^{[k]} \otimes \Phi^{-1}(z) \right)_{z=\Phi(x)}
\]

and let \( \tilde{z} = \Phi(x) \) and \( x = \Phi^{-1}(z) \). Observing that \( \nabla_z^{[0]} \otimes \Phi^{-1}(z) = \Phi^{-1}(z) = \tilde{x} \), i.e. \( G_0(x) = \tilde{x} \), the Taylor formula (27) becomes

\[
x = \tilde{x} + \sum_{k=1}^{\nu} \frac{1}{k!} \tilde{G}_k(\tilde{x})(z - \Phi(\tilde{x}))^{[k]} + r_{\nu+1}(z, \Phi(\tilde{x})) = \tilde{x} + \sum_{k=1}^{\nu} \frac{1}{k!} \tilde{G}_k(\tilde{x})(z - \Phi(\tilde{x}))^{[k]} + r_{\nu+1}(z, \Phi(\tilde{x})),
\]

where

\[
\| r_{\nu+1} \| \leq \frac{\Gamma_{\nu+1}}{(\nu + 1)!} \| z - \Phi(\tilde{x}) \|^{\nu+1}.
\]

Note that \( \Gamma_k = \sup_{z \in \mathbb{R}^n} \| G_k(z) \| = \sup_{x \in \mathbb{R}^n} \| \tilde{G}_k(x) \| \).

The main result proved in the Appendix is that the sequence of derivatives \( \tilde{G}_k(x) \) of the inverse function \( \Phi^{-1}(\cdot) \) can be computed as functions of the matrix coefficients \( F_{h,k}(x) \) recursively defined as

\[
F_{h,k} = (\nabla_x \Phi)^{[k]} = Q^{[k]}(x), \quad k = 1, 2, \ldots
\]

\[
F_{h,k} = \nabla_x \otimes F_{h,k-1} + (\nabla_x \Phi) \otimes F_{h-1,k-1}, \quad h = 1, \ldots , k - 1,
\]

\[
F_{0,k} = 0.
\]

Lemma 9 of the Appendix proves that the sequence of matrices \( \tilde{G}_k(x) \) needed in the Taylor formula (31) can be computed as

\[
\tilde{G}_0 = x,
\]

\[
\tilde{G}_1 = (\nabla_x \Phi)^{-1},
\]

\[
\tilde{G}_k = - \left( \sum_{h=1}^{k-1} \tilde{G}_h F_{h,k} \right) \tilde{G}_1^{[k]}, \quad k = 2, 3, \ldots .
\]

What is proposed in this paper is a family of observers, whose construction is based on the truncation of the Taylor series (31). Each observer in the family is characterized by the degree of the polynomial approximation, and recursively provides at time \( t \) an estimate of the past state \( x(t - n + 1) \), denoted \( \chi(t) \). From this, an estimate \( \hat{x}(t) \) of the current state is obtained as \( \hat{x}(t) = f^{n-1}(\chi(t)) \). The observer starts the operations at a given time \( t_0 \), while the output measurements are available starting from time \( t_0 - n + 1 \). The observer state is \( \chi(t) \), whose initial value \( \chi(t_0) \) is an a priori estimate of \( x(t_0 - n + 1) \). The recursive equations of the proposed observer are the following:

**Polynomial Observer of degree \( \nu \)**

\[
\tilde{G}_1(t) = \left( Q(\chi(t)) \right)^{-1},
\]

\[
\tilde{G}_k(t) = - \left( \sum_{h=1}^{k-1} \tilde{G}_h(t) F_{h,k} (f(\chi(t))) \right) \tilde{G}_1^{[k]}, \quad k = 2, \ldots , \nu
\]
\[
\Delta x_\nu(t + 1) = K(y(t - n + 1) - h(\chi(t)) \\
+ B_h(y(t + 1) - h \circ f'(\chi(t))),
\]
\[
\chi(t + 1) = f(\chi(t)) + \sum_{k=1}^{\nu} \frac{1}{k!} \tilde{G}_k(t) \Delta_{x_\nu}^{[k]}(t + 1),
\]
\[
\hat{x}(t) = f^{-1}(\chi(t)),
\]

Equation (40) provides the observed state \( \hat{x}(t) \). In the observer equations, the observability matrix \( Q(x) \) defined in (2.2) and the matrix functions \( F_{h,k}(x) \) defined in (33) are needed. Note that \( F_{1,1}(x) = Q(x) \). The gain vector \( K \in \mathbb{R}^n \) must be chosen such that all eigenvalues of \( A_b - KC_b \) are in the open unit circle in the complex plane (\( A_b \) and \( C_b \) are the Brunowsky pair defined in (17)).

Remark 3: Comparing equations (36)–(37) with equations (34) it is clear that the matrices \( \tilde{G}_k(t) \) in the summation (39) are the coefficients \( \tilde{G}_k \) of the Taylor approximation of the inverse map evaluated at \( z = f(\chi(t)) \). Note also that, when \( \nu = 1 \) the algorithm (36)–(40) coincides with the observer (24) presented in [12].

Remark 4: Note that in the observer equations (36)–(40), the only matrix inverse to be numerically computed at each time step is the inverse of the observability matrix \( Q(x) \), evaluated at \( f(\chi(t)) \), in (36), independently of the order \( \nu \) of the observer. The matrix functions \( F_{h,k}(x) \) at each time step must be evaluated at \( f(\chi(t)) \). Considering the dimension of the matrices in the summations (37) and (39)

\[
F_{h,k} \in \mathbb{R}^{n \times n^d}, \quad \text{and} \quad \tilde{G}_k \in \mathbb{R}^{n \times n^d}
\]

it is easy to see how the complexity increases by increasing the observer degree. However, the only operations involved, after the evaluation of the coefficients \( F_{h,k} \), are simple products and sums of matrices. The following theorem provides conditions for semiglobal exponential convergence of the observer (36)–(40). For convenience, in the proof the gain vector \( K \) in (38) is chosen such that all eigenvalues of \( A_b - KC_b \) are real and distinct.

