Journal of Mathematical Systems, Estimation, and Control
Vol. 5, No. 3, 1995, pp. 1–27

Lowering the Orders of Derivatives of Controls in Generalized State Space Systems*

E. Delaleau W. Respondek†

Abstract

In this paper we study the problem of lowering the orders of input derivatives in generalized state variables representations of nonlinear systems by generalized state transformations. We give necessary and sufficient conditions for the local and global existence of such transformations. Our conditions are expressed as commutativity of certain vector fields defined in terms of prolonged dynamics. These conditions are restrictive, thereby implying that removing all input derivatives is often impossible in contrast with the linear case where transformations into Kalman dynamics (not involving any input derivatives) always exist. We also consider the problem of lowering the orders of input derivatives in the dynamics and output equations. Our results are illustrated with an engineering example of a crane.

Key words: generalized dynamics, generalized state transformations, input derivatives, Kalman state, prolonged vector fields
AMS subject classifications: 93C10, 93B17, 93B29, 58A30

1 Introduction

In 1977 Williamson [32] studied the problem of designing observers for bilinear systems and he considered transformations depending on derivatives of control. Since then derivatives of controls have been considered in many control problems, like state transformations and nonlinear observers [34, 19, 1, 27], inversion (both linear and nonlinear) [29, 30, 8, 25], canonical forms [27, 10], identification [2] and, recently, equivalence [18, 26]. A

*Received January 23, 1993; received in final form October 21, 1993. Summary appeared in Volume 5, Number 3, 1995.
†The work of W. Respondek was performed when visiting the Laboratoire des Signaux et Systèmes, CNRS–ESE, Gif-sur-Yvette (under support of the French Ministry of Research and Technology).

systematic study of linear and nonlinear systems whose dynamics (and outputs) depend on derivatives of control has been carried out by Fliess in a series of papers [8]–[11] using differential algebra[1]. This approach has led to a new insight into system theory and to a better understanding of control problems such as invertibility [8] (compare also [25]), canonical forms [10] and structure of linear systems [11]. The appearance of input derivatives has been confirmed by some practical studies [32, 15].

According to Fliess [9], a general state variables representation of nonlinear control systems is of the form

$$F(\dot{x}, x, u, \dot{u}, \ldots, u^{(\alpha)}) = 0, \tag{1.1}$$
$$H(y, x, u, \dot{u}, \ldots, u^{(\beta)}) = 0, \tag{1.2}$$

where $u = (u_1, \ldots, u_m)$ and $y = (y_1, \ldots, y_p)$ are respectively the input and the output, and $x = (x_1, \ldots, x_n)$ is the state of the system. Moreover, transformations between two different states x and $\tilde{x}$ may also depend on the input and a finite number of its derivatives and are given by

$$\Phi(x, \tilde{x}, \dot{u}, \ldots, u^{(\gamma)}) = 0.$$

Although, as we mentioned above, the general description (1.1)–(1.2) has been confirmed by some theoretical and practical studies, the classical explicit representation (without derivatives of the input) of the form

$$\dot{x} = f(x, u), \tag{1.3}$$
$$y = h(x, u), \tag{1.4}$$

is still very common and has a lot of advantages. Therefore it is very natural to describe those general nonlinear systems of the form (1.1)–(1.2), explicit with respect to $\dot{x}$ and y, which can be transformed via general state space transformations to (1.3)–(1.4).

This problem was studied by Freedman and Willems [16] in the case where the first derivatives of controls appear. They gave also a stochastic interpretation of the problem. Glad [17] observed that a necessary and sufficient condition to remove $u^{(\alpha)}$, where α is the highest order of derivation of the control variable, is that it appears linearly. Following Glad, the first author gave necessary and sufficient conditions in the multi-input case to lower every α_i by one [4], where α_i is the highest order of derivation of the input u_i. In the present paper we study and solve the problem in its full generality. Namely, given any m-tuple $(\beta_1, \ldots, \beta_m)$ we provide necessary and sufficient conditions for the existence of a (local) generalized state transformation which brings the system into a representation having β_i, $i = 1, \ldots, m$, as the highest input derivatives orders. In particular we

[1] See also [13, 14] for a differential geometric approach.

describe systems for which we can remove all inputs derivatives and thus transform the system into the classical form (1.3)–(1.4). We provide also conditions for global transformations to exist.

Although systems with derivatives of controls, as well as generalized state space transformations, appear naturally in differential algebraic approach (see *e.g.* [7] for rational realization of nonlinear systems), the question of lowering orders of inputs derivatives (removing inputs derivatives) by such transformations is an integrability problem for a system of PDE's and therefore differential geometric framework fits naturally. We extend (prolong) the system and we state the integrability conditions in terms of commutativity of certain vectors fields defined by the extension.

The paper is organized as follows. In Section 2 we state the problem and give the main results, local and global, on lowering the inputs derivatives orders. In Section 3 we consider linear systems and show how our result rediscover the known possibility of removing all input derivatives in the linear case [11]. In this section we also relate our results to the state space linearization problem, where commutativity is crucial as well. In Section 4 we consider the problem of removing input derivatives in the output equations and the dynamics. We also describe some connections of our results with the procedure of realization of input-output differential equations given by van der Schaft [28]. In Section 5 we provide a physical example of crane.

Some results of this paper were announced in [4, 5].

2 Lowering the Orders of Derivatives of Inputs in the Dynamics

Throughout this paper we will use a convenient notation to denote each control variable u_i and its time derivatives up to the order $\delta_i \in I\!N$. Namely

$$u_i^{\langle \delta_i \rangle} = (u_i, \dot{u}_i, \ldots, u_i^{(\delta_i)}),\ i = 1, \ldots, m.$$

Consider an explicit generalized state representation Σ of a multi-input nonlinear dynamics in which each input $u_i,\ i = 1, \ldots, m$, appears with $\alpha_i \geq 0$ as derivation order

$$\Sigma: \qquad \dot{x} = f(x, u_1^{\langle \alpha_1 \rangle}, \ldots, u_m^{\langle \alpha_m \rangle}). \tag{2.1}$$

The state x of this system evolves on a $\mathcal{C}^\infty$-smooth n-dimensional manifold denoted by $\mathcal{X}$ and f is $\mathcal{C}^\infty$-smooth with respect to all its arguments.

Our purpose is to derive conditions which guarantee the existence of generalized coordinates transformations ψ, in the state space, of the form

$$\tilde{x} = \psi(x, u_1^{\langle \alpha_1 - 1 \rangle}, \ldots, u_m^{\langle \alpha_m - 1 \rangle}), \tag{2.2}$$

which lead to a state representation

$$\tilde{\Sigma}: \qquad \dot{\tilde{x}} = \tilde{f}(\tilde{x}, u_1^{\langle\beta_1\rangle}, \ldots, u_m^{\langle\beta_m\rangle}), \tag{2.3}$$

where $\beta_i \leq \alpha_i$ for all $i = 1, \ldots, m$. We look for ψ which is $\mathcal{C}^\infty$-smooth with respect to all its arguments and (locally) invertible with respect to x. Observe that we keep the controls and their time derivatives invariant. In this section we consider dynamics only; systems equipped with outputs will be considered in Section 4.

A natural framework to investigate this problem is differential geometry and so we will work in an *extended state space* denoted by $\mathcal{S}$. The *extension of the state* is made by association of new coordinates $z_{i,j}$ with inputs and their derivatives up to the maximal order appearing in the original state representation,

$$z_{i,j} = u_i^{(j)}, \; i = 1, \ldots, m, \; j = 0, \ldots, \alpha_i. \tag{2.4}$$

Put $K = \sum_{i=1}^m (\alpha_i + 1)$, thus $\mathcal{S} = I\!R^K \times \mathcal{X}$ and the dimension of the extended state space[2] is $N = K + n$.

We rewrite Σ and $\tilde{\Sigma}$ on the extended state space $\mathcal{S}$ respectively as

$$\Sigma^e \quad \left\{ \begin{array}{rcl} \dot{x} & = & f(x, z_{1,0}, \ldots, z_{m,\alpha_m}) \\ \dot{z}_{i,0} & = & z_{i,1} \\ & \vdots & \\ \dot{z}_{i,\alpha_i-1} & = & z_{i,\alpha_i} \\ \dot{z}_{i,\alpha_i} & = & u_i^{(\alpha_i+1)} \end{array} \right. \qquad i = 1, \ldots, m, \tag{2.5}$$

and

$$\tilde{\Sigma}^e \quad \left\{ \begin{array}{rcl} \dot{\tilde{x}} & = & \tilde{f}(\tilde{x}, z_{1,0}, \ldots, z_{m,\beta_m}) \\ \dot{z}_{i,0} & = & z_{i,1} \\ & \vdots & \\ \dot{z}_{i,\alpha_i-1} & = & z_{i,\alpha_i} \\ \dot{z}_{i,\alpha_i} & = & u_i^{(\alpha_i+1)} \end{array} \right. \qquad i = 1, \ldots, m. \tag{2.6}$$

On the extended state space $\mathcal{S}$ we will consider a finite prolongation of the vector field f (resp. $\tilde{f}$), denoted by $\mathcal{F}$ (resp. $\tilde{\mathcal{F}}$), given in local coordinates by

$$\begin{array}{rcl} \mathcal{F} & = & \displaystyle\sum_{i=1}^n f_i \frac{\partial}{\partial x_i} + \sum_{i=1}^m \sum_{j=0}^{\alpha_i-1} z_{i,j+1} \frac{\partial}{\partial z_{i,j}}, \\ (\text{resp. } \tilde{\mathcal{F}} & = & \displaystyle\sum_{i=1}^n \tilde{f}_i \frac{\partial}{\partial \tilde{x}_i} + \sum_{i=1}^m \sum_{j=0}^{\alpha_i-1} z_{i,j+1} \frac{\partial}{\partial z_{i,j}}). \end{array} \tag{2.7}$$

[2]In some cases it is more natural to consider the extended state space $\mathcal{S} = \mathcal{R}^K \times \mathcal{X}$, where $\mathcal{R}^K$ is an open subset of $I\!R^K$.

