

## GENERICITY OF OBSERVABILITY AND THE EXISTENCE OF ASYMPTOTIC OBSERVERS

J. P. GAUTHIER

*Institut Universitaire de France, INSA de Rouen, Department of Mathematics, LMI  
BP 08, Place Emile Blondel, 76131 Mont-Saint-Aignan, France*

I. A. K. KUPKA

*Department of Mathematics, University of Toronto  
100, St Georges Street, Toronto, Ontario, M5S 1A1*

**Abstract.** In this paper, we deal with the genericity of the observability property and the existence of asymptotic observers for nonlinear systems. In the case where the number of outputs is *larger than* the number of inputs and the state space is *compact*, we prove that observability in a very strong sense (more or less, observability for *each* sufficiently differentiable input) is generic. This is obtained by using standard (but not easy) transversality arguments. For the inputs that are *bounded with their derivatives* up to some order, we prove the generic existence of an asymptotic observer with *arbitrary exponential decay* of the error.

**I. Introduction, statement of the results.** We deal with general controlled and observed smooth nonlinear systems:

$$(\Sigma) \quad \dot{x} = f(x, u), \quad y = h(x, u).$$

For the sake of clarity, the precise assumptions that we make on these systems are set and discussed in the next section II.1.

For these systems, our goal is the synthesis of nonlinear observers. A nonlinear observer for  $(\Sigma)$  is another control system  $(O_\Sigma)$  controlled by both the input  $u(t)$  and the output  $y(t)$  of  $(\Sigma)$ . This system  $(O_\Sigma)$  is expected to “estimate” asymptotically the state of  $(\Sigma)$ . In this paper, we will in fact assume that  $(O_\Sigma)$  is controlled by the output  $y(t)$  and the input  $u(t)$  of  $(\Sigma)$  together with some of the derivatives  $u^{(j)}(t)$  of the input. (This last fact seems unavoidable.) Roughly speaking, we require that there is a smooth embedding  $e$  (depending on the inputs)

---

1991 *Mathematics Subject Classification*: 93A99, 93B07, 93B10.

Both authors are partly supported by NSERC GRANT OGP 0036498.

The paper is in final form and no version of it will be published elsewhere.

from the state space  $X$  of  $(\Sigma)$  to the state space  $\widehat{X}$  of  $(O_\Sigma)$  such that the distance  $d(e(x(t)), \widehat{x}(t))$  tends to zero when  $t$  goes to infinity. ( $\widehat{x}$  is the state of  $(O_\Sigma)$ .) If  $d(e(x(t)), \widehat{x}(t))$  tends exponentially to zero, the observer  $(O_\Sigma)$  is said to be “exponential”. Precise definitions and statements of results are given in section IV. There are several ways to deal with the synthesis of nonlinear observers. In the linear case, the problem is solved by the standard Luenberger observer [LU]. In the nonlinear case, a major difficulty, that does not appear in the linear case, is the existence of “bad” inputs that make the system  $(\Sigma)$  unobservable (inputs that are “non-universal” in the sense of [S], i.e. such that couples of different “indistinguishable” initial states exist, i.e. these initial conditions, together with the considered bad input, produce the same output function  $y(t)$  for  $(\Sigma)$ ).

For general nonlinear systems, the problem of synthesis of observers is far from solved. In our opinion, this is due mainly to the presence of these “bad inputs”. Our results in this paper also strengthen this point of view, as the reader will see.

The most general method that we know, to deal with the construction of nonlinear observers, is the “high-gain observer method”. (It is, at least, more general than the output injection method developed in [KI], [KR], [HG1], [HG2], which applies to a very special class of systems only.) This “high-gain observer” method has been initiated in [GHO], [GHK], [GK] and we will use it in this paper. Let us first briefly comment these previous papers: they deal with the single-output case. It could appear as unnatural that the single-output case plays a particular role in the theory. This is completely clarified in this paper. In [GHO], [GK], we dealt with single-output controlled systems, and we showed that the observability property is nongeneric (the strong observability property which is required as soon as one wants to construct observers, i.e. observability for all inputs). When this property holds, it has been shown that systems can be put under a very special canonical form. In the case of systems having this property (of strong observability), we were able to construct nonlinear *exponential observers with arbitrary exponential decay of the estimation error*. In our other paper [GHK], we study the *uncontrolled single output case*. We show that the set of systems such that the state can be reconstructed on the basis of the knowledge of the output plus its  $2d$  first derivatives is generic ( $d =$  dimension of the state space). It results in the fact that the high gain construction applies generically, and hence that *exponential observers generically exist*. Similar results, about the observability property, have been obtained in [T], [A], [TC].

The purpose of this paper is to generalize the previous result to controlled systems for which the *number of outputs is strictly larger than the number of inputs*. For these systems, we will show the following results:

1. *The set of systems such that the state can be reconstructed on the basis of the outputs and their  $k$  first derivatives ( $k$  large enough), for inputs that are also  $k$ -times differentiable, contains a residual set.* If a bound is given a-priori for the inputs and their first derivatives, this set is *open, dense*. (The state space is assumed to be compact.)

2. A bound on the  $k$  first derivatives of the inputs being given, *the set of systems that have an observer working for every such input is open dense.*

Our paper is organized as follows: in section II, we give and discuss our notations, assumptions. We state precisely our observability results and we restrict to an open dense class of systems eliminating too degenerate situations. In section III, we prove our observability results. In section IV, we sketch the proof of our results on the genericity of the existence of observers.

To conclude, let us say that, at the moment when we correct the paper, we are able to prove a number of additional results. In particular, all of our results hold for the class of systems such that  $h$ , the output function, is a function of  $x$  only (with slightly different proofs). Also, let us thank the anonymous referee, whose remarks and suggestions were totally adequate.

**II. Assumptions, notations, preliminaries**

**II.1. Notations, assumptions**

**a. Systems and topology under consideration.** All along the paper, we will deal with nonlinear systems of the form:

$$(\Sigma) \quad \frac{dx}{dt} = f(x, u_0), \quad y = h(x, u_0),$$

where  $x \in X$ , an *analytic compact*  $d_x$ -dimensional connected manifold,  $h : X \times U \rightarrow R^{d_y}$ ,  $u_0 \in U = I^{d_u}$ , where  $I$  is some compact interval of  $R$ .  $f$  and  $h$  are assumed to be jointly  $C^r$  in  $x$  and  $u_0$ , for  $r$  fixed sufficiently large.

The topology that we consider for our systems  $(\Sigma)$  is the topology of  $C^r$  uniform convergence on  $X \times I^{d_u}$ . But for technical reasons, we will need a compact manifold as the set of values of the inputs. By Whitney's extension theorem, for  $r < \infty$ , there is a linear continuous map  $M : C^r(X \times I^{d_u}) \rightarrow C^r(X \times R^{d_u})$  (equipped with the compact-open topology). We consider two compact sets  $K_1$  and  $K_2$ ,  $X \times I^{d_u} \subset \text{int}(K_1) \subset \text{int}(K_2)$ , and a  $C^\infty$  function  $g$  which is 1 on  $K_1$  and 0 outside  $K_2$ , and the map  $M^* : C^r(X \times I^{d_u}) \rightarrow C^r(X \times R^{d_u})$ ,  $M^*(s) = g(x, u_0)M(s)$ .  $M^*$  is also linear continuous. Hence, we see that we can embed  $I^{d_u}$  in a  $d_u$ -dimensional torus  $T$  and that there is a continuous linear map, called again  $M^*$ ,  $M^* : C^r(X \times I^{d_u}) \rightarrow C^r(X \times T)$ . This shows that, if we prove our openness and density results in  $C^r(X \times T)$ , they will also be true in  $C^r(X \times I^{d_u})$ . If they are true for each (sufficiently large)  $r$ , *they are also true in  $C^\infty(X \times I^{d_u})$ .*

We shall denote the class of these systems, endowed with this  $C^r$  topology, by  $(S^r)$ . We denote by  $H^r$  the set of  $C^r$  maps  $h(x, u_0) : X \times T \rightarrow R^{d_y}$  and by  $F^r$  the set of  $C^r$  parametrized vector fields  $f(x, u_0)$  on  $X \times T$ .  $(S^r)$ ,  $H^r$  and  $F^r$  are Banach spaces.

A few comments about our systems are in order:

(i) We need the assumption of analyticity for  $X$  only for technical reasons. If  $X$  is only  $C^\infty$ , since the compact  $C^\infty$  manifold  $X \times T$  has a compatible analytic

structure, our results of genericity of differential observability and existence of observers still hold.

(ii) The compactness assumption for  $X$  can also certainly be avoided. We would certainly get the same results in the Whitney topology. However, for the purpose of construction of observers, compactness is required.