Theorem 3: Consider the system (1) and its observability map defined in (5). Consider system (36)–(40), where \( K \) is such that all eigenvalues of \( A_b - KC_b \) are real and distinct in the interval \((-1,1)\),

Assume that:

- \( H_p \): the system (1) is uniformly globally Lipschitz observable;

- \( H_p \): The functions \( f \) and \( h \) are analytic in all \( \mathbb{R}^n \) and uniformly Lipschitz in all \( \mathbb{R}^n \), i.e.:

\[
\forall \gamma > 0 : \| h(x_1) - h(x_2) \| \leq \gamma \| x_1 - x_2 \|,
\]
\[
\forall \gamma > 0 : \| f(x_1) - f(x_2) \| \leq \gamma \| x_1 - x_2 \|,
\]

(42)

Then, for any pair of real numbers \( \alpha \) and \( \delta \), with \( \alpha \in (0,1) \) and \( \delta > 0 \), there exist an integer \( \nu \), the degree of the observer (36)–(40), and a positive real \( \mu \) such that

\[
\| x(t) - \hat{x}(t) \| \leq \mu \alpha^{-\nu} \| x(t_0 - n + 1) - x(t_0) \|,
\]

(44)

for all pairs \( x(t_0), x(t_0 - n + 1) \in \mathbb{R}^n \) such that \( \| x(t_0 - n + 1) - x(t_0) \| \leq \delta \).

Proof: By assumption \( H_p \), both \( f \) and \( h \) are analytic functions in all \( \mathbb{R}^n \), and therefore the observability map \( \Phi \) is analytic in all \( \mathbb{R}^n \). Thanks to the global observability assumption \( H_p \), the function \( \Phi \) is invertible and analytic in all \( \mathbb{R}^n \), and therefore acts as a global change of coordinates (note that global analyticity of \( \Phi^{-1} \) is also ensured by \( H_p \), see Proposition 2). Let \( \bar{t} \) be the \( \Phi \)-transformed state of the observer (36)–(40), defined by:

\[
\tilde{z}(t) = \Phi(\chi(t)),
\]

(45)

which is an estimate of \( z(t) = y(t - n + 1) \) (recall that \( z(t) = \Phi(\chi(t - n + 1)) \)), and define the error in \( z \)-coordinates as \( e(t) = z(t) - \tilde{z}(t) \). From the observer equations (36)–(40), considering the equivalence \( \tilde{G}_k(t) = G_k(f(\chi(t))) \), the update law of \( \tilde{z}(t) \) takes the form

\[
\tilde{z}(t + 1) = \Phi \left( f(\chi(t)) + \sum_{k=1}^{\nu} \frac{1}{k!} \tilde{G}_k(f(\chi(t))) \Delta_{x_\nu}^{[k]}(t + 1) \right),
\]

(46)

Consider now a sequence \( \bar{z}(t) \) defined by the update law

\[
\bar{z}(t + 1) = \Phi(f(\chi(t))) + \Delta_{\chi}(t + 1),
\]

(47)

where \( \Delta_{\chi}(t + 1) = (y(t - n + 1) - h(\chi(t)) + B_h(y(t + 1) - h \circ f'(\chi(t))) \), as defined in (38). Note that, by definition of the map \( \Phi \), it is
\[ \Phi(f(\chi(t))) = A_{b} \Phi(\chi(t)) + B_{b} h \circ f''(\chi(t)), \]
and \[ h(\chi(t)) = C_{b} \tilde{z}(t). \]  
(48)

Thus, recalling (45), it follows
\[ \dot{z}(t+1) = A_{b} \dot{z}(t) + B_{b} h \circ f''(\chi(t)) + \Delta_{x,y}(t+1) \]
\[ = (A_{b} - KC_{b})\dot{z}(t) + K h(\chi(t)) \]
\[ + B_{b} h \circ f''(\chi(t)) + \Delta_{x,y}(t+1) \]
(49)

and, by the definition of \( \Delta_{x,y}(t+1) \), it is
\[ \dot{z}(t+1) = (A_{b} - KC_{b})\dot{z}(t) + B_{b} y(t+1) \]
\[ + Ky(t-n+1). \]  
(50)

Now, consider the Taylor expansion of \( \Phi^{-1}(\dot{z}(t+1)) \) around \( \Phi(f(\chi(t))) \) (see eq. (31)):
\[ \Phi^{-1}(\dot{z}(t+1)) = f(\chi(t)) \]
\[ + \sum_{k=1}^{\nu} \frac{1}{k!} G_{k}(\chi(t)) \Delta_{x,y}(t+1)^{k} \]
\[ + r_{\nu+1}(t+1) \]
\[ = \chi(t+1) + r_{\nu+1}(t+1), \]  
(51)

where the remainder \( r_{\nu+1}(t+1) \) is such that
\[ \|r_{\nu+1}(t+1)\| \leq \frac{\Gamma_{\nu+1}}{(\nu+1)!} \|\dot{z}(t+1) - \Phi(f(\chi(t)))\|^{\nu+1} \]
\[ \leq \frac{\Gamma_{\nu+1}}{(\nu+1)!} \|\Delta_{x,y}(t+1)\|^{\nu+1}. \]  
(52)

Rewrite equation (51) as
\[ \chi(t+1) = \Phi^{-1}(\dot{z}(t+1)) - r_{\nu+1}(t+1). \]  
(53)

