Observer Design in Convergent Series for a Class of Nonlinear Systems

Zhengtao Ding

Abstract—This paper deals with convergence analysis for power series solutions to a partial differential equation for nonlinear observer design with linear observer error dynamics. This power series solution is used to design the gain matrix for a Luenberger-like observer for nonlinear systems. An explicit domain of convergence around the origin is identified, which is related to the relative sizes of high-order terms in the original nonlinear system with respect to the linearized model. The convergent conditions can provide a guideline for nonlinear observer design with a truncated series for the observer gain.

Index Terms—Convergence analysis, nonlinear systems, observer design, polynomial approximation.

I. INTRODUCTION

Results for observer design of nonlinear systems started to appear in 1970 [1], [2], and observer design for nonlinear systems continue to attract significant attention of control research with many results appeared in literature (for example, see [3]–[7] etc). One significant result was on observer design for nonlinear systems with linear observer error dynamics by introducing output injection [8]. A more general formulation of observers with linear observer error dynamics is reported in [9] based on Lyapunov's auxiliary theorem in which a Luenberger-like observer is presented with the observer gain to be determined by a nonlinear function of the system state. This nonlinear function is a solution to a partial differential equation. This result has attracted significant attention in nonlinear observer design, as evidenced by some of the recent results [10]–[16].

The partial differential equation in [9] depends on the dynamics of the original system and the chosen dynamics of observer errors. Although conditions have been identified for the existence of a solution, a closed-form solution to the partial differential equation is difficult to obtain in general. When a closed-form solution is not available, solutions in power series can be considered, and an iterative method of obtaining high-order polynomials has been introduced [9]. This series can be truncated to a certain order and the truncated series can then be used for computing observer gain in the nonlinear observer design. The existence of a convergent series around the origin to the partial differential equation was established in [10] by exploring the properties of analytic functions. In this paper, we aim at identifying an explicit domain around the origin in which the power series solution to the partial differential equation is guaranteed to converge for a class of nonlinear systems with nonlinear functions in polynomials. The conditions depend on the higher order terms in relation to the first-order one, and on the radius of the state variables with respect to the origin. In establishing the convergence result, we formulate the individual terms in the series as states of a discrete-time system through proper matrix manipulations, and obtain the convergence result based the dynamic of the discrete time system. With the result of the convergence of the series,

Manuscripts received July 04, 2011; revised July 04, 2011; accepted July 04, 2011. Date of publication December 08, 2011; date of current version June 22, 2012. Recommended by Associate Editor X. Xia.

The author with the Control Systems Centre, School of Electrical and Electronic Engineering, University of Manchester, Manchester M13 9PL, U.K. (e-mail: zhengtao.ding@manchester.ac.uk).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2011.2178880

users would be more confident in selecting the order of approximation and identifying the domain for the observer errors to converge. An example is included to demonstrate the notations used in the paper and to reveal some issues in the convergence of the power series in the observer design.

II. PROBLEM FORMULATION

Consider a nonlinear system

$$\dot{x} = f(x) \quad (1)$$

$$y = h(x) \quad (2)$$

where $x \in \mathbb{R}^n$ is the state vector, $y \in \mathbb{R}^m$ is the output, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ are continuous nonlinear functions with $f(0) = 0$, and $h(0) = 0$, and $n > m$.

As shown in [9], an observer can be designed if there exists a nonlinear function $p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that

$$\frac{\partial p(x)}{\partial x} f(x) = Ap(x) + Bh(x) \quad (3)$$

where $A \in \mathbb{R}^{n \times n}$ and $B \in \mathbb{R}^{n \times m}$ are constant matrices with A Hurwitz and $\{A, B\}$ controllable. Let $\xi = p(x)$, and its dynamics are then described by

$$\begin{aligned} \dot{\xi} &= A\xi + By, \\ y &= h(p^{-1}(\xi)). \end{aligned}$$

from which an observer is designed as

$$\begin{aligned} \dot{\hat{\xi}} &= A\hat{\xi} + By \\ \hat{x} &= p^{-1}(\hat{\xi}) \end{aligned} \quad (4)$$

where the invertibility of p is needed.

It is easy to see that the observer error dynamics are linear as

$$\dot{\tilde{\xi}} = A\tilde{\xi}.$$

The observer can also be implemented in the original state as

$$\dot{\hat{x}} = f(\hat{x}) + \left(\frac{\partial p}{\partial x}(\hat{x}) \right)^{-1} B(y - h(\hat{x})). \quad (5)$$

which is in the same structure as the standard Luengerger observer for linear systems by viewing $((\partial p / \partial x)(\hat{x}))^{-1} B$ as the observer gain.

The key step in designing this type of nonlinear observers is to solve the nonlinear function $p(x)$ in (3). Sufficient conditions for the existence of solution of (3) have been given in [9]. In case that a closed-form solution is difficult to obtain, solutions in power series can then be obtained. In this paper, we analyze the convergence of the series solutions to the partial differential (3) for nonlinear observer design.

