NONLINEAR OBSERVER DESIGN IN THE SIEGEL DOMAIN THROUGH COORDINATE CHANGES[1]

Arthur J. Krener* and **MingQing Xiao****

* *Department of Mathematics, University of California, Davis, CA 95616/USA*
** *Department of Mathematics, Southern Illinois University, Carbondale, IL 62901/USA*

Abstract: We extend the method of Kazantzis and Kravaris (Kazantzis and Kravaris, 1998) for the design of an observer to a larger class of nonlinear systems. The extended method is applicable to any real analytic observable nonlinear system. It is based on the solution of a first-order nonlinear PDE. This solution yields a change of state coordinates which linearizes the error dynamics. Under very general conditions, the existence and uniqueness of the solution is proved. Siegel's theorem is obtained as a corollary. The technique is constructive and yields a method for constructing approximate solutions.

Keywords: Nonlinear systems; Nonlinear observers; Linearizable Error Dynamics; Output Injection; Siegel domains.

1. INTRODUCTION

We consider the problem of estimating the current state $x(t)$ of a nonlinear dynamical system, described by a system of first-order differential equations

$$\begin{aligned} \dot{x} &= f(x) \\ y &= h(x), \end{aligned} \tag{1}$$

from the past observations $y(s), s \leq t$. The vector fields $f : \mathbf{R}^n \to \mathbf{R}^n$, and $h : \mathbf{R}^n \to \mathbf{R}^p$ are assumed to be real analytic functions with $f(0) = 0$, $h(0) = 0$. One technique of constructing an observer is to find a nonlinear change of state and output coordinates which transforms the system (1) into a system with linear output map and linear dynamics driven by nonlinear output injection. The design of an observer for such systems is relatively easy (Krener and Respondek, 1985), (Krener and Isidori, 1983), (Bestle and Zeitz, 1983) and the error dynamics is linear in the transformed coordinates. Recently Kazantzis and Kravaris have proposed a simpler method (Kazantzis and Kravaris, 1998). One seeks a change of state coordinates $z = \theta(x)$ such that the dynamics of (1) is linear driven by nonlinear output injection

$$\dot{z} = Az - \beta(y), \tag{2}$$

where A is an $n \times n$ matrix and $\beta : \mathbf{R}^p \to \mathbf{R}^n$ is a real analytic vector field. One does not have to linearize the output map.

Such a θ must satisfy the following first-order partial differential equation:

$$\frac{\partial \theta}{\partial x}(x) f(x) = A\theta(x) - \beta(h(x)). \tag{3}$$

Kazantzis and Kravaris considered the restricted form of this problem where the output injection is linear, $\beta(y) = By$, but their method generalizes immediately to analytic β. There are advantages to using nonlinear β which we discuss later. They

[1] Research supported in part by AFOSR under F49620-95-1-0409 and in part by NSF under 9907998.

showed using a particular form of the Lyapunov Auxiliary Theorem (Liapunov, 1966) that (3) has a unique solution under certain assumptions.

Theorem (Liapunov, 1966). *Assume that* $f : \mathbf{R}^n \to \mathbf{R}^n, h : \mathbf{R}^n \to \mathbf{R}^p$ *and* $\beta : \mathbf{R}^p \to \mathbf{R}^n$ *are analytic vector fields with* $f(0) = 0$, $h(0) = 0$, $\beta(0) = 0$ *and* $F = \frac{\partial f}{\partial x}(0)$, $H = \frac{\partial h}{\partial x}(0)$, $B = \frac{\partial \beta}{\partial x}(0)$. *Let the eigenvalues of* F *be* $(\lambda_1, \dots, \lambda_n)$ *and the eigenvalues of* A *be* $(\mu_1, \dots, \mu_n)$. *If*

(1) 0 *does not lie in the convex hull of* $(\lambda_1, \dots, \lambda_n)$,
(2) there does not exist non-negative integers $m_1, m_2, \dots, m_n$ *not all zero such that* $\sum_{i=1}^{n} m_i \lambda_i = \mu_j$,

then the first-order PDE (3), with initial condition $\theta(0) = 0$, *admits a unique analytic solution* θ *in a neighborhood of* $x = 0$.

