Abstract: In this paper we consider the problem of observer design for dynamical systems with scalar output by linearization of the error dynamics via coordinate change, output injection, and time scaling. We present necessary and sufficient conditions which guarantee the existence of a coordinate change and/or output dependent time scaling such that in the new coordinates and with respect to the new time the system has linear error dynamics. Copyright © 2001 IFAC

Keywords: time scaling, observers, output injection, linearization

1. INTRODUCTION

An important control problem studied extensively is that of observer design for a dynamical system with output. A typical approach to this problem is to find a dynamical system (observer) coupled with the observed system by output injection in such a way that the overall system possesses an invariant asymptotically stable set of a specific structure. Although a solution to the observer design problem in its full generality is not known yet, it is clear that the problem statement is coordinate independent and invariant with respect to time scaling and therefore it is natural to seek conditions, ensuring the existence of an observer, that would also be coordinate independent and invariant with respect to time scaling, similarly to the corresponding properties of asymptotic stability. The problem of transforming a system to observer form via change of state and output coordinates has been intensively studied during the last twenty years, see e.g. (Bestle and Zeitz, 1983), (Gauthier et al, 1992), (Krener and Isidori, 1983), (Krener and Respondek, 1985), (Plestan and Glumineau, 1997), (Xia and Gao, 1989). However, the authors are unaware, except for the reference (Moya et al., 2001), of attempts to enlarge the class of systems admitting an observer exploiting the additional freedom of introducing possible time scaling. In this paper we are going to address this problem. A dual problem of linearization of dynamics with inputs has been considered in (Sampei and Furuta, 1986; Respondek, 1998).
Expressing a system in “physical” coordinates is very natural but sometimes it may complicate the observer design based on the second Lyapunov method. At the same time for the system written in other coordinates that design can be much easier to perform. To be more precise, consider dynamics with output given by the following equations:

\[
\begin{align*}
\dot{\xi}_1 &= s(\xi_1)\xi_2 \\
\dot{\xi}_2 &= s(\xi_1)\xi_3 \\
&\vdots \\
\dot{\xi}_n &= s(\xi_1)(a_1\xi_1 + a_2\xi_2 + \ldots + a_n\xi_n) \\
y &= \xi_1,
\end{align*}
\]

where \(\xi = (\xi_1, \ldots, \xi_n)\) is the state, \(y\) is the scalar output, and \(s(\xi_1)\) is some nonvanishing positive real-valued function. In this case it is possible to linearize the dynamics (1) via time scaling of the form \(d\tau = s(\xi_1(t))dt\). For the linear system written with respect to the new time \(\tau\), the observer design problem can be easily solved using linear techniques. Therefore it is interesting to find conditions which guarantee that a nonlinear system with scalar output

\[
\dot{x} = f(x), \quad y = h(x)
\]
can be put in the form (1).

The next problem which we are going to address in this paper is to find a coordinate change, a time scaling and output injection to transform the system into a form linearized by output injection with the output \(y = x_1\).

It is worth mentioning that in a recent paper (Moya et al., 2001) it was reported that time scaling can significantly simplify the controller design for so-called reaction systems. In particular, in (Moya et al., 2001) it was shown that the model of a reaction system can be written in a coordinate system suitable for time scaling that significantly simplifies the controller/observer design.

In this paper we consider only the case of dynamical systems with scalar output. The more general case of dynamical systems with multiple outputs (and inputs) will be reported elsewhere.

The paper is organized as follows. The problem statement is formulated in Section 2. Section 3 contains the main results. An illustrative example is presented in Section 4. In the last section we draw some conclusions.

2. PROBLEM STATEMENT

Consider nonlinear observed dynamics of the form

\[
\Sigma : \dot{x} = \frac{dx}{dt} = f(x), \quad y = h(x),
\]

where \(x(\cdot) \in \mathbb{R}^n\), and \(y(\cdot) \in \mathbb{R}\) is the observation.

