On local linearization of control systems

Laurent Baratchart, Jean-Baptiste Pomet

ON LOCAL LINEARIZATION OF CONTROL SYSTEMS

LAURENT BARATCHART AND JEAN-BAPTISTE POMET

INRIA, B.P. 93, 06902 Sophia Antipolis cedex, France
Laurent.Baratchart@sophia.inria.fr, Jean-Baptiste.Pomet@sophia.inria.fr

Abstract. We consider the problem of topological linearization of smooth ($\mathbf{C}^\infty$ or $\mathbf{C}^\omega$) control systems, i.e. of their local equivalence to a linear controllable system *via* point-wise transformations on the state and the control (static feedback transformations) that are topological but not necessarily differentiable. We prove that local topological linearization implies local smooth linearization, at generic points. At arbitrary points, it implies local conjugation to a linear system *via* a homeomorphism that induces a smooth diffeomorphism on the state variables, and, except at "strongly" singular points, this homeomorphism can be chosen to be a smooth mapping (the inverse map needs not be smooth). Deciding whether the same is true at "strongly" singular points is tantamount to solve an intriguing open question in differential topology.

1. Introduction

Throughout the paper, *smooth* means of class $\mathbf{C}^\infty$.

In the early works [12, 9, 27], nice necessary and sufficient conditions were obtained for a smooth control system $\dot{x} = f(x,u)$, with state $x \in \mathbb{R}^n$ and control $u \in \mathbb{R}^m$, to be locally smoothly linearizable, *i.e.* locally equivalent to a controllable linear system by means of a diffeomorphic change of variables on the state and the control. The afore-mentioned conditions require certain distributions of vector fields to be integrable, hence locally smoothly linearizable control systems are highly non generic among smooth control systems. Similar results hold for real analytic control systems with respect to real analytic linearizability.

Consider now the *topological* linearizability of a smooth control system, namely the property that it is locally equivalent to a controllable linear system *via* a homeomorphism on the state and the control which may *not*, this time, be differentiable. Obviously, smooth linearizability implies topological linearizability; the extend to which the converse holds will be the main concern of the present paper. We address the real analytic case in the same stroke.

In brief, our goal is to *describe the class of smooth control systems that are locally topologically linearizable, yet not smoothly locally linearizable.* This class in nonempty : the smooth (even real-analytic) scalar system

$$\dot{x} \;=\; u^3 \qquad u \in \mathbb{R},\; x \in \mathbb{R}, \tag{1}$$

gets linearized locally around $(0,0)$ by the homeomorphism $(x,u) \mapsto (x,u^3)$, whereas the conditions for smooth linearizability fail at this point. However, we observe on this example that the conjugating homeomorphism has much more regularity than prescribed *a priori*:

1. it is a smooth (even real-analytic) local diffeomorphism around all points (x,u)

2000 *Mathematics Subject Classification.* 93B18, 34C20, 37C10.

Key words and phrases. Control systems, linearization, topological equivalence, Grobman-Hartman Theorem.

January 20, 2009. To appear in *Journal of Dynamical and Control Systems*.

such that $u \neq 0$,
2. it is triangular and induces a smooth (even real-analytic) diffeomorphism on the state variable (*i.e.* the identity map $x \mapsto x$),
3. it is a smooth (even real-analytic) map that fails to be a diffeomorphism only because its inverse is not smooth.

Theorem 5.2 of the present paper states that this example essentially depicts the general situation. More precisely, if a smooth control system is locally topologically linearizable at some point $(\bar{x}, \bar{u})$ in the state-control space, then
$1'$. in a neighborhood of $(\bar{x}, \bar{u})$, the system is locally smoothly linearizable around each point outside a closed subset of empty interior (an analytic variety of positive co-dimension in the analytic case),
$2'$. around $(\bar{x}, \bar{u})$, there is a triangular linearizing homeomorphism that induces a smooth diffeomorphism on the state variable,
$3'$. the above-mentioned homeomorphism is smooth (although its inverse may not), at least if $\partial f/\partial u$ has constant rank around $(\bar{x}, \bar{u})$ or if $\sup_{x,u} \mathrm{Rank} \partial f/\partial u(x, u) = m$ on every neighborhood of $(\bar{x}, \bar{u})$.

Similar results hold for real-analytic linearization of a real-analytic system.

A homeomorphism satisfying $2'$ will be called *quasi-smooth* (see Definitions 3.9, 5.1), hence our main result is that local topological linearizability implies local quasi-smooth linearizability. A point $(\bar{x}, \bar{u})$ where the first rank condition in $3'$ is satisfied is called *regular*, and at such points local smooth linearizability is equivalent to local topological linearizability (*cf.* Theorem 5.4). A point $(\bar{x}, \bar{u})$ where none of the rank conditions in $3'$ are satisfied is called *strongly singular*. Whether the conclusion of $3'$ continues to hold at strongly singular points raises an intriguing question in differential topology, namely can one redefine the last components of a local homeomorphism whose first few components are smooth so as to obtain a new homeomorphism which is smooth? The answer seems not to be known, see the discussion in section 5.1.

Motivations. They include the following.

1. For systems without controls, *i.e.* ordinary differential equations, local linearization around an equilibrium has generated a sizable literature, see Section 2 for a small sample. It tells us that, even for a real analytic o.d.e., linearizability much depends on the admissible class of transformations (formal, real analytic, C^k or topological). For instance, although analytic linearization requires subtle conditions relying upon a refined analysis of resonances and small divisors, the Grobman-Hartman theorem says nevertheless that topological linearization is always possible at a hyperbolic equilibrium. As one might suspect (this is indeed shown in section 5.4), no naive analog to the Grobman-Hartman theorem can hold for control systems because they feature a family of vector fields rather than a single one. However, it might still be expected that relaxing the smoothness of the allowable transformations increases the class of linearizable control systems. It is in fact hardly so: we knew already from [12, 9, 27] that C^1 linearizability of a smooth control system implies smooth linearizability, and we prove here that for C^0 linearizability this class does not get much bigger. In particular, there are no subtle questions about resonances and one may say that the most prominent feature of a control system is to be, or not to be linearizable, regardless of smoothness.

2. Linearizable control systems are systems with linear dynamics, whose nonlinear character lies in their input-to-state and state-to-output maps only. Such models are advocated in [13, 22] for identification (in the discrete-time case), as their reduced complexity makes them more amenable to standard techniques. It is therefore natural to investigate this class, and topological equivalence is about the weakest possible from the point of view of identification.

3. From a control engineering point of view, it is common practice to design locally stabilizing feedback laws for a given system based on its linear approximation when the latter is controllable... and to a certain extent one believes that the latter and the former *locally* "look alike". It is therefore legitimate to ask about the relationship between them. Since no discriminating topological invariants are known, topological conjugacy might appear as a good candidate. The present paper shows that the relationship is almost never that strong: topological conjugacy to the linear approximation is almost as rare as differential conjugacy.

Incidentally, a system whose linear approximation is not controllable may still happen to be locally topologically linearizable, *i.e.* equivalent to a linear controllable system (which is *not* its linear approximation). This phenomenon is clarified in section 5.3.

Techniques. The conditions for smooth linearizability derived in [12, 9, 27] come up naturally in some sense. Indeed, to any control system, one may associate a sequence of distributions defined via a construction using Lie brackets of vector fields attached to the system; it turns out that the instance of this sequence of distributions for linear systems yields "constant" –hence integrable– distributions that span the entire state space in a finite number of steps if the system is controllable. Since Lie brackets and integrability of distributions are preserved under local *diffeomorphisms*, this translates at once into necessary conditions for smooth linearizability, shown in [12, 9, 27] to be sufficient. In contrast, homeomorphisms do not allow to pull back Lie brackets or tangent vector fields; hence the same conditions need not be necessary for topological linearization, and the proofs in the present paper are more intricate. Specifically, we have to rely upon the notion of orbits of families of smooth vector fields rather than integral manifolds. The proof of Theorem 5.2 uses classical results concerning such orbits, first established in [25], that we recall and slightly expand in Appendix B. Incidentally, the lack of a theory dealing with orbits of $\mathbf{C}^k$ vector fields ($k \in \mathbb{N}$) is the main reason why the results of the present paper restrict to $\mathbf{C}^\infty$ or $\mathbf{C}^\omega$ (*i.e.* real analytic) control systems.

Hopefully our method can be useful to study local topological equivalence to other classes of systems than linear ones; this is not investigated here.