Based on above theorem, Kazantzis and Kravaris proposed a nonlinear observer design method (Kazantzis and Kravaris, 1998), where the state observer is constructed using the coordinate transformation $z = \theta(x)$ and the output injection $\beta(y)$.

Kazantzis and Kravaris Theorem (Kazantzis and Kravaris, 1998). *Assume that* f, h, θ, β *are as in the above theorem and additionally*

3. θ *is a local diffeomorphism,*
4. A *is Hurwitz*

then the local state observer for (1) given by

$$\dot{\hat{x}} = f(\hat{x}) - \left[\frac{\partial \theta}{\partial \hat{x}}(\hat{x})\right]^{-1} (\beta(y) - \beta(h(\hat{x}))). \quad (4)$$

has locally asymptotically stable error dynamics. In z *coordinates, the system is given by (2), the observer is*

$$\dot{\hat{z}} = A\hat{z} - \beta(y) \quad (5)$$

and the error $\tilde{z} = z - \hat{z}$ *dynamics is*

$$\dot{\tilde{z}} = A\tilde{z}. \quad (6)$$

One can show that if the conditions of this theorem hold then (H, F) is an observable pair and (A, B) is a controllable pair. On the other hand if (H, F) is an observable pair then one can choose an invertible T and B so that $A = (TF + BH)T^{-1}$ satisfies *2*, *3* and if the solution of (3) exists for some β such that $\beta(0) = 0$, $\frac{\partial \beta}{\partial x}(0) = B$, then θ is a local diffeomorphism. The size of the neighborhood of 0 on which θ is a diffeomorphism varies with the higher derivatives of β hence the advantage of allowing them to be different from zero.

The approach of Kazantzis and Kravaris has an advantage over that of Krener and Respondek (Krener and Respondek, 1985) and similar attempts to transform the dynamics and output map into observer form. The former uses the Lyapunov Auxiliary Theorem which depends on a nonresonant condition, assumption *2* above, while the latter depends on integrability conditions. The nonresonant condition is generically satisfied while the integrability conditions are generically not satisfied. However, assumption *1* of Kazantzis and Kravaris is quite restrictive, it requires the system to be locally asymptotically stable to the origin in either forward or reverse time. Assumption *1* requires that the eigenvalues of the linear part of $f(x)$ at the origin lie in the *Poincaré domain*, whose definition follows.

Definition 1. *An n-tuple* $\lambda = (\lambda_1, \dots, \lambda_n)$ *of complex numbers belongs to the Poincaré domain if the convex hull of* $(\lambda_1, \dots, \lambda_n)$ *does not contain zero. An n-tuple of complex numbers belongs to the Siegel domain if zero lies in the convex hull of* $(\lambda_1, \dots, \lambda_n)$.

Clearly, requiring the spectrum of F to be in the Poincaré domain rules out many interesting problems, including critical ones where there are eigenvalues on the imaginary axis, (Krener, 1994). In this paper we extend the observer design method of Kazantzis and Kravaris to the Siegel domain (Arnol'd, 1988). We start with a definition.

Definition 2 *Given an* $n \times n$ *matrix* F *with spectrum* $\sigma(F) = \lambda = (\lambda_1, \dots, \lambda_n)$ *and constants* $C > 0$, $\nu > 0$, *we say a complex number* μ *is of type* (C, ν) *with respect to* $\sigma(F)$ *if for any vector* $m = (m_1, m_2, \dots, m_n)$ *of nonnegative integers,* $|m| = \sum m_i > 0$, *we have*

$$|\mu - m \cdot \lambda| \geq \frac{C}{|m|^{\nu}}. \quad (7)$$

Now we are ready to state the main result of this paper.

Main Theorem. *Assume that* $f : \mathbf{R}^n \to \mathbf{R}^n, h : \mathbf{R}^n \to \mathbf{R}^p$ *and* $\beta : \mathbf{R}^p \to \mathbf{R}^n$ *are analytic vector fields with* $f(0) = 0$, $h(0) = 0$, $\beta(0) = 0$ *and* $F = \frac{\partial f}{\partial x}(0)$, $H = \frac{\partial h}{\partial x}(0)$, $B = \frac{\partial \beta}{\partial x}(0)$. *Suppose*

(1) there exists an invertible $n \times n$ *matrix* T *so that* $TFT^{-1} = A - BH$,
(2) there exists a $C > 0, \nu > 0$ *such that all the eigenvalues of* A *are of type* (C, ν) *w.r.t.* $\sigma(F)$.

