A nonlinear canonical form for reduced order observer design

Driss Boutat * Gang Zheng ** Hassan Hammouri ***

* ENSI de Bourges, Institut PRISME, 88 Boulevard de Lahitolle, 18020 Bourges, France
(e-mail: driss.boutat@ensi-bourges.fr).

** INRIA Lille-Nord Europe, 40, avenue Halley, 59650 Villeneuve d'Ascq, France
(e-mail: gang.zheng@inria.fr).

*** Université Lyon 1; CNRS UMR 5007 LAGEP; 43 bd du 11 novembre, 69100 Villeurbanne, France
(e-mail: hammouri@lagep.cpe.fr).

Abstract: This paper presents a nonlinear canonical form which is used for the design of a reduced order observer. Sufficient and necessary geometric conditions are given in order to transform a special class of nonlinear systems to the proposed nonlinear canonical form and the corresponding reduced order observer is analyzed. After that, the main results were generalized to nonlinear systems with inputs.

Keywords:

1. INTRODUCTION

The observer design for nonlinear dynamical systems is an important problem in the field of control theory. For nonlinear systems, several techniques, such as Luenberger observer Luenberger (1971), high gain observer Hammouri and Gautier (1988) Gauthier et al. (1985), Busawon and Kabore (2001), Busawon et al. (1998) and so on Phelps (1991), are proposed to design nonlinear observers for different cases, and a general formalism to design nonlinear observer for arbitrary nonlinear systems is still missing. Since the beginning of the 1980’s, many significant researches were done on the problem of transforming nonlinear dynamical systems into simple normal observable forms, based on which one can apply existing observer techniques, from algebraic and geometric points of view Xia and Gao (1989), Sampei and Furuta (1986), Respondek et al. (2004), Noh et al. (2004), Krener and Respondek (1985), Krener and Isidori (1983), Keller (1987), Glumineau et al. (1996), Bestle and Zeitz (1983), Back et al. (2005), Boutat et al. (2009), Zheng et al. (2006).

Roughly speaking, two classes of observers can be classified. The first class of observers is so-called full order observer which estimates all states of the system, including the measurable states which are the outputs. Obviously this redundant estimations of measurable states are not necessary, and conversely it might increase the complexity of the observer design and practical realization. That is the reason why the second class of observers was born, named as reduced order observer, which, different from full order observers, needs to estimate only unmeasurable states of the studied system. It was firstly introduced for linear systems to reduce the number of dynamical equations by estimating only the unmeasurable states. Then it is generalized for nonlinear dynamical systems by imposing the Lipschitz conditions for nonlinear terms Haddad and Bernstein (1989), Xu (2009), Sundarpandian (2006) and invariant manifold Karagiannis et al. (2008). The problem of designing full order observer and reduced order observer for nonlinear systems with linearizable error dynamics and its application to synchronization problem was analyzed in Nijmeijer and Mareels (1997), in which authors pointed out that the existence of a full order observer with linear error dynamics implies the existence of a reduced order observer with linear error dynamics, however the reverse is not valid. Moreover, there are no results available to provide conditions, under which via a transformation, a reduced order observer with linear error dynamics may be found.

In this paper, we give a new nonlinear canonical form which allows us to design reduced order observers, just like the linear case. And necessary and sufficient conditions are given to guarantee the existence of the proposed nonlinear canonical form.

The paper is organized as follows. In section 2 we give notations and the definition of reduced order observer. Then a nonlinear canonical form is proposed and the corresponding reduced order observer is studied. section 3 is devoted to studying necessary and sufficient geometric conditions which allows us to transform a nonlinear system into the proposed nonlinear canonical form. In section 4, we generalized our results to nonlinear systems with known inputs.

2. A NONLINEAR CANONICAL FORM FOR REDUCED ORDER OBSERVER

Before giving the nonlinear canonical form which will be studied in this paper, let us give a definition of the so-
called reduced order observer. Without loss of generality, consider the following nonlinear dynamical system:

\[
\begin{aligned}
\dot{x}_1 &= F_1(x) \\
\dot{y}_1 &= y_2
\end{aligned}
\]

where \( F_1(x) \) and \( F_2(x) \) are smoothly determined by the choice of \( x_1 \) and the dynamics \( F(x) \) defined in (1).

For (2), we try to design a reduced order observer to estimate \( x_1 \).

**Definition 1.** The dynamical system defined as follows:

\[
\hat{\dot{x}}_1 = \bar{F}_1(\hat{x}_1, x_2)
\]

where \( \hat{\dot{x}}_1 \) is the output of (2), is a symbolically reduced order observer for (2) if

\[
\lim_{t \to \infty} \| \hat{\dot{x}}_1(t) - x_1(t) \| = 0.
\]

Moreover, it is said to be an exponentially reduced order observer if

\[
\| \hat{\dot{x}}_1(t) - x_1(t) \| \leq ae^{-bt} \| \hat{x}_1(0) - x_1(0) \|
\]

for \( t > 0 \), where \( a, b \) are both positive constants.

In what follows, we will present first the nonlinear canonical form which will be studied in this paper, then we will show that an exponentially reduced order observer can be easily designed for the proposed nonlinear canonical form. The sufficient and necessary geometric conditions to transform a generic nonlinear system to such an observable nonlinear canonical form will be detailed in the next section.

Let us consider the following observable nonlinear canonical form:

\[
\begin{align*}
\dot{z}_i &= A_i z_i + \beta_i(y_1) z_o + \rho_i(y) \\
\dot{\xi} &= \alpha_1(y_1) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z, y) \\
y &= (y_1^T, y_2^T)^T = (\xi^T, \eta^T)^T
\end{align*}
\]

where

\[
\begin{align*}
\dot{z}_i &= (z_{i,1}, \ldots, z_{i,r_i})^T \in \mathbb{R}^{r_i} \\
z_o &= (z_{1,1}, \ldots, z_{m,r_m})^T \in \mathbb{R}^m \\
\rho_i &= (\rho_{1,1}, \ldots, \rho_{1,r_1})^T \in \mathbb{R}^{r_1} \\
\xi &= (\xi_1, \ldots, \xi_m)^T \in \mathbb{R}^m \\
\eta &= (\eta_1, \ldots, \eta_p)^T \in \mathbb{R}^p \\
\alpha_2 &= (\alpha_{2,1}, \ldots, \alpha_{2,p})^T \in \mathbb{R}^m
\end{align*}
\]

with \( \sum_{i=1}^m r_i = n - m - p \) and the \( r_i \times m \) matrix \( \beta_i \), the \( m \times m \) matrix \( \alpha_1 \) and the \( r_i \times r_i \) matrix \( A_i \) defined as follows:

\[
A_i = \begin{pmatrix}
0 & \cdots & 0 & 0 \\
1 & \cdots & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots \\
0 & \cdots & 1 & 0 \\
0 & \cdots & 0 & 1
\end{pmatrix}
\]

**Remark 1.** Since the nonlinear canonical form (3) is supposed to be observable, thus \( z_o \) in (3) can be observed from the output \( y_1 \), which implies \( \text{Rank}(\alpha_1(y_1)) = m \).