Organization of the paper. Section 2 recalls classical facts on local linearization of ordinary differential equations. Section 3 introduces conjugation for *control systems* (under a homeomorphism, a diffeomorphism, etc.) and establishes basic properties of conjugating maps. Section 4 reviews (topological, smooth, linear) conjugacy between *linear* control systems after [4, 29]. Section 5 states the main result of the paper (Theorem 5.2), namely that local topological linearizability implies local quasi-smooth linearizability for smooth control systems (smooth meaning either $\mathbf{C}^\infty$ or $\mathbf{C}^\omega$), and discusses the gap between smooth and quasi-smooth linearizability, including geometric characterizations thereof. Section 6 contains the proofs of these results; the proof of Theorem 5.2, given in subsection 6.2, relies upon section 3, results from [25] stated in Appendix B, and technical lemmas from Appendix A.

2. Local linearization for ordinary differential equations

Consider the differential equation

$$\dot{x}(t) = f(x(t)), \tag{2}$$

where $f \in \mathbf{C}^k(U, \mathbb{R}^n)$ with U an open subset of $\mathbb{R}^n$ and $k \in \mathbb{N} \cup \{\infty, \omega\}$, $k \geq 1$.

It is well known (the "flow box theorem", see *e.g.* [2]) that, around each $x_0 \in U$ such that $f(x_0) \neq 0$, there is a change of coordinates of class $\mathbf{C}^k$ that conjugates (2) to the equation $\dot{x}_1 = 1$, $\dot{x}_2 = 0, \ldots, \dot{x}_n = 0$. Hence all differentiable vector fields are equivalent to each other, at points where they do not vanish, *via* a diffeomorphism having the same degree of smoothness (including real analyticity).

At a point $x_0 \in U$ such that $f(x_0) = 0$, *i.e.* at an equilibrium of the dynamical system (2), its linear approximation is the system

$$\dot{x}(t) = Ax(t) - Ax_0 \tag{3}$$

where $A = Df(x_0)$ is the derivative of f at x_0. The equilibrium x_0 is said to be *hyperbolic* if the matrix A has no purely imaginary eigenvalue.

The problem of locally linearizing (2) is that of finding a local homeomorphism $h : V \to W$ around x_0 mapping the trajectories of (2) in V onto trajectories of (3) in W in a time-preserving manner. In other words, if ϕ_t denotes the flow of (2), we should have for each $x \in V$ that

$$h \circ \phi_t(x) = e^{At}\big(h(x) - h(x_0)\big) + h(x_0)$$

provided that $\phi_\rho(x) \in V$ for $0 \leq \rho \leq t$. When this is the case we say that h conjugates (2) and (3), and we speak of topological, $\mathbf{C}^k$, smooth, or analytic linearization depending on the regularity of h and h^{-1}.

Local linearization at an equilibrium is a very old issue. At the beginning of the twentieth century, H. Poincaré already identified the obstructions to the existence of a *formal* change of variables h that removes all the nonlinear terms when f is analytic. These are the so-called resonances, see *e.g.* [8, 2]. In fact, resonant monomials of order ℓ are obstructions to linearizing the Taylor expansion of f at order ℓ and consequently also obstructions to $\mathbf{C}^\ell$ linearization. However, although there exists a formal power series expansion for h when there are no resonant terms, the existence of a *convergent* power series for h (analytic linearization) is a delicate issue. When the eigenvalues of the Jacobian belong to the so-called Poincaré domain, the absence of resonances indeed implies analytic linearizability (the Poincaré theorem). If it is not the case, a famous theorem by Siegel gives additional Diophantine conditions on these eigenvalues to the same conclusion. These conditions are generically satisfied in the measure-theoretic sense [2]. If no eigenvalue of the Jacobian is purely imaginary, it turns out [20] that the absence of resonances is also sufficient for smooth (h, h^{-1} of class $\mathbf{C}^\infty$) but in general not real analytic linearization. This is still valid when f is merely of class $\mathbf{C}^\infty$.

In contrast, if one allows conjugation *via* a topological but not necessarily differentiable homeomorphism, the Grobman-Hartman theorem asserts that every ordinary differential equation with no purely imaginary eigenvalue of the Jacobian (hyperbolicity) can be locally linearized around an equilibrium, that is, resonances are no longer an obstruction. A proof of this classical result can be found in [8]:

Theorem 2.1 (Grobman-Hartman)**.** *Under the assumption that x_0 is a hyperbolic equilibrium point, system* (2) *is topologically conjugate to system* (3) *at x_0.*

In fact, it is proved in [28] that the conjugating homeomorphism h (together with its inverse h^{-1}) can be chosen Hölder-continuous, and even differentiable at x_0 (but not in a neighborhood). This brings additional rigidity to the mapping h.

The above theorem entails that the only invariant under local topological conjugacy, around a hyperbolic equilibrium, is the number of eigenvalues with positive real part in the Jacobian matrix, counting multiplicity. Indeed, as is well-known (*cf.* [1]), the linear system $\dot{x} = Ax$ where A has no pure imaginary eigenvalue is topologically conjugate to $\dot{x} = DX$, where D is diagonal with diagonal entries ± 1, the number of $+1$ being the number of eigenvalues of A with positive real part.

3. Preliminaries on topological equivalence for control systems

3.1. **Control systems and their solutions.** Consider two *control systems* where n, m, n', m' are natural integers :

$$\begin{aligned} \dot{x} &= f(x,u) \ , \quad x \in I\!R^n \ , \quad u \in I\!R^m \ , && (4) \\ \dot{z} &= g(z,v) \ , \quad z \in I\!R^{n'} \ , \quad v \in I\!R^{m'} \ , && (5) \end{aligned}$$

or expanded in coordinates :

$$\begin{aligned} \dot{x}_1 &= f_1(x_1,\ldots,x_n,u_1,\ldots,u_m) & \qquad \dot{z}_1 &= g_1(z_1,\ldots,z_{n'},v_1,\ldots,v_{m'}) \\ &\vdots & &\vdots \\ \dot{x}_n &= f_n(x_1,\ldots,x_n,u_1,\ldots,u_m) & \qquad \dot{z}_{n'} &= g_{n'}(z_1,\ldots,z_{n'},v_1,\ldots,v_{m'}) \end{aligned}$$

where x or z is called the *state* and u or v the *control*.

Although our main results are stated (in section 5) for infinitely differentiable —or real analytic— control systems, their proofs deal with non-smooth objects because the transformations we consider are only assumed to be continuous. This leads us to keep smoothness assumptions to a minimum in the present section. Accordingly, the maps $f_i : I\!R^n \times I\!R^m \to I\!R$ and $g_i : I\!R^{n'} \times I\!R^{m'} \to I\!R$ are assumed to be at least continuous; *any additional regularity assumption will be stated explicitly.* We do not restrict their domains of definition; this is no real loss of generality because they could anyway be extended using partitions of unity (real analyticity plays no role in the present section), and whenever a result is stated, the domain where it holds true is precisely stated and the value of f and g outside this domain does not matter.

If m is zero or f does not depend on u, equation (4) reduces to the ordinary differential equation (2). Of course "genuine" control systems are those whose right hand side does depend on the control.

Definition 3.1. By *a solution of* (4) *that remains in an open set* $\Omega \subset I\!R^{n+m}$, we mean a mapping γ defined on a real interval I, say

$$\begin{array}{rcll} \gamma : & I & \to & \Omega \\ & t & \mapsto & \gamma(t) \ = \ (\,\gamma_{\mathrm{I}}(t)\,,\,\gamma_{\mathrm{II}}(t)\,) \end{array} \tag{6}$$

with $\gamma_{\mathrm{I}}(t) \in I\!R^n$ and $\gamma_{\mathrm{II}}(t) \in I\!R^m$, such that :

- γ is measurable, locally bounded, and γ_{I} is absolutely continuous,
- whenever $[T_1, T_2] \subset I$, we have :

$$\gamma_{\mathrm{I}}(T_2) - \gamma_{\mathrm{I}}(T_1) \ = \ \int_{T_1}^{T_2} f(\,\gamma_{\mathrm{I}}(t)\,,\,\gamma_{\mathrm{II}}(t)\,)\,\mathrm{d}t \ . \tag{7}$$

Solutions of (5) that remain in $\Omega' \subset I\!R^{n'+m'}$ are likewise defined to be mappings

$$\begin{array}{rcll} \gamma' : & I & \to & \Omega' \\ & t & \mapsto & \gamma'(t) \ = \ (\,\gamma'_{\mathrm{I}}(t)\,,\,\gamma'_{\mathrm{II}}(t)\,) \end{array} \tag{8}$$

having the corresponding properties with respect to g.

