131 (2006) MATHEMATICA BOHEMICA No. 4, 411–418

OBSERVABILITY OF NONLINEAR SYSTEMS

H. W. Knobloch, Würzburg

(Received December 9, 2005)

Dedicated to Prof. J. Kurzweil on the occasion of his 80th birthday

Abstract. Observability of a general nonlinear system—given in terms of an ODE $\dot{x} = f(x)$ and an output map $y = c(x)$—is defined as in linear system theory (i.e. if $f(x) = Ax$ and $c(x) = Cx$). In contrast to standard treatment of the subject we present a criterion for observability which is not a generalization of a known linear test. It is obtained by evaluation of "approximate first integrals". This concept is borrowed from nonlinear control theory where it appears under the label "Dissipation Inequality" and serves as a link with Hamilton-Jacobi theory.

Keywords: ordinary differential equations, observability

MSC 2000: 34A34, 93B07

1. Introduction

Observability is a property of a system given by two data: A differential equation and an output (i.e. one or several functions of the state, cf. Sec. 2 for details). It is concerned with the problem of reconstructing the full trajectory from measured output. The reader finds in Sec. 2 of this paper a short review of the basic facts related to observability. A new angle on the subject is then presented based on a source of information which so far seems not to have been exploited: The study of what one may call "approximate first integrals" and which we call—for reasons of agreement with existing definitions—"dissipation equalities" (DE's). A careful explanation is provided in Sec. 3, the origin is the control theoretic concept of "dissipation inequality", see [2] for details. Thereby this paper is a further example of the fruitful interrelation between the fields of ordinary differential equations and control theory which is an important aspect of the work of Jaroslav Kurzweil.

2. Observability

Starting point is the standard state-to-output scenario of control theory which is defined by writing down a dynamic law for the state:

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

and a mapping

$$x \mapsto y = c(x). \tag{2}$$

A system given by (1), (2) is called observable if the following statement is true: Let $x_1(t), x_2(t)$ be two solutions of (1) such that

$$y_1(t) := c(x_1(t)) = c(x_2(t)) =: y_2(t) \tag{3}$$

on some interval $[0, \delta]$, $\delta > 0$, then $x_1(t) = x_2(t)$ for all t. In order to avoid discussions which have nothing to do with the main objective of this paper, we assume sufficient smoothness of f, c as functions of x.

The standard approach to observability questions is based on the fact that together with (3) all time derivatives of y_1, y_2 agree on $[0, \delta]$. These time derivatives can be written as functions of x and these functions are the repeated Lie-derivatives of c which are defined as follows:

$$\begin{aligned} (L_f c)(x) &= c_x(x) f(x), \\ L_f^{(\nu)} c &= L_f(L_f^{(\nu-1)} c), \end{aligned} \tag{4}$$

$\nu \geqslant 2$. Bringing Lie-derivatives into play allows to relate observability questions to algebraic ones, e.g. whether the differentiable map

$$x \mapsto (c(x), (L_f c)(x), (L_f^2 c)(x), \ldots) \tag{5}$$

can be inverted.

We do not aim in this paper at a criterion for (full) observability, instead we propose a method to generate new observable functions of the state from given ones, and this method is not the same as repeated Lie-differentiation. This method has a practical aspect: It may lead to a reconstruction of the state from measured output with less time derivatives (which are ill-posed procedures from a numerical viewpoint) involved than usual. By an observable function $w(x)$ of the state we mean a function with this property:

If $x_1(t)$, $x_2(t)$ are two solutions of (1) and (3) holds true on some interval $[0, \delta]$, then $w(x_1(0)) = w(x_2(0))$.

We say also: $w(x)$ can be observed from $c(x)$. We assume that the output is part of the state, i.e. the system description runs as follows:

$$\begin{aligned} \dot{x}_1 &= f_1(x), \\ \dot{x}_2 &= f_2(x), \\ y &= x_2 \end{aligned} \tag{6}$$

with $x = (x_1, x_2)^T$.