In this paper we will deal with the two following questions. Question 1: when do there exist a (local) diffeomorphism \(z = \varphi(x)\) and a time scaling of the form \(d\tau = s(h(x(t)))dt\), where \(s\) is a nonvanishing real-valued function, such that \(\Sigma\) becomes

\[
\frac{dz}{d\tau} = Az, \quad y = Cz,
\]

where the pair \((A, C)\) is observable? Question 2: when do there exist a (local) diffeomorphism \(z = \varphi(x)\) and a time scaling of the form \(d\tau = s(h(x(t)))dt\), where \(s\) is a nonvanishing real-valued function, such that \(\Sigma\) becomes

\[
\frac{dz}{d\tau} = Az + \gamma(y), \quad y = Cz,
\]

where the pair \((A, C)\) is observable and \(\gamma\) is a vector field whose components depend on \(y = Cz\) only? To answer the first question means to characterize nonlinear dynamics that are linearizable via
a diffeomorphism and an output dependent time scaling, while to answer the second question means to characterize dynamics that are linearizable via a diffeomorphism, an output dependent time scaling, and an output injection.

It is obvious that the first question is equivalent to the following one: when is \( \Sigma \) (locally) equivalent under a diffeomorphism \( z = \varphi(x) \) to
\[
\frac{dz}{dt} = s(y)(Az + Ky), \quad y = Cz = z_1,
\]
where
\[
A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad K = \begin{pmatrix} k_1 \\ \vdots \\ \vdots \\ k_n \end{pmatrix},
\]
and
\[
C = \begin{pmatrix} 1, 0, \ldots, 0 \end{pmatrix}
\]
and \( s(y) \) is a nonvanishing real-valued function?

The second question is obviously equivalent to the following one: when is \( \Sigma \) (locally) equivalent under a diffeomorphism \( z = \varphi(x) \) to
\[
\frac{dz}{dt} = s(y)(Az + \gamma(y)), \quad y = z_1 = Cz,
\]
where
\[
A = \begin{pmatrix} 0 & 1 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 \end{pmatrix}, \quad \gamma(z_1) = \begin{pmatrix} \gamma_1(z_1) \\ \vdots \\ \vdots \\ \gamma_n(z_1) \end{pmatrix},
\]
and
\[
C = \begin{pmatrix} 1, 0, \ldots, 0 \end{pmatrix}
\]
and \( s(y) \) is a nonvanishing real-valued function?

A motivation for our study is the following immediate observation. If \( \Sigma \) is equivalent, via a change of coordinates, to (3), then we can construct the following observer (compare e.g. (Krener and Isidori, 1983), (Krener and Respondek, 1985))
\[
\dot{\hat{z}} = s(y)(A\hat{z} + \gamma(y)) + L(y - C\hat{z})
\]
yielding the error \( e = \hat{z} - z \) that satisfies
\[
\dot{e} = s(y)(A + LC)e
\]
and thus gives the linear equation
\[
\frac{de}{d\tau} = (A + LC)e,
\]
with respect to the new time \( \tau \).

We will work around a fixed initial condition \( x_0 \in \mathbb{R}^n \) and we will assume that the diffeomorphism \( \varphi \) satisfies \( \varphi(x_0) = 0 \). Introduce the following notation
\[
\mathcal{H}^l = \text{span} \{ dh, dL_fh, \ldots, dL_f^{l-1}h \}.
\]

Clearly, a necessary condition for \( \Sigma \) to be equivalent to one of the above discussed forms is the following local observability rank condition (see e.g. (Isidori, 1989) and (Nijmeijer and Van der Schaft, 1990))
\[
\dim \mathcal{H}^n(x_0) = n
\]
and we will assume this throughout the paper.

3. MAIN RESULTS

Following (Krener and Isidori, 1983) (see also (Krener and Respondek, 1985)) define a vector field \( g \) by
\[
L_g L_f^j h = \begin{cases} 0 & \text{if } 0 \leq j \leq n - 2 \\ 1 & \text{if } j = n - 1. \end{cases}
\]

For \( j \geq 2 \) we put \( l_j = \frac{(j-1)j}{2} + 1 \).