Let x^e (resp. $\tilde{x}^e$) denote local coordinates of $\mathcal{S}$, *i.e.*,

$$\begin{aligned} x^e &= (z_{i,j}, x_1, \ldots, x_n,\ i = 1, \ldots, m,\ j = 0, \ldots, \alpha_i) \\ (\text{resp. } \tilde{x}^e &= (z_{i,j}, \tilde{x}_1, \ldots, \tilde{x}_n,\ i = 1, \ldots, m,\ j = 0, \ldots, \alpha_i)). \end{aligned}$$

Σ^e (resp. $\widetilde{\Sigma}^e$) defines on the extended state space $\mathcal{S}$ a control-affine system controlled by $v_i = u_i^{(\alpha_i+1)}$, $i = 1, \ldots, m$, in the following way

$$\begin{aligned} \dot{x}^e &= \mathcal{F} + \sum_{i=1}^{m} v_i \, \frac{\partial}{\partial z_{i,\alpha_i}} \qquad (2.8) \\ (\text{resp. } \dot{\tilde{x}}^e &= \widetilde{\mathcal{F}} + \sum_{i=1}^{m} v_i \, \frac{\partial}{\partial z_{i,\alpha_i}}), \end{aligned}$$

where $\mathcal{F}$ (resp. $\widetilde{\mathcal{F}}$) is given by (2.7). We keep the same name Σ^e (resp. $\widetilde{\Sigma}^e$) for (2.8) because the only difference between (2.5) and (2.8) is to interpret $u_i^{(\alpha_i+1)}$ as the new controls v_i, $i = 1, \ldots, m$.

Observe that (2.2) can be interpreted as a (local) transformation Ψ of $\mathcal{S}$ having the special structure $\Psi(z, x) = (z, \psi(z, x))$, where z denotes global coordinates $z_{i,j}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_j$, of $\mathbb{R}^K$. Preserving the z-coordinates simply means that we do not change u and its derivatives.

It should be noticed that the control vector fields $\frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, are invariant under ψ which does not depend on z_{i,α_i}, $i = 1, \ldots, m$, and therefore they are intrinsically defined in our problem. It turns out that all informations needed to solve the problem are contained in the Lie algebra generated by $\mathcal{F}$ and the $\frac{\partial}{\partial z_{i,\alpha_i}}$'s.

Observe that working locally at z_0 we consider controls which are close to the nominal one corresponding to z_0 via (2.4). A solution for the local version of our problem can be stated as follows.

Theorem 1 *A generalized change of state coordinates ψ of the form* (2.2)*, transforming the representation Σ into $\widetilde{\Sigma}$, exists, locally around*

$$s_0 = (z_0, x_0) \in \mathcal{S},$$

if, and only if,

$$\left[\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}, \mathrm{ad}_{\mathcal{F}}^{r} \frac{\partial}{\partial z_{j,\alpha_j}} \right] = 0 \qquad (2.9)$$

in a neighbourhood of s_0 for any $1 \leq i, j \leq m$ and

$$\begin{aligned} 0 \leq\ & q \leq \alpha_i - \beta_i, \\ 0 \leq\ & r \leq \alpha_j - \beta_j. \end{aligned}$$

Remark 1 *A generalized change of coordinates solving the problem is constructed in the sufficiency part of the proof of this theorem. An example of this construction is given in Section 5.*

Remark 2 *An equivalent way to study the problem of lowering the inputs derivatives orders is to use the infinite prolongation $\mathcal{F}^\infty$ of f defined by (see* [22]*)*

$$\mathcal{F}^\infty = \sum_{k=1}^{n} f_k \frac{\partial}{\partial x_k} + \sum_{j=0}^{\infty} \sum_{i=1}^{m} z_{i,j+1} \frac{\partial}{\partial z_{i,j}}, \tag{2.10}$$

where $z_{i,j} = u_i^{(j)}$, $i = 1, \ldots, m$, $j \geq 0$ (compare (2.4)*). It is immediate to see that* (2.9) *is equivalent to its modification with $\mathcal{F}$ being replaced by $\mathcal{F}^\infty$. In such a modification the infinite sum in* (2.10) *can be treated only formally, since all vector fields $\mathrm{ad}^q_{\mathcal{F}^\infty} \frac{\partial}{\partial z_{i,\alpha_i}}$ (and any of their Lie brackets) depend on a finite number of $z_{i,j}$'s and thus can be easily computed. In* [4] *the condition for lowering the inputs derivatives orders by one (see Corollary 2 below) is given in terms of $\mathcal{F}^\infty$. We want to add that $\mathcal{F}^\infty$ can be given a precise interpretation of a smooth vector field on the infinite-dimensional manifold $J_0^\infty(\mathbb{R}, \mathbb{R}^m) \times \mathcal{X}$, where $J_0^\infty(\mathbb{R}, \mathbb{R}^m)$ denotes the space of infinite jets at zero of $\mathbb{R}^m$-valued smooth functions* [20] *(see, also,* [13, 14, 18]*).*

We state as separate corollaries the two extreme cases, *i.e.*, when we are able to remove all input derivatives and when we are able to lower the order of derivative of every input by one.

Corollary 1 *A generalized change of state coordinates ψ of the form* (2.2)*, transforming Σ into $\widetilde{\Sigma}$, a system without derivatives of control, of the form*

$$\widetilde{\Sigma}: \; \dot{\tilde{x}} = f(\tilde{x}, u),$$

exists if, and only if, Σ satisfies the commutativity condition (2.9) *for $0 \leq q \leq \alpha_i$, $0 \leq r \leq \alpha_j$.*

Corollary 2 *There exists a generalized change of state coordinates ψ of the form* (2.2)*, lowering the highest order of every input derivative of Σ by one, if, and only if, the condition* (2.9) *is satisfied with $0 \leq q, r \leq 1$.*

It follows from Corollary 1 that not for all generalized systems Σ we are able to remove all input derivatives using transformations of the form (2.2). Actually, the restrictive commutativity condition (2.9) implies that this is a very special case. This purely nonlinear phenomenon (compare [11] and Section 3) was observed for the first time and studied by Freedman and Willems [16] (see also [17] and, in the context of state space realization, [28]).

When we compare Corollary 2 with the original statement of Delaleau [4] (given in terms of infinite prolongation $\mathcal{F}^\infty$, compare Remark 2 following Theorem 1), we see that in [4] necessary and sufficient conditions are expressed as (2.9) given for $q = r = 1$ together with the condition that the original dynamics Σ are affine with respect to the highest input derivatives. The latter being necessary conditions for lowering the order of input derivatives [17, 4], can be expressed in an invariant way as

$$\left[\mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_{i,\alpha_i}}, \frac{\partial}{\partial z_{j,\alpha_j}}\right] = 0, \ 1 \le i, j \le m. \tag{2.11}$$

Observe that in the scalar input case, (2.11) takes the form

$$\left[\mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_\alpha}, \frac{\partial}{\partial z_\alpha}\right] = 0, \tag{2.12}$$

where α is the input derivative order and it is the only necessary and sufficient condition for lowering the order of the input derivative by one [17]. Indeed, the condition $\left[\mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_\alpha}, \mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_\alpha}\right] = 0$ (compare Corollary 2) holds automatically. Thus, if a scalar input dynamics are affine with respect to the highest derivative of control variable, then it is always possible to lower its order by one.

Notice that there is no natural order for different collections of β_i's giving the input derivatives orders of the transformed system. Clearly, the best case is $\beta_i = 0, \ i = 1, \dots, m$ (corresponding to removing all inputs derivatives), while the worst one is $\beta_i = \alpha_i, \ i = 1, \dots, m$ (meaning that none of the orders can be lowered). Next to it, however, there are m incomparable cases, $\beta_j = \alpha_j - 1, \ \beta_i = \alpha_i, \ i \neq j$, *i.e.*, when we are able to lower the highest order of derivation of the j^{th} input by one. Clearly, this is possible if, and only if, (compare with (2.12))

$$\left[\mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_{j,\alpha_j}}, \frac{\partial}{\partial z_{i,\alpha_i}}\right] = 0, \ i = 1, \dots, m.$$

Observe that it can happen that we are able to lower the orders of derivation of the j^{th} and k^{th} inputs but, if $\left[\mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_{j,\alpha_j}}, \mathrm{ad}_{\mathcal{F}} \frac{\partial}{\partial z_{k,\alpha_k}}\right] \neq 0$, not both of them simultaneously. In this case there is no collection of β_i's such that $\beta_j \le \alpha_j - 1$ and $\beta_k \le \alpha_k - 1$.

To summarize, given Σ, in general there is no m-tuple $(\beta_1, \dots, \beta_m)$ giving minimal possible orders of derivatives of all inputs simultaneously. For every choice of $\beta_1, \dots, \beta_m$ we can check, using Theorem 1, whether one can lower the input derivatives respectively by $\alpha_i - \beta_i, \ i = 1, \dots, m$, but there is no evident optimal choice. See Section 5 for an illustration of this phenomenon.

Proof of Theorem 1: *(Necessity)* Calculate the prolonged vector field, given by (2.7), associated with the final representation of the system $\widetilde{\Sigma}$. For any k the component $\tilde{f}_k$ does not depend on $z_{i,j}$, $i = 1,\dots,m$, $j = \beta_i + 1,\dots,\alpha_i$, and thus $\mathrm{ad}^q_{\mathcal{F}} \frac{\partial}{\partial z_{i,\alpha_i}} = (-1)^q \frac{\partial}{\partial z_{i,\alpha_i - q}}$, $q = 0,\dots,\alpha_i - \beta_i$. Therefore the condition (2.9) of Theorem 1 holds. Moreover, this condition is invariant under generalized change of coordinates of the form (2.2) and thus it is necessary for the solvability of the problem.

(Sufficiency) For notational convenience we put

$$\begin{array}{lcll} X_i & = & \frac{\partial}{\partial x_i}, & i = 1,\dots,n, \\ Y_{i,j} & = & (-1)^{j-1}\mathrm{ad}^{j-1}_{\mathcal{F}} \frac{\partial}{\partial z_{i,\alpha_i}}, & i = 1,\dots,m,\ j = 1,\dots,\alpha_i - \beta_i + 1, \\ Z_{i,j} & = & \frac{\partial}{\partial z_{i,j-1}}, & i = 1,\dots,m,\ j = 1,\dots,\beta_i, \end{array}$$

where $\beta_i \le \alpha_i$, $i = 1,\dots,m$. Set $L = \sum_{i=1}^m (\alpha_i - \beta_i + 1)$ and $M = \sum_{i=1}^m \beta_i$ (one has $K = L + M$). The X_i's are the n unit vector fields corresponding to the original coordinates of state space $\mathcal{X}$, the $Y_{i,j}$'s are the L vector fields which are involved in condition (2.9) and the $Z_{i,j}$'s are the M vector fields corresponding to the control variables and their time derivatives up to the order[3] $\beta_i - 1$. They form a set of $N = L + M + n$ vector fields on the extended state space. We will denote them by $W_1,\dots,W_N$ and order as follows: the first L of them are the $Y_{i,j}$'s (ordered lexicographically), followed by the $Z_{i,j}$'s (ordered lexicographically) and finally followed by the X_i's.