**b. Statement of the results.** A system  $\Sigma = (f, h)$  and a  $C^{k-1}$  input  $u : [0, T_u] \rightarrow T$  being given, we set  $w(t) = (u(t), u'(t), \dots, u^{(j)}(t), \dots, u^{(k-1)}(t))$  (we will write  $w$  instead of  $w(0)$ ), and we denote by  $R_{k,f}(x(t), w(t), h)$ , the vector formed by the outputs and their  $k - 1$  first derivatives at time  $t$ ,  $R_{k,f}(x(t), w(t), h) = (y(t), y'(t), \dots, y^{(k-1)}(t))$ . Hence, an initial condition  $x$  being given, the map  $R_{k,f}(w, h) : X \rightarrow R^{kd_y}$ ,  $R_{k,f}(w, h)(x) = R_{k,f}(x, w, h) = R_{k,f}(x, w(0), h)$ , is well defined.

A real number  $B > 0$  being given, we denote by  $C^{k-1}(B)$  the set of  $C^{k-1}$  inputs  $u(t)$ , defined on some time interval  $[0, T_u]$ , the  $k - 1$  first derivatives of which are bounded by  $B : \|u^{(j)}(t)\| \leq B, 1 \leq j \leq k - 1, 0 \leq t \leq T_u$ .

Our results are the following, assuming that  $d_y > d_u$ :

**THEOREM II.1.** *The set of systems such that  $R_{k,f}(w, h)$  is an immersion of  $X$  into  $R^{kd_y}$ , for all  $u(\cdot)$   $k - 1$  times differentiable, contains an open, dense subset of  $(S^r)$ , for  $r$  and  $k$  sufficiently large.*

**THEOREM II.2.** *The set of systems such that  $R_{k,f}(w, h)$  is an embedding, for all  $u(\cdot)$   $k - 1$  times differentiable, contains a residual subset of  $(S^r)$ , for  $r$  and  $k$  sufficiently large.*

A bound  $B > 0$  on the derivatives of the inputs being given, we have:

**THEOREM II.3.** *The set of systems such that  $R_{k,f}(w, h)$  is an embedding for all  $u(\cdot) \in C^{k-1}(B)$  is open, dense in  $(S^r)$ , for  $r$  and  $k$  sufficiently large.*

That is, in particular, the set of systems that are observable for all  $C^{k-1}$  inputs is residual. If a bound on the derivatives of the inputs is given a-priori, this set is open, dense.

What we mean by an “observer”, or an “exponential observer”, for a system, relative to a bound  $B$  on the derivatives of the inputs, will be precisely stated in section IV. It is a classical notion from control theory, that can be defined in several ways. More or less, it is another system, whose inputs are the inputs and the outputs of the given system, and the state is expected to estimate asymptotically the state of this system. Our main results on the existence of observers will be the following, as consequences of theorems II.1–3:

**THEOREM II.4.** *The set of systems that have a  $C^k$  observer, relative to the bound  $B > 0$ , contains an open dense set.*

**THEOREM II.5.** *The set of systems that have, relative to every bound  $B > 0$ , a  $C^k$  observer (depending on this bound), contains a residual set.*

Let us again make a few comments on these results:

(i) It can be seen very easily from our paper [GK] that in the case where the number of inputs is equal to the number of outputs, these theorems are all false.

(ii) The set of systems such that the statement of Theorem II.1. is true *is not open*. This again can be derived from our paper [GK].

**c. Notations.**  $(\Sigma)$  being given, we define the  $k$ th dynamical extension  $(\Sigma^k)$  of  $(\Sigma)$  and the  $k$ th dynamical extension  $f^k$  of  $f$  as follows.  $f^k$  is the vector field on  $X \times T \times R^{(k-1)d_u}$  given by

$$f^k(x, u_0, \dots, u_{k-1}) = \sum_{i=1}^{d_x} f_i(x, u_0) \frac{\partial}{\partial x_i} + \sum_{i=0}^{k-2} \sum_{j=1}^{d_u} u_{i+1,j} \frac{\partial}{\partial u_{i,j}},$$

and

$$\Sigma^k = (f^k(x, u_0, \dots, u_{k-1}) + b^k u_k, h(x, u_0)),$$

where  $b^k u_k = \sum_{i=1}^{d_u} b_i^k u_{k,i}$ ,  $b_i^k = \partial/\partial u_{k-1,i}$  and  $u_k = (u_{k,i}) \in R^{d_u}$  is the control variable of  $(\Sigma^k)$ .

We will be led to consider several maps attached to  $f^k$  or  $(\Sigma^k)$ . Denoting by  $L$  the Lie-derivative operator, we define:

$$\begin{aligned} R_{k,f}^* &: X \times T \times R^{(k-1)d_u} \times H^r \rightarrow R^{kd_y} \times T \times R^{(k-1)d_u}, \\ (x, w, h) = (x, u_0, \dots, u_{k-1}, h) &\rightarrow (h(x, u_0), L_{f^k} h(x, w), \dots, (L_{f^k})^{k-1} h(x, w), w), \\ R_{k,f} &: X \times T \times R^{(k-1)d_u} \times H^r \rightarrow R^{kd_y}, \\ (x, w, h) &\rightarrow (h(x, u_0), L_{f^k} h(x, w), \dots, (L_{f^k})^{k-1} h(x, w)). \end{aligned}$$

Notice that  $R_{k,f}^*(x, w, h) = (R_{k,f}(x, w, h), w)$ .

For  $h$  fixed in  $H^r$ , we define:

$$\begin{aligned} R_{k,f}^*(h) &: X \times T \times R^{(k-1)d_u} \rightarrow R^{kd_y} \times T \times R^{(k-1)d_u}, \\ R_{k,f}^*(h)(x, w) &= R_{k,f}^*(x, w, h), \end{aligned}$$

and

$$R_{k,f}(h) : X \times T \times R^{(k-1)d_u} \rightarrow R^{kd_y}, \quad R_{k,f}(h)(x, w) = R_{k,f}(x, w, h).$$

(Of course, these definitions agree with those given in the previous section.)

$\Delta X$  denoting the diagonal in  $X \times X$ , we define also:

$$\begin{aligned} T_{k,f}^* &: X \times X \setminus \Delta X \times T \times R^{(k-1)d_u} \times H^r \rightarrow R^{kd_y} \times R^{kd_y} \times T \times R^{(k-1)d_u}, \\ T_{k,f}^*(x_1, x_2, w, h) &= (R_{k,f}(x_1, w, h), R_{k,f}(x_2, w, h), w), \\ T_{k,f} &: X \times X \setminus \Delta X \times T \times R^{(k-1)d_u} \times H^r \rightarrow R^{kd_y} \times R^{kd_y}, \\ T_{k,f}(x_1, x_2, w, h) &= (R_{k,f}(x_1, w, h), R_{k,f}(x_2, w, h)), \\ T_{k,f}^*(h)(x_1, x_2, w) &= T_{k,f}^*(x_1, x_2, w, h), \\ T_{k,f}(h)(x_1, x_2, w) &= T_{k,f}(x_1, x_2, w, h). \end{aligned}$$

$T^*X^{\otimes s}$  will denote the fiber product of  $s$  copies of the cotangent bundle  $T^*X$  of  $X$ , and  $\Pi : T^*X^{\otimes s} \rightarrow X$ . Then  $T^*X^{\otimes s} = \{(p_1, \dots, p_s) \mid \Pi p_1 = \dots = \Pi p_s\}$ .

The following maps  $V_{k,f}^*$ ,  $V_{k,f}$ ,  $V_{k,f}^*(h)$ ,  $V_{k,f}(h)$  will be also of interest:

$$V_{k,f}^* : X \times T \times R^{(k-1)d_u} \times H^r \rightarrow T^*X^{\otimes(kd_y)} \times T \times R^{(k-1)d_u},$$

$$(x, w, h) = (x, u_0, \dots, u_{k-1}, h) \rightarrow (d_x R_{k,f}(x, w, h), w),$$

where  $d_x$  denotes the differential with respect to the  $x$ -variables only of the components of the map  $R_{k,f}$ :

$$d_x R_{k,f}(x, w, h) = (d_x h(x, u_0), d_x L_{f^k} h(x, w), \dots, d_x (L_{f^k})^{k-1} h(x, w)),$$

$$d_x h(x, u_0) = (d_x h_1(x, u_0), \dots, d_x h_{d_y}(x, u_0)),$$

$$V_{k,f} : X \times T \times R^{(k-1)d_u} \times H^r \rightarrow T^*X^{\otimes(kd_y)},$$

$$(x, w, h) = (x, u_0, \dots, u_{k-1}, h) \rightarrow d_x R_{k,f}(x, w, h),$$

$$V_{k,f}^*(h) : X \times T \times R^{(k-1)d_u} \rightarrow T^*X^{\otimes(kd_y)} \times R^{(kd_u)},$$

$$(x, w) = (x, u_0, \dots, u_{k-1}) \rightarrow V_{k,f}^*(x, w, h),$$

$$V_{k,f}(h) : X \times T \times R^{(k-1)d_u} \rightarrow T^*X^{\otimes(kd_y)},$$

$$(x, w) = (x, u_0, \dots, u_{k-1}) \rightarrow V_{k,f}(x, w, h).$$

In the rest of the paper, we will be interested in the following facts:

(F)  $R_{k,f}^*(h)$  is an embedding from  $X \times T \times R^{(k-1)d_u}$  into  $R^{kd_y} \times T \times R^{(k-1)d_u}$ .