For the convergence analysis, we assume that functions $f(x)$ and $h(x)$ are polynomials of x with finite orders, and we introduce an assumption on the eigenvalues of $(\partial f / \partial x)(0)$ and A . Before that, let us define an extended set of combination of eigenvalues

$$\mathcal{F} := \left\{ \lambda | \lambda = \sum_{i=1}^n q_i \lambda_i \left(\frac{\partial f}{\partial x}(0) \right) \right\}$$

where $\lambda_i((\partial f / \partial x)(0))$, $i = 1, \dots, n$, are the eigenvalues of $(\partial f / \partial x)(0)$, and q_i are any set of non-negative integers which satisfy $\sum_{i=1}^n q_i > 0$.

Assumption 1: All the eigenvalues of $(\partial f / \partial x)(0)$ are distinct real numbers with the same sign, and all the eigenvalues of A are chosen to be distinct and negative real numbers, and they do not belong to the extended set of eigenvalues of $(\partial f / \partial x)(0)$, i.e., $\lambda_i(A) \notin \mathcal{F}$.

Remark 1: Even though functions $f(x)$ and $h(x)$ are polynomials with finite orders, the power series solution to (3) is not of finite order in general. When $f(x)$ and $h(x)$ are other smooth nonlinear functions, they can be approximated by polynomials. It has been shown in literature [9], [10] that the distinctive real eigenvalues made in Assumption 1 are not necessary conditions for the existence of a solution to the partial differential condition. Here, this condition is needed for the convenience of introducing the convergence analysis in this paper.

Assumption 1 allows that $(\partial f/\partial x)(0)$ have all the eigenvalues positive or all negative, and guarantees the existence of a solution to the partial differential (3), following the result shown in [9].

Based on Assumption 1, we can state, without loss of generality, that both $(\partial f/\partial x)(0)$ and A are diagonal.

Remark 2: When the eigenvalues of $(\partial f/\partial x)(0)$ and A are distinct, we can introduce suitable state transforms such that the transformed system and transformed A satisfy Assumption 1 with diagonal $(\partial f/\partial x)(0)$ and A . Let $(\partial f/\partial x)(0) = T_1^{-1} \bar{F}_1 T_1$ where $\bar{F}_1$ is a diagonal matrix. Let $\bar{x} = T_1 x$, and we have

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

where $\bar{f}(\cdot) = T_1 f(T_1^{-1}(\cdot))$, and $\bar{h}(\cdot) = h(T_1^{-1}(\cdot))$. It is easy to verify that $(\partial \bar{f}/\partial \bar{x})(0) = \bar{F}_1$ that is diagonal. For A , we can choose a controllable pair $\{A, B\}$ with A diagonal. If A is not diagonal, but with distinct eigenvalues, a transform can be introduced. If $A = T_2^{-1} \bar{A} T_2$ with $\bar{A}$ diagonal, multiplying both sides of (3) by T_2 , we have

$$T_2 \frac{\partial p(x)}{\partial x} f(x) = \bar{A} T_2 p(x) + T_2 B y$$

which gives

$$\frac{\partial(T_2 p(x))}{\partial x} f(x) = \bar{A}(T_2 p(x)) + T_2 B y.$$

This means that the convergence analysis can be carried out with the diagonal $\bar{A}$ together with $T_2 p(x)$.

III. SOLUTIONS IN POWER SERIES

Series solutions for (3) are discussed in [9]. In this section, we introduce an set of iterative matrix equations, by taking similar notations for series solutions for nonlinear output regulation equations in [17] for the convenience of convergence analysis.

Let us denote a solution $p(x)$ by

$$p(x) = \sum_{i=1}^{\infty} p_i(x) \quad (6)$$

where $p_i : \mathbb{R}^n \rightarrow \mathbb{R}^n$ denotes all the polynomial terms with the order i , i.e., all the terms of $x_1^{a_1} x_2^{a_2} \dots x_n^{a_n}$ with $\sum_{k=1}^n a_k = i$ and $a_k \in \mathbb{N}$. Similarly, we denote

$$f(x) = \sum_{i=1}^{n_f} f_i(x) \quad (7)$$

$$h(x) = \sum_{i=1}^{n_h} h_i(x) \quad (8)$$

with the notations $f_i = 0$ for $i > n_f$, and $h_i = 0$ for $i > n_h$. Substituting the above expressions into (3), we have

$$\sum_{i=1}^{\infty} \frac{\partial p_i(x)}{\partial x} \sum_{i=1}^{n_f} f_i(x) = A \sum_{i=1}^{\infty} p_i(x) + B \sum_{i=1}^{n_h} h_i(x). \quad (9)$$

Comparing the order of the polynomial terms in above equation, we have

$$\begin{aligned}\frac{\partial p_1(x)}{\partial x} f_1(x) - A p_1(x) &= B h_1(x) \\ \frac{\partial p_2(x)}{\partial x} f_1(x) - A p_2(x) &= -\frac{\partial p_1(x)}{\partial x} f_2(x) + B h_2(x) \\ &\vdots \\ \frac{\partial p_i(x)}{\partial x} f_1(x) - A p_i(x) &= -\sum_{j=2}^i \frac{\partial p_{i-j+1}(x)}{\partial x} f_j(x) + B h_i(x) \\ &\dots\end{aligned} \quad (10)$$

An iterative solution can start from (10) for $p_1(x)$ and then follows (11) for $p_i(x)$ for $i = 2, 3, 4$ etc.