The extension procedure (2.4) permits to define coordinates $z_{i,j}$ on the extended state space. In most places it is more convenient to work with double indices (i,j) for the z-coordinates but we have also to use one index and to establish a correspondence between them. Observe that $Y_{i,j} = \frac{\partial}{\partial z_{i,\alpha_i - j + 1}} \pmod{T\mathcal{X}}$, $j = 1,\dots,\alpha_i - \beta_i + 1$ and $Z_{i,j} = \frac{\partial}{\partial z_{i,j-1}}$. Therefore to every W_k, $k = 1,\dots,K$ we can associate a unique vector field $\frac{\partial}{\partial z_{i,j}}$, $i = 1,\dots,m$, $j = 0,\dots,\alpha_i$ such that $W_k = \frac{\partial}{\partial z_{i,j}} \pmod{T\mathcal{X}}$. This correspondence between the W_k's and the $\frac{\partial}{\partial z_{i,j}}$'s allows us to put $\bar{z}_k = z_{i,j}$. We will speak about z-coordinates when using the $z_{i,j}$'s and about $\bar{z}$-coordinates when using the $\bar{z}_k$'s.

For a vector field V on $\mathcal{S}$ we will denote by $\phi^t_V(s)$ its flow, *i.e.*, the solution of the differential equation $\frac{d}{dt}\ \phi^t_V(s) = V(\phi^t_V(s))$ passing through s at $t = 0$. For each s, $t \longmapsto \phi^t_V(s)$ is a curve defined for t in some open interval depending on s. For each t the map $s \longmapsto \phi^t_V(s)$ is a smooth diffeomorphism.

The generalized change of state coordinates ψ is defined as the restriction, to the original state space $\mathcal{X}$, of the (translated) inverse of composition

[3]Notice that there is no vector field $Z_{i,j}$ when $\beta_i = 0$.

of flows of vector fields on $\mathcal{S}$. For an initial condition $s_0 = (\bar{z}_0, x_0)$, consider the map $\Phi_{s_0} : \mathcal{V}_0 \longrightarrow \mathcal{S}$ defined in a neighbourhood $\mathcal{V}_0$ of $0 \in \mathbb{R}^N$ and given for any $(t_1, \ldots, t_N) \in \mathcal{V}_0$ by

$$\Phi_{s_0}(t_1, \ldots, t_N) = \phi^{t_1}_{W_1}(\phi^{t_2}_{W_2}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots)).$$

As we already mentioned we want to preserve the $\bar{z}$-coordinates (controls and their derivatives) and thus we put

$$\Psi_{s_0}(\bar{z}, x) = \Phi^{-1}_{s_0}(\bar{z}, x) + (\bar{z}_0, x_0),$$

which is simply a translation of $\Phi^{-1}_{s_0}$, by a constant vector, of the inverse map of Φ_{s_0} (the translation in the x-part is added to preserve x_0). Observe that from the way we order the components $\bar{z}_k$ of $\bar{z}$ it is clear that $\bar{z}_k$ correspond to t_k.

In the sequel, we will omit the explicit reference to s_0 and denote Φ_{s_0} simply by Φ, keeping in mind that Φ is defined as composition of flows at the point s_0. Similarly, we write Ψ instead of Ψ_{s_0}. We will now prove the five following claims:

(i) Ψ defines a local coordinate system at $s_0 = (\bar{z}_0, x_0) \in \mathcal{S} = \mathbb{R}^K \times \mathcal{X}$.

(ii) Ψ preserves each $z_{i,j}$, $i = 1, \ldots, m$, and $j = 0, \ldots, \alpha_i$, *i.e.*, $\Psi(\bar{z}, x) = (\bar{z}, \tilde{x})$ for a suitable $\tilde{x}(\bar{z}, x)$.

(iii) Ψ maps the vector fields $\mathrm{ad}^j_{\mathcal{F}} \frac{\partial}{\partial z_{i,\alpha_i}}$ into the vector fields $\frac{\partial}{\partial z_{i,\alpha_i - j}}$, $i = 1, \ldots, m$, and $j = 0, \ldots, \alpha_i - \beta_i$.

(iv) All components of Ψ, except for those which are identically equal to z_{i,α_i}, $i = 1, \ldots, m$ (compare step (ii)), do not depend on z_{i,α_i}, $i = 1, \ldots, m$.

(v) In new extended state space coordinates given by $\tilde{x}^e = \Psi(x^e)$, Σ^e takes the form $\widetilde{\Sigma}^e$, *i.e.*, its dynamics $\tilde{f}$ do not depend on $z_{i,j}$, $i = 1, \ldots, m$, and $j = \beta_i + 1, \ldots, \alpha_i$.

It is obvious that statements (i), (iii)–(v) can be proved for Φ^{-1}, instead of Ψ, since the two maps differ by a translation by a constant vector. Because of notation convenience we prove them for Φ^{-1}. In the sequel of the proof we set $t = (t_1, \ldots, t_N)$.

Proof of (i)**:** We will first prove that Φ is a local diffeomorphism by checking its rank at $t = 0 \in \mathbb{R}^N$. For a diffeomorphism $\varphi : \widetilde{\mathcal{S}} \longrightarrow \mathcal{S}$ and a vector field W we denote by $\varphi_* W$ the transformed vector field, *i.e.*, $(\varphi_* W)(y) = (D\varphi)_{|_{\varphi^{-1}(y)}} W(\varphi^{-1}(y))$, where $D\varphi$ stands for the jacobian matrix of φ. The partial derivative of Φ with respect to t_i, $i = 1, \ldots, N$, is

$$\begin{aligned} \frac{\partial \Phi}{\partial t_i}(t) &= D(\phi^{t_1}_{W_1}(\ldots(\phi^{t_{i-1}}_{W_{i-1}})\ldots))_{|\omega_i} \frac{\partial}{\partial t_i}(\phi^{t_i}_{W_i}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots)) \\ &= (\phi^{t_1}_{W_1}(\ldots(\phi^{t_{i-1}}_{W_{i-1}})\ldots))_* \, W_i(\Phi(s_0)), \end{aligned}$$

where $\omega_i = \phi^{t_i}_{W_i}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots)$. This gives $\frac{\partial\Phi}{\partial t_i}(0) = W_i(s_0)$. On the other hand it is easy to see that all W_i are independent since

$$Y_{i,j} = (-1)^{j-1}\mathrm{ad}^{j-1}_{\mathcal{F}}\,\frac{\partial}{\partial z_{i,\alpha_i}} = \frac{\partial}{\partial z_{i,\alpha_i-j+1}} \pmod{T\mathcal{X}}. \tag{2.13}$$

Thus Φ is of full rank N at 0 and Φ is a local diffeomorphism. Finally Φ^{-1}, which is a local diffeomorphism too, from a neighbourhood of $s_0 \in \mathcal{S}$ into $\mathbb{R}^N$, defines a local coordinates system at s_0.

Proof of (ii)**:** Let us see how the composition Φ of flows transforms a point $(\bar{z}, x) \in \mathcal{S}$. It is clear that, for all $i = 1, \ldots, n$,

$$\phi^{t_i}_{X_i}(\bar{z}, x) = (\bar{z}, \tilde{x}), \tag{2.14}$$

for a suitable $\tilde{x} \in \mathcal{X}$. Consider now the remaining vector fields $W_1, \ldots, W_{L+M}$. We have $Y_{i,j} = \frac{\partial}{\partial z_{i,\alpha_i-j+1}} \pmod{T\mathcal{X}}$, $j = 1, \ldots, \alpha_i - \beta_i + 1$, (compare (2.13)) and $Z_{i,j} = \frac{\partial}{\partial z_{i,j-1}} \pmod{T\mathcal{X}}$, $j = 1, \ldots, \beta_i$. Therefore, as we already mentioned, to every W_K, $k = 1, \ldots, K$ (recall that $K = L + M$), there corresponds a unique vector field $\frac{\partial}{\partial z_{i,j}}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i$, such that

$$W_k = \frac{\partial}{\partial z_{i,j}} \pmod{T\mathcal{X}}. \tag{2.15}$$

We have $\bar{z}_k = z_{i,j}$. From (2.15), we see that for any $k = 1, \ldots, K$,

$$\phi^{t_k}_{W_k}(\bar{z}, x) = (\tilde{z}, \tilde{x}), \tag{2.16}$$

where $\tilde{z}_i = \bar{z}_i$, $i \neq k$, $\tilde{z}_k = \bar{z}_k + t_k$ and $\tilde{x}$ is a suitable point of $\mathcal{X}$. Combining (2.14) and (2.16), we see that $\Phi_{(\bar{z}_0,x_0)}(t_1, \ldots, t_N) = (\tilde{z}_1, \ldots, \tilde{z}_K, \tilde{x}) \in \mathbb{R}^K \times \mathcal{X}$, where $\bar{z}_i = \bar{z}_{0i} + t_i$, $i = 1, \ldots, K$. Thus Ψ, which differs from Φ^{-1} by the translation by the vector $(\bar{z}_0, x_0)$ preserves z_i, $i = 1, \ldots, K$ and hence preserves all $z_{i,j}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i$.

Proof of (iii)**:** Because of the condition (2.9) the first L vector fields among the W_k's, which correspond to the $Y_{i,j}$'s, are commuting in a neighbourhood of s_0, *i.e.*, $[W_i,\ W_j] = 0$, $1 \le i, j \le L$, and thus their flows commute too [31]. Hence for any $i = 1, \ldots, L$,

$$\begin{aligned}\Phi(t) &= \phi^{t_1}_{W_1}(\phi^{t_2}_{W_2}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots)) \\ &= \phi^{t_i}_{W_i}(\ldots\phi^{t_{i-1}}_{W_{i-1}}(\phi^{t_{i+1}}_{W_{i+1}}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots))\ldots).\end{aligned}$$

Using this fact we obtain, for $i = 1, \ldots, L$,

$$\begin{aligned}\frac{\partial\Phi}{\partial t_i}(t) &= \frac{\partial}{\partial t_i}\phi^{t_i}_{W_i}(\ldots\phi^{t_{i-1}}_{W_{i-1}}(\phi^{t_{i+1}}_{W_{i+1}}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots))\ldots) \\ &= W_i(\phi^{t_i}_{W_i}(\ldots\phi^{t_{i-1}}_{W_{i-1}}(\phi^{t_{i+1}}_{W_{i+1}}(\ldots(\phi^{t_N}_{W_N}(s_0))\ldots))\ldots)) \\ &= W_i(\Phi(t)).\end{aligned}$$

For the remaining $N-L$ vector fields W_i, $i = L+1,\dots,N$, there is no commutativity assumption and we simply get, like in the proof of (i), $\frac{\partial\Phi}{\partial t_i}(t) = \widetilde{W}_i(\Phi(t))$, where $\widetilde{W}_i$ is the vector field W_i transformed by the action of the jacobian matrix corresponding to composition of flows of $W_1,\dots,W_{i-1}$. Thus $\frac{\partial\Phi}{\partial t}(t) = (W_1,\dots,W_L,\widetilde{W}_{L+1},\dots,\widetilde{W}_N)(\Phi(t))$,. Hence we obtain

$$\mathrm{Id} \quad = \quad (\Phi^{-1})_*(W_1,\dots,W_L,\widetilde{W}_{L+1},\dots,\widetilde{W}_N)$$

meaning that $\mathrm{ad}^j_{\mathcal{F}}\frac{\partial}{\partial z_{i,\alpha_i}}$, which is one of the first L vector fields among the W_i's, is mapped by Φ^{-1} into $\frac{\partial}{\partial z_{i,\alpha_i-j}}$, $i=1,\dots,m$ and $j=0,\dots,\alpha_i-\beta_i$.