Since  $X$  and  $I$  are compact, despite the non-compactness of  $R^{(k-1)d_u}$ , by the special form of  $R_{k,f}^*$ , this is equivalent to:

- (F1)  $R_{k,f}^*(h)$  is one-to-one; and
- (F2)  $R_{k,f}^*(h)$  is an immersion.

The fact (F1) is equivalent to the fact that the map  $T_{k,f}(h) : X \times X \setminus \Delta X \times T \times R^{(k-1)d_u} \rightarrow R^{kd_y} \times R^{kd_y}$  avoids the diagonal in  $R^{kd_y} \times R^{kd_y}$ .

$E$  being a  $d$ -dimensional vector-space,  $d < \infty$ , we denote (for  $k > d$ ) by  $V(k, E)$  the set of  $k$ -tuples  $(v_1, \dots, v_k) \in E^k$  such that all the  $d$ -tuples  $(v_{i_1}, \dots, v_{i_d})$  extracted from  $(v_1, \dots, v_k)$  are linearly dependent.  $V(k, E)$  is an algebraic subvariety of  $E^k$  of codimension  $k - d + 1$ .

The union  $\bigcup_{x \in X} V(kd_y, T_x^*X)$  is a subfiberspace of  $T^*X^{\otimes(kd_y)}$ , denoted by  $\mathbf{V}(kd_y, T^*X)$ .  $\mathbf{V}(kd_y, T^*X)$  is (as  $V(kd_y, T_x^*X)$ ) a finite union of manifolds of codimension  $(kd_y) - d_x + 1$ .

The fact (F2) is equivalent to the fact that the map  $V_{k,f}(h)$  avoids  $\mathbf{V}(kd_y, T^*X)$ :  $R_{k,f}^*(h)$  is immersive at  $(x, w)$  iff  $R_{k,f}(h)$  has maximal rank  $d_x$  with respect to  $x$  only at  $(x, w)$ .

*Several proofs of the “preliminary results” stated in the following sections II.2, II.3 are not obvious. They use basic transversality theorems from [AR], plus several easy but nonstandard facts about linear systems. Due to a lack of place, we leave all these proofs to the reader.*

Also, in the sequel, by “*Abraham’s transversality theorems*”, we mean: Theorem 18.1, p. 46 (openness of nonintersection), Theorem 18.2, p. 47 (openness of transversal intersection) and Theorem 19.1, p. 48 (transversal density), of [AR].

**II.2. Privileged vector fields.**  $X, I$  are given as above (II.1). For every compact subset  $K$  of  $X \times X \setminus \Delta X$ , we claim that there is an open dense subset  $F^{0,r}$  of  $F^r$  whose elements  $f(x, u_0)$  satisfy 1. and 2. below and an open dense subset  $F^{0,r,K}$  whose elements satisfy moreover 3. below:

1. The set  $E_f$  of zeros of  $f(x, u_0)$  on  $X \times T$  is a compact submanifold of  $X \times T$  of dimension  $d_u$ .

2. The set  $E_{f,c}$  of points  $(x_0, u_0)$  of  $E_f$  such that the linearized vector field  $T_X f_{u_0}(x_0) = T_X f(x_0, u_0)$  has cyclic index  $> c$  is a closed finite union of submanifolds of  $X \times T$  of dimension  $\leq d_u - (c + 1)^2 + 1$  at most (negative dimension means the empty set).

3. The set  $F_f$  of zeros of  $(f(x_1, u_0), f(x_2, u_0))$  on  $W_K \times T$  ( $W_K$  some open neighbourhood of  $K$  in  $(X \times X \setminus \Delta X)$ ) is a  $d_u$ -dimensional manifold.

**II.3. Privileged functions.** We keep the notations of the previous section. We choose an element  $f$  of  $F_{0,r}$  which is *analytic* on  $X \times T$  (such elements are dense in  $F^{0,r}$  which is itself open-dense in  $F^r$  by II.2).

There is an open dense subset  $H^{0,r}$  of  $H^r$  whose elements  $h(x, u_0)$  satisfy:

(i) the differential  $d_X h(x, u_0)$  is cocyclic for the linearized  $T_X f(x, u_0)$ , for each  $(x, u_0)$  in  $E_f$ . Equivalently, the linearization of the system  $(\Sigma) = (f, h)$  at  $(x, u_0)$  is observable.

Also, there exists an open dense subset  $H^{0,r,K}$  of  $H^r$  satisfying moreover:

(ii)  $(x_1, x_2) \in W_K$ ,  $(x_1, u_0)$  and  $(x_2, u_0) \in E_f$ , then  $h(x_1, u_0) \neq h(x_2, u_0)$ , where  $W_K$  is some open neighbourhood of  $K$ .

**II.4. Privileged systems**

1. There is an open dense subset  $(S^{0,r})$  of  $(S^r)$  whose elements  $(f, h)$  satisfy 1, 2 of II.2 and  $h$  satisfies (i) of II.3.

2. There is an open dense subset  $(S^{0,r,K})$  of  $(S^r)$  satisfying moreover 3 of II.2 and (ii) of (II.3).

### III. Proof of the observability theorems

**III.1. Openness and density of injectivity away from the diagonal.** In this section, we will in fact prove more than necessary for the theorems stated herein (Theorems II.2–3). However, what we prove (openness and density of injectivity after restriction to every compact subset  $K$  of  $X \times X \setminus \Delta X$ ) is a step to the result that we are not able to prove up to now: the set of systems such that  $R_{k,f}^*(h)$  is an embedding, contains an *open*-dense set (we are able to get the density only). The compact set  $K$  and the open set  $W_K$ , above and below, serve to prove this point.

We fix  $f$  in  $F^{0,r,K}$  of section II.2. We first prove that the set  $H^{r,K,f,k}$  of  $h$  such that  $T_{k,f}(h)$  avoids the diagonal restricted to  $K \times T \times R^{(k-1)d_u}$  is dense (*Fact D<sub>K</sub>*). More precisely, we will first prove that  $T_{k,f} : K \times T \times R^{(k-1)d_u} \times H^{0,r} \rightarrow R^{kd_y} \times R^{kd_y}$  is transverse to  $\Delta_k$ , the diagonal of  $R^{kd_y} \times R^{kd_y}$ . The result will follow for  $k$  large enough by considering dimensions, by the transversal density theorem. With the same reasoning, the map

$$T_k : K \times T \times R^{(k-1)d_u} \times S^{0,r} \rightarrow R^{kd_y} \times R^{kd_y},$$

$$T_k(x_1, x_2, u, \Sigma) = T_{k,f}(h)(x_1, x_2, u)$$

is transverse to  $\Delta_k$  (*Fact D'<sub>K</sub>*).

The set  $H^{0,r,K}$  has been defined in II.3 and is a Banach manifold, as an open set in a Banach space.

We consider  $(x^1, u_0, \dots, u_{k-1}) = (x^1, u)$  and  $(x^2, u_0, \dots, u_{k-1}) = (x^2, u)$ .

Case 1:  $f(x^i, u_0) = 0$ , for both  $i = 1, 2$ . In this case, since  $h \in H^{0,r,K}$ , by II.3 (ii),  $h(x^1, u_0) \neq h(x^2, u_0)$ . Hence,  $T_{k,f}(h)$  avoids  $\Delta_k$  at the point  $(x^1, x^2, u, h)$ . Hence:

$$T_{k,f} \pitchfork_{(x^1, x^2, u, h)} \Delta_k.$$

Case 2:  $f(x^i, u_0) \neq 0$  for one  $i$ , say  $i = 1$ . We can find around  $(x^1, u)$  a coordinate neighbourhood  $U_1$  such that  $f^k$  reads on  $U_1$ :

$$f^k = \frac{\partial}{\partial x_1} + \sum_{j=1}^{d_u} u_{1,j} \sum_{i=1}^{d_x} \Phi_{i,j}(x, u_0) \frac{\partial}{\partial x_i} + \sum_{i=0}^{k-2} \sum_{j=1}^{d_u} u_{i+1,j} \frac{\partial}{\partial u_{i,j}}.$$

Moreover, the  $u$ -coordinates are not changed and the change in the  $x$ -coordinate depends only on  $x$  and  $u_0$ .

$R_{k,f}$ , in these coordinates, reads

$$R_{k,f}(h) = \left( h(x), \frac{\partial h}{\partial x_1} + L_{\Phi}^x h u_1 + \frac{\partial h}{\partial u_0} u_1, \dots \right)$$

where  $L_{\Phi}^x$  means the Lie derivative of  $\Phi$  in the direction of  $x$  only.