If $(\bar{x}, \bar{u})$ is a point in Ω, $\mathcal{U}$ a neighborhood of $\bar{u}$ such that $\{\bar{x}\} \times \mathcal{U} \subset \Omega$, J a real interval, and $\gamma_{\mathrm{II}} : J \to \mathcal{U}$ a measurable and locally bounded map, then, by [6, Ch. 2, Theorem 1.1] and the continuity of f, there exists, on a possibly smaller interval $I \subset J$, a solution γ of (4) that remains in Ω subject to the initial condition $\gamma_{\mathrm{I}}(0) = \bar{x}$. This solution may not be unique without further assumptions on f, for instance that it is continuously differentiable, or merely locally Lipschitz in the first argument.

Remark 3.2. Observe that Definition 3.1 assigns a definite value to $\gamma_{\mathrm{II}}(t)$ for *each* $t \in I$. Of course, since γ_{I} remains a solution to (7) when the control γ_{II} gets redefined over a set of measure 0, one could identify two control functions whose values agree a.e. on I, as is customary in integration theory. However, these values are in any case subject to the constraint that $\gamma(t) \in \Omega$ for *every* $t \in I$, and altogether we find it more convenient to adopt Definition 3.1.

3.2. **Feedbacks.** In the terminology of control, a solution in the sense of Definition 3.1 would be termed *open loop* to emphasize that the value of the control at time t is a function of time only, namely that $\gamma_{\mathrm{II}}(t)$ bears no relation to the state x whatsoever. A central concept in control theory, though, is that of *closed loop* or *feedback* control, where the value of the control at time t is computed from the corresponding value of the state, namely is of the form $\alpha(x(t))$. To make a formal definition of a feedback defined on an arbitrary open set, we need one more piece of notation : if $\Omega \subset \mathbb{R}^n \times \mathbb{R}^m$ is open, we let $\pi_n : \Omega \to \Omega_{\mathbb{R}^n}$ the natural projection that selects the first n components, where $\Omega_{\mathbb{R}^n} = \pi_n(\Omega) \subset \mathbb{R}^n$.

Definition 3.3. Given an open set $\Omega \subset \mathbb{R}^{n+m}$, a *feedback* on Ω is a continuous mapping $\alpha : \Omega_{\mathbb{R}^n} \to \mathbb{R}^m$ such that $(x, \alpha(x)) \in \Omega$ for all $x \in \Omega_{\mathbb{R}^n}$. A $\mathbf{C}^\infty$(resp. $\mathbf{C}^\omega$) feedback on Ω is one of class $\mathbf{C}^\infty$(resp. $\mathbf{C}^\omega$).

A feedback is nothing but a mapping α such that $x \mapsto (x, \alpha(x))$ is a continuous section of the natural fibration $\pi_n : \Omega \to \Omega_{\mathbb{R}^n}$. Of course, there are sets Ω whose topology prevents the existence of any feedback. However, if there is one there are plenty, among which $\mathbf{C}^\infty$feedbacks are uniformly dense. This is the content of the next proposition, that will be used in the proof of Theorem 5.2. To fix notations, let us agree throughout that the symbol $\| \ \|$ designates the Euclidean norm on $\mathbb{R}^\ell$ irrespectively of the positive integer ℓ, while $B(x,r)$ stands for the open ball centered at x of radius r and $\overline{B}(x,r)$ for the corresponding closed ball.

Proposition 3.4. *Let Ω be open in $\mathbb{R}^{n+m}$, and $\alpha : \Omega_{\mathbb{R}^n} \to \mathbb{R}^m$ be a feedback on Ω. To each $\varepsilon > 0$, there is a $\mathbf{C}^\infty$feedback $\beta : \Omega_{\mathbb{R}^n} \to \mathbb{R}^m$ such that $\|\alpha(x) - \beta(x)\| < \varepsilon$ for $x \in \Omega_{\mathbb{R}^n}$.*

Proof. Let $\varnothing = \mathcal{K}_0 \subset \mathcal{K}_1 \cdots \subset \mathcal{K}_k \subset \mathcal{K}_{k+1} \cdots$ be an increasing sequence of compact subsets of $\Omega_{\mathbb{R}^n}$, each of which contains the previous one in its interior, and whose union is all of $\Omega_{\mathbb{R}^n}$. For each $x \in \Omega_{\mathbb{R}^n}$, define an integer

$$k(x) \ \stackrel{\Delta}{=} \ \min\{k \in \mathbb{N};\ x \in \mathcal{K}_k\}\ . \tag{9}$$

To each k, by the continuity of α and the compactness of $\mathcal{K}_k$, there is $\mu_k > 0$ such that

$$x \in \mathcal{K}_k \ \Rightarrow \left\{ \begin{array}{l} \bullet\ B(x,\mu_k) \times \mathrm{Conv}\left\{\, \alpha\big(B(x,\mu_k)\big) \,\right\} \ \subset \ \Omega\,, \\ \bullet\ \forall u_1, u_2 \in \mathrm{Conv}\left\{\, \alpha\big(B(x,\mu_k)\big) \,\right\},\ \|u_1 - u_2\| < \varepsilon\,, \end{array} \right. \tag{10}$$

where the symbol Conv designates the convex hull. In addition, we may assume that the sequence (μ_k) is non increasing.

Denote by $\overset{\circ}{\mathcal{K}}_k$ the interior of $\mathcal{K}_k$, set $\mathcal{D}_k = \mathcal{K}_k \setminus \overset{\circ}{\mathcal{K}}_{k-1}$ for $k \geq 1$, and cover the compact set $\mathcal{D}_k$ with a finite collection $\mathcal{B}_k$ of open balls having the following properties :

- each of these balls is centered at a point of $\mathcal{D}_k$ and is contained in the open set $\overset{\circ}{\mathcal{K}}_{k+1} \setminus \mathcal{K}_{k-2}$ (with the convention that $\mathcal{K}_{-1} = \varnothing$),
- each of these balls has radius at most $\dfrac{\mu_{k+1}}{2}$.

The union $\mathcal{B} = \bigcup_{k \geq 1} \mathcal{B}_k$ is a countable locally finite collection of open balls that covers $\Omega_{\mathbb{R}^n}$, and it has the property that *every ball in $\mathcal{B}$ is included in $B(x, \mu_{k(x)})$*

as soon as it contains x. Let B_j, for $j \in I\!\!N$, enumerate $\mathcal{B}$, and h_j be a smooth partition of unity where h_j has support $\text{supp}h_j \subset B_j$. If we pick $x_j \in B_j$ for each j, the map $\beta : \Omega_{I\!\!R^n} \to I\!\!R^m$ defined by

$$\beta(x) = \sum_{j \in I\!\!N} h_j(x)\alpha(x_j) \tag{11}$$

is certainly smooth. In addition, since by construction x_j belongs to $B(x, \mu_{k(x)})$ whenever $h_j(x) \neq 0$, we get that $\beta(x)$ lies in the convex hull of $\alpha\big(B(x,r)\big)$ for some $r < \mu_{k(x)}$, and therefore, from (10) and (9), that $(x, \beta(x)) \in \Omega$ and $\|\alpha(x) - \beta(x)\| < \varepsilon$. Hence β is a smooth feedback on Ω such that $\|\alpha(x) - \beta(x)\| < \varepsilon$ for all $x \in \Omega_{I\!\!R^n}$. □

3.3. **Conjugacy.** We turn to the notion of conjugacy for control systems, which is the central topic of the paper.

Definition 3.5. Let

$$\begin{array}{rcccl} \chi : & \Omega & \to & \Omega' & \\ & (x,u) & \mapsto & \chi(x,u) & = \ (\,\chi_{\mathrm{I}}(x,u)\,,\,\chi_{\mathrm{II}}(x,u)\,) \end{array} \tag{12}$$

be a bijective mapping between two open subsets of $I\!\!R^{n+m}$ and $I\!\!R^{n'+m'}$ respectively. We say that χ *conjugates* systems (4) and (5) if, for any real interval I, a map $\gamma : I \to \Omega$ is a solution of (4) that remains in Ω if, and only if, $\chi \circ \gamma$ is a solution of (5) that remains in Ω'.

Although this definition makes sense without any regularity assumption, we only consider the case when χ and χ^{-1} are at least continuous. Then Brouwer's invariance of the domain (see *e.g.* [17]) implies that $n' + m' = n + m$ if (4) and (5) are conjugate via such a χ. Proposition 3.6 below asserts that more in fact is true.