Then there exists a unique analytic solution $z = \theta(x)$ *to the PDE (3) locally around* $x = 0$ *and* $\frac{\partial \theta}{\partial x}(0) = T$ *so* θ *is a local diffeomorphism.*

Note: Assumption *2* implies that the eigenvalues of A are distinct from those of F. We shall show the following: assumptions *1,2* imply that (H, F) is an observable pair. On the other hand, if (H, F) is an observable pair then one can let $T = I$ and set the spectrum of A arbitrarily by choice of B. Almost all complex numbers are of type (C, ν) w.r.t. $\sigma(F)$ so assumption *2* is hardly a restriction on A when (H, F) is an observable pair. If A is chosen to be Hurwitz, then the state estimator is given by (4) and the error dynamics is locally asymptotically stable as before.

The paper is organized as follows. Section 2.1 discusses the relationship between the linear part of the nonlinear system (1) and the terms of degree one of the solution (3). A unique formal solution of (3) is given in Section 2.2 and this is shown to be convergent in Section 2.3. We also show in Section 2.1 that (3) has an unique solution for any choice of the eigenvalues of A except for a set of zero measure in $\mathbf{C}^n$. An example is treated in Section 3.

Because of space limitations, we have not included proofs of most results given in section 2 below; they can be found in the full version of the paper, available from the authors.

2. SOLUTION OF THE PDE

2.1 *Terms of Degree One*

If we focus on the terms of degree one in (3), we obtain the equation

$$TF = AT - BH. \qquad (8)$$

We view this as a linear equation for T in terms of given F, H, A, B. Let us state some facts related to these matrices.

- **(A1)** Equation (8) admits a unique solution T if an only if the eigenvalues of F and A are distinct, that is $\sigma(F) \cap \sigma(A) = \emptyset$.
- **(A2)** Suppose $\sigma(F) \cap \sigma(A) = \emptyset$. If T is invertible, then (H, F) is observable and (A, B) is controllable.
- **(A3)** If T is an invertible solution to (8), then A is conjugate to F modified by output injection.
- **(A4)** If $\sigma(F) \cap \sigma(A) = \emptyset$ and A is conjugate to F modified by output injection, then there exists B such that the unique solution to (8) is invertible.

Loosely speaking, a complex number μ is of type (C, ν) with respect to $\sigma(F) = \lambda$ if $|\mu - m \cdot \lambda|$ is never zero and does not approach zero too fast as $|m| \to \infty$. If ν is large enough then the set of μ's which are of type (C, ν) for some $C > 0$ is dense in the complex plane.

Theorem 1. If $C > 0$ and $\nu > \frac{n}{2}$ then

$$\text{meas } \{\mu : \mu \text{ is not of type } (C, \nu)\} \leq k(n, \nu) C^2,$$

where $k(n, \nu)$ is a constant which depends only on n and ν.

If $\nu > \frac{n}{2}$ then the set of points which are not of type (C, ν) for any $C > 0$ is a set of zero measure.

2.2 *The Formal Solution of the PDE*

Assume the hypothesis of the Main Theorem hold, we show that there is an unique solution to the PDE (3) within the class of formal power series. It is convenient to assume that F and A are diagonal, the proof in the general case is similar but much messier. We expand the terms in power series

$$\begin{aligned} f(x) &= Fx + f^{[2]}(x) + f^{[3]}(x) + \ldots, \\ \beta(h(x)) &= BHx + \beta^{[2]}(x) + \beta^{[3]}(x) + \ldots, \\ \theta(x) &= Tx + \theta^{[2]}(x) + \theta^{[3]}(x) + \ldots, \end{aligned}$$

where $f^{[d]}$, $\beta^{[d]}$, and $\theta^{[d]}$ are homogeneous polynomial vector fields of degree d in x. The knowns are f, h, β, T and the unknowns are the higher degree terms $\theta^{[2]}, \theta^{[3]}, \ldots$. The linear terms satisfy (8) by assumption.