Case 2.1:  $u_1 = 0$ . We will choose  $h$  as a function of  $x$  only in a neighbourhood of  $(x, u_0)$ . In this case

$$R_{k,f}(h)(x, u) = \left( h(x), \frac{\partial h}{\partial x_1}, \frac{\partial^2 h}{\partial x_1^2} + L_{\Phi}^x h u_2, \dots, \frac{\partial^n h}{\partial x_1^n} + \Psi_n(h), \dots \right),$$

where  $\Psi_n$  is a differential operator in  $x$  of degree smaller than  $n$ . (See lemma 1, appendix.)

$h$  is then chosen of the form  $h = \alpha(x, u_0) (\sum h_i x_1^i / i!)$  where  $\alpha(x, u_0)$  is a bump function in a neighbourhood of  $(x, u_0)$  in  $X \times T$ .

Given any  $\Psi = (\Psi_0, \dots, \Psi_{k-1})$  in  $R^{kd_y}$ , we see that we can solve inductively the equation  $R_{k,f}(h) = \Psi$  with respect to the  $h_i$  (this is possible since  $\Psi_n(h)$  is a differential operator in  $h$  of degree less than  $n$ ).  $R_{k,f}$  is linear with respect to  $h$ ,



hence it is equal to its tangent map  $T_h R_{k,f}$ . This shows that  $R_{k,f}$  is submersive at the point  $(x^1, u, h)$ .

Case 2.2:  $u_1 \neq 0$ . In this case, we choose  $h$  as a function of  $u_0$  only (up to a bump function). Then

$$R_{k,f}(h) = T_{h_0} R_{k,f}(h) = \left( h(u_0), \frac{\partial h}{\partial u_0}(u_1), \frac{\partial h}{\partial u_0}(u_2) + \frac{\partial^2 h}{\partial u_0^2}(u_1, u_1), \dots \right).$$

The equation  $R_{k,f}(h) = \Psi$  can also be solved inductively for any  $\Psi$ .  $R_{k,f}$  is submersive at the point  $(x_1, u, h)$ .

It is clear that  $R_{k,f}$  being submersive at  $(x^1, u, h)$  implies that

$$T_{k,f} \pitchfork_{(x^1, x^2, u, h)} \Delta_k.$$

Now, applying again Abraham's transversality theorems to

$$T_{k,f} : W_K \times T \times R^{(k-1)d_u} \times H^{0,r,K} \rightarrow R^{kd_y} \times R^{kd_y},$$

we see that the set of  $h$  such that  $T_{k,f}(h) \pitchfork_{(x^1, x^2, u)} \Delta_k$ , for all  $(x^1, x^2, u) \in ((X \times X \setminus \Delta X) \cap K) \times R^k$  (where  $W_K$  has been defined in II.2), is residual, hence dense. But,  $\dim(X \times X \setminus \Delta X) \times T \times R^{(k-1)d_u} = 2d_x + (kd_u)$ ,  $\text{codim } \Delta_k = kd_y$ . Hence, for  $k$  sufficiently large ( $k > 2d_x$ ), transversality means nonintersection. Hence, the fact  $D_K$  is proved, as a consequence of  $(D'_K)$ .

Let  $W_K$  be as in II.2-4, and let  $TW_K$  be the tangent bundle to  $W_K$ . Then  $TW_K \times T$  is a vector bundle over  $W_K \times T$ . Also consider the vector bundle  $W_K \times T \times R^{2d_y}$  (over  $W_K \times T$ ) and  $B$ , the fiber product of these two vector bundles:

$$B = (TW_K \times T) \otimes (W_K \times T \times R^{2d_y}).$$

Sections of  $B$  are maps:

$$(x_1, x_2, u_0) \rightarrow (f_1(x_1, u_0), f_2(x_2, u_0), h_1(x_1, u_0), h_2(x_2, u_0)).$$

Consider  $\Pi : B \rightarrow W_K \times T$ .

Let  $\Pi^k : J^k(\Pi) \rightarrow W_K \times T$  be the analytic vector bundle of  $k$ -jets of  $C^r$  sections of  $B$ . Consider  $(S^r) \times W_K \times T$  and the evaluation map  $\text{ev} : (S^r) \times W_K \times T \rightarrow J^k(\Pi)$ ,  $\text{ev}(\Sigma, x_1, x_2, u_0) = (j^k f(x_1, u_0), j^k f(x_2, u_0), j^k h(x_1, u_0), j^k h(x_2, u_0))$  with  $\Sigma = (f, h)$ ,  $j^k \Sigma = (j^k f, j^k h)$  is the  $k$ -jet extension of  $\Sigma$ .

The restriction  $\text{ev}^\circ$  of  $\text{ev}$  to  $(S^{0,r,K}) \times W_K \times T$  has open dense image  $J^{\circ,k}(\Pi)$  in  $J^k(\Pi)$  since  $\text{ev}$  is a surjective submersion and  $(S^{0,r,K})$  is open dense in  $(S^r)$  by II.4.

$(D'_K)$  says that  $T_k : (S^{0,r,K}) \times W_K \times T \times R^{(k-1)d_u} \rightarrow R^{2kd_y}$ ,  $T_k(\Sigma, x_1, x_2, u) = T_{k,f}(h)(x_1, x_2, u)$  is transverse to  $\Delta_k$ . Clearly,  $T_k$  depends only on the  $k$ -jet  $j^k \Sigma$ , hence  $T_k$  factors as

$$T_k(\Sigma, x_1, x_2, u_0, u_1, \dots, u_{k-1}) = T_k^\circ(\text{ev}^\circ(\Sigma, x_1, x_2, u_0), u_1, \dots, u_{k-1}),$$

$$T_k^\circ : J^{\circ,k}(\Pi) \times R^{(k-1)d_u} \rightarrow R^{2kd_y}.$$

Since  $T_k \pitchfork \Delta_k$  and  $\text{ev}^\circ$  is a submersion onto  $J^{\circ,k}(\Pi)$ ,  $T_k^\circ \pitchfork \Delta_k$ . Hence,  $W_k^\circ = (T_k^\circ)^{-1}(\Delta_k)$  is an analytic submanifold of  $J^{\circ,k}(\Pi) \times R^{(k-1)d_u}$ , of codimension  $kd_y$ .

Otherwise,  $W_k^\circ$  is an analytic, partially algebraic (with respect to  $u_1, \dots, u_k$ ) subset of  $J^{\circ,k}(\Pi) \times R^{(k-1)d_u}$ , hence its projection on  $J^{\circ,k}(\Pi)$  is a semialgebraic subset of  $J^{\circ,k}(\Pi)$  denoted by  $W_k^{\circ*}$ . The codimension of  $W_k^{\circ*}$  is  $k(d_y - d_u) + d_u$ .  $\text{Cl}(W_k^{\circ*})$ , the closure of  $W_k^{\circ*}$  is now a semialgebraic subset of  $J^{\circ,k}(\Pi)$  of the same codimension  $k(d_y - d_u) + d_u$ . It follows from the transversality theorems that the set of  $\Sigma \in S^{0,r,K}$  such that  $\text{ev}^\circ(\Sigma, x_1, x_2, u_0)$  avoids  $\text{Cl}(W_k^{\circ*})$  is open dense for  $(x_1, x_2) \in K$  since  $K$  is compact and  $2d_x + d_u < k(d_y - d_u) + d_u$ ,  $k > 2d_x$ . This set of  $\Sigma$  is, by definition of  $W_k^{\circ*}$ , such that  $T_{k,f}(h)$  avoids  $\Delta_k$  when restricted to  $K \times T \times R^{(k-1)d_u}$ . ■

**Remark.** Taking a countable union of sets  $K_i$  such that  $K_i \subset K_{i+1}$ ,  $\bigcup_{i \in \mathbb{I}} K_i = X \times X \setminus \Delta X$ , shows that the set of  $\Sigma$  such that  $T_{k,f}(h)$  avoids the diagonal is residual.

**III.2. Openness and density of immersivity (Observability theorem II.1).** Again, we fix an  $f$  in  $F^{0,r}$  (see section II).

We want to prove, by II.1.b, that  $V_{k,f}(h)$  avoids  $\mathbf{V}(kd_y, T^*X)$ . We have  $\dim(X \times T \times R^{(k-1)d_u}) = d_x + kd_u$ ,  $\text{codim}(\mathbf{V}(kd_y, T^*X)) = kd_y - d_x + 1$  by II.1. Hence, for  $k(d_y - d_u) \geq 2d_x$ , transversality means nonintersection. We consider any point  $(x, u_0^*, \dots, u_{k-1}^*) = (x, u^*)$ . Again,  $V_{k,f}$  being linear in  $h$ , it is sufficient to prove surjectivity of  $V_{k,f}$  at the point  $(x, u_0^*)$ , in order to obtain density, with the transversality theorems. We will be able to prove this on  $(X \times T \times R^{(k-1)d_u}) \setminus (E_f \times R^{(k-1)d_u})$ . On  $E_f$ , we will proceed differently.