In a more compact manner, system (3) can be rewritten as follows:

\[
\begin{align*}
\dot{z} &= Az + \beta(y_1) z_o + \rho(y) \\
\dot{\xi} &= \alpha_1(y_1) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z, y) \\
y &= (y_1^T, y_2^T)^T = (\xi^T, \eta^T)^T
\end{align*}
\]

where \( A = \text{diag}([A_1, \ldots, A_m]) \), \( \beta = (\beta_1^T, \ldots, \beta_m^T)^T \), \( \rho = (\rho_1^T, \ldots, \rho_m^T)^T \).

Let us denote \( C_i \) the \( 1 \times r_i \) vector defined as \( C_i = [0, \ldots, 0, 1] \) for \( 1 \leq i \leq m \), then we can define a \( m \times (n - m - p) \) matrix \( C \) as follows:

\[
C = \text{diag}[C_1, \ldots, C_m]
\]

which implies \( z_o = C z \).

For (4-7), if we can accurately measure \( y_1 \) and calculate \( \hat{y}_1 \), this allows us to define a "new" output \( Y \) being a function of known output \( y \) and the derivative of \( y_1 \) in (4-7):

\[
Y = \alpha_1^{-1}(y_1)(\hat{y}_1 - \alpha_2(y))
\]

Then we can state the following result.

**Proposition 1.** The following dynamical system:

\[
\begin{align*}
\dot{\hat{z}} &= A\hat{z} + \beta(y_1) C \hat{z} + \rho(y) - K(y_1)(Y - C \hat{z}) \\
K(y_1) &= -\beta(y_1) + \kappa
\end{align*}
\]

where \( K(y_1) = -\beta(y_1) + \kappa \) and \( Y \) is defined in (9), is an exponentially reduced order observer for (4-7), if the chosen \( (n - m - p) \times m \) matrix \( \kappa \) makes \( (A + \kappa C) \) Hurwitz, where \( C \) is defined in (8).

**Proof 1.** Let \( \varepsilon = \hat{z} - z \) be the estimation error. Since \( z_o = C z \), then we can easily derive the dynamic of observation error from (4) and (10) as follows

\[
\varepsilon = [A + (\beta(y_1) + K(y_1)) C] \varepsilon
\]

Since the gain matrix \( K(y_1) \) can be freely chosen, hence without loss of generality we set

\[
K(y_1) = -\beta(y_1) + \kappa
\]

which makes (11) become

\[
\varepsilon = (A + \kappa C) \varepsilon.
\]

Consequently, if \( \kappa \) is chosen in such a way that matrix \((A + \kappa C)\) is Hurwitz, then the exponential of \( \varepsilon \) to \( z \) can be guaranteed.

Let us remark that the proposed reduced order observer (10) is based on the "new" output \( Y \) defined in (9), which clearly shows that the derivative of the real output \( y_1 \) should be calculated according to (9). However, it is well-known that the derivative of noisy signal should be avoided if possible in practice, since derivative operation will amplify the influence of noise. Hence, several techniques can be used to limit the influence of noise when computing
the derivative of noisy signal. The most used way is to pass \( y_1 \) firstly through a low-pass filter and then calculate the derivative of \( y_1 \). It is also possible to calculate the derivative by algebraic method recently proposed in Fliess (2006), Fliess and Sira-Ramirez (2004), Fliess et al. (2008), Mboup et al. (2007) which converting the calculation of derivative to the calculation of integration, which is useful to annihilate noise.

In the following, a more practical observer is proposed based on the following hypothesis.

Hypothesis 1. We assume that the term \( K(y_1)\alpha_1^{-1}(y_1) \) is integrable with respect to \( y_1 \), i.e. we can find a \( \Gamma(y_1) \) such that

\[
\frac{\partial \Gamma(y_1)}{\partial y_1} = K(y_1)\alpha_1^{-1}(y_1)
\]

Based on Hypothesis 1, we can define \( \zeta \) as follows:

\[
\zeta = \hat{z} + \Gamma(y_1)
\]  

(13)

It should be noted that \( \Gamma(y_1) \) defined in Hypothesis 1 can be considered as a signal of filtered \( y_1 \), and thus limit the influence of noise on \( y_1 \).

Inserting \( Y \) defined in (9) into (10), (10) is rewritten into

\[
\hat{z} + K(y_1)\alpha_1^{-1}(y_1)y_1 = (A + \beta(y_1)C + K(y_1)C)\hat{z} + \rho(y) + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y)
\]  

(14)

In order to avoid the derivative of \( y_1 \), we take the new variable \( \zeta \) into account, then a more practical reduced order observer can be derived from (14) as follows

\[
\zeta = (A + \beta(y_1)C + K(y_1)C)(\zeta - \Gamma(y_1)) + \rho(y) + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y)
\]

\[
= (A + \kappa C)\zeta + \rho(y) - (A + \kappa C)\Gamma(y_1) + K(y_1)\alpha_1^{-1}(y_1)\alpha_2(y)
\]  

(15)

with \( \Gamma(y_1) \) defined in (13).

Remark 2. The new and more practical reduced order observer defined in (15) aims at limiting the influence of noise on the output, and it is based on Hypothesis 1. The estimation of \( z \) can be computed according to (13).

Corollary 1. Concerning the nonlinear canonical form with known inputs \( u \in \mathbb{R}^d \) as follows:

\[
\dot{z} = Az + \beta(y_1)z_0 + \rho(y) + \epsilon(y,u)
\]

\[
\dot{\xi} = \alpha_1(y_1)z_0 + \alpha_2(y)
\]

\[
\dot{\eta} = \mu(z,y)
\]

\[
y = (y_T, z_2)_T = (\xi_T, \eta_T)_T
\]

(16)  

(17)  

(18)  

(19)

an exponentially reduced order observer can be designed of the form:

\[
\dot{\hat{z}} = A\hat{z} + \beta(y_1)C\hat{z} + \rho(y) + \epsilon(y,u) - K(y_1)(Y - C\hat{z})
\]

with \( Y \) is defined in (9), and \( K(y_1) = -\beta(y_1) + \kappa \) where \( \kappa \) is chosen in such a way that \( (A + \kappa C) \) is Hurwitz with \( C \) defined in (8). Moreover, following the same procedure of deriving (15), a more practical reduced order observer can be deduced as well.