For the first-order term, we have $p_1(x) = P_1 x$ and $f_1(x) = F_1 x$ with $P_1, F_1 \in \mathbb{R}^{n \times n}$ being constant matrices, and $h_1(x) = H_1 x$ with $H_1 \in \mathbb{R}^{n \times m}$ being a constant matrix. Hence, the (10) leads to a matrix equation

$$P_1 F_1 - A P_1 = B H_1.$$

The solution of this equation is guaranteed by the condition specified in Assumption 1.

Remark 3: When $\{A, B\}$ is controllable and $\{F_1, H_1\}$ is observable, P_1 is invertible, which implies the invertibility of $p(x)$ around a neighborhood of the origin.

For the high-order terms, we need to introduce a few notations. Let us introduce a base function for polynomials such that we can write

$$p_i(x) = P_i v_i(x).$$

where $v_i(x)$ contains all the unique polynomial terms of order i .

Adopting some notations from [17], we define

$$\begin{aligned}v^{(1)} &= x = [x_1, \dots, x_n]^T \\ &\text{and} \\ v^{(i)} &= v^{(i-1)} \otimes x, \text{ for } i > 1\end{aligned}$$

where $\otimes$ denotes Kronecker product. Hence, v_i can be formed by taking the unique elements from $v^{(i)}$, and it can be denoted by

$$v_i = M_i v^{(i)}.$$

Note that M_i has more columns than rows for $i > 1$, and therefore it is not invertible. However, since the unique entries in $v^{(i)}$ are contained in v_i , we can write

$$v^{(i)} = N_i v_i.$$

It can be seen that $M_i N_i = I$, from their definitions.

With the notations v_i , we can write $f_i(x) := F_i v_i(x)$ and $h_i(x) := H_i v_i(x)$ with F_i and H_i are constant matrices with proper dimensions.

Furthermore, we denote matrix notations Σ_i and $Q_{i,j}$ for partial differential operations by

$$\frac{\partial v_i(x)}{\partial x} F_1 x := \Sigma_i v_i(x), \quad (12)$$

$$\frac{\partial v_{i-j}(x)}{\partial x} f_{j+1}(x) := -Q_{i,j} v_i(x) \quad (13)$$

for $j = 1, \dots, n_f - 1$. From [17, Lemma 4.6], we have the explicit expressions of Σ_i and $Q_{i,j}$ as

$$\Sigma_i := M_i \bar{\Sigma}_i N_i \quad (14)$$

$$Q_{i,j} := M_{i-j} \bar{Q}_{i,j} N_i \quad (15)$$

where

$$\bar{\Sigma}_i = \left[\sum_{k=1}^i I_n^{(k-1)} \otimes F_1 \otimes I_n^{(i-k)} \right] \quad (16)$$

$$\bar{Q}_{i,j} = - \left[\sum_{k=1}^{i-j} I_n^{(k-1)} \otimes F_j \otimes I_n^{(i-j-k)} \right] \quad (17)$$

and the superscript (k) denotes k times Kronecker products, i.e., $I_n^{(3)} = I_n \otimes I_n \otimes I_n$.

With the notations introduced, we can rewrite the (11) as

$$P_i \Sigma_i v_i(x) - A P_i v_i(x) = \sum_{j=1}^{n_f-1} P_{i-j} Q_{i,j} v_i(x) + B H_i v_i(x)$$

and therefore

$$P_i \Sigma_i - A P_i = \sum_{j=1}^{n_f-1} P_{i-j} Q_{i,j} + B H_i \quad (18)$$

for $i \in \mathbb{N}$. If we have solutions of P_i form (18), we will then have a solution for the partial differential (3). Based on Assumption 1, the solution of P_i is guaranteed.

Theorem 1: There exists a unique solution for the matrix (18) if Assumption 1 is satisfied.

Proof: The existence of unique solution can be established based on the results shown [9].