(1) *First case,  $f(x, u_0^*) \neq 0$ .* We choose a coordinate system around  $(x, u_0^*)$  in which  $f^k$  reads again (see III.1)

$$(III.2.f0) \quad f^k = \frac{\partial}{\partial x_1} + \sum_{j=1}^{d_u} u_{1,j} \sum_{i=1}^{d_x} \Phi_{i,j}(x, u_0) \frac{\partial}{\partial x_i} + \sum_{i=0}^{k-2} \sum_{j=1}^{d_u} u_{i+1,j} \frac{\partial}{\partial u_{i,j}}.$$

The change of coordinates in  $X \times T \times R^{(k-1)d_u}$  leaves the  $u$  coordinates unchanged and the change of coordinates in  $x$  depends only on  $x$  and  $u_0$ .

If  $u_1^*$  is different from zero, let us assume that  $(u_1^*)^1$  is nonzero and proceed as follows:

$$(III.2.f1) \quad (L_{f^k})^n h = \frac{\partial^n h}{(\partial x_1)^n} + \frac{\partial^n h}{(\partial u_0)^n}(u_1, \dots, u_1) + \Psi_n(x, u)(h),$$

where  $\Psi_n$  is a polynomial differential operator of degree  $n$  in  $L_{\Phi}^x$ ,  $\partial/\partial x_1$ , and of degree  $n - 1$  in  $\partial/\partial u_0$  (this is obtained with an obvious induction).

In that case, we choose a bump function  $\alpha(x, u_0)$ , and

$$h(x, u_0) = \left[ h_0(x) + \sum h_i(x)(u_0^1 - u_0^{*1})^i / i! \right] \alpha(x, u_0)$$

with

$$h_0(x) = \sum_{i=0}^{k-1} \sum_{j=2}^{d_x} \frac{x_1^i}{i!} x_j b_{i,j} + \sum_{i=1}^{k-1} \frac{x_1^{i+1}}{i+1!} b_{i,1}.$$

We can choose  $h_n$  such that

$$d_x \left( \frac{\partial^n h}{(\partial u_0)^n}(u_1, \dots, u_1) + \Psi_n(x, u)(h) \right) = 0$$

at the point  $(x, u_0^*)$  under consideration: this is possible since this equation can be rewritten as

$$d_x \left( (u_1^1)^n \frac{\partial^n h}{(\partial u_0^1)^n} + \Psi_n(x, u)(h) \right) = 0,$$

and  $\Psi_n(x, u)(h)$  does not depend on the derivatives  $\partial^i h / (\partial u_0^1)^i$ ,  $i \geq n$  by III.2.f1. Hence,

$$d_x (L_{f^k})^n h(x, u^*) = d_x \frac{\partial^n h}{(\partial x_1)^n} = (b_{n,1}, \dots, b_{n,d_x}).$$

These quantities can be chosen arbitrarily,  $V_{k,f}$  is submersive.

If  $u_1^* = 0$ , in the above expression III.2.f1, one can see that, because of the term  $\sum_{j=1}^{d_u} u_{1,j} \sum_{i=1}^{d_x} \Phi_{i,j}(x, u_0) \partial / \partial x_i$  in III.2.f0,  $\partial^n h / (\partial x_1)^n$  can appear, but multiplied by components of  $u_1^*$ , that are zero. Hence, the equation  $d_x(\partial^n h / (\partial x_1)^n + \Psi_n(x, u)(h)) = b_n$  can be solved inductively for arbitrary  $b_n$ .

Setting  $h(x) = [h_0(x_2, \dots, x_n) + \dots + (x_1^n / n!) h_n(x_2, \dots, x_n) + \dots] \alpha(x, u_0)$ ,  $\alpha(x, u_0)$  a bump function, we get

$$\frac{\partial^n h}{(\partial x_1)^n} = (h_n(x_2, \dots, x_n) + x_1 h_{n+1}(x_2, \dots, x_n) + \dots) \quad \text{near } (x, u_0^*),$$

$$d_x \frac{\partial^n h}{(\partial x_1)^n} \Big|_{x=0} = h_{n+1}(0) dx_1 + \sum_{j=2}^n \frac{\partial h_n}{\partial x_j}(0) dx_j.$$

The  $h_i$  can be taken affine.

$V_{k,f}$  is submersive on  $(X \times T \times R^{(k-1)d_u}) \setminus (E_f \times R^{(k-1)d_u})$ . The set of  $h$  such that  $V_{k,f}(h)$  avoids  $\mathbf{V}(kd_y, T^*X)$  restricted to  $(X \times T \times R^{(k-1)d_u}) \setminus (E_f \times R^{(k-1)d_u})$  is residual.

Consider

$$V_k : S^{0,r} \times X \times T \times R^{(k-1)d_u} \rightarrow T^*X^{\otimes kd_y},$$

$$V_k(\Sigma, x, u_0, u_1, \dots, u_k) = V_{k,f}(h)(x, u_0, u_1, \dots, u_k).$$

With the same proof, we have that  $V_k$  is transverse to  $\mathbf{V}(kd_y, T^*X)$  for points  $(x, u_0)$  such that  $f(x, u_0) \neq 0$  (fact  $I_k$ ).

**(2) Second case,  $f(x, u_0^*) = 0$  and rest of the proof.** In the sequel, we work as follows:  $\varrho$  is the first integer such that  $u_\varrho^* \neq 0$ . We assume that  $(u_\varrho^*)^1 \neq 0$ . We fix a coordinate system around  $x$ .

We consider the vector bundle  $\Pi_{B\Sigma} : B\Sigma \rightarrow X \times T$  which is the fiber product over  $X \times T$  of  $TX \times T$ ,  $TX$  (the tangent bundle of  $X$ ) and  $X \times T \times R^{d_y}$ . We have  $B\Sigma = (TX \times T) \otimes (X \times T \times R^{d_y})$ . Sections of  $B\Sigma$  are maps  $(x, u_0) \rightarrow (f(x, u_0), h(x, u_0))$ . Let  $\Pi_{B\Sigma}^k : J^k(\Pi_{B\Sigma}) \rightarrow X \times T$  be the analytic vector bundle of  $k$ -jets of  $C^r$  sections of  $B\Sigma$ . Consider the evaluation map  $\text{ev} : (S^r) \times X \times T \rightarrow J^k(\Pi_{B\Sigma})$ ,  $\text{ev}(\Sigma, x, u_0) = (j^k f(x, u_0), j^k h(x, u_0))$  with  $\Sigma = (f, h)$  and  $j^k \Sigma = (j^k f, j^k h)$  is again the  $k$ -jet of  $\Sigma$ . Again, the restriction  $\text{ev}^\circ$  of  $\text{ev}$  to  $((S^{0,r}) \times X \times T)$  has open dense image  $J^{\circ k}(\Pi_{B\Sigma}) \subset J^k(\Pi_{B\Sigma})$  since  $\text{ev}$  is a surjective submersion and  $(S^{0,r})$  is open dense in  $(S^r)$  by II.4.

Let  $J_E^{\circ k}(\Pi_{B\Sigma})$  be the analytic subset of  $J^{\circ k}(\Pi_{B\Sigma})$  formed by elements  $(j^k \Sigma(x, u_0))$  such that  $f(x, u_0) = 0$ . By our appendix, lemmas 2, 3, we know that:

- (i)  $V_k$  depends only on the  $k$ -jet of  $\Sigma$ ;
- (ii) The  $n$ th component of  $V_k$  is, when  $f(x, u_0) = 0$ , of the form

$$\sum_{i_1, \dots, i_r=1}^{d_u} d_x (L_f^x)^{n-q\varrho} \frac{\partial^q h}{(\partial u_0^{i_1})^{s_1} \dots (\partial u_0^{i_r})^{s_r}} (u_\varrho^{i_1})^{t_1} \dots (u_\varrho^{i_r})^{t_r} \\ + \sum_{i>0} \delta_{i,n} d_x \frac{\partial^q h}{(\partial u_0^{i_1, i})^{q_{1,i}} \dots (\partial u_0^{i_r, i})^{q_{r,i}}} d_x ((L_f^x)^{n-q\varrho-1-i} f) + Z_n,$$

where  $\delta_{i,n}$ ,  $Z_n$  are smooth expressions with respect to  $(x, u_0)$ , algebraic with respect to  $(u_1, \dots, u_{k-1})$ , depending only on the  $n$ th jets of  $f$  and on the  $n$ th jets with respect to  $x$  of  $\partial^t h / \partial u_0^t$ ,  $0 \leq t < q$ , with  $q = [n/\varrho]$ , integer part of  $n/\varrho$ .