Proof of (iv)**:** From (iii), we know that $\Phi_*^{-1}(\frac{\partial}{\partial z_{i,\alpha_i}}) = \frac{\partial}{\partial z_{i,\alpha_i}}$ and thus the components of Φ^{-1} do not depend on z_{i,α_i}, $i=1,\dots,m$ (except for those which are identically equal to z_{i,α_i}).

Proof of (v)**:** For any $i=1,\dots,m$ and $j=\beta_i+1,\dots,\alpha_i$ compute

$$\begin{aligned}
[\widetilde{\mathcal{F}}\ ,\tfrac{\partial}{\partial z_{i,j}}] &= [\Phi_*^{-1}(\mathcal{F})\ ,\Phi_*^{-1}(\mathrm{ad}^{\alpha_i-j}_{\mathcal{F}}\tfrac{\partial}{\partial z_{i,\alpha_i}})],\\
&= \Phi_*^{-1}([\mathcal{F}\ ,\mathrm{ad}^{\alpha_i-j}_{\mathcal{F}}\tfrac{\partial}{\partial z_{i,\alpha_i}}]),\\
&= \Phi_*^{-1}(\mathrm{ad}^{\alpha_i-j+1}_{\mathcal{F}}\tfrac{\partial}{\partial z_{i,\alpha_i}}),\\
&= \tfrac{\partial}{\partial z_{i,j-1}},
\end{aligned}$$

where the first and the last equalities follow from (iii), whereas the second one does from a fundamental property of the Lie bracket (see [31]). From the equality $[\frac{\partial}{\partial z_{i,j}},\widetilde{\mathcal{F}}] = \frac{\partial}{\partial z_{i,j-1}}$ we conclude that $\tilde{f}$, which is the $\tilde{x}$-part of $\widetilde{\mathcal{F}}$, does not depend on $z_{i,j}$, $i=1,\dots,m$, $j=\beta_i+1,\dots,\alpha_i$. We have $\Psi(\bar{z},x) = (\bar{z},\tilde{x})$ where $\tilde{x}$, being a function of $(\bar{z},x)$, does not depend on z_{i,α_i}, $i=1,\dots,m$ (compare (iv)). The identity with respect to $\bar{z}$-components follows from (ii).

The generalized change of coordinates ψ in the original state space $\mathcal{X}$ is now easy to define. Namely, put $\psi(\bar{z},x) = \tilde{x}(\bar{z},x)$. Coming back to the original notations we get the desired form (2.2) of

$$\psi = \psi(x,u_1^{\langle\alpha_1-1\rangle},\dots,u_m^{\langle\alpha_m-1\rangle}).$$

Moreover, (v) implies that ψ transforms Σ into $\widetilde{\Sigma}$ yielding the dynamics $\dot{\tilde{x}} = \tilde{f}(\tilde{x},u_1^{\langle\beta_1\rangle},\dots,u_m^{\langle\beta_m\rangle})$ with input derivatives orders lowered respectively by $\alpha_i-\beta_i$, $i=1,\dots,m$. □

We end this section by studying global aspects of the problem of lowering the orders of input derivatives. To formulate the problem, assume that we are given a nonlinear control system Σ, evolving on a smooth manifold $\mathcal{X}$, whose dynamics depend on input derivatives and is given, in local coordinates $x = (x_1, \ldots, x_n)$ of $\mathcal{X}$, by (2.1). Together with Σ, we consider its prolongation Σ^e, whose state space is $\mathcal{S} = I\!R^K \times \mathcal{X}$ and whose dynamics are given by (2.5). If we consider $\mathcal{S}$ as a trivial fibre bundle over $I\!R^K$ then (2.1) describes the dynamics along the fibres.

Now, the problem of global lowering the inputs derivatives orders is to find a global fibre bundle diffeomorphism $\Psi = (\mathrm{id}_{I\!R^K}, \psi)$ of $I\!R^K \times \mathcal{X}$, where ψ does not depend on z_{i,α_i}, $i = 1, \ldots, m$, and Ψ transforms the form of Σ^e given by (2.5) into that of $\widetilde{\Sigma}^e$ given by (2.6). As we already mentioned we do not change the control variables (and their derivatives) and hence we do not transform the base $I\!R^K$.

Having defined the problem of global lowering the orders of inputs derivatives we state and prove our global result.

Theorem 2 *The problem of global lowering the orders of input derivatives, by a global transformation $\Psi = (\mathrm{id}_{I\!R^K}, \psi)$ of Σ^e into $\widetilde{\Sigma}^e$, is solvable if, and only if, Σ^e satisfies the commutativity condition* (2.9) *of Theorem 1 everywhere, for $i, j = 1, \ldots, m$, $0 \le q \le \alpha_i - \beta_j$, $0 \le r \le \alpha_j - \beta_j$, and moreover the vector fields $\mathrm{ad}_{\mathcal{F}}^{j} \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i - \beta_i$, are complete.*

Proof: *(Necessity)* By Theorem 1, the condition (2.9) is necessary for the problem of local lowering the orders of input derivatives orders and so is for the global problem. For $\widetilde{\Sigma}^e$ we have $\mathrm{ad}_{\widetilde{\mathcal{F}}}^{j} \frac{\partial}{\partial z_{i,\alpha_i}} = \frac{\partial}{\partial z_{i,\alpha_i - j}}$, $j = 0, \ldots, \beta_i$ and since the $z_{i,j}$'s form global coordinates on the base $I\!R^K$, the completeness of $\mathrm{ad}_{\mathcal{F}}^{j} \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i - \beta_i$, follows.

(Sufficiency) Recall from the proof of Theorem 1 the notation

$$\begin{array}{rcll} Y_{i,j} & = & (-1)^{j-1} \mathrm{ad}_{\mathcal{F}}^{j-1} \frac{\partial}{\partial z_{i,\alpha_i}}, & i = 1, \ldots, m, \; j = 1, \ldots, \alpha_i - \beta_i + 1, \\ Z_{i,j} & = & \frac{\partial}{\partial z_{i,j-1}}, & i = 1, \ldots, m, \; j = 1, \ldots, \beta_i \end{array}$$

and keep denoting the $Y_{i,j}$'s by $W_1, \ldots, W_L$ and the $Z_{i,j}$'s by $W_{L+1}, \ldots, W_K$, where $K = L + M$. Define a map $\Psi : \mathcal{S} \longrightarrow \mathcal{S}$ by

$$\Psi(t, x) = \phi_{W_1}^{t_1}(\phi_{W_2}^{t_2}(\ldots(\phi_{W_K}^{t_K}(0, x))\ldots)),$$

for any $(t_1, \ldots, t_K, x) = (t, x) \in I\!R^K \times \mathcal{X} = \mathcal{S}$. Since the vector fields $Y_{i,j}$ are complete, as clearly are the $Z_{i,j}$, we see that Ψ is a globally defined smooth map. Observe that $W_k = \frac{\partial}{\partial \bar{z}_k}$ (mod $T\mathcal{X}$), $k = 1, \ldots, K$ and thus repeating the arguments used in the proof of Theorem 1 (step (ii)) we see that Ψ is an identity on the base $I\!R^K$ (composition of flows of X_i, as used in the proof of Theorem 1, is irrelevant since they are identity on the base too). To prove that Ψ is bijective assume that $\Psi(t,x) = \Psi(\bar{t},\bar{x})$. When projected on the base $I\!R^K$, every $Y_{i,j}$ and every $Z_{i,j}$ is of the form $W_k = \frac{\partial}{\partial \bar{z}_k}$ for a suitable $k = 1,\ldots,K$ and hence $t = \bar{t}$. This means that $t_i = \bar{t}_i$, $i = 1,\ldots,K$ and the uniqueness of solutions of nonautonomous differential equations defined by each W_i, $i = 1,\ldots,K$, on $\mathcal{X}$ implies that $x = \bar{x}$. To prove that Ψ is surjective choose $(\tau, y) \in \mathcal{S}$. To find (t,x) such that $\Psi(t,x) = (\tau, y)$ we take $\tau = t$ and $(t,x) = \phi_{W_K}^{-\tau_K}(\ldots(\phi_{W_1}^{-\tau_1}(0,y))\ldots)$. Thus, by completeness, the inverse of Ψ always exists and is smooth. Repeating the argument used in the proof of Theorem 1 we see that $\Psi(\bar{z},x)$ maps globally Σ^e into $\widetilde{\Sigma}^e$. □

3 Linear and Linearizable Systems

In this section we show how our results, when applied to linear systems, rediscover the known facts about removing all input derivatives [11]. Consider a generalized linear system of the form

$$\Sigma_L : \qquad \dot{x} = Ax + \sum_{i=1}^{m} \sum_{j=0}^{\alpha_i} B_{i,j}\, u_i^{(j)},$$

where $x \in I\!R^n$, $u_i^{(j)} \in I\!R$. We study the problem of transforming Σ_L, via a generalized change of coordinates of the form

$$\tilde{x} = \psi(x, u_1^{\langle \alpha_1 - 1\rangle}, \ldots, u_m^{\langle \alpha_m - 1\rangle}), \tag{3.17}$$

into a Kalman representation, *i.e.*, a linear control system of the form

$$\Sigma_K : \qquad \dot{\tilde{x}} = \tilde{A}\tilde{x} + \tilde{B}u, \tag{3.18}$$

where $u = (u_1, \ldots, u_m) \in I\!R^m$. Of course, it is natural to look for (3.17) within the class of linear transformations (depending on the control and its derivatives). It is known from recent results in the linear theory [11] obtained using module theory approach that transforming Σ_L into Σ_K is always possible (compare also [4]). We can deduce this fact from the results of Section 2.