Then
\[ e(t+1) = z(t+1) - \dot{z}(t+1) \]
\[ = z(t+1) - \Phi(\chi(t+1)) \]
\[ = z(t+1) - \Phi(\Phi^{-1}(\dot{z}(t+1))) \]
\[ = z(t+1) - \Phi^{-1}(\dot{z}(t+1)) \]
\[ + r_{\nu+1}(t+1) \]
\[ = z(t+1) - \dot{z}(t+1) \]
\[ + \left( \int_{0}^{1} (\nabla \chi(s))_{\xi(s)} ds \right) r_{\nu+1}(t+1) \]  
(54)

where
\[ \xi(s) = \Phi^{-1}(\dot{z}(t+1)) - s r_{\nu+1}(t+1). \]  
(55)

Now, rewrite (19) for \( z(t+1) \) as follows
\[ z(t+1) = (A_{b} - KC_{b})z(t) + B_{b} y(t+1) \]  
\[ + Ky(t-n+1). \]  
(56)

Using this and (50) yields
\[ z(t+1) - \dot{z}(t+1) = (A_{b} - KC_{b})(z(t) - \dot{z}(t)), \]  
(57)

and using (54)
\[ e(t+1) = (A_{b} - KC_{b})e(t) + \tilde{r}(t+1); \]  
(58)

where
\[ \tilde{r}(t+1) = \left( \int_{0}^{1} (\nabla \chi(s))_{\xi(s)} ds \right) r_{\nu+1}(t+1). \]  
(59)

By the Uniform Lipschitz Observability assumption
\[ \|\nabla \chi(s)\| \leq \gamma_{\Phi}. \forall \chi \in \mathbb{R}^{n}, \]
and therefore
\[ \|\tilde{r}(t+1)\| \leq \gamma_{\Phi}\|r_{\nu+1}(t+1)\|. \]  
(60)

Recalling the bound (52) on \( r_{\nu+1}(t+1) \), and observing that, from definition (38),
\[ \|\Delta_{x,y}(t+1)\| \leq \gamma_{\Phi} (\|K\| + \gamma_{\Phi}^{\gamma}) \|x(t-n+1) - \chi(t)\| \]
\[ \leq \gamma_{\Phi} (\|K\| + \gamma_{\Phi}^{\gamma}) \gamma_{\Phi} \|\dot{z}(t) - z(t)\| \]  
(61)

it follows
\[ \|\tilde{r}(t+1)\| \leq \beta(\nu) \frac{\|e(t)\|^{\nu+1}}{(\nu+1)!}, \]  
(62)

where
\[ \beta(\nu) = \gamma_{\Phi} \Gamma_{\nu+1} ((\|K\| + \gamma_{\Phi}^{\gamma}) \gamma_{\Phi} \nu)^{\nu+1}. \]  
(63)

By assumption, \( K \) is such that all eigenvalues of \( A_{b} - KC_{b} \) are real and distinct. Let \( K_{a} \) denote the gain that assigns a set of eigenvalues \( \lambda = \{\lambda_{1}, \ldots, \lambda_{n}\} \), and let \( \lambda_{\text{Max}} = \max_{i=1,\ldots,n} \{\|\lambda_{i}\|\} \). Choose \( \lambda \) such that \( \lambda_{\text{Max}} < \alpha \). Let \( V_{\lambda} \) denote the Vandermonde matrix
\[ V_{\lambda} = \begin{bmatrix} 1 & \lambda_{1} & \cdots & \lambda_{1}^{n-1} \\ \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_{n} & \cdots & \lambda_{n}^{n-1} \end{bmatrix}, \]  
(64)

which is nonsingular, being \( \lambda_{i} \neq \lambda_{j}, i \neq j \), and such that
\[ V_{\lambda}(A_{b} - K_{a} C_{b}) V_{\lambda}^{-1} = \Lambda = \text{diag}\{\lambda\}. \]  
(65)
It is easy to verify that
\[ K_{\lambda} = -V^{-1}_{\lambda} \begin{bmatrix} \lambda_{1n} \\ \vdots \\ \lambda_{nn} \end{bmatrix}. \]  
(66)

Note that, being \( \lambda_{\text{Max}} < 1 \), it is \( \sqrt{n} \leq \| V_{\lambda} \| < n \) and \( K_{\lambda} \| \leq \| V^{-1}_{\lambda} \| \sqrt{n} \). Consider the following transformation of the error:
\[ \eta(t) = V_{\lambda} e(t). \]  
(67)

Then
\[ \eta(t+1) = \Lambda \eta(t) + V_{\lambda} \bar{f}(t+1) \]  
(68)

which implies that:
\[
\| \eta(t+1) \| \leq \lambda_{\text{Max}} \| \eta(t) \| + \| V_{\lambda} \| \| \bar{f}(t+1) \| \\
\leq \lambda_{\text{Max}} \| \eta(t) \| \\
+ \| V_{\lambda} \| \beta(\nu) \frac{\| V^{-1}_{\lambda}(\eta(t)) \|^{\nu+1}}{(\nu + 1)!} \\
\leq \lambda_{\text{Max}} \| \eta(t) \| \\
+ \| V_{\lambda} \| \| V^{-1}_{\lambda} \|^{\nu+1} \beta(\nu) \frac{\| \eta(t) \|^{\nu+1}}{(\nu + 1)!} \\
\leq (\lambda_{\text{Max}} + \| V_{\lambda} \| \| V^{-1}_{\lambda} \|^{\nu+1}) \beta(\nu) \frac{\| \eta(t) \|^{\nu}}{(\nu + 1)!} \| \eta(t) \|. 
\]  
(69)

Now, choose the degree of the observer \( \nu \) such that, for the given \( \alpha \in (0, 1) \) and \( \delta > 0 \), it is
\[
\frac{\Gamma_{\nu+1}}{(\nu + 1)!} \left( \| V_{\lambda} \| \| V^{-1}_{\lambda}(\| K_{\lambda} \| + \gamma_{\phi}^{\nu+1}) \|^{\nu+1} \frac{\Gamma_{\nu+1}}{(\nu + 1)!} \right) \leq \delta (\alpha - \lambda_{\text{Max}}). 
\]  
(70)