Proposition 3.6. *If the map* χ *in* (12) *is a homeomorphism that conjugates* (4) *to* (5)*, then* $n = n'$*,* $m = m'$*, and* χ_{I} *depends only on* x*:*

$$\chi(x,u) \;=\; (\,\chi_{\mathrm{I}}(x)\,,\,\chi_{\mathrm{II}}(x,u)\,)\ . \tag{13}$$

Moreover, $\chi_{\mathrm{I}} : \Omega_{I\!\!R^n} \to \Omega'_{I\!\!R^n}$ *is a homeomorphism. Here, one should recall the notation* $\Omega_{I\!\!R^n}$ *that was introduced before Definition 3.3.*

Proof. Let $\bar{x}$, $\bar{u}$, $\bar{u}'$ be such that $(\bar{x}, \bar{u})$ and $(\bar{x}, \bar{u}')$ belong to Ω. Let further $x(t)$ be a solution[1] to (4) with $x(0) = \bar{x}$ and

$$\begin{array}{l} u(t) = \bar{u} \text{ if } t \leq 0, \\ u(t) = \bar{u}' \text{ if } t > 0\ . \end{array}$$

By conjugacy, $z(t) = \chi_{\mathrm{I}}(x(t), u(t))$ is a solution to (5) with v given by $v(t) = \chi_{\mathrm{II}}(x(t), u(t))$, for $t \in (-\epsilon, \epsilon)$ and some $\epsilon > 0$. In particular $\chi_{\mathrm{I}}(x(t), u(t))$ is continuous in t so its values at 0^+ and 0^- are equal. Hence $\chi_{\mathrm{I}}(\bar{x}, \bar{u}) = \chi_{\mathrm{I}}(\bar{x}, \bar{u}')$ so that $\chi_{\mathrm{I}} : \Omega_{I\!\!R^n} \to \Omega'_{I\!\!R^{n'}}$ is well defined and continuous. Similarly, $\big(\chi^{-1}\big)_{\mathrm{I}}$ induces a continuous inverse $\Omega'_{I\!\!R^{n'}} \to \Omega_{I\!\!R^n}$. By invariance of the domain $n = n'$. □

In view of this proposition, we will only consider conjugacy between systems having the same number of states and inputs. Hence the distinction between (n, m) and (n', m') from now on disappears.

Remark 3.7. In the literature, there seems to be no general agreement on what should be called a solution of a control system, nor on the concept of equivalence. We discuss and compare some notions in use in section 3.5.

[1] This solution is not necessarily unique since here f and g are merely assumed to be continuous.

Remark 3.8. Taking into account the triangular structure of χ in Proposition 3.6, one may describe conjugacy as resulting from a change of coordinates in the state-space (upon setting $z=\chi_{\mathrm{I}}(x)$) and then feeding the system with a function both of the state and of a new control variable v (upon setting $u=(\chi^{-1})_{\mathrm{II}}(z,v)$), in such a way that the correspondence $(x,u)\mapsto(z,v)$ is invertible. In the language of control, this is known as a *static feedback transformation*, and two systems conjugate in the sense of Definition 3.10 would be termed *equivalent under static feedback*.

This notion has received considerable attention (see for instance [11]), albeit only in the *differentiable* case (i.e. when χ is a diffeomorphism). Differentiability has the following advantage : when χ_{I} and $(\chi_{\mathrm{I}})^{-1}$ are differentiable, χ conjugates systems (4) and (5) on some domain if, and only if

$$g(\chi_{\mathrm{I}}(x),\chi_{\mathrm{II}}(x,u)) \;=\; \frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\,f(x,u) \tag{14}$$

holds true on this domain. Hence one may replace Definition 3.10, which is based on *solutions* to (4) and (5), by the equality above expressing the way in which χ transforms the *equations*. Note that the differentiability of χ_{II} is not required.

Various degrees of regularity for χ give rise to corresponding notions of conjugacy in Definition 3.10 below.

Definition 3.9. For $k\in I\!N\cup\{\infty,\omega\}$, $k\geq 1$, a map χ as in (13) is called a *quasi-*$\mathbf{C}^k$ *diffeomorphism* if and only of it is $\mathbf{C}^0$ homeomorphism and χ_{I} is a $\mathbf{C}^k$ diffeomorphism $\Omega_{I\!R^n}\to\Omega'_{I\!R^n}$, *i.e.* χ_{I} and ${\chi_{\mathrm{I}}}^{-1}$ are of class $\mathbf{C}^k$.

Definition 3.10. Let $k\in I\!N\cup\{\infty,\omega\}$, $k\geq 1$.

Systems (4) and (5) are *topologically* (resp. $\mathbf{C}^k$, resp. *quasi-*$\mathbf{C}^k$) *conjugate over the pair* Ω,Ω' if there exists a homeomorphism (resp. $\mathbf{C}^k$ diffeomorphism, resp. quasi-$\mathbf{C}^k$ diffeomorphism) $\chi:\Omega\to\Omega'$ that conjugates the two systems.

System (4) is *locally* topologically ($\mathbf{C}^k$, quasi-$\mathbf{C}^k$) conjugate to system (5) at $(\bar{x},\bar{u})\in I\!R^{n+m}$ if[2] the two systems are topologically ($\mathbf{C}^k$, quasi-$\mathbf{C}^k$) conjugate over a pair Ω, Ω', where Ω is a neighborhood of $(\bar{x},\bar{u})$.

Remark 3.11. All definitions are invariant under linear time re-parameterization, namely : *if* $\chi:\Omega\to\Omega'$ *conjugates systems* (4) *and* (5), *then for any* $\lambda\in I\!R$ *(if* $\lambda<0$, *this reverses time) the map* χ *also conjugates the systems*

$$\dot{x}=\lambda f(x,u) \qquad \textit{and} \qquad \dot{z}=\lambda g(z,v)\ .$$

Indeed, this is trivial for $\lambda=0$, otherwise, if $t\mapsto(x(t),u(t))$ is a solution of $\dot{x}=\lambda f(x,u)$ on a time-interval $[t_1,t_2]$, and $\tilde{x}(t)$ and $\tilde{u}(t)$ denote respectively $x(t/\lambda)$ and $u(t/\lambda)$, then $t\mapsto(\tilde{x}(t),\tilde{u}(t))$ is a solution of (4) on $[\lambda t_1,\lambda t_2]$, hence χ sends $(\tilde{x}(t),\tilde{u}(t))$ to $(\tilde{z}(t),\tilde{v}(t))$ satisfying $\dot{\tilde{z}}(t)=g(\tilde{z}(t),\tilde{v}(t))$. Consequently, χ maps $(x(t),u(t))$ to $(z(t),v(t))=((\tilde{z}(\lambda t),\tilde{v}(\lambda t)))$, which is a solution of $\dot{z}=\lambda g(z,v)$.

In case there is no control (*i. e.* $m=m'=0$) so that neither u nor χ_{II} appear in (12), Definition 3.10 coincides with the usual notion of local conjugacy for ordinary differential equations.

3.4. **Properties of conjugating maps.** Below we derive some technical facts about conjugacy and feedback that are fundamental to the proof of Theorem 5.2, although they are not needed to understand the result itself.

In the proof of Proposition 3.6, we only used conjugacy on a very small class of solutions, namely those corresponding to piecewise constant controls with a single

[2]It would be more natural to say that system (4) **at** $(\bar{x},\bar{u})\in I\!R^{n+m}$ is locally conjugate to system (5) **at** $(\bar{x}',\bar{u}')\in I\!R^{n+m}$ if the two systems are conjugate over a pair Ω, Ω', where Ω is a neighborhood of $(\bar{x},\bar{u})$ and Ω' is a neighborhood of $(\bar{x}',\bar{u}')$. However, prescribing $(\bar{x}',\bar{u}')$ would increase notational burden and add no relevant information.

discontinuity. This raises the question whether smaller classes of solutions than prescribed in Definition 3.1 are still sufficiently rich to check for conjugacy. Under mild conditions on f and g, as we will see in the forthcoming proposition, conjugacy essentially holds if it is granted for a class of inputs that locally uniformly approximates piecewise continuous functions, and this fact will be of technical use in the proof of Lemma 6.3. To fix terminology, we agree that a function $I \to \mathbb{R}^m$, where I is a real interval, is called *piecewise continuous* if it is continuous except possibly at *finitely many* interior points of I where it has limits from both sides and is either right or left continuous. If in addition the function is constant (resp. affine, resp. $\mathbf{C}^\infty$) on every open interval not containing a discontinuity point, we say that it is *piecewise constant* (resp. *piecewise affine*, resp. *piecewise* $\mathbf{C}^\infty$).