IV. CONVERGENCE ANALYSIS OF POWER SERIES SOLUTIONS

In this section, we will analyze the convergence issue of the power series solutions obtained from the iterative polynomial method. As shown in the previous section that P_i can be obtained from (18) iteratively for $i = 1, \dots, N$, with N denoting any big positive integer. Hence, we can write

$$p(x) = \sum_{i=1}^{\infty} P_i v_i(x).$$

We introduce a number of notations for the convergence analysis. We use σ_1 to denote the largest singular value and σ_n for the smallest singular values, regardless the actual dimension of the matrix, and similarly, λ_1 and λ_n for the eigenvalues of a matrix with largest and smallest modules. The notation $\text{vec}(A)$ is to denote the vector formed by stacking all the vectors of A . We have used $\|\cdot\|$ to denote 2-norm for a vector or its induced norm for a matrix, and $\|\cdot\|_{\infty}$ for the infinity norm for a matrix $A \in R^{n \times m}$ as

$$\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^m |a_{i,j}|.$$

We also define $\mu_n \in R^n$ with all the elements equal 1, i.e., $\mu_3 = [1, 1, 1]^T$, and for a matrix $A = \{a_{i,j}\} \in R^{n \times m}$ we denote $A^+ = \{|a_{i,j}|\}$. Furthermore, for $A, B \in R^{n \times m}$, we define $A \leq B$ if $a_{i,j} \leq b_{i,j}$ for $i = 1, \dots, n$ and $j = 1, \dots, m$.

For the convergence of this series, we have the following theorem.

Theorem 2: The series $\{p_i(x)\}$ for $i \in \mathbb{N}$ converges in l_2 for $|x_i| < d_x$, $i = 1, \dots, n$, if

$$|\lambda_1(S)| < 1 \quad (19)$$

where $S \in R^{(n_f-1) \times (n_f-1)}$ and

$$S = \begin{bmatrix} 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ddots & 1 \\ s_{n_f-1} & s_{n_f-2} & \dots & s_1 \end{bmatrix} \quad (20)$$

with

$$s_j = \frac{\|F_{j+1}\|_{\infty}}{\sigma_n(F_1)} d_x^j$$

for $j = 1, \dots, n_f - 1$.

We need a few technical results for the proof of this theorem.

In the following lemma, we list a number of results on matrices and singular values that are needed later on.

Lemma 3: The following facts hold:

3.1 For $A \in R^{n \times m}$, $\sigma_1(I \otimes A) = \sigma_1(A \otimes I) = \sigma_1(A)$ and $\sigma_n(I \otimes A) = \sigma_n(A \otimes I) = \sigma_n(A)$.

3.2 For $A \in R^{n \times n}$, $\sigma_n(A) = |\lambda_n(A)|$ if A is a normal matrix.

3.3 For A, X and B in proper dimensions such that AXB exists, $\text{vec}(AXB) = (B^T \otimes A)\text{vec}(X)$.

3.4 For $A \in R^{n \times m}$, $\|A \otimes I\|_{\infty} = \|I \otimes A\|_{\infty} = \|A\|_{\infty}$.

3.5 $A \in R^{n \times m}$, $A^+ \mu_m \leq \|A\|_{\infty} \mu_n$.

3.6 For A and B in proper dimensions such that AB exists, $(AB)^+ \leq A^+ B^+$, and if A is diagonal, $(AB)^+ \leq \sigma_1(A) B^+$.

Proof: The results shown in 3.1 to 3.4 can be directly found or derived from the results in [18]. The rest of the results in the lemma follow directly from the definition of A^+ .

Lemma 4: For the matrices M_i and N_i , the following facts hold:

4.1 $N_i M_i \bar{v}^{(i)} = \bar{v}^{(i)}$ with $\bar{v}^{(i)} := \mu_{n_i}$.

4.2 Let D be a diagonal matrix with proper dimensions such that $M_i D N_i$ can be evaluated, and D invertible. The matrix $M_i D N_i$ is diagonal too, and $(M_i D N_i)^{-1} = M_i D^{-1} N_i$.

Proof: From the definitions, we have $v^{(i)} = N_i v_i = N_i M_i v^i$. Therefore, 4.1 can be obtained by setting $x_j = 1$ for $j = 1, \dots, n$. The results in 4.2 can be established by exploiting the positions of "1" in matrices M_i and N_i .

We are ready to prove Theorem 2.

Proof of Theorem 2: We first establish a relationship between P_i and P_j with $j = i - 1, \dots, i - n_f + 1$. Let us assume that $i > n_h$ which implies $H_i = 0$. From (18), we then have

$$P_i \Sigma_i - A P_i = \sum_{j=1}^{n_f-1} P_{i-j} Q_{i,j}. \quad (21)$$

Using the notation of $\text{vec}(\cdot)$, and the result 3.3 in Lemma 3, we have

$$\text{vec}(P_i) = \Omega_i^{-1} \sum_{j=1}^{n_f-1} [Q_{i,j}^T \otimes I] \text{vec}(P_{i-j}) \quad (22)$$

where

$$\Omega_i = [\Sigma_i^T \otimes I - I \otimes A].$$

From (16) and 4.2 of Lemma 4, we know that Σ_i is diagonal and furthermore that Ω_i is diagonal, as A is diagonal.