Hence,  $V_k$  factors as

$$V_k(\Sigma, x, u_0, u_1, \dots, u_{k-1}) = V_k^\circ(\text{ev}^\circ(\Sigma, x, u_0), u_1, \dots, u_{k-1}), \\ V_k^\circ : J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u} \rightarrow T^*X^{\otimes kd_y}.$$

Set  $\mathbf{W}(kd_y, T^*X) = (V_k^\circ)^{-1}(\mathbf{V}(kd_y, T^*X)) \subset J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$ . Fact  $I_k$  above says that at a point  $(\Sigma, x, u_0, u_1, \dots, u_{k-1})$  such that  $f(x, u_0) \neq 0$ ,  $V_k$  is transverse to  $\mathbf{V}(kd_y, T^*X)$ . At such a point  $(\text{ev}^\circ(\Sigma, x, u_0), u_1, \dots, u_{k-1})$ ,  $V_k^\circ$  must be transverse to  $\mathbf{V}(kd_y, T^*X)$ . This subset of  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  is nothing but the open set  $C(J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u})$ , the complement of  $J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  in  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$ . Therefore  $\mathbf{W}^1(kd_y, T^*X) = \mathbf{W}(kd_y, T^*X) \cap C(J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u})$  is a semi-analytic, partially algebraic (with respect to  $u_1, \dots, u_{k-1}$ ) subset of  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$ . Since by II.1,  $\mathbf{V}(kd_y, T^*X)$  has codimension  $kd_y - d_x + 1$ ,  $\mathbf{W}^1(kd_y, T^*X)$  has codimension  $kd_y - d_x + 1$  in  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$ .

$\mathbf{W}(kd_y, T^*X)$  is also an analytic, partially algebraic subset of  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  and  $\mathbf{W}(kd_y, T^*X) = \mathbf{W}^1(kd_y, T^*X) \cup \mathbf{W}^2(kd_y, T^*X)$  where  $\mathbf{W}^2(kd_y, T^*X) = \mathbf{W}(kd_y, T^*X) \cap (J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u})$ . We will compute the codimension of  $\mathbf{W}^2(kd_y, T^*X)$ :

The codimension of the analytic subset  $J_E^{\circ k}(\Pi_{B\Sigma})$  of  $J^{\circ k}(\Pi_{B\Sigma})$  is  $d_x$ . The codimension of  $\mathbf{W}^2$  in  $J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  is given by lemma 4 of our appendix:

For a fixed  $\varrho$ , corresponding to the point  $z = (\text{ev}^\circ(\Sigma, x, u_0), u_1, \dots, u_{k-1})$  we set  $A' = T_X f(x, u_0)$ ,  $V'_0 = d_x h(x, u_0)$  ( $'$  means transpose). By II.3(i),  $V_0$  is cyclic for  $A$ . Hence, the fact that  $z \in \mathbf{W}^2$  means that  $(V_0, A)$  meets the assumptions of appendix, lemma 4.

It follows that the codimension of this set  $\mathbf{W}^2_\varrho$  in  $J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  is larger than  $(d_y \varrho [k/\varrho]) - d_x + 1$ , hence than  $kd_y - d_x + 1 - (rd_y)$ , with  $k = q\varrho + r$ .

$\mathbf{W}^2(kd_y, T^*X) = \bigcup_{\varrho < d_x} \mathbf{W}^2_\varrho$ , since for  $\varrho \geq d_x$ ,  $\mathbf{W}^2_\varrho$  is empty.

$\mathbf{W}^2_\varrho$  is defined by  $(u_1, \dots, u_{r-1}) = 0$  and the set of equations of lemma 4. The codimension of  $\mathbf{W}^2_\varrho$  is therefore at least  $(r-1)d_u + kd_y - d_x + 1 - (rd_y)$ , since lemma 4 gives  $kd_y - d_x + 1 - (rd_y)$  for the codimension in the jets of  $h$  for  $(x, u_0, f)$  fixed.

We have  $\text{codim}(\mathbf{W}^2_\varrho(kd_y, T^*X)) \geq kd_y - d_x + 1 - (\varrho - 1)(d_y - d_u)$  in  $J_E^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$ .

Finally,  $\text{codim}(\mathbf{W}^2(kd_y, T^*X)) \geq kd_y - d_x(d_y - d_u) - d_x + 1$  in  $J^{\circ k}(\Pi_{B\Sigma}) \times R^{(k-1)d_u}$  and the same is true for  $\text{codim}(\mathbf{W}(kd_y, T^*X))$ .

Let  $p(\mathbf{W})$  be the projection of  $\mathbf{W}(kd_y, T^*X)$  on  $J^{\circ k}(\Pi_{B\Sigma})$ . By partial algebraicity,  $p(\mathbf{W})$  is semi-algebraic, of codimension  $kd_y - d_x(d_y - d_u) - d_x + 1 - (k-1)d_u = (k-d_x)(d_y - d_u) + d_u - d_x + 1$ .  $\text{Cl}(p(\mathbf{W}))$ , the closure of  $p(\mathbf{W})$  is still semi-algebraic, with the same codimension. For  $k$  large enough (for  $k \geq 3d_x$ ), since  $d_y > d_u$ ,  $\dim(X \times T) < \text{codim}(\text{Cl}(p(\mathbf{W})))$ .  $\text{Cl}(p(\mathbf{W}))$  is a closed set, hence the subset formed by  $\Sigma \in S^{0,r}$  such that  $j^k \Sigma(x, u_0)$  avoids  $\text{Cl}(p(\mathbf{W}))$  is open and dense. This implies that the set of  $\Sigma \in S^{0,r}$  such that  $V_k(\Sigma)$  avoids  $\mathbf{V}(kd_y, T^*X)$  contains this open-dense set, or the set of  $\Sigma \in S^{0,r}$  such that  $R_k, R_k(\Sigma) = R_{k,f}(h)$  is an immersion for all  $(u_1, \dots, u_{k-1})$  contains an open-dense set. ■

### III.3. Proof of observability theorems II.2, II.3

1. **Theorem II.2.** The remark at the end of section III.1 plus the result of section III.2 show that the set of  $\Sigma$  such that  $R_{k,f}^*(h)$  is an injective immersion contains a residual subset of  $(S^r)$ . As observed in II.1.b, this is sufficient to get the result. ■

2. **Theorem II.3.** The set of embeddings from  $X \times T \times I_B^{(k-1)d_u}$  is open: this is proved for instance in [HIR] in the case of embeddings from compact manifolds with boundaries. We leave the reader to check that it also works for maps from compact “corner” manifolds such as  $X \times T \times I_B^{(k-1)d_u}$ .

Otherwise, the map

$$(S^r) \rightarrow C^r(X \times T \times I_B^{(k-1)d_u}, R^{kd_y} \times R^{(k-1)d_u}), \quad (\Sigma) \rightarrow \overline{R}_{k,f}^*(h),$$

where  $\overline{R}_{k,f}^*(h)$  is the restriction of  $R_{k,f}^*(h)$  to  $X \times T \times I_B^{(k-1)d_u}$  is continuous (by compactness).

Hence, the set of  $(\Sigma)$  in  $(S^r)$  such that  $\overline{R}_{k,f}^*(h)$  is an embedding is open. It is dense by theorem II.2. ■

#### IV. Construction of the observer

**IV.1. Statement of the results.**  $X, I$  and  $\Sigma = (f, h)$  being given on  $X \times I^{d_u}$ , a  $C^k$  observer of  $\Sigma$  relative to the bound  $B$  on the derivatives of the control is a triple  $O_\Sigma = (R^n, F_{y,u,v}, \tau)$  where:

(A)  $F_{y,u,v}$  is a parametrized family of  $C^r$  vector fields on  $X$ , depending smoothly on  $y, u, v$ ,

(B)  $\tau : X \times I^{d_u} \times R^{(k-1)d_u} \rightarrow R^n \times R^{kd_u}$  is an embedding of the form (setting now  $u = (u_0, \dots, u_{k-1})$ )

$$\tau(x, u) = \tau(x, u_0, u_1, \dots, u_{k-1}) = (\phi(x, u_0, u_1, \dots, u_{k-1}), u_0, u_1, \dots, u_{k-1}),$$

for some mapping  $\phi : X \times I^{d_u} \times R^{(k-1)d_u} \rightarrow R^n$ ,

(C) the differential equation

$$(IV.0) \quad (\dot{x}, \dot{u}) = f^k(x, u) + b^k u_k, \quad \dot{z} = F_{h(x), u, u_k}(z)$$

(where  $f^k$  and  $b^k$  have been defined in (II.1.b)) satisfies for any  $(x_0, z_0) \in X \times R^n$ , any  $u_0(t)$   $k$  times differentiable with all the  $k$  first derivatives  $(u_1(t), \dots, u_{k-1}(t), u_k(t))$  bounded by  $B$ , for any time  $t \geq 0$ ,

$$\lim_{t \rightarrow +\infty} \|\phi(x(u, u_k, t, x_0), u(t)) - Z(y, u, u_k, t, z_0)\| = 0,$$

where  $x(u, u_k, t, x_0)$  and  $Z(y, u, u_k, t, z_0)$  denote the solutions of (IV.0) at time  $t$ , associated to the control  $(u, u_k)$  and the initial conditions  $(x_0, z_0)$ .