The degree d part of (3) is

$$\frac{\partial \theta^{[d]}}{\partial x}(x) F x - A\theta^{[d]}(x) = -\tilde{\beta}^{[d]}(x) \qquad (9)$$

where

$$\tilde{\beta}^{[d]}(x) = \beta^{[d]}(x) + T f^{[d]}(x) + \sum_{j=2}^{d-1} \frac{\partial \theta^{[j]}}{\partial x}(x) f^{[d+1-j]}(x). \qquad (10)$$

Let $\mathbf{e}^k$ denote the k^{th} unit vector in z space and $x^m = x_1^{m_1} \ldots x_n^{m_n}$. Then the above terms can be expanded as

$$\begin{aligned} \tilde{\beta}^{[d]}(x) &= \sum_{k=1}^{n} \sum_{|m|=d} \tilde{\beta}_{k,m} \mathbf{e}^k x^m, \\ \theta^{[d]}(x) &= \sum_{k=1}^{n} \sum_{|m|=d} \theta_{k,m} \mathbf{e}^k x^m \end{aligned}$$

and we obtain the equations

$$(\mu_k - m \cdot \lambda)\theta_{k,m} = \tilde{\beta}_{k,m}. \qquad (11)$$

These equations have unique solutions because $m \cdot \lambda - \mu_k \neq 0$.

The formal approach yields a method for constructing an observer with approximately linear error dynamics. Start by choosing an T, A, B satisfying the linear equation (8). Then successively solve (9) up to some degree d. At each step $\beta^{[j]}$ can be chosen to make $\theta^{[j]}$ smaller and thereby try to keep $\theta(x)$ close to its globally invertible linear part Tx. The approximate solution

$$\theta(x) = Tx + \theta^{[2]}(x) + \theta^{[3]}(x) + \ldots + \theta^{[d]}(x)$$
$$\beta(y) = By + \beta^{[2]}(y) + \beta^{[3]}(y) + \ldots + \beta^{[d]}(y)$$

transforms the system (1) into

$$\dot{z} = Az - \beta(y) + O(x)^{d+1}$$

so the observer (4) has approximately linearizable error dynamics. The error is $O(x, \hat{x})^{d+1}$. When implementing the method, the matrices F, A need not be diagonal but this makes solving (9) very straight-forward.

2.3 *Convergence of the Formal Solution*

Let $|x| = \max\{|x_1|, \ldots, |x_n|\}$. We write

$$f(x) = Fx + \bar{f}(x)$$
$$\beta(y) = BHx + \bar{\beta}(x)$$

where $AT - TF = BH$. We first show that the following sequence of PDEs

$$A\theta_2(x) - \frac{\partial}{\partial x}\theta_2(x)Fx = T\bar{f}(x) + \bar{\beta}(x)$$
$$A\theta_k(x) - \frac{\partial}{\partial x}\theta_k(x)Fx = \frac{\partial}{\partial x}\theta_{k-1}(x)\bar{f}(x)$$

admits a sequence of analytical solutions $\theta_2(x)$, $\theta_3(x)$, ...in some neighborhood of the origin. Then we show that the sum

$$Tx + \theta_2(x) + \theta_3(x) + \ldots$$

converges to an analytic function which solves (3).

We define a positive real function $b_k : [0, 1) \to [0, \infty)$ to be

$$b_k(q) := \max_{d \in \mathbf{Z}_+, d \geq k} \left[C^{-1} d^\nu q^{\frac{d}{2}} \right],$$

where $C > 0$ and $\nu > 0$ are given.

Theorem 2. Let $P(x)$ be a real analytic function in $|x| < r$ with $P(0) = 0$. Suppose all of the eigenvalues of A are of type (C, ν) with respect to $\sigma(F)$. Then the first order PDE

$$A\theta(x) - \frac{\partial\theta(x)}{\partial x}Fx = P(x) \tag{12}$$

admits a unique analytic solution $\theta(x)$ in $|x| < r$ with $\theta(0) = 0$.