From (22), using the result 3.6, we have

$$\begin{aligned} \text{vec}(P_i^+) &\leq (\Omega_i^{-1})^+ \sum_{j=1}^{n_f-1} [(Q_{i,j}^T)^+ \otimes I] \text{vec}(P_{i-j}^+) \\ &\leq \sigma_1((\Omega_i^{-1})^+) \sum_{j=1}^{n_f-1} [(Q_{i,j}^T)^+ \otimes I] \text{vec}(P_{i-j}^+) \end{aligned}$$

$$\begin{aligned}
&\leq \sigma_1((\Omega_i^+)^{-1}) \sum_{j=1}^{n_f-1} [(Q_{i,j}^T)^+ \otimes I] \text{vec}(P_{i-j}^+) \\
&= \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} [(Q_{i,j}^T)^+ \otimes I] \text{vec}(P_{i-j}^+)
\end{aligned}$$

which implies, with the result 3.3 in Lemma 3, that

$$\begin{aligned}
P_i^+ &\leq \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} P_{i-j}^+ Q_{i,j}^+ \\
&\leq \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} P_{i-j}^+ M_{i-j} \bar{Q}_{i,j}^+ N_i.
\end{aligned} \quad (23)$$

Multiplying both sides of (23) by $M_i \bar{v}^{(i)}$ and using the results 4.1, 3.5 and 3.6, we have

$$\begin{aligned}
P_i^+ M_i \bar{v}^{(i)} &\leq \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} P_{i-j}^+ M_{i-j} \bar{Q}_{i,j}^+ N_i M_i \bar{v}^{(i)} \\
&= \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} P_{i-j}^+ M_{i-j} \bar{Q}_{i,j}^+ \bar{v}^{(i)} \\
&\leq \sigma_n^{-1}(\Omega_i^+) \sum_{j=1}^{n_f-1} P_{i-j}^+ M_{i-j} \|\bar{Q}_{i,j}^+\|_\infty \bar{v}^{(i-j)}.
\end{aligned}$$

Let us evaluate $\sigma_n^{-1}(\Omega_i^+)$ and $\|\bar{Q}_{i,j}^+\|_\infty$. When F_1 has all the eigenvalues negative, we have

$$\Omega_i = -(\Sigma^T)^+ \otimes I + I \otimes A^+. \quad (24)$$

Assuming i is big enough, we have

$$\Omega_i^+ = (\Sigma_i^T)^+ \otimes I - I \otimes A^+ \quad (25)$$

and

$$\sigma_n(\Omega_i^+) \geq \sigma_n(\Sigma_i) - \sigma_1(A) \geq i\sigma_n(F_1) - \sigma_1(A).$$

Similarly, when F_1 has all the eigenvalues positive, we have

$$\Omega_i^+ = (\Sigma^T)^+ \otimes I + I \otimes A^+$$

and it can be obtained that

$$\sigma_n((\Omega_i)^+) \geq \sigma_n(\Sigma) \geq i\sigma_n(F_1).$$

Hence, combining both the cases, we have

$$\sigma_n((\Omega_i)^+) \geq i\sigma_n(F_1) - \sigma_1(A).$$

From (17), it can be obtained that

$$\begin{aligned}
\|\bar{Q}_{i,j}^+\|_\infty &= \left\| \sum_{k=1}^{i-j} I_n^{(i-1)} \otimes F_{j+1} \otimes I_n^{(i-1-k)} \right\|_\infty \\
&\leq \sum_{k=1}^{i-j} \|I_n^{(i-1)} \otimes F_{j+1} \otimes I_n^{(i-1-k)}\|_\infty \\
&\leq \sum_{k=1}^{i-j} \|F_{j+1} \otimes I_n^{(i-1-k)}\|_\infty \\
&\leq (i-j) \|F_{j+1}\|_\infty.
\end{aligned}$$

Therefore, we have

$$P_i^+ M_i \bar{v}^{(i)} \leq \sum_{j=1}^{n_f-1} \frac{(i-j) \|F_{j+1}\|_\infty}{i\sigma_n(F_1) - \sigma_1(A)} P_{i-j}^+ M_{i-j} \bar{v}^{(i-j)}$$

$$\begin{aligned}
&= \sum_{j=1}^{n_f-1} \frac{\|F_{j+1}\|_\infty}{\sigma_n(F_1)} P_{i-j}^+ M_{i-j} \bar{v}^{(i-j)} \\
&\quad \times \left(1 + \frac{\sigma_1(A)}{i\sigma_n(F_1) - \sigma_1(A)} \right).
\end{aligned} \quad (26)$$

Note that $\lim_{i \rightarrow \infty} (\sigma_1(A)/i\sigma_n(F_1) - \sigma_1(A)) = 0$, and that there exists a positive real ϵ such that $|x_j| < d_x - \epsilon, j = 1, \dots, n$. Considering these factors together with the result shown in (26), we can establish a bound of $p_i(x)$ for $i > i^*$, with i^* being sufficiently large, as

$$\begin{aligned}
\|p_i(x)\| &\leq \|P_i^+ M_i (v^{(i)})^+\| \\
&\leq \|P_i^+ M_i \bar{v}^{(i)}\| \left(\max_{j=1, \dots, n} |x_j| \right)^i \\
&< \sum_{j=1}^{n_f-1} \frac{\|F_{j+1}\|_\infty}{\sigma_n(F_1)} \|P_{i-j}^+ M_{i-j} \bar{v}^{(i-j)}\| d_x^i \\
&\leq \sum_{j=1}^{n_f-1} \frac{\|F_{j+1}\|_\infty}{\sigma_n(F_1)} \|P_{i-j}^+ M_{i-j} \bar{v}^{(i-j)}\| d_x^j.
\end{aligned}$$