This inequality can be satisfied for a sufficiently large \( \nu \), thanks to the assumption \( H_{p} \). From this, thanks to the definition of \( \beta(\nu) \) given in (63), it follows
\[
\lambda_{\text{Max}} + \| V_{\lambda} \| \| V^{-1}_{\lambda} \|^{\nu+1} \beta(\nu) \frac{\| V_{\lambda} \| \gamma_{\phi}^{\nu+1}}{(\nu + 1)!} \leq \alpha. 
\]  
(71)

It is easy to see that if at a given time \( t \) it is \( \| x(t) - x(t-n+1) \| \leq \delta \), then
\[
\| \eta(t) \| = \| V_{\lambda} e(t) \| = \| V_{\lambda}(\Phi(x(t-n+1)) \\
- \Phi(x(t))) \| \leq \| V_{\lambda} \| \gamma_{\phi}, 
\]  
(72)

so that
\[
\| \eta(t) \|^{\nu+1} = \| \eta(t) \|^{\nu} \| \eta(t) \| \leq (\| V_{\lambda} \| \gamma_{\phi})^{\nu} \| \eta(t) \|. 
\]  
(73)

From this, thanks to inequality (69) it follows
\[
\| \eta(t+1) \| \leq \left( \lambda_{\text{Max}} + (\| V_{\lambda} \| \| V^{-1}_{\lambda} \|^{\nu+1}) \beta(\nu) \frac{\gamma_{\phi}^{\nu+1}}{(\nu + 1)!} \right) \| \eta(t) \|, 
\]  
(74)

and thanks to (71)
\[
\| \eta(t+1) \| \leq \alpha \| \eta(t) \|. 
\]  
(75)

Thus, being \( \alpha \in (0, 1) \), if at time \( t \) it is \( \| \eta(t) \| \leq \| V_{\lambda} \| \gamma_{\phi} \delta \), then also at time \( t+1 \) it is \( \| \eta(t+1) \| \leq \| V_{\lambda} \| \gamma_{\phi} \delta \). Therefore, if at \( t_{0} \) \( \| \chi(t_{0}) - x(t_{0} - n + 1) \| \leq \delta \), then
\[
\| \eta(t_{0}) \| \leq \| V_{\lambda} \| \gamma_{\phi} \delta, \]  
and thanks to (75)
\[
\| \eta(t) \| \leq \alpha^{t-t_{0}} \| \eta(t_{0}) \|, \quad t \geq t_{0}. 
\]  
(76)

This implies
\[
\| e(t) \| \leq \alpha^{t-t_{0}} \| V_{\lambda} \| \| V^{-1}_{\lambda} \| \| e(t_{0}) \|, 
\]  
(77)

and, being \( e(t) = \| \Phi(x(t-n+1)) - \Phi(x(t)) \| \), it follows
\[
\| x(t-n+1) - x(t) \| \leq \alpha^{t-t_{0}} \| V_{\lambda} \| \| V^{-1}_{\lambda} \| \| x(t_{0} - n + 1) - x(t_{0}) \|. 
\]  
(78)

and finally
\[
\| x(t) - \hat{x}(t) \| \leq \alpha^{t-t_{0}} \| V_{\lambda} \| \| V^{-1}_{\lambda} \| \| x(t_{0} - n + 1) - x(t_{0}) \|, 
\]  
(79)

which is the thesis (44), with \( \mu = \gamma_{\phi}^{t-t_{0}} \gamma_{\phi} \gamma_{\phi}^{-1} \| V_{\lambda} \| \| V^{-1}_{\lambda} \| . \]

### 4. Simulation Results

The polynomial observer (36)-(40) has been tested on many systems. In general, the theoretical convergence results have been confirmed by the simulation tests. Here, some results concerning a chaos synchronization problem have been presented.

Consider the well known Henon map [20]
\[
\begin{bmatrix} x_{1}(t+1) \\ x_{2}(t+1) \end{bmatrix} = \begin{bmatrix} 1 - ax_{1}^{2}(t) + x_{2}(t) \\ bx_{1}(t) \end{bmatrix}, 
\]  
(80)
Polynomial Approximation of the Inverse Observability Map

which, for some values of the parameters $a$ and $b$, defines a discrete time chaotic system. The canonical values of the parameters of the Hénon map are $a = 1.4$ and $b = 0.3$, which produce a chaotic behavior of the system.

The polynomial observer (36)-(40) has been applied to the problem of estimating the state of the Hénon system when the constant parameter $b$ is unknown. Only the first state variable $x_1(t)$ is assumed to be measured. This problem is known as a chaos synchronization problem with uncertain parameters (see e.g. [15], [25]).

The unknown parameter $b$ is modeled as a constant state variable, i.e. $x_3(t + 1) = x_3(t) = b$. The system equations considered for the observer construction are the following

$$
\begin{bmatrix}
    x_1(t+1) \\
    x_2(t+1) \\
    x_3(t+1)
\end{bmatrix} =
\begin{bmatrix}
    1 - 1.4x_1^2(t) + x_2(t) \\
    x_1(t)x_3(t) \\
    x_3(t)
\end{bmatrix}
$$

$$
y(t) = x_1.
$$

The polynomial observers of orders $\nu = 1, 2, 3$ and 4 have been applied. The matrix functions $F_{h,k}(x)$, defined in (33), where $k = 1,\ldots,4$ and $h = 1,\ldots, k$, have been computed using the symbolic toolbox of MATLAB, by reproducing exactly the recursive computations in (33).