Proposition 1 *There always exists a global linear transformation of the form*

$$\tilde{x} = P\,x + \sum_{i=1}^{m} \sum_{j=0}^{\alpha_i - 1} R_{i,j}\, u_i^{(j)},$$

with P invertible, bringing Σ_L into a Kalman state representation Σ_K.

Proof: Consider the extension Σ_L^e of Σ_L. By a direct computation we get that every $(-1)^j \mathrm{ad}_{\mathcal{F}}^j \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i$ is of the form $\frac{\partial}{\partial z_{i,\alpha_i - j}} + F_i^j \frac{\partial}{\partial x}$ for a suitable constant vector F_i^j. Therefore all $\mathrm{ad}_{\mathcal{F}}^j \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $j = 0, \ldots, \alpha_i$, having constant components, commute and thus satisfy the condition (2.9) of Theorem 1 for $\beta_i = 0$, $i = 1, \ldots, m$. Obviously they are complete and from Theorem 2 we deduce the existence of a global transformation Ψ bringing Σ_L into Σ_K. Moreover Ψ, as defined in the proof of Theorem 2, is linear with respect to all its arguments. □

One can observe that the conditions of theorems 1 and 2 for the local and global solvability of the problem remind respectively those which describe local and global state space linearization, *i.e.*, linearization via a (local) diffeomorphism of the state space. Indeed, if we consider the control system (2.8) then (2.9) form a part of the commutativity conditions (see [3, 24]) which describe state space linearizable systems, whereas the completeness of $\mathrm{ad}_{\mathcal{F}}^q \frac{\partial}{\partial z_{i,\alpha_i}}$ (recall that $\frac{\partial}{\partial z_{i,\alpha_i}}$ are the control vector fields of (2.8)) appears in [3, 24] in the solution of the global state space linearization.

This issue can be clarified if we consider the problem of transforming Σ, given by (2.1), into a Kalman linear representation Σ_K given by (3.18). Assume $f(s_0) = 0$, otherwise we have to add a constant vector to the right side of (3.18).

Proposition 2

(i) Σ *is locally transformable, at* $s_0 = (z_0, x_0)$, *via a generalized state space transformation* ψ *of the form* (2.2), *into a Kalman linear system* Σ_K *if, and only if, the extension* Σ^e, *given by* (2.8), *of* Σ *satisfies*

$$\left[\mathrm{ad}_{\mathcal{F}}^q \frac{\partial}{\partial z_{i,\alpha_i}} , \mathrm{ad}_{\mathcal{F}}^r \frac{\partial}{\partial z_{j,\alpha_j}} \right] = 0, \tag{3.19}$$

in a neighbourhood of s_0, *for any* $1 \leq i, j \leq m$ *and*

$$\begin{array}{ccccc} 0 & \leq & q & \leq & \alpha_i + n + 1, \\ 0 & \leq & r & \leq & \alpha_j + n + 1, \end{array}$$

and

$$\dim \operatorname{span} \{ \mathrm{ad}_{\mathcal{F}}^{\alpha_i + q} \frac{\partial}{\partial z_{i,\alpha_i}}(s_0),\ 1 \leq i \leq m,\ 1 \leq q \leq n \} = n. \tag{3.20}$$

(ii) Σ *is globally transformable to* Σ_K *if, and only if, it satisfies* (3.19) *and* (3.20) *everywhere, the vector field* $\mathcal{F}$ *is complete, and moreover* $\mathcal{X}$ *is simply connected.*

Remark 3 *Now the role of* (3.19) *is clear. If it is fulfilled for* $0 \le q \le \alpha_i$, $0 \le r \le \alpha_j$ *then we are able to remove all input derivatives and we get a system of the form*

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u). \tag{3.21}$$

The assumption $\left[\mathrm{ad}_{\mathcal{F}}^{\alpha_i+1} \frac{\partial}{\partial z_{i,\alpha_i}}, \mathrm{ad}_{\mathcal{F}}^{\alpha_j} \frac{\partial}{\partial z_{j,\alpha_j}}\right] = 0$ *implies that* $z_{i,0}$, $i = 1, \ldots, m$ *appear linearly and thus* (3.21) *is an affine system with respect to controls* $u_i = z_{i,0}$, *i.e., of the form*

$$\dot{\tilde{x}} = g_0(\tilde{x}) + \sum_{i=1}^{m} u_i g_i(\tilde{x}). \tag{3.22}$$

Now we observe that $\mathrm{ad}_{\mathcal{F}}^{\alpha_i+1} \frac{\partial}{\partial z_{i,\alpha_i}} = \sum_{k=1}^{n} g_{i,k} \frac{\partial}{\partial \tilde{x}_k}$, *where* $g_{i,k}(\tilde{x})$ *are components of* $g_i(\tilde{x})$, *and therefore* (3.19) *for* $\alpha_i + 1 \le q \le \alpha_i + n + 1$ *and* $\alpha_j + 1 \le r \le \alpha_j + n + 1$, *together with* (3.20), *represent the standard state space linearization conditions (see e.g.* [21, Theorem 5.3] *or* [24, Theorem 3.1]*).*

Notice that if we apply linearization directly to Σ^e, *as it satisfies the commutativity condition* (3.19) *and*

$$\dim \operatorname{span}\{\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}(s_0),\, 1 \le i \le m,\, 0 \le q \le \alpha_i + n\} = K + n,$$

then we end up with a linear system Σ_K^e *(extension of* Σ_K*) but we may change* z*-coordinates. We want to preserve them and that is the reason to transform* Σ^e *into* Σ_K^e *in two steps (removing input derivatives and then applying the linearization to the* $\tilde{x}$*-part of the system).*

Remark 4 *If we drop the simple connectedness assumption then the conditions* (ii) *in Proposition 2 imply the global transformation of* Σ *into a Kalman linear system evolving on* $\mathbb{R}^p \times T^{n-p}$, *where* T^{n-p} *denotes a* $(n-p)$*-dimensional torus. Observe that if we are interested in global lowering the orders of (removing, in particular) input derivatives only we do not have to put any topological assumption on* $\mathcal{X}$. *Indeed,* $\mathcal{X}$ *can be any smooth manifold and the map* Ψ, *as defined in the proof of Theorem 2, gives a global diffeomorphism of* $\mathbb{R}^K \times \mathcal{X}$. *In the problem of simultaneous removing derivatives and linearization we are looking for a global diffeomorphism* $\Psi : \mathbb{R}^K \times \mathcal{X} \longrightarrow \mathbb{R}^K \times \mathbb{R}^n$ *and hence we have to put a topological assumption on* $\mathcal{X}$.

Proof of Proposition 2: Necessity of (i) is obvious. To prove sufficiency we just perform the two-steps procedure as described in Remark 3.

Necessity of (ii) is obvious. To prove sufficiency, observe that the vector fields $\frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, are complete and that the Lie algebra $\mathcal{L}$ generated by $\mathcal{F}$ and $\frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, is finite dimensional (the latter follows, for instance, from local linearizability (i)). Hence all vector fields belonging to $\mathcal{L}$ are complete [23], in particular so are $\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $q = 0, \ldots, \alpha_i + n$. Now, by Theorem 2, we apply a global diffeomorphism to remove all inputs derivatives and we arrive at (3.21). Just as above we conclude that (3.21) is actually of the form (3.22). By applying global linearization results [3, 24] (commutativity, completeness and simple connectedness of $\mathcal{X}$) we get a global diffeomorphism transforming (3.22) into (3.18). □

Observe that under (3.19), satisfied for $0 \leq q \leq \alpha_i + n + 1$, $0 \leq r \leq \alpha_j + n + 1$, and (3.20) (*i.e.*, in the case of locally linearizable systems) the completeness of $\mathcal{F}$ and that of $\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}$, $i = 1, \ldots, m$, $q = 0, \ldots, \alpha_i + n$, are equivalent. Actually, in the statement of Proposition 2 (ii) the former can be replaced by the latter. If we study the global problem of lowering the orders of derivatives of inputs, then such equivalence is not present ($\mathcal{L}$ need not be finite dimensional) and in Theorem 2 we use the completeness of $\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}$.

4 Lowering the Orders of Input Derivatives in the Dynamics and Output Equations

In this section we discuss the problem of lowering the input derivatives orders simultaneously in the dynamics and in the output equations

$$y = h(x, u_1^{\langle \alpha_1 \rangle}, \ldots, u_m^{\langle \alpha_m \rangle}), \tag{4.23}$$

where h is a $\mathbb{R}^p$-valued function, smooth with respect to all its arguments.

It is known that already in the linear case this is, in general, not possible and requires some additional conditions [11]. Let h be a smooth vector-valued function and f a smooth vector field. Then $L_f\, h$ stands for the Lie derivative of h along f.

Theorem 3 *There exist a generalized change of state coordinates of the form* (2.2) *transforming locally* Σ*, with output equations* (4.23)*, into* $\tilde{\Sigma}$*, with output equations*

$$y = \tilde{h}(\tilde{x}, u_1^{\langle \beta_1 \rangle}, \ldots, u_m^{\langle \beta_m \rangle}), \tag{4.24}$$

if, and only if, Σ *satisfies the commutativity conditions* (2.9) *and moreover*

$$L_{\mathrm{ad}_{\mathcal{F}}^{q} \frac{\partial}{\partial z_{i,\alpha_i}}}\, h \equiv 0, \tag{4.25}$$

$i = 1, \ldots, m$ *and* $0 \leq q \leq \alpha_i - \beta_i - 1$.

Proof: *(Necessity)* By Theorem 1 the commutativity conditions (2.9) are necessary in order to lower the orders of input derivatives in the dynamics. To prove necessity of condition (4.25) consider the final representation of the extension and the output given respectively by (2.6) and (4.24). The components $\tilde{f}_k$ of $\widetilde{\mathcal{F}}$ do not depend on $z_{i,j}$, $i = 1, \dots, m$, $j = \beta_i + 1, \dots, \alpha_i$, hence $(-1)^q \mathrm{ad}^q_{\widetilde{\mathcal{F}}} \frac{\partial}{\partial z_{i,\alpha_i}} = \frac{\partial}{\partial z_{i,\alpha_i - q}}$, $i = 1, \dots, m$, $q = 0, \dots, \alpha_i - \beta_i - 1$, and thus $L_{\mathrm{ad}^q_{\widetilde{\mathcal{F}}} \frac{\partial}{\partial z_{i,\alpha_i}}} \tilde{h} = (-1)^q L_{\frac{\partial}{\partial z_{i,\alpha_i - q}}} \tilde{h} \equiv 0$, $i = 1, \dots, m$, $q = 0, \dots, \alpha_i - \beta_i - 1$, since $\tilde{h}$ does not depend on $z_{i,j}$, $i = 1, \dots, m$, $j = \beta_i + 1, \dots, \alpha_i$. The condition (4.25) is invariant under change of coordinates of the form (2.2) and thus it is necessary for lowering the input derivatives orders in the output equations.