Example 1. Let us consider the following nonlinear system which is already of the form (3) as follows

\[
\begin{align*}
\dot{z}_{1,1} &= y_1z_{1,2} + z_{2,1}, \dot{z}_{1,2} = z_{1,1} + \rho_1(y_1) \\
\dot{z}_{2,1} &= y_2^2z_{1,2}, \dot{z}_{2,2} = z_{2,1} \\
\dot{\xi}_1 &= z_{1,2}, \dot{\xi}_2 = z_{2,1} \\
\dot{\eta} &= \mu(z, \xi) \\
y_1 &= (\xi_1, \xi_2)_T \\
y_2 &= \eta
\end{align*}
\]  

(20)

then one has

\[
z_0 = \begin{pmatrix} z_{1,2} \\ z_{2,1} \end{pmatrix}, C_1 = (0, 1), C_2 = 1, C = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{pmatrix}
\]

\[
A = \begin{pmatrix} 0 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \beta(y_1) = \begin{pmatrix} y_1 \\ 1 \end{pmatrix}, \gamma_1(y) = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}
\]

It can be checked that if one takes \( \kappa = \begin{pmatrix} -3 & -1 \end{pmatrix} \), then

\( (A + \kappa C) \) is Hurwitz.

Hence, \( K(y_1) = \begin{pmatrix} y_1 - 8 & -2 \\ 0 & y_1^2 - 2 \end{pmatrix} \) and it is easy to get

\[
\Gamma(y_1) = -\begin{pmatrix} \frac{y_1^2}{2} - 8y_1 - 2y_1 \\ -4y_1 - 2y_1 \\ \frac{y_1^3}{3} - 2y_1 - y_1 \end{pmatrix}
\]

Consequently Hypothesis 1 is satisfied and we can design a more practical reduced order observer in the form of (15) for (20).

3. TRANSFORMATION TO THE NONLINEAR CANONICAL FORM

In the last section, we defined a new nonlinear canonical form, and its associated reduced order observer is discussed as well. This section is devoted to deducing necessary and sufficient geometric conditions which allows us to transform a nonlinear system into the proposed nonlinear canonical form (3).

Let us consider a class of nonlinear systems, where the generic nonlinear system (2) can be decomposed into the following form:

\[
\dot{x} = F_1(x, \zeta, \vartheta) = f(x, \zeta, \vartheta)
\]

\[
\dot{\zeta} = F_{21}(x, \zeta, \vartheta) = \gamma_1(\zeta)H(x) + \gamma_2(\zeta, \vartheta)
\]

\[
\dot{\vartheta} = F_{22}(x, \zeta, \vartheta) = \epsilon(x, \zeta, \vartheta)
\]

\[
y = (\zeta^T, \vartheta^T)_T
\]

(21)  

(22)  

(23)  

(24)

where \( x \in \mathbb{R}^{n-m-p}, \zeta \in \mathbb{R}^m, \vartheta \in \mathbb{R}^p, y \in \mathbb{R}^{m+p}, f : \mathbb{R}^n \rightarrow \mathbb{R}^{n-m-p}, H(x) = (H_1(x), \ldots, H_m(x))^T \) are linearly independent. Moreover, we suppose system (21-24) is observable, and the observability indices Krener and Respondek (1985), Marino and Tomei (1995) for \( H(x) = (H_1(x), \ldots, H_m(x))^T \) with respect to (21) are respectively noted as \( (r_1, \ldots, r_m) \), such that \( r_1 \geq r_2 \geq \cdots \geq r_m \geq 1 \) and \( \sum_{i=1}^m r_i = n - m - p \).

Let us denote

\[
\theta_{i,j} = dL_f^{-1}H_i
\]

(24). It is possible by reordering \( H_i \) for \( 1 \leq i \leq m \).
for \(1 \leq i \leq m\) and \(1 \leq j \leq r_i\). Because of the observability of system (21-24), then the codistribution \(\text{span}\{\theta_{j,i} \mid 1 \leq i \leq m, 1 \leq j \leq r_i\}\) is of rank \(n - m - p\).

Let us note

\[ F = (f^T, f^T_2)^T \]
\[ F_2 = (f^T_2, f^T_2)^T \]

and the decomposition of \(F_2\) into \(F_{21}\) and \(F_{22}\) makes the rank of \(\gamma_1(\zeta)\) equal to \(m\).

We define For \(1 \leq i \leq m\), let us denote \(\tau_{i,1}\) vectors determined by the following equations:

\[
\begin{align*}
\theta_{i,r}(\tau_{i,1}) &= 1, \quad \text{for} \quad 1 \leq i \leq m \\
\theta_{i,k}(\tau_{i,1}) &= 0, \quad \text{for} \quad 1 \leq k \leq r_i - 1 \\
\theta_{j,k}(\tau_{i,1}) &= 0, \quad \text{for} \quad j < i \quad \text{and} \quad 1 \leq k \leq r_i \\
\theta_{j,k}(\tau_{i,1}) &= 0, \quad \text{for} \quad j > i \quad \text{and} \quad 1 \leq k \leq r_j
\end{align*}
\]  

Then, by induction we can define the following family of vector fields from \(\tau_{i,1}\) as follows

\[
\tau_{i,j} = [\tau_{i,j-1}, f] \quad \text{for} \quad 1 \leq i \leq m \quad \text{and} \quad 2 \leq j \leq r_i
\]

As we will prove in the following theorem that a necessary condition to transform (21-24) into (4-7) is

\[
[\tau_{i,j}, \tau_{s,l}] = 0
\]

for \(1 \leq i \leq m, 1 \leq j \leq r_i, 1 \leq s \leq m\) and \(1 \leq l \leq r_s\). Suppose that this condition is satisfied, then we can construct \(m\) vector fields \(\sigma_1, \ldots, \sigma_m\) and \(p\) vector fields \(v_1, \ldots, v_p\) such that \([\tau_{i,j}, \sigma_k, v_l]\) forms a basis for \(1 \leq i \leq m, 1 \leq j \leq r_i, 1 \leq k \leq m\) and \(1 \leq l \leq p\), by the following equations:

\[
d\zeta_k(\sigma_k) = \delta^k_i, \quad d\zeta_l(v_l) = 0, \quad \text{for} \quad 1 \leq i \leq m, 1 \leq k \leq l \leq p
\]

and

\[
[\tau_{i,j}, \sigma_k] = [\sigma_s, \sigma_k] = [\tau_{i,j}, v_l] = [v_l, v_l] = [\sigma_s, v_l] = 0
\]

where \(\delta^k_i\) represents Kronecker delta, i.e. \(\delta^k_i = 1\) if \(i = k\), otherwise \(\delta^k_i = 0\).