The map $\Phi$ is as follows:

$$
\Phi(x) = \begin{bmatrix}
    1 - 1.4(1 - 1.4x_1^2 + x_2)^2 + x_3x_1 \\
    1 - 1.4x_1^2 + x_2 \\
    x_1
\end{bmatrix}.
$$

$$
F_{1,1}(x) = Q(x) = \begin{bmatrix}
    7.84(1 - 1.4x_1^2 + x_2)x_1 + x_3 \\
    -2.8x_1 \\
    1
\end{bmatrix}
$$

$$
F_{1,2}(x) = \begin{bmatrix}
    -32.928x_1^2 + 7.84 + 7.84x_2 \\
    -2.800 \\
    0 \\
\end{bmatrix}
$$

In matrix $F_{1,3}(x)$ the only nonzero elements are $[F_{1,3}]_{1,1} = -65.856x_1$ and $[F_{1,3}]_{1,2} = [F_{1,3}]_{1,4} = [F_{1,3}]_{1,10} = 7.84$. Matrices $F_{h,k}$ for $h > 1$ are rather big sparse matrices, and can not be easily reported. Note that matrices $F_{k,k}$ can be computed as $F_{1,1}^k$ (see (33)).

The gain matrix used in the simulations here reported is the following

$$
K = \begin{bmatrix}
    -0.0017 & 0.0428 & -0.3600
\end{bmatrix},
$$

and is such to assign eigenvalues $\lambda_1 = 0.1$, $\lambda_2 = 0.12$ and $\lambda_3 = 0.14$ to the Brunowsky pair $(A_B, C_B)$ of dimension $n = 3$. The Figs 1-3 report simulation results where all observers for $\nu = 1, 2, 3$ and 4 provide convergent state estimates. The true and observed initial states are

$$
x(0) = \begin{bmatrix}
    -0.6 \\
    -0.1 \\
    0.3
\end{bmatrix}^T
$$

$$
\hat{x}(0) = \begin{bmatrix}
    -0.3 \\
    -0.2 \\
    0.2
\end{bmatrix}^T,
$$

It can be seen that higher order observers produce faster transients.

In many cases, the first order observer does not converge, while higher order ones are convergent. For instance, when the initial states are as follows

$$
x(0) = \begin{bmatrix}
    -0.6 \\
    -0.1 \\
    0.3
\end{bmatrix}^T
$$

$$
\hat{x}(0) = \begin{bmatrix}
    0 \\
    0 \\
    0.2
\end{bmatrix}^T,
$$

the first order observer does not converge.

5. Conclusions

A new technique for the observer construction for analytic discrete-time nonlinear systems has been presented in this paper. The method is based on the Taylor...
polynomial approximation of the inverse of the observability map. The degree \( \nu \) chosen for the approximation is the most important design parameter of the observer. Thanks to the analyticity assumption, it is shown that the degree \( \nu \) can be suitably chosen to ensure the error convergence at any desired exponential rate, when an upper bound for the norm of the initial observation error is assigned (semiglobal exponential convergence of the observation error). The resulting observation algorithm has a prediction-correction structure, where the correction term is a polynomial function of the output prediction error. It must be noted that when \( \nu = 1 \) the proposed observer coincides with the one presented in [12]. Computer simulations support the theoretical convergence results.

Acknowledgment

This work is supported by MiUR (Italian Ministry of University and Research) and by CNR (National Research Council of Italy).

References

Polynomial Approximation of the Inverse Observability Map

This appendix describes the mathematical tools of the Kronecker algebra, used in this paper to represent the formulas of the Taylor series of vector functions of several variables and of their inverse functions. Here, the symbol \( I_n \) denotes the identity matrix in \( \mathbb{R}^{n \times n} \).

The Kronecker product of two matrices \( M \) and \( N \) of dimensions \( p \times q \) and \( r \times s \) respectively, is the \((p \cdot r) \times (q \cdot s)\) matrix

\[
M \otimes N = \begin{bmatrix}
m_{11}N & \cdots & m_{1q}N \\
\vdots & \ddots & \vdots \\
m_{p1}N & \cdots & m_{pq}N
\end{bmatrix},
\]

where the \( m_{ij} \) are the entries of \( M \). The Kronecker power of a matrix \( M \) is recursively defined as

\[
M^{[0]} = 1, \quad M^{[i]} = M \otimes M^{[i-1]}, \quad i \geq 1. \quad (89)
\]

Note that if \( M \in \mathbb{R}^{p \times q} \), then \( M^{[i]} \in \mathbb{R}^{p^i \times q^i} \). Some useful properties of the Kronecker product are the following:

\[
(A + B) \otimes (C + D) = A \otimes C + A \otimes D + B \otimes C + B \otimes D \quad (90)
\]

\[
A \otimes (B \otimes C) = (A \otimes B) \otimes C \quad (91)
\]

\[
(A \cdot C) \otimes (B \cdot D) = (A \otimes B) \cdot (C \otimes D) \quad (92)
\]

\[
\det(A \otimes B) = \det(B \otimes A) \quad (93)
\]

Repeated application of the properties (91) and (92) provides the identities

\[
(AB)^{[i]} = A^{[i]}B^{[i]} \quad \text{and} \quad (A^{[k]})^{-1} = (A^{-1})^{[k]} \quad (94)
\]

Using (92), given \( A \in \mathbb{R}^{n \times n} \) and \( B \in \mathbb{R}^{m \times m} \), the following properties can be easily derived

\[
(A \otimes B)^{-1} = A^{-1} \otimes B^{-1} \quad (95)
\]

\[
A \otimes B = (A \otimes I_m)(I_n \otimes B) \quad (96)
\]

From the last, thanks to (93), it follows \( \det(A \otimes B) = \det(A \otimes I_m) \det(I_n \otimes B) \). Using (93) it is \( \det(I_n \otimes B) = \det(B \otimes I_n) \). Recognizing the block diagonal structure of \( I_n \otimes A \) and of \( I_n \otimes B \) it follows

\[
\det(A \otimes B) = (\det(A))^m(\det(B))^n. \quad (97)
\]

A quick survey on the Kronecker algebra can be found in the Appendix of [9]. See [21] for more properties.