(Sufficiency) Apply the local change of coordinates Ψ defined in the proof of Theorem 1. In new coordinates $\tilde{x}^e = \Psi(x^e)$ we have

$$(-1)^j \mathrm{ad}^j_{\widetilde{\mathcal{F}}} \frac{\partial}{\partial z_{i,\alpha_i}} = \frac{\partial}{\partial z_{i,\alpha_i - j}},$$

$i = 1, \dots, m$, $j = 0, \dots, \alpha_i - \beta_i$, and thus (4.25) yields

$$L_{\frac{\partial}{\partial z_{i,\alpha_i - j}}} \tilde{h} \equiv 0,$$

i.e., $\tilde{h}$ does not depend on $u_i^{(j)}$, $i = 1, \dots, m$, $j = \beta_i + 1, \dots, \alpha_i$. A generalized change of coordinates, transforming (2.1), (4.23) into respectively (2.3), (4.24), is given as the restriction of Ψ to $\mathcal{X}$. □

We end this section by discussing some relations of our results with a nice study of realization of nonlinear higher order differential equations in inputs and outputs given by van der Schaft [28].

Consider Σ given by (2.1) and assume that its dynamics

$$\dot{x} = f(x, u_1^{\langle \alpha_1 \rangle}, \dots, u_m^{\langle \alpha_m \rangle})$$

describes the input-output, *i.e.*, external, behaviour for the inputs u_i, $i = 1, \dots, m$ and outputs $y_i = x_i$, $i = 1, \dots, n$. Assume that in (2.1) $\alpha_i = 1$, $i = 1, \dots, m$. This case is also considered in [28, Section 4] and in [16]. We can therefore rewrite (2.1) as

$$\dot{y} - f(y, u, \dot{u}) = 0. \tag{4.26}$$

Now according to [28] compute the maximal invariant manifold N^* of

$$\frac{d}{dt}\begin{pmatrix} y \\ u \end{pmatrix} = \begin{pmatrix} \dot{y} \\ \dot{u} \end{pmatrix}, \quad \frac{d}{dt}\begin{pmatrix} \dot{y} \\ \dot{u} \end{pmatrix} = \begin{pmatrix} v_1 \\ v_2 \end{pmatrix},$$

contained in $R = 0$, where $R(y, \dot{y}, u, \dot{u}) = \dot{y} - f(y, u, \dot{u})$. The invariant manifold N^* is given by (4.26) and we get the following driven state space system

$$\frac{d}{dt}\begin{pmatrix} y \\ u \\ \dot{u} \end{pmatrix} = \begin{pmatrix} f(y,u,\dot{u}) \\ \dot{u} \\ v \end{pmatrix} \tag{4.27}$$

with the output

$$w = (y, u)^{\mathrm{T}}.$$

Observe that the driven state space system (4.27) coincides exactly with the extension Σ^e, defined by (2.8), of (2.1). The next step of the realization procedure of [28] is to check whether all distributions S_i, $i \geq 1$, of the algorithm giving S^*, the minimal conditioned invariant distribution containing all control vector fields of (4.27) (see *e.g.* [21]), are involutive. Clearly, $S^1 = \operatorname{span}\{\frac{\partial}{\partial \dot{u}_i},\ i = 1, \dots, m\}$ is involutive whereas the form of (4.27) implies that involutivity of S^2 is equivalent to $\left[\operatorname{ad}_{\mathcal{F}} \frac{\partial}{\partial \dot{u}_i}, \operatorname{ad}^q_{\mathcal{F}} \frac{\partial}{\partial \dot{u}_j}\right] = 0$, for $q = 0, 1$, and $i, j = 1, \dots, m$. The latter is just the commutativity condition (2.9) satisfied for $\alpha_i = 1$, $\beta_i = 0$, $i = 1, \dots, m$. From Theorem 1 we thus conclude the existence of new coordinates $\tilde{x} = \psi(x, y)$ such that

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u). \tag{4.28}$$

By local invertibility of ψ with respect to $x = y$ we get

$$y = \tilde{\psi}(\tilde{x}, u) \tag{4.29}$$

and (4.28)–(4.29) yields a state space realization of (4.26). In the case $\alpha_i \leq 1$ we get analogous correspondence between the realization procedure [28] and our approach with the only modification that in (4.27) we take the extension $\frac{d^2}{dt^2}u_i = v_i$ for such i that $\alpha_i = 1$ only. Also in this case the involutivity of S_i, the distributions of the S^*-algorithm, coincides with the commutativity (2.9).

However, if there exists i such that $\alpha_i > 1$, then the solvability conditions, and consequently solutions, of both problems differ. Indeed, consider (2.1), and assume that it describes the input-output behaviour for the inputs u_i, $i = 1, \dots, m$ and the outputs $y_i = x_i$, $i = 1, \dots, n$. We thus have

$$\dot{y} - f(y, u_1^{\langle \alpha_1 \rangle}, \dots, u_m^{\langle \alpha_m \rangle}) = 0. \tag{4.30}$$

If the commutativity conditions (2.9) are satisfied for $\beta_i = 0$, $i = 1, \dots, m$, then by Theorem 1 there exist new coordinates

$$\tilde{x} = \psi(x, u_1^{\langle \alpha_1 - 1 \rangle}, \dots, u_m^{\langle \alpha_m - 1 \rangle})$$

such that (2.1) becomes

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u). \tag{4.31}$$

By local invertibility of ψ with respect to $x = y$ we get

$$y = \tilde{\psi}(\tilde{x}, u_1^{\langle\alpha_1 - 1\rangle}, \dots, u_m^{\langle\alpha_m - 1\rangle}), \tag{4.32}$$

where ψ depends nontrivially on $u_i^{(\alpha_i - 1)}$ since so does f on $u_i^{(\alpha_i)}$. Therefore the realization (4.31)–(4.32) of (4.30) is more general than that considered in [28] (see, however, realizations of nonproper linear systems [33], [11]).

On the other hand, apply the realization procedure of [28] to (4.30). A driven state space realization of (4.30) is given by

$$\begin{aligned} \frac{d}{dt}\, y &= f(y, u_1^{\langle\alpha_1\rangle}, \dots, u_m^{\langle\alpha_m\rangle}) \\ \frac{d^{\alpha_i+1}}{dt^{\alpha_i+1}}\, u_i &= v_{n+i},\ i = 1, \dots, m, \end{aligned} \tag{4.33}$$

together with the output $w = (y, u)^{\mathrm{T}}$. Observe once again that (4.33) coincides with the prolongation, defined by (2.8), of (2.1). Compute now S^*, the minimal conditioned invariant distribution of the system (4.33), equipped with the output $w = (y, u)^{\mathrm{T}}$, containing the control vector fields. Recall that we consider the case in which for at least one i we have $\alpha_i > 1$, and thus there exist $r \geq 1$ output components, say $y_{k_1}, \dots, y_{k_r}$, such that

$$P^* \cap \operatorname{span}\{\, \mathrm{d}y_{k_1}, \dots,\ \mathrm{d}y_{k_r}\} = 0,$$

where $P^* = (S^*)^{\perp}$. Now observe that $y_{k_1}, \dots, y_{k_r}$ will serve, according to the realization procedure [28], as controls. Recall that in [28] (4.30) is considered as a differential equation in the external variables $w = (y, u)^{\mathrm{T}}$ and that a part of the realization problem is to split w into an input and an output part. In particular, if all

$$\alpha_i > 1$$

and

$$S^* = S^2 = \operatorname{span}\{\frac{\partial}{\partial u_i^{(\alpha_i)}}, \operatorname{ad}_{\mathcal{F}} \frac{\partial}{\partial u_i^{(\alpha_i)}}, i = 1, \dots, m\}$$

then

$$\begin{aligned} &P^* + \operatorname{span}\{\, \mathrm{d}y_{k_i}, i = 1, \dots, m\} = \\ &\qquad P^* + \operatorname{span}\{\, \mathrm{d}u_i,\ \mathrm{d}y_j, i = 1, \dots, m, j = 1, \dots, n\} \end{aligned}$$

for a suitable nonunique choice of k_i, $i = 1, \dots, m$. Hence, in the realization of (4.30) constructed according to [28], y_{k_i}, $i = 1, \dots, m$, will serve as controls whereas the remaining y_j and all original controls u_i will serve as outputs. To summarize, if we want to realize (4.30), satisfying $\alpha_i > 1$ for some $i = 1, \dots, m$, via the procedure of [28] then, under the involutivity of

S^i, we get a realization not involving input derivatives but some components of y must serve as controls. If we want to keep the original inputs and outputs of (4.30) then, assuming that the commutativity conditions (2.9) are satisfied for $\beta_i = 0$, we get a realization (4.31)–(4.32) whose dynamics do not depend on input derivatives but the outputs do.

In the above analysis we considered the generalized state x as the output of (2.1). Now assume that (2.1) describes the dynamics whereas the output is given by

$$y_i = h_i(x, u_1^{\langle\alpha_1\rangle}, \ldots, u_m^{\langle\alpha_m\rangle}),\ i = 1, \ldots, p,$$

where h_i are smooth $\mathbb{R}$-valued functions. Assume that $\alpha_i = 1$, $i = 1, \ldots, m$, and consider the problem of realization of

$$\left\{\begin{array}{rcl} \dot{x} - f(x,u,\dot{u}) & = & 0 \\ y - h(x,u,\dot{u}) & = & 0 \end{array}\right. \tag{4.34}$$

According to [28, Section 5], $w = (y,u)^{\mathrm{T}}$ forms the vector of external variables whereas x that of internal ones. A driven state space realization of (4.34) takes the form (compare [28])

$$\left\{\begin{array}{rclll} \frac{d}{dt}\, x_i & = & f_i(x,u,\dot{u}), & i = 1, \ldots, n, \\ \frac{d^2}{dt^2}\, u_i & = & v_i, & i = 1, \ldots, m, \end{array}\right. \tag{4.35}$$

with the outputs

$$\left\{\begin{array}{rclll} w_i & = & h_i(x,u,\dot{u}), & i = 1, \ldots, p, \\ w_{p+i} & = & u_i, & i = 1, \ldots, m, \end{array}\right. \tag{4.36}$$

Now, according to [28], we have to compute S^*, the minimal conditioned invariant distribution for (4.35), equipped with the outputs (4.36), containing the control vector fields. We have $S^1 = \mathrm{span}\left\{\frac{\partial}{\partial \dot{u}_i},\ i = 1, \ldots, m\right\}$. Compute $H = \langle \mathrm{d}h,\ \frac{\partial}{\partial \dot{u}}\rangle$ and assume that rank $H =$ const. If rank $H = k > 0$ then dim $(S^1 \cap \ker \mathrm{d}h) = m - k$ and k original output components must serve as inputs of the realization, *i.e.*, if we want to realize the equations (4.34) as an input-state-output system not involving derivatives of inputs then, assuming $k > 0$, we are not able to keep the original specifications of external variables w into inputs and outputs.