Following from Theorem 2, we immediately have

Corollary 1. *Suppose all of the eigenvalues of A are of type (C, ν) w.r.t. $\sigma(F)$. The following PDEs*

$$A\theta_2(x) - \frac{\partial\theta_2}{\partial x}(x)Fx = T\bar{f}(x) + \bar{\beta}(x), \tag{13}$$
$$A\theta_k(x) - \frac{\partial\theta_k}{\partial x}(x)Fx = \frac{\partial\theta_{k-1}}{\partial x}(x)\bar{f}(x), \tag{14}$$

admit analytic solutions in $|x| < r$, satisfying $\theta_2(0) = 0$, and $\theta_k(0) = 0$, for $k = 3, 4, \ldots$.

The next step is to prove that

$$\theta_2(x) + \theta_3(x) + \ldots + \theta_k(x) + \ldots$$

converges near the origin and solves the PDE (3).

Theorem 3. There exists $0 < r_1 < r$ such that if $P(x)$ is analytic in $|x| < r_1$ where $|P(x)| \leq N$ then

$$\left| \frac{\partial P}{\partial x}(x)\bar{f}(x) \right| \leq N$$

in $|x| < r_1$. Note that r_1 does not depend on N.

In the definition of type (C, ν), without loss of generality we can assume that ν is a positive integer since if ν is not, we can replace it by a larger integer.

Theorem 4. Let $r_2 := r_1/n$, where r_1 is given in Lemma 6. Let $\theta_k(x)$ be the solution of

$$A\theta_k(x) - \frac{\partial\theta_k}{\partial x}(x)Fx = \frac{\partial\theta_{k-1}}{\partial x}(x)\bar{f}(x).$$

Then if $|\theta_{k-1}(x)| \leq N$ for $|x| < r_2$, we have

$$|\theta_k(x)| \leq \frac{NP(|x_1| + |x_2| + \cdots + |x_n|)}{C\left(r_1 - (|x_1| + |x_2| + \cdots + |x_n|)\right)^{\nu+1}},$$

for $|x| < r_2$, where P is a polynomial of degree ν with coefficients depending only on r_1.

Let $r_3 := r_2/2$ and

$$\hat{N} := \max_{|x| \leq r_3} \frac{P(|x_1| + \cdots + |x_n|)}{C(r_1 - (|x_1| + \cdots + |x_n|))^{\nu+1}}.$$

and

$$M := \max_{|x| \leq r} \sum_{d=2}^{\infty} \left(|\beta^{[d]}(x)| + |Tf^{[d]}(x)| \right).$$

Theorem 5. Let $\theta_k(x)$ be the solution of

$$A\theta_k(x) - \frac{\partial\theta_k}{\partial x}(x)Fx = \frac{\partial\theta_{k-1}}{\partial x}(x)\bar{f}(x) \quad \theta_k(0) = 0.$$

Then for any $|x| \leq qr_3$ with $0 < q < 1$ we have

$$|\theta_k(x)| \leq b_k(q)\hat{N}^{k-2}M.$$

Corollary 2. *When q is small enough, the series*

$$\theta_2(x) + \theta_3(x) + \ldots + \theta_k(x) + \ldots$$

converges in $|x| \leq qr_3$, where $\theta_d(x)$ for $d = 2, 3, \ldots$ is the solution of (14).

From Corollary 2, we know that series

$$\theta_2(x) + \theta_3(x) + \ldots + \theta_d(x) + \ldots \qquad (15)$$

defines an analytic function in $|x| \leq qr_3$.

(Siegel's Theorem) *Assume that $f : \mathbf{R}^n \to \mathbf{R}^n$ is an analytic vector field with $f(0) = 0, \frac{\partial f}{\partial x}(0) = F$. Suppose, for some $C > 0, \nu > 0$, the eigenvalues of F are of type (C, ν) with respect to $\sigma(F)$. Then there is an analytic solution in some neighborhood of the origin of the first-order PDE:*

$$\frac{\partial \theta}{\partial x}(x)f(x) = F\theta(x)$$

with initial condition $\theta(0) = 0$. Moreover $z = \theta(x)$ is a local analytic diffeomorphism around $x = 0$ which transforms the differential equation