The Kronecker formalism is also useful for representing derivatives of matrix functions of several variables, and provides a compact notation for the Taylor polynomial approximation of nonlinear vector functions. Let \( M(x) \) be a smooth matrix function of \( x \in \mathbb{R}^n \), i.e., \( M : \mathbb{R}^n \to \mathbb{R}^{r \times c} \), and let \( \nabla x \otimes M : \mathbb{R}^n \to \mathbb{R}^{r \times n \times c} \) denote the following matrix of derivatives:

\[
\nabla x \otimes M = \begin{bmatrix}
\frac{\partial M}{\partial x_1} & \cdots & \frac{\partial M}{\partial x_n}
\end{bmatrix} \quad (98)
\]
This matrix Jacobian is a formal Kronecker product of the vector \( \nabla_x = [\partial/\partial x_1 \cdots \partial/\partial x_n] \) with the matrix \( M(x) \). Repeated derivatives can be recursively defined as

\[
\nabla_{x}^{[i]} \otimes M = \nabla_{x} \otimes (\nabla_{x}^{[i]} \otimes M), \quad i \geq 1.
\]

(99)

Let \( \mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^m \) be a smooth nonlinear map. In this case \( \nabla_x \otimes \mathcal{F} \) is the standard Jacobian of \( \mathcal{F} \), and therefore throughout the paper the following symbols will be indifferently used

\[
\left( \frac{\partial}{\partial x} \right) \mathcal{F}, \quad \nabla_x \otimes \mathcal{F} = \frac{d \mathcal{F}}{dx} = \nabla_x \mathcal{F}.
\]

(100)

The last symbol provides the most compact notation. Note that for matrix functions the use of \( \nabla_x \otimes \) is necessary.

A useful property of the operator \( \nabla_x \) is the following:

\[
\nabla_x \otimes (A(x) \cdot B(x)) = (\nabla_x \otimes A)(I_n \otimes B) + A(\nabla_x \otimes B),
\]

(101)

where \( A, B \) are generic matrix functions.

Let \( \mathcal{F}(x) \) and \( H(z) \) be functions \( \mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^n \) and \( H: \mathbb{R}^n \rightarrow \mathbb{R}^{r \times c} \). The composition \( H \circ \mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^{r \times c} \) is defined as

\[
H \circ \mathcal{F}(x) = H(\mathcal{F}(x)).
\]

(102)

**Proposition 4:** If \( H(z) \) and \( \mathcal{F}(x) \) are differentiable, then

\[
\nabla_x \otimes (H \circ \mathcal{F}) = ((\nabla_z \otimes H) \circ \mathcal{F})(((\nabla_x \mathcal{F}) \otimes I_c).
\]

(103)

where the term

\[
(\nabla_z \otimes H) \circ \mathcal{F}(x).
\]

(104)

denotes the derivative of \( H \) w.r.t. \( z \) computed at \( z = \mathcal{F}(x) \).

**Proof:** Observe that

\[
\nabla_x \otimes (H \circ \mathcal{F}) = \begin{bmatrix} \frac{\partial(H \circ \mathcal{F})}{\partial x_1} & \cdots & \frac{\partial(H \circ \mathcal{F})}{\partial x_n} \end{bmatrix},
\]

(105)

and that

\[
\frac{\partial(H \circ \mathcal{F})}{\partial x_i} = \sum_{j=1}^n \frac{\partial H(z)}{\partial z_j} \frac{\partial \mathcal{F}(x)}{\partial x_i},
\]

(106)

Substitution of this into the right hand side of (105) provides equation (103).

Throughout the paper, the symbol \( S_{x_0}(\rho) \) will denote the open ball of radius \( \rho \) centered at \( x_0 \in \mathbb{R}^n \), i.e. \( S_{x_0}(\rho) = \{ x \in \mathbb{R}^n : \| x - x_0 \| > \rho \} \). The symbol \( \overline{S}_{x_0}(\rho) \) will denote the closure of \( S_{x_0}(\rho) \).

Using the Kronecker notation, the Taylor Theorem for vector functions of several variables can be written as follows:

**Theorem 5:** Let \( \mathcal{F}: \mathbb{R}^n \rightarrow \mathbb{R}^m \) be a function defined on the closed ball, \( \overline{S}_{x_0}(\rho) \), for some \( x_0 \in \mathbb{R}^n \) and \( \rho > 0 \). If, for a given integer \( \nu \), all derivatives \( \nabla_x^{[i]} \otimes \mathcal{F}, \ k = 1, \ldots, \nu + 1 \), exist and are continuous at every point of \( S_{x_0}(\rho) \), then the following identity holds for all \( x \in S_{x_0}(\rho) \):

\[
\mathcal{F}(x) = \sum_{k=0}^\nu \frac{1}{k!} \left( \nabla_x^{[k]} \otimes \mathcal{F}(x) \right)_{x_0} (x - x_0)^{[k]} + R_{\nu+1}(x, x_0)(x - x_0)^{[\nu+1]},
\]

(107)

where the remainder satisfies the inequality

\[
\|R_{\nu+1}(x)\| \leq \frac{1}{(\nu + 1)!} \left( \sup_{x \in S_{x_0}(\rho)} \| \nabla_x^{[\nu+1]} \otimes \mathcal{F}(x) \| \right).
\]

(108)

The function \( \mathcal{F} \) is said to be analytic in the set \( S_{x_0}(\rho) \) if the Taylor series is convergent for all \( x \in S_{x_0}(\rho) \). Defining

\[
\varphi_k(x_0, \rho) = \sup_{x \in S_{x_0}(\rho)} \| \nabla_x^{[k]} \otimes \mathcal{F}(x) \|,
\]

(109)

a sufficient condition for the analyticity of \( \mathcal{F} \) in \( S_{x_0}(\rho) \) is the following

\[
\lim_{k \to \infty} \frac{\varphi_k(x_0, \rho)}{k!} = 0.
\]

(110)
Theorem 6: Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a smooth function defined on all $\mathbb{R}^n$. If, for any choice of $x_0 \in \mathbb{R}^n$ and of $\rho \in \mathbb{R}^+$ the sequence $\varphi_k(x_0, \rho)$ defined in (109) is such that
\[
\lim_{k \to \infty} \frac{\varphi_k(x_0, \rho)\rho^k}{k!} = 0,
\] (111)
then $F$ is analytic in all $\mathbb{R}^n$.