Since  $X$  is compact, the projection set  $E_t$  of  $Z_t = Z(y, u, u_k, t, z_0)$  on  $\phi(X, u(t))$  is well defined:

$$E_t = \{x^* \in \phi(X, u(t)) \mid \|Z_t - x^*\| = \inf_{x \in X} \|Z_t - \phi(x, u(t))\|\}.$$

$(\phi_{u(t)})^{-1}(E_t)$  estimates  $x(t)$  asymptotically in the sense that, for any metric  $\delta$  on  $X$  (compatible with the topology of  $X$ ):

$$\lim_{t \rightarrow +\infty} \delta(x(t), (\phi_{u(t)})^{-1}(E_t)) = 0$$

(by compactness of  $X, I$  and boundedness of  $(u_1, \dots, u_{k-1})$ ).

**Remark.** The practical interest of an observer is that it allows asymptotic estimation of the state  $x(t)$  of the system  $\Sigma$  on the basis of the knowledge of the outputs and the inputs, without differentiating the outputs. Differentiating inputs has physical meaning since usually, the inputs (with their derivatives) are chosen by the user and therefore are perfectly known.

Our results are the following, setting  $I_B = \{v \in R^{d_u} \mid \|v\| \leq B\}$ :

**THEOREM IV.1.** *If  $R_{k,f}^*(h)$  is an embedding for some bound  $B > 0$  on the derivatives of the control,*

$$R_{k,f}^*(h) : X \times I^{d_u} \times I_B^{(k-1)d_u} \rightarrow R^{kd_y} \times R^{kd_u},$$

*there does exist a  $C^k$  observer for  $\Sigma$  (relative to  $B$ ), with  $n = kd_y$ .*

I.e. since in section III we have shown that, for any bound  $B$ , the set of  $\Sigma \in (S^r)$  such that  $R_{k,f}^*(h)$  is an embedding (for  $k$  large enough) is open, dense, *this proves our main theorems II.4–5.*

Our proof, moreover, is *constructive*. It is a slight improvement of the “High-gain” technique developed in our papers [GHO], [GHK], [GK]. Moreover, the estimation rate can be made exponential arbitrary, i.e.:

For any  $\theta$  real positive, one can construct an observer such that

$$(IV.1) \quad \|\phi(x(u, u_k, t, x_0), u(t)) - Z(y, u, u_k, t, z_0)\| \leq ke^{-\theta t}.$$

Or, for any Riemann metric  $\delta$  on  $X$ ,

$$(IV.2) \quad \delta(x(t), (\phi_{u(t)})^{-1}(E_t)) \leq k\delta e^{-\theta t}.$$

**IV.2. Proof of theorem IV.2 and construction of the observer.** We consider  $R_{k,f}(h) : X \times I^{d_u} \times R^{(k-1)d_u} \rightarrow R^{kd_y}$ , which is assumed to be such that  $R_{k,f,u}(h) = R_{k,f}(h)(x, u)$  is an embedding for each  $u = (u_0, \dots, u_{k-1}) \in I^{d_u} \times R^{(k-1)d_u}$  (possibly for  $u \in I^{d_u} \times I_B^{(k-1)d_u}$  only).

We consider the maps  $\phi^i = (L_{f^k})^k h^i$  and  $\psi^{i,j} = (L_{b^k j})(L_{f^k})^{k-1} h^i$ , where  $h^i$  is any component of  $h$ ,  $i = 1, \dots, d_y$ :

$$\phi^i, \psi^{i,j} : X \times I^{d_u} \times R^{(k-1)d_u} \rightarrow R.$$

$\Theta$  being one of these maps, consider

$$\bar{\Theta} : W = R_{k,f}^*(h)(X \times I^{d_u} \times R^{(k-1)d_u}) \subset R^{kd_y} \times R^{kd_u} \rightarrow R,$$

$$\Theta(x, u) = \bar{\Theta} \circ (R_{k,f}^*(h))(x, u)$$

(recall that  $R_{k,f}^*(h)(x, u) = (R_{k,f}(h)(x, u), u)$ ).

$\bar{\Theta}$  is well defined since  $R_{k,f}^*(h)$  is an embedding. Since  $W$  is closed, by the Whitney extension theorem (see [AR], p. 120),  $\bar{\Theta}$  can be extended to a  $C^r$  map defined on all of  $R^{kd_y} \times R^{kd_u}$ .

Consider  $W_B$ , the closed subset of  $W$  obtained when the derivatives  $u_1, \dots, \dots, u_{k-1}$  of the control are restricted to be bounded by  $B$ . Then  $\bar{\Theta}$  can be made globally Lipschitz on all of  $R^{kd_y} \times R^{kd_u}$  just by multiplication by a  $C^\infty$  function which is equal to 1 on a neighbourhood of  $W_B$  and equal to zero at infinity, without modifying its values on  $W_B$ .

Consider the following system on  $R^{kd_y}$ :

$$\bar{\Sigma}^k : \begin{cases} \dot{z}_i = z_{i+1} \quad (z_i = (z_i^1, \dots, z_i^{d_y})), \quad i \leq k-1, & \text{or} \quad \dot{z} = H(z, u, u_k). \\ \dot{z}_k = \bar{\phi}(z, u) + \bar{\psi}(z, u)u_k \end{cases}$$

$\Sigma^k$  embeds into  $\bar{\Sigma}^k$  in the following sense: for any input  $u(t)$ ,  $k$  times differentiable, such that  $|u(i)(t)| \leq B$ , the trajectories of  $\Sigma^k$  and  $\bar{\Sigma}^k$  are mapped to one another by  $(x(t), u(t)) \rightarrow R_{k,f}(h)(x(t), u(t))$ . (Of course,  $(L_{b^k})(L_{f^k})^l h^i = 0$  when  $l < k-1$ .)

We consider now the candidate “High-gain, Luenberger-type” observer:

$$\dot{\mu} = H(\mu, u, u_k) - K_\lambda(C_\mu - y),$$

where  $K_\lambda \in R^{kd_y} \times \dots \times R^{kd_y}$  ( $d_y$  times),  $K_\lambda = \Delta_\lambda K$ ,  $\Delta_\lambda = \text{block-diag}(\lambda \text{Id}^{d_y}, \dots, \lambda^k \text{Id}^{d_y})$ ,  $\text{Id}^{d_y}$  the identity  $d_y \times d_y$  matrix,  $\lambda$  to be chosen later,  $C = (\text{Id}^{d_y}, 0, \dots, 0)$ ,  $C : R^{kd_y} \times \dots \times R^{kd_y}$  ( $d_y$  times)  $\rightarrow R^{d_y}$ .

One can prove that, provided that  $\lambda$  is chosen large enough and that  $K$  is chosen such that  $A - KC$  is a stable matrix, where  $A$  is the  $d_y$ -antishift matrix:

$$A = \begin{bmatrix} 0 & \text{Id}^{d_y} & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \text{Id}^{d_y} \\ 0 & 0 & \dots & 0 \end{bmatrix}$$

(such a  $K$  exists by observability of the linear pair  $(C, A)$  and by the standard linear theory) the relation IV.1 above holds for arbitrary real  $\theta$ . This will end the proof.

The proof of this last fact is nothing but a restatement in this case of the basic proof of [GHO], [GHK], [GK].

**Appendix.** All the proofs in this appendix are computational. The proof of lemma 4 is not that easy, but it is elementary. Everything is left to the reader.

LEMMA 1. *If*

$$f^k = \frac{\partial}{\partial x_1} + \sum_{j=1}^{d_u} u_{1,j} \sum_{i=1}^{d_x} \Theta_{i,j}(x, u_0) \frac{\partial}{\partial x_i} + \sum_{i=0}^{k-2} \sum_{j=1}^{d_u} u_{i+1,j} \frac{\partial}{\partial u_{i,j}},$$

then

$$(L_{f^k})^r(h) = \frac{\partial^r h}{(\partial x_1)^r} + \Phi_u(h)u_1 + \Psi_u(h),$$

where  $\Phi_u$  is a differential operator in  $x$  of degree  $\leq r$  and  $\Psi_u$  is a differential operator in  $x$  of degree  $< r$ .