In order to avoid the trivial case, we will assume throughout \( n \geq 2 \). We have the following results.

**Theorem 1.** \( \Sigma \) is, locally around \( x_0 \), equivalent under a diffeomorphism \( z = \varphi(x) \) to the system (2) if and only if in a neighborhood of \( x_0 \) it satisfies
\[
\begin{align*}
\text{(i)} & \quad dL_g L_f^j h = l_j \lambda dL_f h \mod \mathcal{H}^l, \quad \text{for some smooth function } \lambda; \\
\text{(ii)} & \quad [\text{ad}_{\tilde{g}} f, \text{ad}_{\tilde{g}} g] = 0, \quad \text{for } 0 \leq i < j \leq n, \quad \text{where } \tilde{f} = \frac{j}{2} f, \quad \tilde{g} = s^{n-1} g, \quad \text{and } s = \exp \sigma, \quad \text{with } \sigma \text{ being a solution of } \\
L_{\text{ad}_{\tilde{g}} g} \sigma &= \begin{cases} 0 & \text{if } 0 \leq j \leq n - 2 \\ (-1)^{n-1} \lambda & \text{if } j = n - 1. \end{cases}
\end{align*}
\]

**Theorem 2.** \( \Sigma \) is, locally around \( x_0 \), equivalent under a diffeomorphism \( z = \varphi(x) \) to the system (3) if and only if in a neighborhood of \( x_0 \) it satisfies the conditions (i) and (ii) of Theorem 1 with \( i \) and \( j \) in item (ii) satisfying \( 0 \leq i < j \leq n - 1 \)