3. Dissipation equalities and the algebraic generation of observable elements

The role of dissipation inequalities in nonlinear H_∞ theory is well understood, cf. [2]. These are relations connecting integrals for functions of the state and of the inputs of a control system with the boundary values of these quantities. Typically one can pass from the integral form of a classical dissipation inequality to a pointwise form which can be written as a Hamilton-Jacobi partial differential equation. A new way of employing Hamilton-Jacobi theory into the study of systems with two inputs—control and disturbance—is proposed in [1]. These are the main features of what is called a DE ("dissipation equality") in [1]:

(i) It is a relation—not an inequality but an approximate equality—, the formal structure being the same as of traditional dissipation inequalities. The length of the time interval is $\delta \ll 1$, however one cannot pass to the limit $\delta \to 0$ as in the standard H_∞ theory since the control functions involved are of the open loop type introduced in [1] (cf. Sec. 1, formula (3)). The limit for $\delta \to 0$ of such a function is not a point but a set.

(ii) The basic mathematical tool in setting up a DE is integration of a Hamilton-Jacobi partial differential equation along the classical "method of characteristics". However in applying the results of [1] to concrete situations no explicit integration of such a differential equation is required. Instead one has to find parameters such that certain algebraic relations are satisfied.

The main result of this section is based on a special case of the main result from Sec. 9 in [1], quoted as Proposition 9.1 henceforth. One arrives at this situation if the two integers m and n_2 which appear in [1] coincide and if the n_2-dimensional parameter b is equal to 0. And $b = 0$ implies that the control function $\tilde{r}(t)$ which appears in the dynamic law is equal to zero. To begin with we wish to restate Proposition 9.1 for this special case:

First observation. If $\tilde{r}(t) = 0$ the system description underlying the considerations in [1] reads

$$\dot{x}_1 = p_1(x) + G_1(x)w(x), \quad \dot{x}_2 = p_2(x), \tag{7}$$

$x(0) = x^{(0)}$, and does not depend upon the coefficients of $\tilde{r}(t)$. These coefficients are matrices $B_1(x)$ (of type $n_1 \times n_2$) and $B_2(x)$ (of type $n_2 \times n_2$). In the present context the symbols B_1, B_2, $(B_i)_x$ (with $(B_2)_{x_1} = 0$) take the place of the values of the $B_i(x)$ and their partial derivatives at $x = x^{(0)}$ and can be regarded as “free” parameters which can be used to meet e.g. the hypotheses of Proposition 9.1. In the sequel θ is a fixed parameter $\neq 0$, δ, π are positive parameters subject to restrictions

$$\delta \ll 1, \quad \pi \gg 1. \tag{8}$$

We assume throughout this paper that $w(x)$ is a scalar ($\Rightarrow G_1(x)$ is a n_1-dimensional column vector). We write p_i, G_i, w for $p_i(x^{(0)})$, $G_1(x^{(0)})$, $w(x^{(0)})$, and use the abbreviation

$$P := (p_2)_{x_1}, \quad \ddot{x}_2 = L_f^2 x_2 a t x = x^{(0)}. \tag{9}$$

Note that because of $\tilde{r}(t) = 0$ the n_2-dimensional vector which carried the symbol ξ_0 in [1] is now equal to $\frac{1}{2}\ddot{x}_2 + \mathcal{O}(\delta)$.

Next we turn to a discussion of the hypotheses (ii)–(iv) of Proposition 9.1:

(iv) is satisfied trivially because of $b = 0$.

(ii) concerns parameters λ_0, y_0 which in the present context can be chosen freely in order to meet the two conditions listed in [1]. It is easily confirmed by inspection that this can be done if

$$\hat{l}^T P B_1 \neq 0, \tag{10}$$

$\hat{l}$ defined by (13) below.

(iii) follows from

$$\hat{l}^T P p_1 \neq 0, \tag{11}$$

as is also easily confirmed.

We now are in the position to write down the statement of Proposition 9.1 for the special case under consideration. We write $(PG_1, 0, \ldots, 0)$ for the $n_2 \times n_1$ matrix whose first column is the n_1-dimensional vector PG_1 and the remaining columns are 0. We put $x^{(0)} = (x_1^{(0)}, x_2^{(0)})^T$.