Proposition 3.12 (Conjugacy from restricted classes of inputs)**.** *Assume that f and g are continuous $\mathbb{R}^n \times \mathbb{R}^m \to \mathbb{R}^n$ and locally Lipschitz-continuous with respect to their first argument*[3]*. Let $\chi : \Omega \to \Omega'$ be a homeomorphism between two open subsets of $\mathbb{R}^{n+m}$, and denote by Ω_{II} and Ω'_{II} respectively the open subsets of $\mathbb{R}^m$ obtained by projecting Ω and Ω' onto the second factor. Let further $\mathcal{C}$ and $\mathcal{C}'$ be collections of locally bounded measurable functions $\mathbb{R} \to \mathbb{R}^m$ whose restrictions $\mathcal{C}|_J$ and $\mathcal{C}'|_J$ to any compact interval J contain in their respective closures, for the topology of uniform convergence, the set of all piecewise continuous functions $J \to \Omega_{\mathrm{II}}$ and $J \to \Omega'_{\mathrm{II}}$ respectively. If χ maps every solution* (6) *of* (4) *such that $\gamma_{\mathrm{II}}(t) \in \mathcal{C}|_I$ to a solution of* (5) *while, conversely, χ^{-1} maps every solution* (8) *of* (5) *such that $\gamma'_{\mathrm{II}}(t) \in \mathcal{C}'|_I$ to a solution of* (4)*, then the restriction of χ to any relatively compact open subset $O \subset \Omega$ conjugates systems* (4) *and* (5) *over the pair O, $\chi(O)$.*

Proof. Let us first show that

$$\left.\begin{array}{l}\text{for any solution } \gamma : I \to \Omega \text{ of (4) } \textit{such that } \gamma_{\mathrm{II}} \textit{ is} \\ \textit{piecewise continuous}, \chi \circ \gamma \text{ is a solution to (5).}\end{array}\right\} \tag{15}$$

Since the property of being a solution is local with respect to time, we may suppose that I is a compact interval. Then, there is an open set $\mathcal{O}$ and a compact set $\mathcal{K}$ such that $\gamma(I) \subset \mathcal{O} \subset \mathcal{K} \subset \Omega$. By the hypothesis on $\mathcal{C}$, there exists a sequence of functions $\gamma_{\mathrm{II},k} : I \to \mathbb{R}^m$ converging uniformly to γ_{II} such that $\gamma_{\mathrm{II},k} \in \mathcal{C}|_I$. Define for each $k \in \mathbb{N}$ a time-varying vector field X^k by $X^k(t,x) = f(x, \gamma_{\mathrm{II},k}(t))$. By the continuity of f, this sequence converges uniformly on compact subsets of $I \times \mathbb{R}^n$ to $X(t,x) = f(x, \gamma_{\mathrm{II}}(t))$; moreover, since γ_{II} is bounded (being piecewise continuous) $\gamma_{\mathrm{II},k}$ is also bounded, thus the local Lipschitz character of $f(x,u)$ with respect to x implies by compactness that $X(t,x)$ and $X^k(t,x)$ are themselves locally Lipschitz with respect to x on $I \times \mathcal{O}_{\mathbb{R}^n}$. Pick $t_0 \in I$ and apply Lemma A.3 with $I = [t_1, t_2]$, $x_0 = \gamma_{\mathrm{I}}(t_0)$, and $\mathcal{U} = \mathcal{O}_{\mathbb{R}^n}$. This yields, say for $k > K$, that the solution $\gamma_{\mathrm{I},k}$ to the Cauchy problem

$$\dot{\gamma}_{\mathrm{I},k}(t) = X^k(t, \gamma_{\mathrm{I},k}(t))\,, \qquad \gamma_{\mathrm{I},k}(t_0) = \gamma_{\mathrm{I}}(t_0)\,,$$

maps I into $\mathcal{O}_{\mathbb{R}^n}$ and that the sequence $(\gamma_{\mathrm{I},k})_{k>K}$ converges uniformly on I to γ_{I}. Hence, if we let

$$\gamma_k(t) = (\gamma_{\mathrm{I},k}(t), \gamma_{\mathrm{II},k}(t)),$$

the sequence $(\gamma_k)_{k>K}$ converges to γ, uniformly on I. In particular $\gamma_k(I) \subset \mathcal{K} \subset \Omega$ for k large enough.

Now, since $\gamma_k : I \to \Omega$ is a solution to (4) with $\gamma_{\mathrm{II},k} \in \mathcal{C}|_I$, it follows from the hypothesis that $\chi \circ \gamma_k$ is a solution to (5) that remains in Ω', *i.e.* with the notations

[3]This means that each $(\bar{x}, \bar{u}) \in \Omega$ has a neighborhood $\mathcal{N}$ such that $\|f(x',u) - f(x,u)\| \leq c\,\|x' - x\|$ for some constant c whenever (x,u) and (x',u) lie in $\mathcal{N}$.

of (12) we have, for k large enough,

$$\chi_{\rm I}\circ\gamma_k(t)\ -\ \chi_{\rm I}\circ\gamma_k(t_0)\ =\ \int_{t_0}^t g(\chi\circ\gamma_k(s))\,{\rm d}s,\ t\in I. \tag{16}$$

By the continuity of χ, the convergence of $\gamma_k(t)$ to $\gamma(t)$, and the fact that g remains bounded on the compact set $\chi(\mathcal{K})$, we can apply the dominated convergence theorem to the right hand-side of (16) to obtain in the limit, as $k\to\infty$, that

$$\chi_{\rm I}\circ\gamma(t)\ -\ \chi_{\rm I}\circ\gamma(t_0)\ =\ \int_{t_0}^t g(\chi\circ\gamma(s))\,{\rm d}s,\ t\in I.$$

Thus $\chi\circ\gamma : I\to I\!R^{n+m}$ is a solution to (5) that remains in Ω', thereby proving (15).

The next step is to observe from (15) that, since piecewise constant controls are in particular piecewise continuous, the proof of Proposition 3.6 applies to show that $\chi:\Omega\to\Omega'$ has a triangular structure of the form (13).

With (15) and (13) at our disposal, let us now prove the proposition in its generality. Choose an arbitrary open subset $\mathcal{O}$ with compact closure $\overline{\mathcal{O}}$ in Ω, and fix two compact subsets $\mathcal{K}$ and $\mathcal{K}_1$ of Ω such that

$$\mathcal{O}\ \subset\ \overline{\mathcal{O}}\ \subset\overset{\circ}{\mathcal{K}}\subset\ \mathcal{K}\ \subset\overset{\circ}{\mathcal{K}}_1\subset\ \mathcal{K}_1\ \subset\ \Omega.$$

where $\overset{\circ}{\mathcal{K}}$ stands for the *interior* of $\mathcal{K}$.

Let $\gamma : I\to\mathcal{O}$ be a solution of (4). We need to prove that $\chi\circ\gamma$ is a solution to (5) and again, since the property of being a solution is local with respect to time, we may suppose that I is compact. Notations being as in (6), it follows by definition of a solution that $\gamma_{\rm II}$ is a bounded measurable function $I\to I\!R^m$. We shall proceed as before in that we again approximate γ by a sequence γ_k of trajectories of (4) that are mapped by χ to trajectories of (5). This time, however, the approximation process is slightly more delicate, because it is no longer granted by the hypothesis on $\mathcal{C}$ but it will rather depend on general point-wise approximation properties to measurable functions by continuous ones.

By the compactness of $\mathcal{K}$, there is $\varepsilon_{\mathcal{K}}>0$ such that

$$(x,u)\in\mathcal{K}\ \Rightarrow\ B\big((x,u)\,,\,\varepsilon_{\mathcal{K}}\big)\subset\overset{\circ}{\mathcal{K}}_1\,. \tag{17}$$

Let $u_{\gamma_{\rm I}} : I\to I\!R^m$ be an auxiliary function with the following properties :

(i) $u_{\gamma_{\rm I}}$ is piecewise constant on I,

(ii) $(\xi(t),u_{\gamma_{\rm I}}(t))\in\overset{\circ}{\mathcal{K}}_1$ for all $t\in I$ and every map $\xi : I\to I\!R^n$ that satisfies

$$\sup_{t\in I}\|\xi(t)-\gamma_{\rm I}(t)\|<\varepsilon_{\mathcal{K}}/2. \tag{18}$$

Such a function $u_{\gamma_{\rm I}}$ certainly exists. Indeed, by definition of a solution, $\gamma_{\rm I}$ is absolutely continuous thus *a fortiori* continuous $I\to I\!R^n$, and therefore we know for each $t\in I$ that the set

$$\gamma_{\rm I}^{-1}\big(B(\gamma_{\rm I}(t),\varepsilon_{\mathcal{K}}/2)\big)$$

is an open neighborhood of t in I, hence a disjoint union of open intervals in I one of which contains t; call this particular interval U_t. By the compactness of I, we may cover the latter with finitely many intervals U_{t_j} for $1\le j\le\nu$. Let now $j(t)$ denote, for each $t\in I$, the smallest index $j\in\{1,\dots,\nu\}$ such that $t\in U_{t_j}$. Then, the map

$$u_{\gamma_{\rm I}}(t)=\gamma_{\rm II}(t_{j(t)})$$

clearly satisfies (i), and since $(\gamma_{\rm I}(t_{j(t)}),\gamma_{\rm II}(t_{j(t)}))\in\mathcal{O}\subset\mathcal{K}$, it follows from (17) and the fact that $\|\gamma_{\rm I}(t)-\gamma_{\rm I}(t_{j(t)})\|<\varepsilon_{\mathcal{K}}/2$ by definition of $j(t)$ that $u_{\gamma_{\rm I}}$ also satisfies (ii).