**Remark 1.** Notice that the system
\[
L_{\text{ad}_{\tilde{g}} g} \sigma &= \begin{cases} 0 & \text{if } 0 \leq j \leq n - 2 \\ (-1)^{n-1} \lambda & \text{if } j = n - 1. \end{cases}
\]
of first order partial differential equations on the state space \( \mathbb{R}^n \) is actually a first order ordinary
Proof of Theorem 1. Sufficiency. From the definition of \( g \), it follows that \( \sigma \), which is a solution of
\[
L_{ad^j g}^\sigma = \begin{cases} 
0 & \text{if } 0 \leq j \leq n-2 \\
(-1)^{n-1} \lambda & \text{if } j = n-1,
\end{cases}
\]
satisfies \( d\sigma \in \mathcal{H}^1 \). Therefore
\[
ad^j \tilde{g} = s^{n-j-1} ad^j g \mod \text{span} \{ g, \ldots, ad^{l-1} g \},
\]
for any \( 0 \leq j \leq n-1 \). We thus conclude that
\[
L_{ad^j g}^\sigma = \begin{cases} 
0 & \text{if } 0 \leq j \leq n-2 \\
(-1)^{n-1} \lambda & \text{if } j = n-1.
\end{cases}
\]
(5)
It is well known (see e.g. (Nijmeijer and Van der Shaft, 1990) and (Isidori, 1989)) that (ii) implies that we can find a local diffeomorphism \( z = \varphi(x) \) such that \( \tilde{g} = \frac{\partial}{\partial z_n} \) and
\[
\dot{f}(z) = Az + Kz_1,
\]
where
\[
A = \begin{pmatrix} 0 & 1 & \ldots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \vdots & \ddots & 1 & \vdots \\ 0 & \ldots & 0 & 0 \end{pmatrix},
\]
and
\[
K = \begin{pmatrix} k_1 \\ \vdots \\ \vdots \\ k_n \end{pmatrix},
\]
which simply means that \( z = \varphi(x) \) linearizes the control system \( \dot{x} = f(x) + \tilde{g}(x)u \). Moreover, by (5) we have that in \( z \)-coordinates, \( h = z_1 \). Since \( f = s(z_1)\bar{f} \), it follows that in \( z \)-coordinates \( \Sigma \) reads as
\[
\frac{dz}{dt} = s(z_1)(Az + Kz_1), \quad y = z_1.
\]
\[\text{Necessity. Assume that there exists } z = \varphi(x) \text{ bringing } \Sigma \text{ into}
\]
\[
\dot{z}_1 = s(z_1)(z_2 + k_1 z_1), \quad y = z_1
\]
\[
\vdots
\]
\[
\dot{z}_{n-1} = s(z_1)(z_n + k_{n-1} z_1)
\]
\[
\dot{z}_n = s(z_1)k_n z_1.
\]
We have \( h = z_1 \), \( L_f h = s(z_2 + k_1 z_1) \), and \( L^2_f h = s^2 z_3 + s's^j \dot{z}_2 z_j \), where \( \forall m \) stands for the derivative with respect to \( z_1 \). It is straightforward to prove by an induction argument that
\[
L_f^j h = s^j z_{j+1} + l_j s^j \dot{z}_2 z_j + z_j a_j(\bar{z}_1) + b_j(z_1, \ldots, z_{j-1}).
\]
It thus follows that
\[
g = \frac{1}{s^n - \frac{\partial}{\partial z_n}} \text{ and that}
\]
\[
L_f^j h = s^j k_n z_1 + l_n s^j \dot{z}_2 z_n + z_n a_n(\bar{z}_1) + b_n(z_1, \ldots, z_{n-1}).
\]
Hence \( L_f L_f^j h = s a_n + l_n s' \dot{z}_2 + a_n(z_1) \) and
\[
dL_f L_f^j h = l_n s' \dot{z}_2 \mod \mathcal{H}_1,
\]
which gives \( \lambda = \frac{s'}{s} \). The system
\[
L_{ad^j g}^\sigma = \begin{cases} 
0 & \text{if } 0 \leq j \leq n-2 \\
(-1)^{n-1} \lambda & \text{if } j = n-1,
\end{cases}
\]
is thus the ordinary differential equation
\[
\sigma'(z) = \frac{s'(z_1)}{s(z_1)} = (\log s(z_1))'.
\]
(6)
We have \( \sigma(z_1) = \log s(z_1) + d \), where \( d \) is a constant, which gives a 1-parameter family of solutions \( s_c(z_1) = c \exp \sigma(z_1) \), where \( c \in \mathbb{R}, c \neq 0 \), is a multiplicative constant. It is clear that \( \bar{f} = \frac{1}{s} f = \frac{1}{s} Az \) and \( \bar{g} = (s_c)^{n-1} g = c^{n-1} \frac{\partial}{\partial z_n} \) satisfy (ii), for \( 1 \leq i < j \leq n \), which, actually, is another way of expressing the fact that \( \dot{z} = f(z) + \tilde{g}(z)u = \frac{1}{s} Az + c^{n-1} \frac{\partial}{\partial z_n} u \) is a linear system for any \( c \in \mathbb{R}, c \neq 0 \). \[\square\]

Proof of Theorem 2. The proof follows the same line. The only difference being that the commutation relation (ii), satisfied for \( 1 \leq i < j \leq n-1 \), is a necessary and sufficient condition for the observed dynamics \( \dot{x} = f(x) \) and \( y = h(x) \) to be equivalent to the nonlinear observer form \( \dot{z} = Az + \gamma(Cz), \quad y = Cz \) (see (Krener and Isidori, 1983)). \[\square\]

Notice that in both cases \( \sigma \) is calculated via (6) up to an additive constant, which means that \( s = \exp \sigma \) is calculated up to a multiplicative
constant. This is in agreement with the obvious observation that if a time rescaling \( \frac{d\tau}{d\tau} = s(z_1) \) works then any rescaling \( \frac{d\tau}{d\tau} = cs(z_1) \), where \( c \neq 0 \), works as well.