Proposition 3.1. *Let S, B_2 be invertible matrices of the type $n_2 \times n_2$, S symmetric, and*

$$\widehat{P} := P - \frac{1}{2}(PG_1, 0, \dots, 0), \tag{12}$$
$$\hat{l} := S\widehat{P}(p_1 - B_1 B_2^{-1} p_2).$$

Hypotheses:

(i) p_1, G_1 *and the* n_2 *columns of* B_1 *are linearly independent* n_1*-dimensional vectors,*

(ii)

$$\hat{l}^T PG_1 = 0, \quad (10), (11) \textit{ hold true.} \tag{13}$$

Conclusion:

$$\frac{1}{8}\ddot{x}_2^T SPG_1 w - (1-\theta\pi)\Big\{\frac{1}{8}\hat{l}^T S^{-1}\hat{l} + \frac{1}{4}\hat{l}^T \ddot{x}_2 - \frac{1}{8}\hat{l}^T Pp_1 + \frac{1}{2}k^T \frac{1}{\delta^4}\int_0^\delta t^2 (x_2(t) - x_2^{(0)})\,\mathrm{d}t\Big\} = \mathcal{O}_1(\pi^{-1}) + \mathcal{O}_2(\delta), \tag{14}$$

where

$$k^T = \hat{l}^T PB_1 B_2^{-1}. \tag{15}$$

The two terms on the rhs (= right hand side) of (14) admits estimates $K_1\pi^{-1}$, $K_2\delta$. The K_i depend upon the rhs of (7) and upon the B_i, K_2 in addition may depend upon π.

In order to bring the statement of the proposition into a form which is more suited for applications we eliminate B_2 from (12), (15). We deal with the following question: Given k, d, can one find B_2 such that

$$k^T = \hat{l}^T PB_1 B_2^{-1}, \tag{16}$$
$$B_2^{-1} p_2 = d.$$

In other words: One wants to know whether two linear matrix equations can be solved subject to the side condition that the solution should be invertible. The answer is provided by

Lemma 3.1. *Given four vectors a, b, c, d of the same dimension ν. Then one can find a $\nu \times \nu$ matrix X satisfying*

$$\begin{aligned} a^T X &= b^T, \\ Xc &= d, \\ \det(X) &\neq 0 \end{aligned} \tag{17}$$

if

$$a^T d = b^T c \neq 0. \tag{18}$$

P r o o f. By standard arguments after the problem has been reduced (orthogonal transformation!) to the case $a^T = (\alpha, 0, \dots, 0)$. □

(16) can be considered as a system of the form (17), the unknown X now being B_2^{-1}. Condition (18) reads then

$$k^T p_2 = \hat{l} P B_1 d \neq 0. \tag{19}$$

In passing we note that (10) is implied by (19) and hence will not be mentioned anymore. We assume from now on that the $n_2 \times n_2$ matrix

$$\widehat{P} B_1 \text{ is invertible.} \tag{20}$$

The linear map (cf. (12), (16))

$$d \mapsto S\widehat{P} p_1 - S(\widehat{P} B_1) d = \hat{l} \tag{21}$$

is then invertible and one can express d in terms of $\hat{l}$. Furthermore we have—in view of (12), (13)—

$$\hat{l}^T \widehat{P} = \hat{l}^T P. \tag{22}$$

One can now eliminate d from (19) with the help of (21), (22) and obtains then a condition which does not contain B_2 explicitly but guarantees that (16) can be fullfilled by a proper choice of B_2. It runs as follows:

$$k^T p_2 = \hat{l}^T P p_1 - \hat{l}^T S^{-1} \hat{l} \neq 0. \tag{23}$$

Finally we specialize S as

$$S = (\alpha I + \hat{l}\hat{l}^T), \quad \alpha > 0 \tag{24}$$

I the n_2-dimensional unit matrix. S is symmetric and invertible and we have

$$SPG_1 = \alpha PG_1, \qquad \hat{l}^T S^{-1}\hat{l} = \frac{\|\hat{l}\|^2}{\alpha + \|\hat{l}\|^2} \tag{25}$$

(under the constraint $\hat{l}^T PG_1 = 0$!). Thereby we can now express the lhs of (14) completely in terms of the parameter α, $\hat{l}$, k which are free subject to the side conditions

$$\alpha > 0, \quad \hat{l}^T PG_1 = 0, \quad \hat{l}^T Pp_1 \neq 0 \quad \text{and (23)}. \tag{26}$$

In additions (20) and hypothesis (i) of Proposition 3.1 must hold true.

We are now in a position to state the main result—Proposition 3.2—of this section and let $x^{(0)}$ vary in an open and connected subset $\mathcal{X}$ of the x space. Parameter α and B_1 may depend upon $x^{(0)}$ but we suppress mostly this argument as we do it with p_i, G_1, P. Note that also $\ddot{x}_2 = L_f^2 x_2$ is actually a function of $x^{(0)}$.