LEMMA 2. *Consider  $R_{k,f}(h) : X \times T \times R^{(k-1)d_u} \rightarrow R^{kd_y}$ . Consider  $(x, u_0)$  and let  $\varrho$  be the first integer  $> 0$  such that  $u_\varrho \neq 0$ ,  $(u_1, \dots, u_{\varrho-1}) = 0$ , possibly  $u_0 \neq 0$ . For any integer  $n$ , one has  $(L_{f^k})^n h(x, u_0, \dots, u_{k-1}) = \sum M_{n,s}$ , a sum of monomials with positive integer coefficients, where, writing  $u_i^j$  for the  $j$ th component of  $u_i$ :*

$$M_{n,s} = d_x^\lambda \frac{\partial^m h}{\partial u_0^{i_1} \dots \partial u_0^{i_m}} (V_1, \dots, V_\lambda) u_{\beta_1}^{i_1} \dots u_{\beta_s}^{i_s},$$

$$V_i = (L_f^x)^{j_{r_i,i}} \frac{\partial^{\nu_{r_i,i}}}{\partial u_0^{\alpha_1^{r_i,i}} \dots \partial u_0^{\alpha_{\nu_{r_i,i}}^{r_i,i}}} \dots (L_f^x)^{j_{1,i}} \frac{\partial^{\nu_{1,i}}}{\partial u_0^{\alpha_1^{1,i}} \dots \partial u_0^{\alpha_{\nu_{1,i}}^{1,i}}} f$$

with

$$n = \lambda + \sum_{i=1}^{\lambda} \sum_{t=1}^{r_i} j_{t,i} + \sum_{i=1}^s \beta_i, \quad s = m + \sum_{i=1}^{\lambda} \sum_{t=1}^{r_i} \nu_{t,i}.$$



$L_f^x g$  denotes, in local coordinates, the Lie derivative of each component of the vector field  $g$  with respect to  $f$ .

LEMMA 3. Under the assumptions and notations of lemma 2, assume that  $(x, u_0) \in E_f$ . One has

$$\begin{aligned} d_x(L_f k)^n h &= Z_n \left( d_x^{*r_1} h, d_x^{*r_2} \frac{\partial h}{\partial u_0}, \dots, d_x^{*r_{q-1}} \frac{\partial^{q-1} h}{\partial u_0^{q-1}} \right) \\ &+ \sum_{i>0} \delta_{i,n} d_x \frac{\partial^q h}{(\partial u_0^{i_1,i})^{q_{1,i}} \dots (\partial u_0^{i_r,i})^{q_{r,i}}} d_x((L_f^x)^{n-q\varrho-1-i} f) \\ &+ \sum_{i_1, \dots, i_r=1}^{d_u} d_x \frac{\partial^q h}{(\partial u_0^{i_1})^{s_1} \dots (\partial u_0^{i_r})^{s_r}} d_x((L_f^x)^{n-q\varrho-1} f)(u_{\varrho}^{i_1})^{t_1} \dots (u_{\varrho}^{i_r})^{t_r} \end{aligned}$$

with  $q_{1,i} + \dots + q_{r,i} = q, s_1 + \dots + s_r = q, t_1 + \dots + t_r = q$  and with  $q = [n/\varrho]$ , the integer part of  $n/\varrho$ ;  $r_i \leq n, Z_n = \delta_{i,n} = 0$  if  $n < \varrho$ . Here  $d_x^{*r_1} h$  denotes the set of derivatives of  $h$  with respect to  $x$ , from order 0 to order  $r_1$ .

This, in particular, shows that  $d_x(L_f k)^n h(x, u_0, \dots, u_{k-1})$  depends only on the  $s$ th jets with respect to  $x$  of  $\partial^t h / \partial u_0^t$ , with  $t \leq q = [n/\varrho]$  and  $s(q) = 1, s(t) = n$  for  $t < q$ .

There are monomials corresponding to  $\beta_1 = \dots = \beta_s = \varrho$ , they correspond to the higher order Lie derivative with respect to  $f$ :

$$(AL3) \quad d_x \frac{\partial^q h}{(\partial u_0^{i_1})^{s_1} \dots (\partial u_0^{i_r})^{s_r}} d_x((L_f^x)^{n-q\varrho-1} f)(u_{\varrho}^{i_1})^{t_1} \dots (u_{\varrho}^{i_r})^{t_r},$$

and other terms for  $\beta_i > \varrho$ :

$$d_x \frac{\partial^q h}{(\partial u_0^{i_1})^{s_1} \dots (\partial u_0^{i_r})^{s_r}} d((L_f^x)^{n-q\varrho-1-i} f)(u_{\beta_1}^{i_1})^{t_1} \dots (u_{\beta_r}^{i_r})^{t_r}, \quad i > 0.$$

The term (AL3) is also:

$$(AL3.1) \quad d_x(L_f^x)^{n-q\varrho} \frac{\partial^q h}{(\partial u_0^{i_1})^{s_1} \dots (\partial u_0^{i_r})^{s_r}} (u_{\varrho}^{i_1})^{t_1} \dots (u_{\varrho}^{i_r})^{t_r}.$$

LEMMA 4. Let  $R^{(d_x d_y)}$  be the set of  $d_y$ -tuples of vectors of  $R^{d_x}$ . We consider  $R^N$  for some fixed  $N$  and also some fixed algebraic expressions  $Z_{i,j}(V_0, \dots, V_{j-1}, W), V_i \in R^{(d_x d_y)}, W \in R^N, Z_j \in R^{(d_x d_y)}, t_{i,j}^k(W) \in R$ . Let  $h, s$  be also fixed integers and  $A$  a fixed  $d_x \times d_x$  matrix;  $C$  denotes a nonzero projective linear form on  $R^{d_x}$ .

Let  $\mathbf{W}_s^h$  be the set of  $(C, V_0, \dots, V_{s-1}, W), V_i \in R^{(d_x d_y)}, W \in R^N$  such that:

- (i)  $V_0$  is a  $d_y$ -tuple of cyclic vectors for  $A$ ;
- (ii)  $CA^i V_0 = 0, i = 0, \dots, h - 1$ ;
- $C(Z_{1,j}(V_0, \dots, V_{j-1}, W) + V_j) = 0,$
- $C(Z_{2,j}(V_0, \dots, V_{j-1}, W) + t_{2,1}^j V_j + AV_j) = 0, \dots,$

$$C(Z_{h,j}(V_0, \dots, V_{j-1}, W) + t_{h-1,1}^j V_j + \dots + t_{h-1,h-2}^j A^{h-2} V_j + A^{h-1} V_j) = 0;$$

$$j = 1, \dots, s-1.$$

Let  $W_s^h$  be the projection in the direction of  $C$  of  $\mathbf{W}_s^h$ , i.e.

$$W_s^h = \{(V_0, \dots, V_{s-1}, W) \mid \exists C \neq 0, (C, V_0, \dots, V_{s-1}, W) \in \mathbf{W}_s^h\}.$$

Then  $W_s^h$  is a semialgebraic subset of  $R^{(sd_y d_x)} \times R^N$  of codimension  $(d_y sh) - d_x + 1$ .

### References

- [A] D. Aeyels, *Generic observability of differentiable systems*, SIAM J. Control Optim. 19 (1981), 1–15.
- [AR] R. Abraham and J. Robbin, *Transversal Mappings and Flows*, Benjamin, 1967.
- [GHK] J. P. Gauthier, H. Hammouri and I. Kupka, *Observers for nonlinear systems*, IEEE CDC Conference, Brighton, 1991, 1483–1489.
- [GHO] J. P. Gauthier, H. Hammouri and S. Othman, *A simple observer for nonlinear systems, application to bioreactors*, IEEE Trans. Automat. Control 37 (1992), 875–880.
- [GK] J. P. Gauthier and I. Kupka, *Observability and observers for nonlinear systems*, SIAM J. Control Optim. 32 (1994), 975–995.
- [HG1] H. Hammouri and J. P. Gauthier, *Bilinearization up to output injection*, Systems Control Letters 11 (1988), 139–149.
- [HG2] —, —, *Global time varying linearization up to output injection*, SIAM J. Control 6 (1992), 1295–1310.
- [HIR] M. W. Hirsch, *Differential Topology*, Springer, 1976.
- [KI] A. Krener and A. Isidori, *Linearization by output injection and nonlinear observers*, Systems Control Letters 3 (1983), 47–52.
- [KR] A. Krener and W. Respondek, *Nonlinear observers with linearizable error dynamics*, SIAM J. Control Optim. 23 (1985), 197–216.
- [L] S. Lojasiewicz, *Triangulation of semi-analytic sets*, Ann. Scuola Norm. Sup. PISA, 1964, 449–474.
- [LU] D. G. Luenberger, *Observers for multivariable systems*, IEEE Trans. Automat. Control 11 (1966), 190–197.
- [S] H. Sussmann, *Single input observability of continuous time systems*, Math. Systems Theory 12 (1979), 263–284.
- [T] F. Takens, *Detecting strange attractors in turbulence*, in: Dynamic Systems and Turbulence, Warwick 1980, Springer, Berlin, 1981, 366–381.
- [TC] K. Tchon, *On solvability of several affine systems*, Systems Control Letters 4 (1984), 373–379.
- [TOU] J. C. Tougeron, *Idéaux de fonctions différentiables*, Springer, 1972.
- [WH] H. Whitney, *Analytic extensions of differentiable functions defined in closed sets*, Trans. Amer. Math. Soc. 36 (1934), 63–89.