Next, recall that $\gamma_{\rm II}$ is a bounded measurable function $I \to I\!R^m$ so, by Lusin's theorem [21, Theorem 2.23] applied component-wise, there is, for every integer $k \geq 1$, a continuous function $h_k : I \to I\!R^m$ that coincides with $\gamma_{\rm II}$ outside some set $\mathcal{T}_k \subset I$ of Lebesgue measure strictly less than $1/k^2$, and in addition such that

$$\sup_{t\in I} \|h_k(t)\| \leq \sqrt{m} \sup_{t\in I} \|\gamma_{\rm II}(t)\|. \tag{19}$$

Put $E_k = \{t \in I;\ (\gamma_{\rm I}(t), h_k(t)) \notin \overset{\circ}{\mathcal{K}}\}$. Since h_k is continuous E_k is compact, and since $\gamma(I) \subset \mathcal{O} \subset \overset{\circ}{\mathcal{K}}$ it is clear that $E_k \subset \mathcal{T}_k$ hence E_k has Lebesgue measure strictly less than $1/k^2$. Consequently, by the outer regularity of Lebesgue measure, E_k can be covered by finitely many open real intervals $I_{k,1}, \ldots, I_{k,N_k}$ whose lengths add up to no more than $1/k^2$.

We now define the sequence of functions $\gamma_{{\rm II},k}$ on I by setting, for $k \geq 1$,

$$\begin{array}{rcl} \gamma_{{\rm II},k}(t) & = & h_k(t) \text{ if } t \in I \setminus \bigcup_{j=1}^{N_k} I_{k,j}, \\ \gamma_{{\rm II},k}(t) & = & u_{\gamma_{\rm I}}(t) \text{ if } t \in \bigcup_{j=1}^{N_k} I_{k,j}. \end{array} \tag{20}$$

By construction $\gamma_{{\rm II},k}$ is piecewise continuous, and uniformly bounded independently of k in view of (19) and the fact that $u_{\gamma_{\rm I}}$, being piecewise constant, is bounded. Moreover, as $\sum_{k\geq 1} 1/k^2 < \infty$, the measure of the set $\cup_{j=1}^{N_k} I_{k,j}$ is the general term, indexed by k, of a convergent series, hence almost every $t \in I$ belongs at most to finitely many of these sets so that $\gamma_{{\rm II},k}$ converges point-wise a.e. to $\gamma_{\rm II}$ on I as $k \to \infty$.

Redefine now $X^k(t,x) = f(x, \gamma_{{\rm II},k}(t))$, $X(t,x) = f(x, \gamma_{\rm II}(t))$, and observe from what we just said and the continuity of f that $X^k(t,x)$ converges to $X(t,x)$ when $k \to \infty$, locally uniformly with respect to $x \in \mathcal{O}_{I\!R^n}$, as soon as $t \notin E$ where $E \subset I$ is a set of zero measure which is independent of k. Moreover, again from the boundedness of $\gamma_{{\rm II},k}$, $\gamma_{\rm II}$ and the local Lipschitz character of f, we have that $X^k(t,x)$, $X(t,x)$ are locally Lipschitz with respect to x. Pick $t_0 \in I$ and apply Lemma A.3 with $\mathcal{U} = \mathcal{O}_{I\!R^n}$, $I = [t_1, t_2]$, and $x_0 = \gamma_{\rm I}(t_0)$. We get, say for $k > K$, that the solution $\gamma_{{\rm I},k}$ to the Cauchy problem

$$\dot\gamma_{{\rm I},k}(t) = X^k(t, \gamma_{{\rm I},k}(t))\,, \qquad \gamma_{{\rm I},k}(t_0) = \gamma_{\rm I}(t_0)\,,$$

is defined over I, maps the latter into $\mathcal{O}_{I\!R^n}$, and that the sequence $(\gamma_{{\rm I},k})_{k>K}$ converges uniformly on $[t_1, t_2]$ to $\gamma_{\rm I}$.

We claim that $\gamma_k(t) = (\gamma_{{\rm I},k}(t), \gamma_{{\rm II},k}(t))$ lies in $\overset{\circ}{\mathcal{K}}_1$ for all $t \in I$ when k is so large that

$$\sup_{t\in I} \|\gamma_{{\rm I},k}(t) - \gamma_{\rm I}(t)\| < \varepsilon_{\mathcal{K}}/2. \tag{21}$$

Indeed, if $t \in \cup_j I_{k,j}$, this follows automatically from definition (20) by property (ii) of $u_{\gamma_{\rm I}}$; if $t \notin \cup_j I_{k,j}$, then $(\gamma_{\rm I}(t), h_k(t)) \in \overset{\circ}{\mathcal{K}}$ by the very definition of $\cup_j I_{k,j}$, and since $\gamma_k(t) = (\gamma_{{\rm I},k}(t), h_k(t))$ in this case, we deduce from (17) and (21) that $\gamma_k(t) \in \overset{\circ}{\mathcal{K}}_1$. *This proves the claim.*

Altogether, we have shown that $\gamma_k : I \to \overset{\circ}{\mathcal{K}}_1$ is a solution of (4) as soon as k is large enough, with $\gamma_{{\rm II},k}$ a piecewise continuous function on I by construction. By (15), we now deduce that, for k large enough, $\gamma'_k = \chi \circ \gamma_k$ is a solution of (5) that stays in Ω'. Let us block-decompose γ'_k into

$$\gamma'_{{\rm I},k}(t) \;=\; \chi_{\rm I}(\,\gamma_{{\rm I},k}(t)\,)\;, \quad \gamma'_{{\rm II},k}(t) \;=\; \chi_{\rm II}(\,\gamma_{{\rm I},k}(t)\,,\,\gamma_{{\rm II},k}(t)\,)\;,$$

where we have taken into account the triangular structure of χ. That $\gamma'_k : I \to \Omega'$ is a solution of (5) means exactly that

$$\gamma'_{\mathrm{I},k}(t) \;-\; \gamma'_{\mathrm{I},k}(t_0) \;=\; \int_{t_0}^{t} g(\,\gamma'_{\mathrm{I},k}(s)\,,\,\gamma'_{\mathrm{II},k}(s)\,)\mathrm{d}s,\ t\in I. \tag{22}$$

Due to the continuity of χ, the functions $\gamma'_{\mathrm{I},k}$ and $\gamma'_{\mathrm{II},k}$ respectively converge uniformly and point-wise almost everywhere to $\gamma'_{\mathrm{I}} = \chi_{\mathrm{I}}\circ\gamma_{\mathrm{I}}$ and $\gamma'_{\mathrm{II}} = \chi_{\mathrm{II}}\circ\gamma$ on I. Since g is bounded on the compact set $\chi(\mathcal{K}_1)$ that contains $\gamma_k(I)$ for k large enough, we get on the one hand, by dominated convergence, that the right-hand side of (22) converges, as $k\to\infty$, to $\int_{t_0}^{t} g(\,\gamma'_{\mathrm{I}}(s)\,,\,\gamma'_{\mathrm{II}}(s)\,)\mathrm{d}s$, and on the other hand that the left-hand side converges to $\gamma'_{\mathrm{I}}(t)-\gamma'_{\mathrm{I}}(t_0)$. Therefore $(\gamma'_{\mathrm{I}},\gamma'_{\mathrm{II}}) = \chi_\circ\gamma : I\to\Omega'$ is a solution of (5).