Define

$$\bar{p}_i = \|P_{i-1}^+ M_{i-1} \bar{v}^{(i)}\| d_x^i.$$

We then have that

$$\bar{p}_i < \sum_{j=1}^{n_f-1} \frac{\|F_{j+1}\|_\infty}{\sigma_n(F_1)} d_x^j \bar{p}_{i-j}. \quad (27)$$

Let us define a discrete-time system with state variable $z \in \mathbb{R}^{n_f-1}$ with

$$z(i) = [\bar{p}_{i-n_f+1}, \dots, \bar{p}_{i-1}, \bar{p}_i]^T.$$

From (27), we have

$$\bar{p}_{i+1} < \sum_{j=1}^{n_f-1} \frac{\|F_{j+1}\|_\infty}{\sigma_n(F_1)} d_x^j \bar{p}_{i-j+1}$$

and therefore we obtain that

$$z(i+1) \leq S z(i). \quad (28)$$

From (28), there exist positive definite matrices $U, W \in \mathbb{R}^{(n_f-1) \times (n_f-1)}$ such that

$$S^T U S - U = -W. \quad (29)$$

Let

$$J_k = \sum_{i=k}^{\infty} z^T(i) W z(i)$$

for an integer $k \geq \max\{n_f, n_h, i^*\}$. From (28) and (29), we have

$$\begin{aligned}
J_k &= \sum_{i=k}^{\infty} z^T(i) (U - S^T U S) z(i) \\
&\leq \sum_{i=k}^{\infty} z^T(k) (S^T)^{i-k} (U - S^T U S) S^{i-k} z(k) \\
&\leq \sum_{i=k}^{\infty} z^T(k) \\
&\quad \times [(S^T)^{i-k} U S^{i-k} - (S^T)^{i-k+1} U S^{i-k+1}] z(k) \\
&= z^T(k) (S^T)^{i-k} U S^{i-k} z(k).
\end{aligned}$$

With $\|h_i(x)\| < \bar{h}_i(x)$, we have

$$\sum_{i=k}^{\infty} \|h_i(x)\|^2 < \frac{1}{w_{n_f-1, n_f-1}} J_k$$

where w_{n_f-1, n_f-1} is the last diagonal element of W , and therefore we can conclude $h_i(x)$ converges in l_2 .

Fig. 1. Unmeasured state variable and its estimates from observers.

V. AN EXAMPLE

Consider a nonlinear system

$$\begin{aligned}\dot{x}_1 &= 2.1x_1 + x_2 + 0.3x_1^2 - 0.5x_1^3 \\ \dot{x}_2 &= -x_1 \\ y &= x_2.\end{aligned}\quad (30)$$

For this system, we have

$$\begin{aligned}F_1 &= \begin{bmatrix} 2.1 & 1 \\ -1 & 0 \end{bmatrix}, F_2 = \begin{bmatrix} 0.3 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \\ F_3 &= \begin{bmatrix} -0.5 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} \\ H_1 &= [0 \quad 1].\end{aligned}$$

With

$$A = \begin{bmatrix} -3 & 1 \\ -2 & 0 \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

it can be easily seen that Assumption 1 is satisfied as the eigenvalues of A are negative and those of F_1 are positive. Also it can be shown that $\{A, B\}$ and $\{F_1, B_1\}$ are controllable and observable, respectively, which ensures the invertibility of P_1 . Indeed, we obtain that

$$\begin{aligned}P_1 &= \begin{bmatrix} 0.1352 & 0.3104 \\ 0.3791 & 1.0665 \end{bmatrix} \\ P_2 &= \begin{bmatrix} -0.0081 & 0.0032 & -0.0042 \\ -0.0212 & 0.0083 & -0.0093 \end{bmatrix}.\end{aligned}$$

Based on Theorem 2, the convergence region is obtained as a parallelogram specified by the following:

$$\begin{aligned}|2.6491x_1 + 1.9334x_2| &< 0.3077 \\ |1.9334x_1 + 2.6491x_2| &< 0.3077.\end{aligned}\quad (31)$$

Observers in the form of (5) are implemented with $p(x) = P_1 v_1$ for the first-order approximation of $p(x)$ and then with $p = P_1 v_1 + P_2 v_2$ for the second-order approximation. The simulation results are shown in Figs. 1 and 2. The second-order approximation shows improvements over the first-order approximation.

Fig. 2. Enlargement of Fig. 1.