4. EXAMPLE

Consider the following system

\[
\begin{align*}
\dot{x}_1 &= x_2 \\
\dot{x}_2 &= a e^{x_1} x_1 + b e^{x_1} x_2 + x_2^2 \\
y &= x_1,
\end{align*}
\]

where \( a, b \in \mathbb{R} \). Obviously, this system is globally observable and satisfies the local observability rank condition everywhere, in particular, around the origin \( x_1 = x_2 = 0 \). Equation (4) takes the form

\[ L_y h = 0 \quad L_y L_f h = 1 \]

and has the solution \( g = \frac{\partial}{\partial x_2} \). Our next step is to verify condition (i) and to find \( \lambda \). For the second order system (7) we have \( l_2 = 2, L_f h = x_2 \) and \( L_y L_f h = b e^{x_1} + 2 x_2 \). Thus the condition (i) is equivalent to

\[ b e^{x_1} dx_1 + 2 dx_2 = 2 \lambda dx_1 \]

and is satisfied for \( \lambda = 1 \).

The next step is to find \( \sigma \). Simple calculations give

\[ ad_f g = g = \frac{\partial}{\partial x_2}, \quad ad_f g = [f, g] = - \frac{\partial}{\partial x_1} - (2x_2 + be^{x_1}) \frac{\partial}{\partial x_2} \]

and therefore \( \sigma \) is a solution of the following system of equations

\[ \frac{\partial \sigma}{\partial x_1} = \lambda = 1, \quad \frac{\partial \sigma}{\partial x_2} = 0 \]

which has a solution \( \sigma = x_1 + d \) where \( d \) is an arbitrary constant. Take \( d = 0 \), then \( \tilde{f} = e^{-x_1} f, \quad \tilde{g} = e^{x_1} g \).

Next we should check condition (ii) of Theorem 1. In view of Remark 2 it is sufficient to check two conditions

\[ [\tilde{g}, [\tilde{f}, g]] = 0 \]

and

\[ [\tilde{g}, ad_f^2 \tilde{g}] = 0 \]

Simple calculations give

\[ [\tilde{f}, \tilde{g}] = - \frac{\partial}{\partial x_1} - (be^{x_1} + x_2) \frac{\partial}{\partial x_2} \]

and therefore the condition (8) is satisfied.

To check the condition (9) one has to verify that

\[ ad_f^2 \tilde{g} = b \frac{\partial}{\partial x_1} + ((a + b^2)e^{x_1} + bx_2) \frac{\partial}{\partial x_2} \]

and

\[ ad_f^3 \tilde{g} = -(a + b^2) \frac{\partial}{\partial x_1} - ((2ab + b^3)e^{x_1} + (a + b^2)x_2) \frac{\partial}{\partial x_2} \]

Thus the condition (9) is satisfied as well, and according to Theorem 1 there is a coordinate change which transforms (7) to a form (2).

Indeed, it is not difficult to verify that the coordinate change

\[ \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ e^{-x_1} x_2 - bx_1 \end{pmatrix} \]

transforms the system (7) to the following one

\[ \begin{align*}
\dot{z}_1 &= e^{x_1} (z_2 + bz_1) \\
\dot{z}_2 &= e^{x_1} z_1 \\
y &= z_1
\end{align*} \]

5. CONCLUSIONS

In this paper we considered the problem of transforming, via coordinate change and time scaling, dynamics with output to a form which admits an observer with a linear error dynamics. The class of admissible time scalings used in this paper is given by equation

\[ dr = s(y(t)) dt, \]

where \( s \) is a real-valued positive function which depends only on the output of the observed dynamical system. This condition is necessary for practical implementation of an observer designed by the proposed technique. Moreover, designing the observer it should be additionally verified that \( t \mapsto \tau \) is a diffeomorphism \( [0, \infty) \to [0, \infty) \), that is, \( s(y(t)) \) is a nonvanishing function of time defined everywhere on \( [0, \infty) \).

We proposed necessary and sufficient conditions ensuring the existence of an appropriate (local) coordinate change and/or output-dependent time scaling such that in the new coordinates and with respect to the new time the system has one of the two canonical forms (either linear or linear up to a nonlinear additive output injection) for which the observer design problem can be easily solved.
6. REFERENCES