The remaining case, *i.e.*, $k = 0$, gives a nice connection between the realization procedure [28] and Theorem 3 of this Section. Compute S^*, the minimal conditioned invariant distribution containing $S^1 = \mathrm{span}\,\{\frac{\partial}{\partial \dot{u}}\}$, for the system (4.35) with the output (4.36). Recall that $k = \mathrm{rank}\ \langle \mathrm{d}h,\ \frac{\partial}{\partial \dot{u}}\rangle$ and thus $k = 0$ simply means that $\langle \mathrm{d}h,\ \frac{\partial}{\partial \dot{u}}\rangle = L_{\frac{\partial}{\partial \dot{u}}}\, h \equiv 0$ or, equivalently, $S^1 \cap \ker \mathrm{d}h = S^1$. Hence $S^2 = S^1 + \mathrm{span}\,\{\mathrm{ad}_{\mathcal{F}}\, \frac{\partial}{\partial \dot{u}}\}$, where $\mathcal{F}$ denotes the right hand side of (4.35), and $S^2 = S^*$. Obviously S^1 is involutive, whereas the form of (4.35) implies that involutivity of S^2 is equivalent to $\left[\mathrm{ad}_{\mathcal{F}}^{q}\frac{\partial}{\partial \dot{u}_i}, \mathrm{ad}_{\mathcal{F}}\frac{\partial}{\partial \dot{u}_j}\right] = 0$ for $q = 0, 1$, and $i, j = 1, \ldots, m$, which is just commutativity condition (2.9) satisfied for $\alpha_i = 1$, $\beta_i = 0$, $i = 1, \ldots, m$. This and $L_{\frac{\partial}{\partial \dot{u}}} h \equiv 0$ yield, according to Theorem 3, new coordinates $\tilde{x} = \psi(x, u)$ such that the dynamics and the output take respectively the form

$$\dot{\tilde{x}} = \tilde{f}(\tilde{x}, u) \tag{4.37}$$

and

$$y = \tilde{h}(\tilde{x}, u). \tag{4.38}$$

To summarize, if the distribution S^2 is involutive and $S^1 \cap \ker \mathrm{d}h = S^1$, which is equivalent to (2.9), for $\alpha_i = 1$, $\beta_i = 0$, and $L_{\frac{\partial}{\partial \dot{u}}} h \equiv 0$, respectively, then the realization procedure of [28] gives the same coordinates change as Theorem 3. This results in realizing (4.34) as the system (4.37)–(4.38) in which the original specification of x, u and y as the internal variables, the inputs, and the outputs, respectively, is kept.

5 Example: Simplified Model of a Crane

We now consider an example borrowed from [15] the interest of which is manifold. This is a physical system —a crane— in which input derivatives appear. It illustrates the impossibility for some nonlinear systems to admit a Kalman state representation and the impossibility of finding the best value for the m-tuple $(\beta_1, \ldots, \beta_m)$. The state equations are

$$\left\{ \begin{array}{lcl} \dot{x}_1 & = & x_2 \\ \dot{x}_2 & = & -\frac{g \sin x_1}{R} - \frac{2x_2}{R}\dot{R} - \frac{\cos x_1}{R}\ddot{D} \end{array} \right. , \tag{5.39}$$

where the inputs are R, the length of the rope and D, the trolley position, and the state variables (x_1, x_2) are the angle $x_1 = \theta$, between the rope and vertical axis, and its time derivative $x_2 = \dot{\theta}$ (see [15] for a precise discussion of the choice of variables). We see that $\dot{R}$ and $\ddot{D}$ appear linearly in those equations. Setting $u_1 = R$, $u_2 = D$ and keeping the same notations as in the whole paper (and especially in the proof of Theorem 1), we have $\alpha_1 = 1$ and $\alpha_2 = 2$.

The state space of this system is two-dimensional, $\mathcal{X} =]-\frac{\pi}{2}, +\frac{\pi}{2}[\times \mathbb{R}$ and we extend it to $\mathcal{S}$ of dimension seven ($K = \alpha_1 + 1 + \alpha_2 + 1 = 5$), by setting $z_{1,0} = u_1^{(0)}, z_{1,1} = u_1^{(1)}, z_{2,0} = u_2^{(0)}, z_{2,1} = u_2^{(1)}, z_{2,2} = u_2^{(2)}$. The extended state space is $\mathcal{S} = \mathbb{R} \times]0, +\infty[\times \mathbb{R}^3 \times \mathcal{X}$. The Lie bracket $\left[\mathrm{ad}_{\mathcal{F}}\frac{\partial}{\partial z_{1,1}}, \frac{\partial}{\partial z_{2,2}}\right]$ (resp. $\left[\frac{\partial}{\partial z_{1,1}}, \mathrm{ad}_{\mathcal{F}}\frac{\partial}{\partial z_{2,2}}\right]$) vanishes everywhere on $\mathcal{S}$, and then there exists a generalized change of coordinates leading to a representation with $(\beta_1, \beta_2) = (0, 2)$ (resp.$(\beta_1, \beta_2) = (1, 1)$) at any point of

$\mathcal{S}$. However, $\left[\mathrm{ad}_{\mathcal{F}}\frac{\partial}{\partial z_{1,1}}, \mathrm{ad}_{\mathcal{F}}\frac{\partial}{\partial z_{2,2}}\right] = -\frac{\cos x_1}{(z_{1,0})^2}\frac{\partial}{\partial x_2}$ does not vanish in any open subset of $\mathcal{S}$ and no generalized change of state coordinate exists with $(\beta_1, \beta_2) = (0, 1)$. As we already discussed in Section 2, we have two incomparable solutions to the problem of lowering the orders of the input derivatives.

In both cases we will compute the vector fields $W_1, \ldots, W_7$, the flows of which permit to construct change of coordinates Ψ_{s_0} whose restriction to $\mathcal{X}$ gives a generalized change of coordinates ψ leading to the desired representation. We keep the same notations as used in the proof of Theorem 1. Let $s_0 = (z^0, x^0) \in \mathcal{S}$ be the initial condition. In both cases all involved vector fields are complete and therefore we will construct global coordinates (compare also Theorem 2) which can be centered at any s_0. For convenience we choose $z^0_{1,1} = z^0_{2,0} = z^0_{2,1} = z^0_{2,2} = x^0_1 = x^0_2 = 0$ and a fixed $z^0_{1,0} > 0$ because $R > 0$. Let $s = (\bar{z}, x)$ be a point of $\mathcal{S}$ in a neighbourhood of s_0.

1^{st} case: Removing $\dot{R}$ from the representation: In this case we will lower by one the order of derivation of the first control variable R. Using the notations introduced in the proof of Theorem 1, we have

$$\begin{array}{lllllll}
W_1 = Y_{1,1} = \frac{\partial}{\partial z_{1,1}}, & \bar{z}_1 = z_{1,1} \\
W_2 = Y_{1,2} = -\frac{2x_2}{z_{1,0}}\frac{\partial}{\partial x_2} + \frac{\partial}{\partial z_{1,0}}, & \bar{z}_2 = z_{1,0} \\
W_3 = Y_{2,1} = \frac{\partial}{\partial z_{2,2}}, & \bar{z}_3 = z_{2,2} \\
W_4 = Z_{2,1} = \frac{\partial}{\partial z_{2,0}}, & \bar{z}_4 = z_{2,0} \\
W_5 = Z_{2,2} = \frac{\partial}{\partial z_{2,1}}, & \bar{z}_5 = z_{2,1} \\
W_6 = \frac{\partial}{\partial x_1} & \\
W_7 = \frac{\partial}{\partial x_2} &
\end{array}$$

Thus

$$\Phi_{s_0}(t_1, \ldots, t_7) = \left(t_1, \bar{z}^0_2 + t_2, t_3, t_4, t_5, t_6, \frac{(\bar{z}^0_2)^2 t_7}{(\bar{z}^0_2 + t_2)^2}\right),$$

(recall that $\bar{z}^0_2 \neq 0$),

$$\Phi^{-1}_{s_0}(s) = \left(\bar{z}_1, \bar{z}_2 - \bar{z}^0_2, \bar{z}_3, \bar{z}_4, \bar{z}_5, x_1, \frac{x_2(\bar{z}_2)^2}{(\bar{z}^0_2)^2}\right),$$

and

$$\Psi_{s_0}(\bar{z}, x) = \left(\bar{z}_1, \bar{z}_2, \bar{z}_3, \bar{z}_4, \bar{z}_5, x_1, \frac{x_2(\bar{z}_2)^2}{(\bar{z}^0_2)^2}\right).$$

This leads to the new state coordinates (on the original state space)

$$\left\{\begin{array}{lll} \tilde{x}_1 & = & x_1 \\ \tilde{x}_2 & = & \frac{x_2 \bar{z}_2^2}{(\bar{z}^0_2)^2} \end{array}\right. . \tag{5.40}$$

The final representation of the system is (compare [4, 15])

$$\left\{ \begin{array}{lcl} \dot{\tilde{x}}_1 & = & \frac{R(0)^2 \tilde{x}_2}{R^2} \\ \dot{\tilde{x}}_2 & = & -\frac{gR \sin \tilde{x}_1 + R\ddot{D} \cos \tilde{x}_1}{(R(0))^2} \end{array} \right. , \tag{5.41}$$

where $R(0)$ can be chosen arbitrarily. With an appropriate rescaling of R we can always take $R(0) = 1$.