Proof: The proof is a direct consequence of the remainder expression (108).

Let $F^{-1} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denote the inverse map of the function $F$, i.e.
\[
z = F(x) \iff x = F^{-1}(z),
\] (112)
so that $F^{-1} \circ F$ is the identity map in $\mathbb{R}^n$: $F^{-1}(F(x)) = x$. Let us define the following notation for the derivatives of the inverse function $F^{-1}$:
\[
G_k(z) = \nabla_z^{-k} \odot F^{-1}(z), \quad k = 0, 1, \ldots
\] (113)
A useful equivalent recursive definition is
\[
G_0(z) = F^{-1}(z), \quad G_{k+1}(z) = \nabla_z \odot G_k(z), \quad k = 0, 1, \ldots
\] (114)

Lemma 7: The derivatives of the inverse mapping $F^{-1}$, i.e. the matrix functions $G_k(z)$, $k = 1, 2, \ldots$, defined in equation (113), satisfy the following chain of identities:
\[
(G_1 \odot F)F_{1,1} = I_n,
\] (115)
\[
\sum_{h=1}^k (G_h \odot F)F_{h,k} = 0, \quad k = 2, 3, \ldots
\] (116)
where
\[
F_{h,k} = (\nabla_z \odot F)^{[k]}, \quad k = 1, 2, \ldots
\]
\[
F_{h,k} = \nabla_x \odot F_{h,k-1} + (\nabla_x F) \odot F_{h-1,k-1},
\] (117)
\[
F_{0,k} = 0.
\]

Proof: The property $(G_1 \circ F)F_{1,1} = I_n$, with $F_{1,1} = \nabla_x F$, it will be proved in Lemma 8. From this
\[
\nabla_x \odot ((G_1 \odot F)F_{1,1}) = 0.
\] (118)
The computation of the derivative, using (101), gives
\[
\nabla_x \odot ((G_1 \odot F)F_{1,1}) = (G_2 \circ F)(\nabla_x \odot F_{1,1})(I_n \odot F_{1,1}) + (G_1 \circ F)(\nabla_x \odot F_{1,1}) = 0.
\] (119)
From the property $(A \odot B)(C \odot D) = (AC) \odot (BD)$ it follows
\[
(\nabla_x F \odot I_n)(I_n \otimes F_{1,1}) = \nabla_x F \odot F_{1,1} = (\nabla_x F)^{[2]} = F_{2,2}.
\] (120)
Moreover, from (117), it is $F_{1,2} = \nabla_x F_{1,1}$. Therefore equation (119) can be written as
\[
(G_2 \circ F)F_{2,2} + (G_1 \circ F)F_{1,2} = 0,
\] (121)
so that (116) is proved for $k = 2$. The identity (116) for $k > 2$ will be proved by induction. Then, assume that (116) holds for some $k \geq 2$, and prove that it holds also for $k + 1$. If (116) is true, then
\[
\nabla_x \odot \left( \sum_{h=1}^k (G_h \odot F)F_{h,k} \right) = 0.
\] (122)
Differentiating, and using (101), yields
\[
\sum_{h=1}^k (G_{h+1} \circ F)((\nabla_x F) \odot I_n^{h})(I_n \otimes F_{h,k}) + (G_h \circ F)(\nabla_x \odot F_{h,k}) = 0.
\] (123)
Reorganizing the subscripts one gets
\[
(G_{k+1} \circ F)((\nabla_x F) \odot I_n^{k})(I_n \otimes F_{k,k}) + (G_k \circ F)(\nabla_x \odot F_{k,k}) = 0,
\] (124)
that can be rewritten as
\[
\sum_{h=1}^{k+1} (G_h \circ F)\tilde{F}_{h,k+1} = 0,
\] (125)
where
\[
\tilde{F}_{k+1,k+1} = \left((\nabla_x F) \odot I_n^{k}\right)(I_n \otimes F_{k,k}),
\]
\[
\tilde{F}_{h,k+1} = \nabla_x \odot F_{h,k} + ((\nabla_x F) \odot I_n^{h-1})(I_n \otimes F_{h-1,k}), \quad 1 \leq h \leq k.
\] (126)
Exploiting the property $(A \odot B)(C \odot D) = (AC) \odot (BD)$
\[
\tilde{F}_{k+1,k+1} = (\nabla_x F \odot F_{k,k}),
\]
\[
\tilde{F}_{h,k+1} = \nabla_x \odot F_{h,k} + (\nabla_x F \odot F_{h-1,k}), \quad 1 \leq h \leq k.
\] (127)
Considering that \( F_{1,1} = \nabla_x \mathcal{F} \), it follows \( \tilde{F}_{k+1,k+1} = (\nabla_x \mathcal{F})^{k+1} \). The comparison with (117) shows that \( \tilde{F}_{h,k} = F_{h,k} \), for all pairs \( h,k \), and therefore
\[
\sum_{h=1}^{k+1} (G_h \circ F) F_{h,k+1} = 0. \tag{128}
\]
This equation completes the induction and concludes the proof.