Let us note

\[
\theta = (\theta_{1,1}, \ldots, \theta_{1,r_1}, \ldots, \theta_{m,1}, \ldots, \theta_{m,r_m})
\]

and

\[
\tau = (\tau_{1,1}, \ldots, \tau_{1,r_1}, \ldots, \tau_{m,1}, \ldots, \tau_{m,r_m}, \sigma_{1,1}, \ldots, \sigma_{m,1}, \ldots, \sigma_{m,v_1}, \ldots, \sigma_{m,v_p})
\]

Set \(\Lambda = \theta(\tau)\). Due to the observability rank condition, this matrix is invertible, hence we can define the following multi 1-forms

\[
\omega = \Lambda^{-1}\theta = \begin{pmatrix} \omega_1 \\ \omega_2 \end{pmatrix}
\]

where \(\omega_2 = (d\zeta_1, \ldots, d\zeta_m, d\theta_1, \ldots, d\theta_p)^T\) and \(\omega_1\) is the rest of \(\omega\).

Now we are ready to claim our main result.

**Theorem 1.** There exists a diffeomorphism \((z, \xi, \eta) = \phi(x, \zeta, \theta)\) which transforms the dynamical system (21-24) into the nonlinear canonical form (4-7) if and only if the following conditions are satisfied

1. \([\tau_{i,j}, \tau_{s,l}] = 0\), for \(1 \leq i \leq m, 1 \leq j \leq r_i, 1 \leq s \leq m\) and \(1 \leq l \leq r_s\);
2. \(\theta_{j,1}(\tau_{i,k}) = 0\), for \(j > i, 1 \leq i \leq m, r_j + 1 \leq k \leq r_i\);
3. \([\tau_{i,j}, \tilde{F}] = \sum_{i=1}^r V_{i,j}(y_1)\tau_{i,j} + \sum_{k=1}^m W_k(y_1)\sigma_k\) for \(1 \leq i \leq m, 1 \leq j \leq r_i\) and \(1 \leq k \leq p\);
4. \([\tau_{i,j}, F_2] \in \ker \omega_1\) for \(1 \leq i \leq m\) and \(1 \leq j \leq r_i - 1\);

where \(\tilde{F} = \left(\begin{array}{c} f \\ F_{21} \end{array}\right)\), \(V_{i,j}(y_1)\) and \(W_k(y_1)\) are smooth functions of \(y_1\) defined in (24), \(\sigma_k\) is defined by (26-27) and \(\omega_1\) is defined in (28).

**Proof 2.** Necessity: Indeed, if (21-24) can be transformed into (4-7) via the diffeomorphism \((z, \xi, \eta) = \phi(x, \zeta, \theta)\), then \(\tau_{i,j} = \frac{\partial \phi}{\partial \zeta_j}\) for \(1 \leq i \leq m, 1 \leq j \leq r_i\), \(\sigma_k = \frac{\partial \phi}{\partial \zeta_k}\) for \(1 \leq k \leq m\) and \(v_l = \frac{\partial \phi}{\partial \zeta_l}\) for \(1 \leq l \leq p\). And it is easy to check that all conditions of Theorem 1 are satisfied.

**Sufficiency:** Consider the multi 1-forms \(\omega\) defined in (28), we have \(\omega(\tau) = I_{(n-m) \times (n-m)}\), which implies \(\omega(\tau_{i,j})\), \(\omega(\sigma_k)\) and \(\omega(v_l)\) are constant. Therefore,

\[
d\omega(\tau_{i,j}, \sigma_k) = L_{\tau_{i,j}}\omega(\tau_{s,l}) - L_{\tau_{s,l}}\omega(\tau_{i,j}) - \omega([\tau_{i,j}, \tau_{s,l}])
\]

thus, we can calculate \(m\) vector fields \(\sigma_1, \ldots, \sigma_m\) and \(p\) vector fields \(v_1, \ldots, v_p\), such that \([\tau_{i,j}, \sigma_k, v_l]\) forms a basis, satisfying (26) and (27). Following the same principle, we have

\[
d\omega(\tau_{i,j}, \sigma_k) = -\omega([\tau_{i,j}, \sigma_k]), \quad d\omega(\tau_{i,j}, v_l) = -\omega([\tau_{i,j}, v_l])
\]

\[
d\omega(\sigma_k, v_l) = -\omega([\sigma_k, v_l])
\]

Since \(\omega\) is an isomorphism, this implies the equivalence between

\[
[\tau_{i,j}, \tau_{s,l}] = 0
\]

and

\[
d\omega = 0\]

According to theorem of Poincaré Abraham and Marsden (1966), \(d\omega = 0\) implies that there exists a local diffeomorphism \((z, \xi, \eta) = \phi(x, \zeta, \theta)\) such that \(\omega = d\phi\). We note \(\omega_1 = \frac{\partial \phi}{\partial \zeta_1}\) for \(1 \leq i \leq 2\).

Since condition (1) in Theorem 1 is satisfied, it implies \(\phi_*(\tau_{i,j}) = \frac{\partial}{\partial \zeta_i}\), \(\phi_*(\sigma_k) = \frac{\partial}{\partial \zeta_k}\) and \(\phi_*(v_l) = \frac{\partial}{\partial \zeta_l}\) for \(1 \leq i \leq m, 1 \leq j \leq r_i, 1 \leq k \leq m\) and \(1 \leq l \leq p\).

Now let us clarify the affect of this transformation on \(f(x, \zeta, \theta)\) defined in (21). By the diffeomorphism \(\phi(x, \zeta, \theta)\), we have

\[
\phi_*(F) = \phi_*(f) + \phi_*(F_2) = \left(\begin{array}{c} \phi_*(f) + \phi_*(F_2) \\ \phi_*(f) + \phi_*(F_2) \end{array}\right)
\]

Then, for \(1 \leq i \leq m\) and \(1 \leq j \leq r_i - 1\), we get

\[
\frac{\partial \phi_*(F)}{\partial \zeta_{i,j}} = \phi_*(\tau_{i,j}, \tau_{s,l}) = \phi_*(\tau_{i,j}, v_l)
\]

since condition (4) \([\tau_{i,j}, F_2] \in \ker \omega_1\) for \(1 \leq i \leq m\) and \(1 \leq j \leq r_i - 1\) implies \(\omega_1(\tau_{i,j}, F_2) = 0\).
Finally, by the condition (3), we have
\[
\omega_1(F) =Az + \varrho(y,z_o)
\]
Then, let us prove that the diffeomorphism \( \phi(x,\zeta,\nu) \) will transform (22) to (5).