VI. CONCLUSION

In this paper, we have established an explicit expression on the domain of convergence of a series solution to the observer gain design for nonlinear dynamic systems in relation to nonlinearities in the original nonlinear system. This result is obtained by analyzing the relationship between the sizes of subsequent terms and then establishing a dynamic model in discrete-time. Further studies are needed on how this presented result can be used for determining the proper order of approximation and on possible domain of attraction of the observer errors. Note that the domain of attraction of the observer errors are different from the domain of the convergence of the power series for the observer gain design.

REFERENCES

- [1] F. Thau, "Observing the states of nonlinear dynamical systems," *Int. J. Control*, vol. 18, pp. 471–479, 1973.
- [2] S. Kou, D. Ellitt, and T. Tarn, "Exponential observers for nonlinear dynamical systems," *Inform. and Control*, vol. 29, pp. 204–216, 1975.
- [3] D. Bestie and M. Zeitz, "Canonical form observer design for non-linear time-variable systems," *Int. J. Control*, vol. 38, pp. 47–52, 1983.
- [4] X. Xia and W. Gao, "Nonlinear observer design by canonical form," *Int. J. Control*, vol. 47, pp. 1081–1100, 1988.
- [5] A. Krener and W. Kang, "Locally convergent nonlinear observers," *SIAM J. on Control and Optimiz.*, vol. 42, no. 1, pp. 155–177, 2003.
- [6] S. Lee, "Robust reduced order observer for lipschitz nonlinear systems," *IEICE Trans. Fundam. of Electron., Commun. and Comput. Sci.*, vol. E92-A, no. 6, pp. 1530–1534, 2009.
- [7] Z. Ding, "Differential stability and design of reduced-order observers for nonlinear systems," *IET Control Theory and Applic.*, vol. 5, no. 2, pp. 315–322, 2011.
- [8] A. J. Krener and A. Isidori, "Linearization by output injection and nonlinear observers," *Sys. Control Lett.*, vol. 3, pp. 47–52, 1983.
- [9] M. Kazantzis and C. Kravaris, "Nonlinear observer design using Lyapunov's auxiliary theorem," *Syst. and Control Lett.*, vol. 34, pp. 147–241, 1998.
- [10] A. Krener and M. Xiao, "Nonlinear observer design in the siegel domain," *SIAM J. on Control and Optimiz.*, vol. 41, no. 13, pp. 932–953, 2002.
- [11] C. Kravaris, "Modular design of discrete-time nonlinear observers for state and disturbance estimation," *Eur. J. Control*, vol. 15, no. 2, pp. 120–130, 2009.
- [12] N. Kazantzis, "Map invariance and the state reconstruction problem for nonlinear discrete-time systems," *Eur. J. Control*, vol. 15, no. 2, pp. 105–119, 2009.
- [13] I. Karafyllis and C. Kravaris, "From continuous-time design to sampled-data design of observers," *IEEE Trans. Autom. Control*, vol. 54, no. 9, pp. 2169–2174, Sep. 2009.

- [14] W. Kang, "Moving horizon numerical observers of nonlinear control systems," *IEEE Trans. Autom. Control*, vol. 51, no. 2, pp. 344–350, Feb. 2006.
- [15] M. Xiao, "The global existence of nonlinear observers with linear error dynamics: A topological point of view," *Syst. and Control Lett.*, vol. 55, no. 10, pp. 849–858, 2006.
- [16] Y. Shen, Y. Huang, and J. Gu, "Global finite-time observers for Lipschitz nonlinear systems," *IEEE Trans. Autom. Control*, vol. 56, no. 2, pp. 418–424, Feb. 2011.
- [17] J. Huang, *Nonlinear Output Regulation Theory and Applications*. Philadelphia, PA: SIAM, 2004.
- [18] R. A. Horn and C. A. Johnson, *Topics in Matrix Analysis*. Cambridge, U.K.: Cambridge Univ. Press, 1991.

Modeling and Control of Systems with Active Singularities Under Energy Constraints: Single- and Multi-Impact Sequences

Joseph Bentsman, Boris M. Miller, Evgeny Ya. Rubinovich, and
Sudip K. Mazumder

Abstract—A controller synthesis setting for systems with controlled singularities under incomplete information is extended to the cases characterized by a constraint on the total physical impulse as well as on the time interval between the adjacent engagement phases. The use of this setting for the optimal control law calculation is illustrated on the ball/racket system representative of a large class of controlled impact problems, where there are two separate bodies interacting through controlled viscoelastic force.

Index Terms—Control over observations, controlled singularities, energy constraints, impulse control, mechanical systems.