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

into its linear part

$$\dot{z} = Fz$$

Proof: Apply the Main Theorem with $\beta = 0$ and $T = I$. □

3. AN EXAMPLE

As discussed in the introduction, Kazantzis and Kravaris (Kazantzis and Kravaris, 1998) considered only linear output injection but there are distinct advantages to considering *nonlinear output injection* $\beta(y)$. It is desirable that θ be a diffeomorphism over as large a range as possible because this is the domain of convergence of the observer. Nonlinear output injection can make θ a global diffeomorphism. To see this, we consider a Van der Pol oscillator:

$$\ddot{x} + (x^2 - 1)\dot{x} + x = 0$$
$$y = x$$

which is equivalent to the planar system

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} - \begin{bmatrix} 0 \\ x_1^2 x_2 \end{bmatrix}$$
$$y = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$$

Ly letting $x_1 = x, x_2 = \dot{x}$. Now we have

$$f(x) = \begin{bmatrix} x_2 \\ -x_1 + x_2 - x_1^2 x_2 \end{bmatrix}, \quad h(x) = x_1,$$
$$F = \begin{bmatrix} 0 & 1 \\ -1 & 1 \end{bmatrix}, \qquad H = \begin{bmatrix} 1 & 0 \end{bmatrix}.$$

We look for a nonlinear coordinate transformation $z = \theta(x)$ such that in the new coordinates z, the system can be described in the form

$$\dot{z} = Az - \beta(y).$$

Let us choose A and β to be

$$A = \begin{bmatrix} b_1 & 1 \\ b_2 - 1 & 1 \end{bmatrix}, \qquad \beta(y) = \begin{bmatrix} b_1 y + \frac{y^3}{3} \\ b_2 y + \frac{y^3}{3} \end{bmatrix}.$$

where b_1, b_2 are constants such that $1 + b_1 < 0, b_1 - b_2 + 1 > 0$. Clearly, A is stable since $trace(A) = 1 + b_1 < 0$ and $det(A) = b_1 - b_2 + 1 > 0$. Moreover $A = F + BH$ with $B = [b_1, b_2]'$. The solution of (3) in this case is given by

$$\theta(x) = \begin{bmatrix} x_1 \\ x_2 + \frac{x_1^3}{3} \end{bmatrix}.$$

Note that θ is *globally invertible* on $\mathbf{R}^2$.

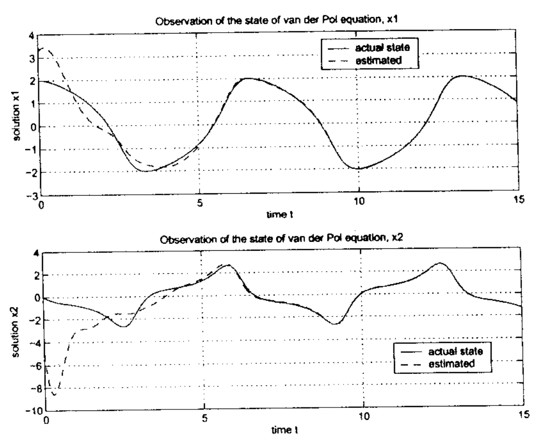

Fig. 1. Observation of Van der Pol Oscillator

4. REFERENCES

Arnol'd, V. I. (1988). *Geometrical Methods in the Theory of Ordinary Differential Equations.* Springer-Verlag, Berlin. New York.

Bestle, D. and M. Zeitz (1983). Canonical form observer design for nonlinear time-variable systems. *Int. J. Control* **38**, 419–431.

Kazantzis, N. and C. Kravaris (1998). Nonlinear observer design using lyapunov's auxiliary theorem. *Systems & Control Letters* **34**, 241–247.

Krener, A. J. (1994). Nonlinear stabilizability and detectability. In: *In Systems and Networks: Mathematical Theory and Applications* (R. Mennicken U. Helmke and J. Saurer, Eds.). pp. 231–250. Akademie Verlag. Berlin.

Krener, A. J. and A. Isidori (1983). Linearization by output injection and nonlinear observers. *Systems & Control Letters* **3**, 47–52.

Krener, A. J. and W. Respondek (1985). Nonlinear observers with linearizable error dynamics. *SIAM J. Control and Optimization* **23**, 197–216.

Liapunov, A. M. (1966). *Stability of Motion.* Academic Press. New York and London.