As we know
\[
\frac{\partial}{\partial z_i} H_j \circ \phi^{-1} = dH_j(\tau_{i,k}) = \theta_{j,1}(\tau_{i,k})
\]
According to the definition of \( \tau_{i,1} \) in (25), we get
\[
\theta_{j,1}(\tau_{i,k}) = \theta_{j,1}(\tau_{i,k-1}) = \cdots = \theta_{j,k}(\tau_{i,1}) = 0
\]
for \( j < i \) and \( 1 \leq k \leq r_i \). Following the same procedure, we have \( \theta_{j,1}(\tau_{i,k}) = 0 \), for \( j > i \) and \( 1 \leq k \leq r_j \). Combined with condition (2) in Theorem 1, we have
\[
\theta_{j,1}(\tau_{i,k}) = \begin{cases} 1, & i = j, k = r_i \\ 0, & \text{otherwise} \end{cases}
\]
which implies (22) can be written as
\[
\dot{z} = \gamma_1(\xi) z_o + \gamma_2(y) \tag{29}
\]
via \( z = \phi(x,\zeta,\nu) \). Hence, by setting \( \omega_2 = d\varphi_2 \) where \( \varphi_2 = \mathbf{1}_{(m+p) \times (m+p)} \), then we get
\[
\phi_*(F) = \left( \gamma_1(y_1) z_o + \gamma_2(y) \right) \tag{30}
\]
Finally, by the condition (3), we have
\[
\frac{\partial \phi_*}{\partial z_i} \bigg|_{F_i} = \phi_\star \left( \left[ \tau_{i,r_i}, F \right] \right)
\]
\[
= \phi_\star \left( \sum_{j=1}^{r_i} V_{i,j}(y_1) \tau_{i,j} + \sum_{k=1}^{m} W_k(y_1) \sigma_k \right) \tag{31}
\]
\[
= \sum_{j=1}^{r_i} V_{i,j}(y_1) \frac{\partial}{\partial z_i} + \sum_{k=1}^{m} W_k(y_1) \frac{\partial}{\partial \xi_k}
\]
which means \( W(y_1) = \alpha(y_1) \) and \( \nu(y, z_o) \) in (29) can be decomposed as:
\[
\varrho(y, z_o) = \beta(y_1) z_o + \rho(y)
\]
Thus we proved that (21-24) can be transformed to form (4-7) via \( \phi \).

**Remark 3.** As explained in the above proof, Condition (1,2,4) of Theorem 1 is used to determine the diffeomorphism \( (z,\xi,\eta) = \phi(x,\zeta,\nu) \), which transforms (21-22) to the form:
\[
\begin{align*}
\dot{z} &= Az + \varrho(y,z_o) \\
\dot{\xi} &= \alpha_1(y_1) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z,\xi,\eta) \\
y &= (\xi^T,\eta^T)^T
\end{align*}
\]
Condition (3) guarantees that the above form can be written in (4), i.e. \( \varrho(y, z_o) = \beta(y_1) z_o + \rho(y) \).

**Remark 4.** The satisfactions of Conditions (1) and (3) of Theorem 1 imply that (21-24) can be transformed into
\[
\begin{align*}
\dot{z} &= Az + \varrho(y,z_o) \\
\dot{\xi} &= \alpha_1(y_1) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z,\xi,\eta) \\
y &= (\xi^T,\eta^T)^T
\end{align*}
\]
If we can accurately measure \( y \) and calculate \( \dot{y} \), then classical observers with linear error dynamics, such as

Luenerberger observer or high-gain observer, can be always applied, where \( y \) and \( \dot{y} \) are both considered as known inputs for observers. The efficient method to calculate the nth order derivative of \( y \) can be referred to Fliess (2006), Fliess and Sira-Ramirez (2004), Fliess et al. (2008), Mboup et al. (2007), even corrupted with high frequency noise. If we want that \( \dot{y} \) will not be involved into observers, then nonlinear canonical form (4-7) should be studied. Hence all conditions of Theorem 1 need to be satisfied, and an exponentially reduced order observer of the form (15) without involving \( \dot{y} \) might be designed, provided that Hypothesis 1 is fulfilled.

**Remark 5.** If system (21) is with inputs \( u \in \mathbb{R}^q \) of the following form
\[
\dot{x} = (\mathbf{F}(x,\zeta,\nu) + \sum_{i=1}^{q} g_i(x,\zeta,\nu) u_i)
\]
then it can be transformed into (16), if Theorem 1 is valid and also the following condition is satisfied:
\[
[\tau_{i,j},g_k] = 0
\]
for \( 1 \leq i \leq m, 1 \leq j \leq r_i - 1 \) and \( 1 \leq k \leq q \).

The reason is that, if \( [\tau_{i,j},g_k] = 0 \), then
\[
\frac{\partial}{\partial z_{i,j}} \phi_*(g_k) = \phi_*(g_k) = 0
\]
which implies \( \varphi(g) = \nu(y, u) \), and thus (31) is transformed into (16).

Moreover, if Theorem 1 is valid and also the following condition is satisfied:
\[
[\tau_{i,j},g_k] = 0
\]
for \( 1 \leq i \leq m, 1 \leq j \leq r_i \) and \( 1 \leq k \leq q \), then it is easy to prove that (31) can be transformed into (16) with \( \nu(y, u) = \nu(y) u \).

**Remark 6.** In the case where \( m = 1 \) for (21-24), there exists a diffeomorphism \( \phi(x) = z \) which transforms (21) in the following canonical form:
\[
\begin{align*}
\dot{z} &= Az + \beta(y_1) + \rho(y) \\
\dot{\xi} &= \alpha_1(y_1) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z,\xi,\eta) \\
y &= (\xi^T,\eta^T)^T
\end{align*}
\]
which is an extension of the linear canonical form modulo an injection output studied in Krneren and Isidori (1983). An example in this form can be found in Nijmeijer and Marcelis (1997).