Our general hypothesis is now

(a) $p_2 \neq 0$,
(b) $\ddot{x}_2$, PG_1, Pp_1 are linearly independent,
(c) p_1, G_1, and the columns of B_1 are linearly independent,
(d) (20) is satisfied, $\widehat{P}$ defined by (12).

As a preparation we need

Lemma 3.2. *Let* (a)–(d) *hold true and* $n_2 \geqslant 5$. *Let* $x^{(0)}$, $x'^{(0)}$ *be two points in* $\mathcal{X}$. *Then one can find* $\hat{l}$, $\hat{l}'$ *such that*

$$\begin{aligned} \hat{l}^T PG_1 &= \hat{l}'^T (PG_1)' = 0, \\ \hat{l}^T Pp_1 &= \hat{l}'^T (Pp_1)' \neq 0 \\ \hat{l}^T \ddot{x}_2 &= \hat{l}^T (\ddot{x}_2)' \end{aligned} \tag{27}$$

where the arguments on the lhs are $x^{(0)}$, *on the rhs are* $x'^{(0)}$. *Furthermore, we have*

$$\frac{\|\hat{l}\|^2}{\alpha + \|\hat{l}\|^2} = \frac{\|\hat{l}'\|^2}{\alpha' + \|\hat{l}'\|^2}, \tag{28}$$

$\alpha = \alpha(x^{(0)})$, $\alpha' = \alpha(x'^{(0)})$.

P r o o f. The statement is clear, if $x^{(0)} = x'^{(0)}$. If this is not the case we choose $x^\star \in \mathcal{X}$ and connect this point with $x^{(0)}$, $x'^{(0)}$ by a curve which is situated in $\mathcal{X}$ and parametrized by the same parameter $\tau, 0 \leqslant \tau \leqslant 1$. The two points which correspond to the same parameter τ are then considered in the place of $x^{(0)}$, $x'^{(0)}$. Using the IFT (implicit function theorem) and a standard analytic continuation argument one sees that one can meet the requirements of the lemma by appropriate $\hat{l}(\tau), \hat{l}'(\tau)$. $\square$

Proposition 3.2. *Hypotheses:* $n_2 \geqslant 5$, (a)–(d) *hold true and* $\ddot{x}_2^T PG_1 \neq 0$ *hold true on* $\mathcal{X}$. *Conclusion:* $w(x)$ *can be observed from* x_2 *on* $\mathcal{X}$.

P r o o f. Let $x^{(0)}$, $x'^{(0)}$ be two points in $\mathcal{X}$ such that the solution $x(t), x'(t)$ with the initial value $x^{(0)}$, $x'^{(0)}$ respectively satisfy $x_2(t) = x'_2(t)$. We choose $\alpha(x) > 0$ such that

$$\alpha(x)\ddot{x}_2^T PG_1|_{x=x^{(0)}} = \alpha(x)\ddot{x}_2^T PG_1|_{x=x'^{(0)}} \neq 0. \tag{29}$$

Next we employ Lemma 3.2 and construct $\hat{l}$, $\hat{l}'$ such that the DE set up for $x^{(0)}$ and the one set up for $x'^{(0)}$ have the same coefficients. This is possible since these coefficients are determined by the same linear equations, cf. (23), (25), (27), (28), note that $p_2 = L_f x_2$ is observable. We subtract the two DE's to obtain

$$\alpha(x)\ddot{x}_2^T PG_1|_{x=x^{(0)}} - \alpha(x)\ddot{x}_2^T PG_1|_{x=x'^{(0)}} = \mathcal{O}_1(\pi^{-1}) + \mathcal{O}_2(\delta). \tag{30}$$

Now the lhs of the last relation does not depend upon π, δ and the coefficient of w is the same and $\neq 0$, cf. (29). Hence we must have $w(x^{(0)}) = w(x'^{(0)})$. $\square$

References

[1] *Knobloch, H. W.*: Disturbance Attenuation in Control Systems, Part II: Proofs and Applications. Contributions to Nonlinear Control Theory, F. Allgöwer, H. W. Knobloch, Shaker Verlag, Herzogenrath, 2006. To appear.

[2] *D. Flockerzi*: Dissipation Inequalities and Nonlinear H_∞-Theory. Contributions to Nonlinear Control Theory, F. Allgöwer, H. W. Knobloch, Shaker Verlag, Herzogenrath, 2006. To appear.

Author's address: *Hans Wilhelm Knobloch*, Universität Würzburg, Mathematisches Institut, Lehrstul Math. II, Am Hubland, Würzburg, D-97094, Germany.