2^{nd} case: Removing $\ddot{D}$ from the representation: In this case we will lower by one the order of derivation of the second control variable D. Calculate

$$\begin{array}{lllll} W_1 = Y_{1,1} = \frac{\partial}{\partial z_{1,1}}, & & \bar{z}_1 = z_{1,1} \\ W_2 = Y_{2,1} = \frac{\partial}{\partial z_{2,2}}, & & \bar{z}_2 = z_{2,2} \\ W_3 = Y_{2,2} = -\frac{\cos x_1}{z_{1,0}} \frac{\partial}{\partial x_2} + \frac{\partial}{\partial z_{2,1}}, & & \bar{z}_3 = z_{2,1} \\ W_4 = Z_{1,1} = \frac{\partial}{\partial z_{1,0}}, & & \bar{z}_4 = z_{1,0} \\ W_5 = Z_{2,1} = \frac{\partial}{\partial z_{2,0}}, & & \bar{z}_5 = z_{2,0} \\ W_6 = \frac{\partial}{\partial x_1} & & \\ W_7 = \frac{\partial}{\partial x_2} & & \end{array}$$

Thus

$$\Phi_{s_0}(t_1, \ldots, t_7) = \left(t_1, t_2, t_3, \bar{z}_4^0 + t_4, +t_5, t_6, -\frac{\cos t_6}{\bar{z}_4^0 + t_4} t_3 + t_7 \right),$$

(recall that $\bar{z}_4^0 \neq 0$),

$$\Phi_{s_0}^{-1}(s) = \left(\bar{z}_1, \bar{z}_2, \bar{z}_3, \bar{z}_4 - \bar{z}_4^0, \bar{z}_5, x_1, x_2 + \frac{\bar{z}_3 \cos x_1}{\bar{z}_4} \right),$$

and

$$\Psi_{s_0}(\bar{z}, x) = \left(\bar{z}_1, \bar{z}_2, \bar{z}_3, \bar{z}_4, \bar{z}_5, x_1, x_2 + \frac{\bar{z}_3 \cos x_1}{\bar{z}_4} \right).$$

This leads to the new state coordinates (on the original state space)

$$\left\{ \begin{array}{lcl} \tilde{x}_1 & = & x_1 \\ \tilde{x}_2 & = & x_2 + \frac{\bar{z}_3 \cos x_1}{\bar{z}_4} \end{array} \right. .$$

The final representation of the system is (compare [15, 4])

$$\left\{ \begin{array}{lcl} \dot{\tilde{x}}_1 & = & \tilde{x}_2 - \frac{\dot{D} \cos \tilde{x}_1}{R} \\ \dot{\tilde{x}}_2 & = & -\frac{g \sin \tilde{x}_1 + 2\dot{R}\tilde{x}_2 + \dot{D}\tilde{x}_2 \sin \tilde{x}_1}{R} \\ & & + \frac{\dot{R}\dot{D} \cos \tilde{x}_1 + (\dot{D})^2 \cos \tilde{x}_1 \sin \tilde{x}_1}{(R)^2} \end{array} \right. .$$

The two representations of the dynamics are not equivalent from the practical point of view.

As we have just seen it is impossible to remove $\dot{R}$ and $\ddot{D}$ simultaneously by a generalized change of coordinates but after removing $\dot{R}$ we arrive at (5.41) and we can get rid of $\ddot{D}$ by simply introducing the new control variable $d = \ddot{D}$. This yields the system (see (5.41), where for simplicity we put $R(0) = 1$)

$$\left\{ \begin{array}{rcl} \dot{\tilde{x}}_1 & = & \tilde{x}_2/R^2 \\ \dot{\tilde{x}}_2 & = & -dR\cos\tilde{x}_1 - gR\sin\tilde{x}_1 \end{array} \right. , \qquad (5.42)$$

controlled by R and d and thus no derivatives of controls are involved any more. The reason for which this procedure works is that the dynamics (5.41) do not depend on D and $\dot{D}$. The engineering interpretation of the substitution $d = \ddot{D}$ is clear: instead of controlling the position D we control the acceleration d. A geometric interpretation of this procedure replacing some controls derivatives $u_i^{(\delta_i)}$ by new controls $\hat{u}_i$ (which, although lowers inputs derivatives, is of a different nature than generalized state transformations) will be discussed in [6].

6 Conclusion

Generalized state representations of nonlinear systems are studied and the problem of lowering the inputs derivatives orders by generalized coordinate changes is considered. The obtained restrictive commutativity conditions imply that generalized systems for which we can remove input derivatives (or even lower their orders) exhibit a very special structure (compare also [16], [17], [28]). The discrete-time version of the problem, *i.e.*, that of removing delays in discrete-time systems [12], is still open.

References

[1] J. Birk and M. Zeitz. Extended luenberger observers for nonlinear multivariable systems, *Int. J. Control* **47** (1988), 1823–1836.

[2] M. Cohen de Lara and J. Lévine. Local identification of nonlinear systems, bifurcations and normal forms. In *Analysis of Controlled Dynamical Systems*, pages 316–325, (B . Bornard, B. Bride, J.-P. Gauthier, and I. Kupka, eds.). Boston: Birkhäuser, 1991.

[3] W. Dayawansa, M. Boothby, and D. L. Elliot. Global state and feedback equivalence of nonlinear systems, *Systems & Control Letters* **6** (1985), 229–234.

[4] E. Delaleau. Lowering orders of input derivatives in generalized state representations of nonlinear systems. In *Nonlinear Control Systems*

Design 1992, Selected papers from the 2nd IFAC Symposium, Bordeaux, France, 24 – 26 June 1992, pages 347–351, (M. Fliess, ed.). Oxford: Pergamon Press, 1993.

[5] E. Delaleau and W. Respondek. Lowering the orders of derivatives of controls in generalized state space systems. In *Proc. 31st Control and Decision Conference, Tucson, AZ, December 1992*, pages 3663–3668, 1992.

[6] E. Delaleau and W. Respondek. Lowering the inputs derivatives orders by generalized state transformations and pure preintegration. In preparation.

[7] S. Diop. Rational system equivalence and generalized realization theory. In *Nonlinear Control Systems Design 1992, Selected papers from the 2nd IFAC Symposium, Bordeaux, France, 24 – 26 June 1992*, pages 153–158, (M. Fliess, ed.). Oxford: Pergamon Press, 1993.

[8] M. Fliess. Some remarks on nonlinear invertibility and dynamic state-feedback. In *Theory and Applications of Nonlinear Control Systems*, pages 115–121, (C. I. Byrnes and A. Lindquist, eds.). Amsterdam: North Holland, 1986.

[9] M. Fliess. Automatique et corps différentiels, *Forum Math.* **1** (1989), 227–238.

[10] M. Fliess. Generalized controller canonical forms for linear and nonlinear dynamics, *IEEE Trans. Automat. Control* **AC–35** (1990), 994–1001.

[11] M. Fliess. Some basic structural properties of generalized linear systems, *Systems & Control Letters* **15** (1990), 391–396.

[12] M. Fliess. Reversible linear and nonlinear dynamics, *IEEE Trans. Automat. Control* **AC–37** (1992), 1144–1153.

[13] M. Fliess, J. Lévine, P. Martin, and P. Rouchon. Towards a new differential geometric setting in nonlinear control. In *Internat. Geometrical Coll.*, Moscow, May 1993. To appear.

[14] M. Fliess, J. Lévine, P. Martin, and P. Rouchon. Linéarisation par bouclage dynamique et transformations de Lie-Bäcklund, *C. R. Acad. Sci. Paris* (1993). To appear.

[15] M. Fliess, J. Lévine, and P. Rouchon. Generalized state variable representation for a simplified crane description, *Int. J. Control* **58** (1993), 277–283.

[16] M. I. Freedman and Jan C. Willems. Smooth representation of systems with differentiated inputs, *IEEE Trans. Automat. Control* **AC–23** (1978), 16–21.

[17] S. T. Glad. Nonlinear state space and input output descriptions using differential polynomials. In *New Trends in Nonlinear Control Theory*, pages 182–189, (J. Descusse, M. Fliess, A. Isidori, and D. Leborgne, eds.). New York: Springer-Verlag, 1988.

[18] B. Jakubczyk. Remarks on equivalence and linearization of nonlinear systems. In *Nonlinear Control Systems Design 1992, Selected papers from the 2nd IFAC Symposium, Bordeaux, France, 24 – 26 June 1992*, pages 143–148, (M. Fliess, ed.). Oxford: Pergamon Press, 1993.

[19] H. Keller. Non-linear observers design by transformation into a generalized observer canonical form, *Int. J. Control* **46** (1987), 1915–1930.

[20] I. S. Krasilchik, V. V. Lychagin, and A. M. Vinogradov. *Geometry of Jet Spaces and Nonlinear Partial Differential Equations.* New York: Gordon and Breach, 1986.

[21] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems.* Berlin: Springer-Verlag, 1990.

[22] P. J. Olver. *Application of Lie Groups to Differential Equations.* New York: Springer-Verlag, 1986.

[23] R. S. Palais. *A Global Formulation of the Lie Theory of Transformations Groups.* New York: Amer. Math. Soc., 1986.

[24] W. Respondek. Global aspect of linearization, equivalence to polynomial forms and decomposition of nonlinear systems. In *Algebraic and Geometric Methods in Nonlinear Control Theory*, pages 257–284, (M. Fliess and M. Hazewinkel, eds.). Amsterdam: D. Reidel Publishing Company, 1986.

[25] W. Respondek. Right and left invertibility of nonlinear control systems. In *Nonlinear Control Theory*, pages 133–176, (H. J. Sussmann, ed.). New York: Marcel Dekker, 1990.

[26] J. Rudolph. Viewing input-output system equivalence from differential algebra, *Journal of Mathematical Systems, Estimation, and Control* **4(3)** (1994), 353-383.

[27] J. Rudolph, J. Birk, and M. Zeitz. Dependence on time derivatives of the input in the nonlinear controller canonical form. In *New Trends in Systems Theory*, pages 636–643, (G. Conte, A. Perdon, and W. Wyman, eds.). Boston: Birkhäuser, 1991.

[28] A. J. van der Schaft. On realization of nonlinear systems described by higher-order differential equations, *Math. Systems Theory* **19** (1987), 239–275.

[29] L. M. Silverman. Inversion of multivariable linear systems, *IEEE Trans. Automat. Control* **AC–14** (1969), 270–276.

[30] S. N. Singh. A modified algorithm for invertibility in nonlinear systems, *IEEE Trans. Automat. Control* **AC–26** (1981), 595–598.

[31] M. Spivak. *A Comprehensive Introduction to Differential Geometry.* Berkeley: Publish or Perish, 1970.

[32] D. Williamson. Observation of bilinear systems with application to biological control, *Automatica* **13** (1977), 243–254.

[33] L. A. Zadeh and C. A. Desoer. *Linear systems Theory: The State Space Approach.* New York: McGraw-Hill, 1963.

[34] M. Zeitz. Observability canonical (phase-variable) form for non-linear time-variable systems, *Int. J. Control* **15** (1984), 949–958.

Laboratoire des Signaux et Systèmes, C.N.R.S. – E.S.E, Plateau de Moulon, F–91192 Gif-sur-Yvette Cedex, France

Institute of Mathematics, Polish Academy of Sciences, Śniadeckich 8, 00–950 Warsaw, Poland

Communicated by Michel Fliess