**Corollary 2.** There exists a diffeomorphism \( (z,\xi,\eta) = \phi(x,\zeta,\nu) \) which transforms the dynamical system (21-24) into
\[
\begin{align*}
\dot{z} &= Az + \beta(y_1) z_o + \rho(y) \\
\dot{\xi} &= \alpha_1(y_1) B(y) z_o + \alpha_2(y) \\
\dot{\eta} &= \mu(z,\xi,\eta) \\
y &= (\xi^T,\eta^T)^T
\end{align*}
\]
where \( A, \beta, \gamma \) are defined in (4-7) with
we can rewrite (37) as follows

\[ B(y) = \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
0 & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
b_{m-1,1}(y) & b_{m-1,2}(y) & \cdots & 1 & 0 \\
b_{m,1}(y) & b_{m,2}(y) & \cdots & b_{m,m-1}(y) & 1
\end{pmatrix} \] (36)

where \( b_{i,j}(y) \) is a function of \( y \) for \( 2 \leq i \leq m \) and \( 1 \leq j \leq i-1 \), if and only if the following conditions are satisfied

1. Conditions (1, 3, 4) of Theorem 1 are satisfied;
2. \( \theta_{j,1}(\tau_{i,k}) = \begin{cases} 
0, & \text{for } j > i, 1 \leq i \leq m, r_{j} + 1 \leq k \leq r_{i} - 1 \\
b_{i,j}(y), & \text{for } j > i, 1 \leq i \leq m, k = r_{i} 
\end{cases} \)

where \( b_{i,j}(y) \) is a function of \( y \).

Proof 3. The proof of Corollary 2 follows the same argument as that of Theorem 1. Hence we only explain Condition (2) of Corollary 2.

Indeed, Condition (2) of Corollary 2 can be interpreted as, for \( j > i, \)

\[ \frac{\partial}{\partial z_{i,r}} H_{j} \circ \phi^{-1} = \theta_{j,1}(\tau_{i,i}) = b_{j,i}(y) \]

and \( \frac{\partial}{\partial z_{i,k}} H_{j} \circ \phi^{-1} = 0 \) for \( j > i, 1 \leq i \leq m \) and \( r_{j} + 1 \leq k \leq r_{i} - 1 \). This, with the definition of \( \tau_{i,1} \) in 25, yields

\[ \frac{\partial}{\partial z_{0}} H \circ \phi^{-1} = \begin{pmatrix}
1 & 0 & \cdots & 0 & 0 \\
b_{2,1}(y) & 1 & \cdots & 0 & 0 \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
b_{m-1,1}(y) & b_{m-1,2}(y) & \cdots & 1 & 0 \\
b_{m,1}(y) & b_{m,2}(y) & \cdots & b_{m,m-1}(y) & 1
\end{pmatrix} = B(y) \]

and \( \frac{\partial}{\partial z_{i,k}} H \circ \phi^{-1} = 0 \) for \( k \neq r_{i} \). By integration, we can prove that the differentiable \( \phi \) transforms (22) to (33) with an invertible matrix \( B(y) \) defined in (36).

Remark 7. Since matrix \( B(y) \) defined in (36) is invertible, the proposed observer of the form (10) is still valid, where the "new" output \( Y \) should be redefined as follows

\[ Y = B^{-1}(y) \sigma^{-1}(y_{1})(y_{1} - \alpha_{2}(y)) \]

The following example highlights the validity of the proposed results.

Example 2. Let us consider the following nonlinear system

\[ \begin{align*}
\dot{x}_{1} &= x_{5}x_{1} + x_{3}x_{2} - x_{2} = x_{1} \\
\dot{x}_{3} &= x_{1} + x_{4} = x_{2} \\
\dot{x}_{5} &= x_{1} + x_{3}x_{2} + x_{4}x_{5} - x_{6} \\
y_{1} &= x_{4}, y_{2} = x_{5}, y_{3} = x_{6}
\end{align*} \] (37)

By setting

\[ x = \begin{pmatrix} x_{1} \\ x_{2} \\ x_{3} \end{pmatrix}, \quad \zeta = \begin{pmatrix} x_{4} \\ x_{5} \end{pmatrix}, \quad \theta = x_{6} \]

we can rewrite (37) as follows

\[ \begin{align*}
\dot{x} &= \begin{pmatrix} \dot{x}_{1} \\ \dot{x}_{2} \\ \dot{x}_{3} \end{pmatrix} = \begin{pmatrix} \zeta_{2}x_{1} + x_{3}x_{2} \\ x_{1} \end{pmatrix} \\
\dot{\zeta} &= \begin{pmatrix} \dot{\zeta}_{1} \\ \dot{\zeta}_{2} \end{pmatrix} = \begin{pmatrix} \dot{x}_{4} \\ \dot{x}_{5} \end{pmatrix} = \begin{pmatrix} x_{2} \\ x_{3} \end{pmatrix} \\
\dot{\theta} &= x_{4} + x_{3}\dot{\zeta}_{1} + \zeta_{1}\zeta_{2}\theta \\
y_{1} &= (\zeta_{1}, \zeta_{2})^{T}, y_{2} = \theta
\end{align*} \] (38)

which is of the form (4-7) with \( H(x) = (x_{2}, x_{3})^{T}, f = (\zeta_{2}x_{1} + x_{3}x_{2}, x_{1})^{T}, F_{2} = (x_{2}, x_{3}, \zeta_{1} + \zeta_{2})^{T} \) and \( \bar{F} = (f, F_{2})^{T} = (\zeta_{2}x_{1} + x_{3}x_{2}, x_{1}, x_{2}, x_{3})^{T}. \) Then we can define the following 1-forms:

\[ \theta_{1,1} = dx_{2}, \theta_{1,2} = dx_{1} \text{ and } \theta_{2,1} = dx_{3} \]

\[ d\zeta_{1} = dx_{4}, d\zeta_{2} = dx_{5} \text{ and } d\theta_{1} = dx_{6} \]

According to (25), (26) and (27), we obtain

\[ \begin{align*}
\tau_{1,1} &= \frac{\partial}{\partial x_{1}} \tau_{1,2} = \frac{\partial}{\partial x_{2}} + \frac{\partial}{\partial x_{3}} + x_{5}\frac{\partial}{\partial x_{1}}, \tau_{2,1} = \frac{\partial}{\partial x_{3}} \\
\sigma_{1} &= \frac{\partial}{\partial x_{4}} + \frac{\partial}{\partial x_{5}} + x_{2}\frac{\partial}{\partial x_{1}}, v_{1} = \frac{\partial}{\partial x_{6}}
\end{align*} \]