**Lemma 8:** The derivatives of the inverse mapping \( \mathcal{F}^{-1}(z) \) computed in \( z = \mathcal{F}(x) \) satisfy the following chain of identities
\[
G_1 \circ \mathcal{F} = (\nabla_x \mathcal{F})^{-1}, \tag{129}
\]
\[
G_{k+1} \circ \mathcal{F} = (\nabla_x \circ (G_k \circ \mathcal{F}))(G_1 \circ \mathcal{F}) (\nabla_x \mathcal{F}) \circ \mathcal{I}_p. \tag{130}
\]

**Proof:** Since \( \mathcal{F}^{-1} \circ \mathcal{F} \) is the identity map, then \( \nabla_x (\mathcal{F}^{-1} \circ \mathcal{F}) = \mathcal{I}_n \). From this, using the chain rule
\[
\nabla_x (\mathcal{F}^{-1} \circ \mathcal{F}) = ((\nabla_x \mathcal{F}^{-1}) \circ \mathcal{F})(\nabla_x \mathcal{F}) = (G_1 \circ \mathcal{F})(\nabla_x \mathcal{F}) = \mathcal{I}_n, \tag{131}
\]
and therefore equation (129) is proved.

Since for any differentiable matrix function \( H(z) \), \( H : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n} \) it is
\[
\nabla_x \circ (H \circ \mathcal{F}) = ((\nabla_x \circ H) \circ \mathcal{F})((\nabla_x \mathcal{F}) \circ \mathcal{I}_p), \tag{132}
\]
see Proposition 4, it follows
\[
\nabla_x \circ (G_k \circ \mathcal{F}) = ((\nabla_x \circ G_k) \circ \mathcal{F})((\nabla_x \mathcal{F}) \circ \mathcal{I}_p) = (G_{k+1} \circ \mathcal{F})((G_1 \circ \mathcal{F})^{-1} \circ \mathcal{I}_p). \tag{133}
\]
From this
\[
G_{k+1} \circ \mathcal{F} = (\nabla_x \circ (G_k \circ \mathcal{F}))(G_1 \circ \mathcal{F})^{-1} \circ \mathcal{I}_p, \tag{134}
\]
Thanks to property (95) it is
\[
((G_1 \circ \mathcal{F})^{-1} \circ \mathcal{I}_p)^{-1} = (G_1 \circ \mathcal{F}) \circ \mathcal{I}_p, \tag{135}
\]
and the theorem is proved.

Let
\[
\bar{G}_k(x) = G_k \circ \mathcal{F}(x) = (\nabla_x^{[k]} \circ \mathcal{F}^{-1}(z))_{z = \mathcal{F}(x)}. \tag{136}
\]
Note that \( \bar{G}_0(z) = \mathcal{F}^{-1}(z) \) and that \( \bar{G}_0(x) = x \). Denoting \( z = \mathcal{F}(x) \) and \( \bar{z} = \mathcal{F}(\bar{x}) \), the Taylor series expansion of \( \mathcal{F}^{-1}(z) \) around the point \( \bar{z} \) can be written as
\[
x = x + \sum_{k=1}^{\infty} \frac{1}{k!} \bar{G}_k(x)(z - \bar{z})^{[k]} + \mathcal{O}(\|z - \bar{z}\|^{k+1}). \tag{137}
\]
It is known that if \( \mathcal{F} \) is analytic in a set \( \tilde{S}_\rho(\rho) \) and admits the inverse function \( \mathcal{F}^{-1} \), then \( \mathcal{F}^{-1} \) is analytic too. The coefficients \( \bar{G}_k(x_0) \) of the Taylor series expansion of \( \mathcal{F}^{-1}(z) \) around a point \( z_0 = \mathcal{F}(x_0) \) can be recursively computed using the algorithm of Lemma 8, without the explicit knowledge of a closed form of \( \mathcal{F}^{-1}(z) \). However, the computations involved in the iterations require the derivates of inverse matrices (consider for instance \( \tilde{G}_2 = (\nabla_x \circ (\nabla_x \mathcal{F})^{-1})((\nabla_x \mathcal{F})^{-1} \circ \mathcal{I}_n) \)). The following Theorem shows that the coefficients \( \bar{G}_k \) can be computed without the differentiation of inverse matrix functions.

**Lemma 9:** The matrix functions \( \bar{G}_k : \mathbb{R}^n \rightarrow \mathbb{R}^{n \times n} \) defined in (136) can be recursively computed as follows:
\[
\bar{G}_0 = x, \tag{138}
\]
\[
\bar{G}_k = (\nabla_x \mathcal{F})^{-1}, \tag{139}
\]
where the matrix coefficients \( F_{h,k} \) are recursively computed as
\[
F_{k,k} = (\nabla_x \mathcal{F})^{[k]}, \quad k = 1, 2, \ldots, \tag{139}
\]
\[
F_{h,k} = \nabla_x \circ F_{h,k-1} + (\nabla_x \mathcal{F}) \circ F_{h-1,k-1}, \quad h = 1, \ldots, k - 1, \tag{139}
\]
\[
F_{h,k} = 0. \tag{139}
\]
**Proof:** The proof is a direct consequence of Lemma 7. Equations (138) are simply derived from the identities (115)–(116), proved in Lemma 7. In particular, rewrite (116) as
\[
\bar{G}_k F_{h,k} + \sum_{h=1}^{k} \bar{G}_h F_{h,k} = 0, \tag{140}
\]
Solving this equation for \( \bar{G}_{h+k} \) and considering that, thanks to (94)
\[
(F_{k,k})^{-1} = ((\nabla_x \mathcal{F})^{[k]}^{-1})^{-1} = ((\nabla_x \mathcal{F})^{-1})^{[k]} = \bar{G}_k^{[k]} \tag{141}
\]
gives back the equations (138).

**Remark 5:** Theorem 9 shows that, given an analytic function \( \mathcal{F} : \mathbb{R}^n \rightarrow \mathbb{R}^n \), with nonsingular Jacobian \( \nabla_x \mathcal{F} \) at a given \( x_0 \in \mathbb{R}^n \), the coefficient of the Taylor series expansion of the inverse function \( \mathcal{F}^{-1} \) around \( z = \mathcal{F}(x_0) \) exists and can be recursively computed as functions of the derivatives of the direct map \( \mathcal{F} \) computed at \( x_0 \), as long as \( \nabla_x \mathcal{F}(x_0) \) is nonsingular.