I. INTRODUCTION

Dynamical systems with active, or controlled, singularities is a new class of systems [1] characterized by the presence of *active, or controlled, constraints* capable of radically changing the attainability set of the post-impact system state. The engagement phase of the system with such constraint is termed *active singularity*, and the system motion in the domain of constraint violation—the *singular* motion phase. Based on these concepts, the corresponding rigorous technique for modeling and controller synthesis for systems with active singularities of elastic

type under full state accessibility was developed in [1]. [2] extended the setting of [1] to admit an incomplete observation in the singular motion phase. To accommodate the output feedback optimal controller synthesis in the singular phase, [2] introduced the novel dynamic mode of system interaction with the constraints, termed *temporal multi-impact*, characterized by the very short duration isolated set of temporally sequenced control signals. This mode was shown to vastly increase attainability set of the post-temporal-multi-impact system state in comparison to the single-impact mode considered in [1]. To adequately represent temporal multi-impact, [2] also introduced a new mode of system behavior—the *interlaced singular phase*—an engagement phase exhibited by the system under a temporal multi-impact, and generalized the framework of [1] to encompass this mode. [2] also presented an extensive example of calculation of the single temporal multi-impact control laws for the open loop and the observation-based optimal ball stopping in an elastic ball/racket system.

Practical realization of controlled singularities [1], [2] in a number of systems is, however, characterized by a constraint on the total physical impulse that accounts for the energy depletion during an engagement phase, as well as on the time interval between adjacent engagement phases to secure the energy recovery. The optimization problem that arises in this case is fundamentally different from the magnitude constraint one considered in [2]. Integrating the material given in [3] and [4] into a succinct presentation, this paper extends the framework of [1] and [2] to permit the design of the observations-based optimal control laws under energy constraints. The latter are shown on a rather non-trivial ball/racket system example to require the single-impact and/or the temporal multi-impact finite control sequences to attain the desired control objective. The applications that stand to benefit from controller design for this case include, among others, power systems under faults and ultra-fast microgrids [8], mobile sensor networks, impact actuators [10], and robotic manipulators [5]–[7], [9].

II. GENERAL PROBLEM STATEMENT

Motion in the Regular and the Singular Phases: Let the controlled dynamical system be described by the state vector $x(t) = (x_p(t), x_v(t))$, $x_p(t) \in R^n$, $x_v(t) \in R^n$, where vectors x_p and x_v are referred to as the sets of *generalized positions* and *generalized velocities*, respectively, and $t \in [0, T]$, where T is sufficiently large.

Suppose that system motion includes interaction with constraint that undergoes the elastic deformation parametrized by some coefficient $\mu > 0$, so that for finite μ the constraint would admit a system motion, although inhibited, within the domain occupied by it. Let the constraint-free domain be given by $\{(x_p, t) : G(x_p, t) > 0\}$ where $G : R^n \times [0, T] \rightarrow R$ is a sufficiently smooth function. Following [1], the system motions in the domain occupied by the constraint and in the constraint-free domain will be referred to as the *singular* and the *constraint-free* motion phases, respectively.

Let the system motion be described by a system

$$\begin{aligned} \dot{x}_p(t) &= F_p^r(x_p(t), x_v(t), t), \dot{x}_v(t) = F_v^r(x_p(t), x_v(t), u(t), t) \\ &+ \mu F_v^s(x_p(t), x_v(t), w_1^\mu(\xi, t), t, \mu) I\{t : G(x_p(t), t) \leq 0\} \\ &+ \mu F_v^{rs}(x_p(t), x_v(t), w_2^\mu(\xi, t), t, \mu) \end{aligned} \quad (1)$$

where $u(t) \in U \subset R^r$ is a generalized control in the constraint-free phase that is unbounded in R^r , but restricted in the integral sense: $\int_0^T |u(t)| dt \leq M < \infty$, $M = \text{const}$, $F_p^r(x_p, x_v, u, t)$ and $F_v^r(x_p, x_v, u, t)$ are the generalized forces in the constraint-free phase, $w_1^\mu(\xi, t)$ is an external control signal in the singular phase, $\mu F_v^s(x_p, x_v, w_1^\mu(\xi, t), t, \mu)$ is a generalized controlled force in the

Manuscript received March 07, 2010; revised September 02, 2010; accepted January 25, 2012. Date of publication February 03, 2012; date of current version June 22, 2012. This work was supported by the U.S. National Science Foundation under Grants ECS-0501407 and DMI-0900138, by Australian Research Council Grant DP0988685, and by the Russian Foundation for Basic Research under grants 10-08-01070-a, 10-01-00710. Recommended by Associate Editor H. Ito.

J. Bentsman is with Department of Mechanical Science and Engineering, University of Illinois at Urbana-Champaign, Urbana, IL 61801 USA (e-mail: jbentsma@illinois.edu).

B. M. Miller is with the School of Mathematical Sciences, Monash University, Clayton, 3800, Vic., Australia and also with the Institute for Information Transmission Problems, Moscow, Russia (e-mail: boris.miller@sci.monash.edu.au).

E. Y. Rubinovich is with the Institute of Control Sciences, Moscow 117997, Russia (e-mail: rubinvch@ipu.rssi.ru).

S. K. Mazumder is with the Department of Electrical and Computer Engineering, University of Illinois at Chicago, Chicago, IL 60607 USA (e-mail: mazumder@ece.uic.edu).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TAC.2012.2186401