It is easy to check that \( [\tau_{j,1}, \tau_{k,s}] = 0 \) for \( 1 \leq i \leq 2, 1 \leq j \leq r_{1}, 1 \leq s \leq 2 \) and \( 1 \leq l \leq r_{2} \) with \( r_{1} = 2 \) and \( r_{2} = 1. \) Moreover we have

\[ \begin{align*}
[\tau_{1,2}, \bar{F}] &= \zeta_{2}\tau_{1,2} + \sigma_{2} + \sigma_{1} \\
[\tau_{2,1}, \bar{F}] &= \sigma_{2}
\end{align*} \]

and

\[ \theta_{2,1}(\tau_{1,2}) = 1 \]

In order to calculate the diffeomorphism, let us consider

\[ \Lambda = \theta_{\tau} = \begin{pmatrix}
0 & 1 & 0 & 0 & 0 & 0 \\
1 & x_{5} & 0 & x_{2} & 0 \\
0 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1
\end{pmatrix} \]

which yields

\[ \omega = \Lambda^{-1}\theta = \begin{pmatrix}
dx_{1} - x_{5}dx_{2} - x_{2}dx_{5} \\
dx_{2} \\
dx_{4} \\
dx_{5} \\
dx_{6}
\end{pmatrix} \]

which gives \( \omega_{1} = \begin{pmatrix}
dx_{1} - x_{5}dx_{2} - x_{2}dx_{5} \\
dx_{2} \\
dx_{4} + dx_{3} \\
dx_{5} \\
dx_{6}
\end{pmatrix} \).

Then \( \omega_{1} [\tau_{1,1}, F_{2}] = 0, \) implying \( [\tau_{1,1}, F_{2}] \in \ker \omega_{1}. \) Thus all conditions of Corollary 2 are satisfied, and we have

\( (z_{1,1}, z_{1,2}, z_{2,1}, \xi_{1}, \xi_{2}, \eta) = (x_{1} - x_{2}x_{5}, x_{2}, x_{3} - x_{2}x_{4}, x_{5}, x_{6})^{T} \) which transforms (37) into

\[ \begin{align*}
\dot{z}_{1,1} &= 0, \dot{z}_{1,2} = z_{1,1} + \xi_{2}z_{1,2}, \dot{z}_{2,1} = 0 \\
\dot{\xi}_{1} &= z_{1,2}, \dot{\xi}_{2} = z_{2,1} + z_{1,2}, \eta = \xi_{1} + \xi_{2} \\
y_{1} &= (\zeta_{1}, \zeta_{2})^{T}, y_{2} = \eta
\end{align*} \]

4. SYSTEM WITH INPUTS AND HIGH-GAIN REDUCED ORDER OBSERVER

In this section, we will generalize the main results to systems with inputs. At the same time, a corresponding high-gain reduced order observer is given.

Let us now consider the dynamical system with inputs in the following form:
\[ \dot{x} = F_1(x, \zeta, \vartheta, u) = f(x, \zeta, \vartheta, u) \]  
(39)
\[ \dot{\zeta} = F_{21}(x, \zeta, \vartheta) = \gamma_1(\zeta)H(x) + \gamma_2(\zeta, \vartheta) \]  
(40)
\[ \dot{\vartheta} = F_{22}(x, \zeta, \vartheta) = \varepsilon(x, \zeta, \vartheta, u) \]  
(41)
\[ y = (\zeta^T, \vartheta^T)^T \]  
(42)
where \( x, \zeta, f, \alpha_1, \alpha_2 \) and \( H \) are defined as those in previous section, \( u \in \mathbb{R}^q \) is assumed to be known. Moreover, we assume as well that system (39-42) is observable and the matrix \( \gamma_1(y) \) is of rank \( m \). Suppose that the observability indices for \( H(x) = (H_1(x), \ldots, H_m(x))^T \) with respect to (39) are respectively noted as \( (r_1, \ldots, r_m) \), such that \( r_1 \geq r_2 \geq \cdots \geq r_m \geq 1 \) and \( \sum_{i=1}^n r_i = n - m - p \).

Let us note
\[ F = (f^T, F_{21}^T) \]  
(43)
\[ F_2 = (F_{21}^T, F_{22}^T) \]  
(44)
For the given system (39-42), we are trying to seek a diffeomorphism \( (z, \xi, \eta) = \phi(x, \zeta, \vartheta, u) \) such that (39-42) can be transformed into the following canonical form:
\[ \begin{align*}
\dot{z} &= A(y, u)z + \beta(y_1)z_0 + \rho(y) \\
\dot{\xi} &= \alpha_1(y_1)z_0 + \alpha_2(y) \\
\dot{\eta} &= \mu(z, \xi, \eta, u) \\
y &= (\xi^T, \eta^T)^T
\end{align*} \]  
(45)
where \( z, z_0, \beta, \gamma \) are defined as those in previous section, and \( A(y, u) = \text{diag}[A_1(y, u), \ldots, A_m(y, u)] \) with
\[ A_i = \begin{pmatrix}
0 & \cdots & 0 & 0 & 0 \\
0 & \cdots & 0 & 0 & 0 \\
\vdots & \ddots & \vdots & \vdots & \vdots \\
0 & \cdots & 0 & 0 & 0 \\
0 & \cdots & 0 & 0 & 0
\end{pmatrix}
\]  
(46)
where \( \mu_{ij}(y, u) \neq 0 \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq r_i - 1 \).

Before to state our result, for system (39-40) we denote \( \theta_{ij} = dL_i^T H_i \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq r_i \). Then, following the same fashion we can define the frame \( \tau_{ij} \) through (25) and \( \sigma_k, \nu_l \) by (26-27) for \( 1 \leq k \leq m \) and \( 1 \leq l \leq p \).