This way we have shown that χ maps any solution of (4) that stays in a relatively compact open subset $\mathcal{O}$ of Ω to a solution of (5) that stays in Ω'. This achieves the proof, for the converse is obtained symmetrically upon swapping f and g, $\mathcal{C}$ and $\mathcal{C}'$, and replacing χ by χ^{-1}. □

The triangular structure of conjugating homeomorphisms asserted by Proposition 3.6 is to the effect that any such homeomorphism $\chi:\Omega\to\Omega'$ is a fiber preserving map from the bundle $\Omega\to\Omega_{\mathbb{R}^n}$ to the bundle $\Omega'\to\Omega'_{\mathbb{R}^n}$. Since feedbacks are naturally associated to sections of these bundles by Definition 3.3, χ gives rise to a natural transformation from feedbacks on Ω to feedbacks on Ω'. This transformation will prove important enough to deserve a notation : to any feedback α on Ω, we associate a feedback $\chi\,\blacksquare\,\alpha$ on Ω' by the formula

$$\chi\,\blacksquare\,\alpha(z) \;\triangleq\; \chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),\alpha(\chi_{\mathrm{I}}^{-1}(z)))\,. \tag{23}$$

We leave it to the reader to check that the properties of an action are satisfied, and in particular that

$$\chi^{-1}\,\blacksquare\,(\chi\,\blacksquare\,\alpha) \;=\; \alpha\ . \tag{24}$$

Naturally associated to a control system (4) and a feedback α is the following continuous vector field f_α on $\Omega_{\mathbb{R}^n}$:

$$f_\alpha(x) \;=\; f(x,\alpha(x))\ . \tag{25}$$

If the homeomorphism χ in (13) conjugates system (4) to system (5), then it is clear that χ_{I} maps the solutions of the ordinary differential equation $\dot x = f_\alpha(x)$ to the solutions of the ordinary differential equation $\dot z = g_{\chi\,\blacksquare\,\alpha}(z)$. Indeed if $x(t)$ is a solution of the former, then $(x(t),\alpha(x(t)))$ is a solution of the control system (4) in the sense of Definition 3.1 so the conjugacy assumption implies that $(\chi_{\mathrm{I}}(x(t)),\chi_{\mathrm{II}}(x(t),\alpha(x(t))))$ is a solution of (5), and setting $z(t)=\chi_{\mathrm{I}}(x(t))$ one clearly has $\chi_{\mathrm{II}}(x(t),\alpha(x(t)))) = \chi\,\blacksquare\,\alpha(z(t))$; hence $z(t)$ is a solution to $\dot z = g_{\chi\,\blacksquare\,\alpha}(z)$ because $(z(t),\chi\,\blacksquare\,\alpha(z(t)))$ is a solution of (5).

Now, if α_1 and α_2 are two feedbacks on Ω, and the two vector fields f_{α_1} and f_{α_2} are defined on $\Omega_{\mathbb{R}^n}$ by (25), we denote their difference by $\delta f_{\alpha_1,\alpha_2}$:

$$\delta f_{\alpha_1,\alpha_2} \;=\; f_{\alpha_1} \;-\; f_{\alpha_2}\ . \tag{26}$$

Such vector fields are similar to the difference vector fields used in [15], except that we consider arbitrary feedbacks instead of constant ones. To us, these vector fields will play an essential role. The next proposition states that a homeomorphism that conjugates two control systems also conjugates the integral curves of such difference vector fields.

Proposition 3.13 (preservation of difference vector fields)**.** *Suppose that f and g in* (4) *and* (5) *are continuous and locally Lipschitz continuous with respect to their first argument. Assume they are locally topologically conjugate at* $(0,0)$ *over the* *pair Ω, Ω'. Then, notations for $\chi_{\rm I}$ and $\chi_{\rm II}$ being as in Proposition 3.6, we have for every pair of feedbacks α_1, α_2 on Ω that $\chi_{\rm I}$ conjugates any solution of*

$$\dot{x} \;=\; \delta f_{\alpha_1,\alpha_2}(x) \tag{27}$$

that remains in $\Omega_{I\!R^n}$ to a solution of

$$\dot{z} \;=\; \delta g_{\chi\blacksquare\alpha_1,\chi\blacksquare\alpha_2}(z) \tag{28}$$

that remains in $\Omega'_{I\!R^n}$.

It is perhaps worth emphasizing that the solutions of (27) and (28) need not be unique since α is merely assumed to be continuous.

Proof. Let $\eta : [t_1, t_2] \to \Omega_{I\!R^n}$ be an integral curve of $\delta f_{\alpha_1,\alpha_2}$, and set

$$u_1(t) \;=\; \alpha_1(\eta(t)) \;\;,\qquad u_2(t) \;=\; \alpha_2(\eta(t)) \;. \tag{29}$$

Let further $\widehat{f} : I\!R^{n+m} \to I\!R^n$ be bounded, continuous and Lipschitz continuous with respect to its first argument, and coincide with f on some compact neighborhood of

$$\eta([t_1,t_2]) \times \Big(\alpha_1(\eta([t_1,t_2])) \bigcup \alpha_2(\eta([t_1,t_2]))\Big) \;.$$

Such a $\widehat{f}$ is easily obtained upon multiplying f by a function of class $\mathbf{C}^\infty$with compact support. For $\ell \in I\!N$, let η^ℓ be the solution to the Cauchy problem

$$\eta^\ell(t) \;=\; \eta(t_1) \;+\; \int_{t_1}^{t} G_\ell(\tau, \eta^\ell(\tau))\mathrm{d}\tau \;, \tag{30}$$

with

$$\begin{aligned} G_\ell(t,x) \;&=\; 2\,\widehat{f}(x,u_1(t)) \\ &\qquad\quad \text{if } t \in [t_1 + \tfrac{j}{\ell}(t_2-t_1), t_1 + (\tfrac{j}{\ell} + \tfrac{1}{2\ell})(t_2-t_1)), \\ G_\ell(t,x) \;&=\; -2\,\widehat{f}(x,u_2(t)) \\ &\qquad\quad \text{if } t \in [t_1 + (\tfrac{j}{\ell} + \tfrac{1}{2\ell})(t_2-t_1), t_1 + \tfrac{j+1}{\ell}(t_2-t_1)), \\ G_\ell(t_2,x) \;&=\; -2\,\widehat{f}(x,u_2(t_2)), \qquad 0 \le j \le \ell-1. \end{aligned} \tag{31}$$

The definition of η^ℓ is valid because, since $G_\ell(t,x)$ is bounded and locally Lipschitz with respect to the variable x, the solution to (30) uniquely exists.

From Lemma A.4 applied to the case where $X^{1,\ell}(t,x) = \widehat{f}(x,u_1(t))$ and $X^{2,\ell}(t,x) = \widehat{f}(x,u_2(t))$ are in fact independent of ℓ, any accumulation point of the sequence (η^ℓ), say η^∞, is a solution to

$$\dot{\eta}^\infty(t) \;=\; \widehat{f}(\eta^\infty(t),u_1(t)) \;-\; \widehat{f}(\eta^\infty(t),u_2(t)) \;\;,\qquad \eta^\infty(t_1) \;=\; \eta(t_1) \;.$$

Since $\widehat{f}$ is locally Lipschitz continuous with respect to its first argument, the solution to this Cauchy problem is unique and, since f and $\widehat{f}$ coincide at all points $(\eta(t), u_1(t))$ and $(\eta(t), u_2(t))$, this entails $\eta^\infty = \eta$. Thus (η^ℓ) converges uniformly to η on $[t_1,t_2]$ and, for ℓ large enough, η^ℓ remains a solution of (30) if $\widehat{f}$ is replaced by f in (31). Moreover, $\eta^\ell([t_1,t_2]) \subset \Omega_{I\!R^n}$ for ℓ large since the same is true of η. Since χ conjugates the two systems, hence also by Remark 3.11 the systems where f and g are multiplied by 2 or -2, the map $\chi_{\rm I} \circ \eta^\ell : [t_1,t_2] \to \Omega'_{I\!R^n}$ is, for ℓ large enough, a solution to

$$\chi_{\rm I} \circ \eta^\ell(t) \;=\; \chi_{\rm I} \circ \eta(t_1) \;+\; \int_{t_1}^{t} \widetilde{G}_\ell(\tau, \chi_{\rm I} \circ \eta^\ell(\tau))\mathrm{d}\tau \tag{32}$$

with

$$\begin{aligned}
\widetilde{G}_\ell(t,z) &= 2\,g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),u_1(t))) \\
&\qquad \text{if } t\in[t_1+\tfrac{j}{\ell}(t_2-t_1),t_1+(\tfrac{j}{\ell}+\tfrac{1}{2\ell})(t_2-t_1)), \\
\widetilde{G}_\ell(t,z) &= -2\,g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),u_2(t))) \\
&\qquad \text{if } t\in[t_1+(\tfrac{j}{\ell}+\tfrac{1}{2\ell})(t_2-t_1),t_1+\tfrac{j+1}{\ell}(t_2-t_1)), \\
\widetilde{G}_\ell(t_2,z) &= -2\,g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),u_2(t_2))).
\end{aligned} \tag{33}$$