**Theorem 2.** There exists a diffeomorphism \( (z, \xi, \eta) = \phi(x, \zeta, \vartheta, u) \) which transforms (39-42) into (43-46) if there exists a family of \( \mu_{ij}(y, u) \) such that the frame \( \tilde{\tau}_{ij} \) constructed as follows:
\[ \tilde{\tau}_{ij} = \tau_{ij}, \quad \text{for} \ 1 \leq i \leq m \]
\[ \tilde{\tau}_{i,j+1} = \frac{1}{\mu_{ij}(y, u)}[\tilde{\tau}_{ij}, f], \quad \text{for} \ 1 \leq j \leq r_i - 1 \]
satisfies the following conditions:

1. \( [\tilde{\tau}_{ij}, \tilde{\tau}_{ik}] = 0 \), for \( 1 \leq i \leq m, 1 \leq j \leq r_i, 1 \leq s \leq m \) and \( 1 \leq l \leq r_s \);
2. \( [\tilde{\tau}_{r_i, r_{i-1}}] = \sum_{j=1}^{\rho_{1}} V_{ij}(y_1)\tilde{\tau}_{ij} + \sum_{k=0}^{\rho_{2}} W_{ik}(y_1)\sigma_k, \) for \( 1 \leq i \leq m, 1 \leq j \leq r_i \) and \( 1 \leq k \leq m; \)
3. \( \theta_{ij, 1}(\tilde{\tau}_{ij}) = 0 \), for \( j > i, 1 \leq i \leq m \) and \( r_i + 1 \leq j \leq r_s; \)
4. \( [\tilde{\tau}_{ij}, F_2] \in \ker \omega_1 \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq r_i - 1; \)
5. \( [\tilde{\tau}_{ij}, \partial_\alpha] = 0 \), for \( 1 \leq i \leq m, 1 \leq j \leq r_i \) and \( 1 \leq k \leq q; \)

where \( \tilde{F} = \begin{pmatrix} f \\ F_{21} \end{pmatrix} \), \( V_{ij}(y_1) \) and \( W_{ik}(y_1) \) are smooth functions of \( y \) defined in (24), \( \sigma_k \) is defined by (26-27) and \( \omega_1 \) is defined in (28).

**Proof.**

**Necessity:** If (39-42) can be transformed into (43-46), then it is easy to show that, for \( 1 \leq i \leq m \), we have \( \tau_{ij, 1} = \frac{\partial}{\partial \sigma_{i,1}}, \sigma_k = \frac{\partial}{\partial \sigma_k} \) for \( 1 \leq k \leq m \) and \( \nu_l = \frac{\partial}{\partial \nu_l} \) for \( 1 \leq l \leq p \). Then we have \( \tau_{ij, 1} = \frac{\partial}{\partial \sigma_{i,1}} \). By construction we obtain \( \tau_{ij, 1} = \frac{\partial}{\partial \sigma_{i,1}} \) for \( 2 \leq j \leq r_i \). And we can check that Condition (1-5) of Theorem 2 are satisfied.

** Sufficiency:** The proof of sufficiency of Theorem 1 can be adapted easily from that of Theorem 1, and the following gives only the sketch.

Indeed, Condition (1) of Theorem 2 implies that there exists a diffeomorphism \( (z, \xi, \eta) = \phi(x, \zeta, \vartheta, u) \) where \( \omega = d\phi(x, \zeta, \vartheta, u) \) such that \( \phi_*(\tilde{\tau}_{ij}) = \frac{\partial}{\partial \sigma_{i,1}} \) for \( 1 \leq i \leq m \) and \( 1 \leq j \leq r_i \). Then, if condition (4) is satisfied, we have
\[ \partial \phi_*(F) = \phi_*([\tilde{\tau}_{ij}, F]) = \begin{pmatrix} \omega_1(\tilde{\tau}_{ij}, f) + \omega_1 \tilde{\tau}_{ij, F_2} \\ [\omega_2 \tilde{\tau}_{ij, f} + \omega_2 \tilde{\tau}_{ij, F_2}] \end{pmatrix} \]  
(47)
By integrating we obtain:
\[ \omega_1(F) = A(y, u)z + \rho(y, z_0) \]
As shown in the proof of Theorem 1, it can be proved
\[ \phi_*(F) = \begin{pmatrix} A(y, u)z + \rho(y, z_0) \\ \alpha_1(y_1)z_0 + \alpha_2(y) \\ \mu(z, \xi, \eta, u) \end{pmatrix} \]  
(48)
Since in the current case, the diffeomorphism depends on \( u \) as well, thus we have
\[ \left( \begin{array}{c}
\dot{z} \\
\dot{\xi} \\
\dot{\eta}
\end{array} \right) = \phi_*(F) + \sum_{k=1}^q \frac{d\phi}{du_k} \dot{u}_k \]  
(49)
Note \( \omega = \omega_x + \omega_u \) where \( \omega_x \) and \( \omega_u \) represent respectively the differential in the state and the input space, then the total differential of \( \omega \) can be noted as \( d\omega = d\omega_x + d\omega_u \). Since \( \omega(\tilde{\tau}_{ij}) = \frac{\partial}{\partial \sigma_{i,1}} \) and \( \omega_u = \frac{\partial}{\partial \nu_{ij}} \), then we have the following equation:
For $1 \leq i \leq m$, non-linear canonical form (43-46), which will be described as a diffeomorphism which transforms (39-42) into (43-46).

Given dynamical system of the form (39-42), i.e. we can find a symmetric positive definite for every $\epsilon > 0$. Then it is proved in Busawon et al. (1998) that (48) can be written as

$$
\begin{array}{l}
\dot{\tilde{z}} = \left(\begin{array}{c}
\dot{\tilde{y}} \\
\dot{\tilde{\eta}}
\end{array}\right) = \left(\begin{array}{c}
A(y,u)z + \phi(y,z_0) \\
\alpha_1(y)z_0 + \alpha_2(y) \\
\mu(z,\xi,\eta,u)
\end{array}\right)
\end{array}
$$

The rest of proof is similar to that of Theorem 1, and thus to be omitted.

Remark 8. Zheng et al. (2006) gave sufficient and necessary geometric conditions to determine $\mu_i(y,u)$ for $i = 1$ and $1 \leq j \leq r_i$ and $i \leq k \leq q$, which implies (48) can be written as

$$
\begin{array}{l}
\dot{\tilde{z}} = \left(\begin{array}{c}
\dot{\tilde{y}} \\
\dot{\tilde{\eta}}
\end{array}\right) = \left(\begin{array}{c}
A(y,u)z + \phi(y,z_0) \\
\alpha_1(y)z_0 + \alpha_2(y) \\
\mu(z,\xi,\eta,u)
\end{array}\right)
\end{array}
$$

The proof of the convergence and its corresponding conditions of such an observer can be referred to Busawon et al. (1998).

5. CONCLUSION

A non-linear canonical form was studied in this paper. We firstly gave a set of sufficient and necessary geometric conditions which transform a special class of non-linear systems to the proposed non-linear canonical form. And then a reduced order observer was proposed. Finally we generalized our main results to non-linear systems with inputs. The proposed normal forms are more generic since they contain a redundant dynamic $\vartheta$, which in return enables to design a more robust observer.

REFERENCES