Since $(\chi_{\mathrm{I}}\circ\eta^\ell)$ converges uniformly to $\chi_{\mathrm{I}}\circ\eta$ by the continuity of χ, replacing g by a bounded and continuous $\widehat{g}:\mathbb{R}^{n+m}\to\mathbb{R}^n$ that coincides with g on a compact neighborhood of

$$\chi_{\mathrm{I}}\circ\eta([t_1,t_2])\times\Big(\chi_{\mathrm{II}}(\eta([t_1,t_2]),\alpha_1(\eta([t_1,t_2])))\bigcup\chi_{\mathrm{II}}(\eta([t_1,t_2]),\alpha_2(\eta([t_1,t_2])))\Big)$$

does not affect the validity of (32)-(33) for ℓ large enough. Lemma A.4 now implies that all accumulation points of the sequence $(\chi_{\mathrm{I}}\circ\eta^\ell)$ in the uniform topology on $[t_1,t_2]$ are solutions of

$$\dot{z}=g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),u_1(t)))\;-\;g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),u_2(t))).$$

Because $\chi_{\mathrm{I}}\circ\eta$ is such an accumulation point, it is by (29) a solution to

$$\dot{z}\;=\;g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),\alpha_1(\chi_{\mathrm{I}}^{-1}(z)))\;-\;g(z,\chi_{\mathrm{II}}(\chi_{\mathrm{I}}^{-1}(z),\alpha_2(\chi_{\mathrm{I}}^{-1}(z)))\;,$$

which is nothing but (28). □

3.5. **Alternative notions of conjugacy and equivalence.**

3.5.1. *Transformations in functional spaces.* Following [7], one may view the control system (4) as a flow on the product space $\mathbb{R}^n\times\mathcal{U}$, where $\mathcal{U}$ is a functional space of admissible controls whose dynamics is induced by the time-shift. Transformations on $\mathbb{R}^n\times\mathcal{U}$ then naturally arise; they involve the *future and the past of the control*, unlike the mere homeomorphisms on finite dimensional spaces that we consider here. The corresponding notion of equivalence is obviously rather weak. In [3], a "Grobman-Hartman theorem" theorem is proved in this setting, i.e. generic control systems (4) are locally conjugate to a linear system via this kind of transformation. With the much stronger notion of equivalence that we use here, we shall see (section 5.4) that "almost" no system is conjugate to a linear system.

Let us also mention [5], where control systems are maps $(x(0),u(.))\mapsto x(.)$ that satisfy certain axioms, without reference to differential equations, and where the notion of topological equivalence involves transformations on the product $\mathbb{R}^n\times\mathcal{U}$.

3.5.2. *x-conjugacy.* Let us call *x-solution* of system (4) any map $t\mapsto\gamma_{\mathrm{I}}(t)$ *such that there exists* a map γ_{II} for which $\gamma=(\gamma_{\mathrm{I}},\gamma_{\mathrm{II}})$ is a solution in the sense of Definition 3.1; the set of x-solutions is the projection on the x factor of the set of solutions. Let then x-conjugacy be defined in the same way as Definition 3.10 defines conjugacy, except that we replace solutions by x-solutions and the homeomorphism χ that acts on state and control with a homeomorphism $x\mapsto z=h(x)$ on the state only.

In the literature, both notions are used (without the prefix "x-"). For instance [29], devoted to the topological classification of *linear* control systems (see section 4.2) relies on x-conjugacy. We favor Definitions 3.5 and 3.10 of conjugacy and solutions because results have to be stated locally with respect both to x and u for nonlinear control systems.

Conjugacy implies x-conjugacy: use Proposition 3.6, take $h=\chi_{\mathrm{I}}$ and ignore χ_{II}. The converse is not true in general, as the reader may check easily.

4. The case of linear control systems

4.1. **Kronecker indices.** A *linear* control systems is a special instance of (4), of the form

$$\dot{x} \;=\; Ax \;+\; Bu \tag{34}$$

where A and B are constant $n \times n$ and $n \times m$ matrices respectively. When dealing with linear systems, it is natural to consider an equivalence relation similar to that of Definition 3.10, but where χ is restricted to be a linear isomorphism :

Definition 4.1. Two linear systems

$$\dot{x} \;=\; Ax \;+\; Bu \qquad \text{and} \qquad \dot{z} \;=\; \widetilde{A}z \;+\; \widetilde{B}v$$

are *linearly conjugate* if and only if any of the following two equivalent properties is satisfied :

(1) There is a nonempty open set $\Omega \subset I\!R^{n+m}$, and a linear isomorphism χ of $I\!R^{n+m}$ whose restriction $\Omega \to \chi(\Omega)$ conjugates the two systems in the sense of Definition 3.10.
(2) There exist matrices $P \in I\!R^{n\times n}$, $Q \in I\!R^{m\times m}$ and $K \in I\!R^{n\times m}$, with P and Q invertible, such that

$$\begin{array}{rcl} \widetilde{A} &=& P(A-BK)P^{-1} \;, \\ \widetilde{B} &=& PBQ^{-1} \;. \end{array} \tag{35}$$

Since, by Proposition 3.6, a linear conjugating homeomorphism is necessarily of the form $(x,u) \mapsto (Px, Kx+Qu)$, the equivalence between properties (1) and (2) follows at once from differentiating the solutions. Provided it exists, Ω plays absolutely no role in this context since (35) implies that the two systems are in fact linearly conjugate on all of $I\!R^{n+m}$.

Linear conjugacy actually defines an equivalence relation on linear control systems or equivalently on pairs (A,B), for which (35) can be read as "(A,B) is equivalent to $(\widetilde{A},\widetilde{B})$". The classification of linear systems under this equivalence relation is well-known [4], and goes as follows. Each equivalence class contains a pair (A_c, B_c) of the form (block matrices) :

$$A_c \;=\; \begin{pmatrix} A_0^c & 0 & \cdots & 0 \\ 0 & A_1^c & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & A_m^c \end{pmatrix} \;, \quad B_c \;=\; \begin{pmatrix} 0 & \cdots & 0 \\ b_1^c & \ddots & \vdots \\ 0 & \ddots & 0 \\ \vdots & 0 & b_m^c \end{pmatrix} \tag{36}$$

where

$$A_i^c \;=\; \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & & & 0 \\ & & & \ddots & 1 \\ 0 & \cdots & & \cdots & 0 \end{pmatrix}_{(\kappa_i\times\kappa_i)} , \qquad b_i^c \;=\; \begin{pmatrix} 0 \\ \vdots \\ \vdots \\ 0 \\ 1 \end{pmatrix}_{(\kappa_i\times 1)} , \quad 1 \le i \le m. \tag{37}$$

The integers $(\kappa_1, \ldots, \kappa_m)$ are called the controllability indices of the control system, also known as the Kronecker indices of the matrix pencil (A,B), while A_0^c is a square matrix of dimension $n-(\kappa_1+\cdots+\kappa_m)$ that may be assumed in Jordan canonical form. Note that $\kappa_1+\cdots+\kappa_m \le n$, and if $\kappa_1+\cdots+\kappa_m = n$ there is no A_0^c ; also, it may well happen that $\kappa_i = 0$, in which case A_i^c and b_i^c are empty and do not occur in (36) to the effect that there are less than m blocks beyond A_0^c. Normalizing so that

$$\kappa_1 \geq \cdots \geq \kappa_m \geq 0,$$

and ordering the Jordan blocks arbitrarily, there is one and only one such normal form per equivalence class. A complete set of invariants is then the list of Kronecker indices and the spectral invariants of the matrix A_0^c.

With the natural partition $z = (Z_0, Z_1, \ldots, Z_m)$ corresponding to the block decomposition (36), the control system associated to the pair (A_c, B_c) reads

$$\dot{Z}_0 \;=\; A_0 Z_0 \;, \quad \dot{Z}_1 \;=\; A_1 Z_1 \;+\; u_1 b_1^c \;, \quad \cdots \quad , \quad \dot{Z}_m \;=\; A_m Z_m \;+\; u_m b_m^c \;,$$

where Z_0 is missing if $\kappa_1 + \cdots + \kappa_m = n$ and Z_i is missing if $\kappa_i = 0$. Because it is not influenced at all by the controls, Z_0 is sometimes called the non-controllable part of the state. In this paper, we are only interested in controllable linear systems, namely :

Definition 4.2. A linear control system (34) is said to be *controllable* if, and only if, the following two equivalent properties are satisfied :

(1) There is no bloc A_0^c in the associated normal form (36).
(2) Kalman's criterion for controllability :

$$\operatorname{Rank}(B, AB, \ldots, A^{n-1}B) = n.$$

To see the equivalence of the two properties, observe that the $n - \kappa_1 - \cdots - \kappa_m$ first rows of the matrix P that puts (A, B) into canonical form (i.e. $z = Px$) form a basis of the smallest dual subspace that annihilates the columns of B and at the same time is invariant under right multiplication by A, i.e. they are a basis of the left kernel of $(B, AB, \ldots, A^{n-1}B)$. For controllable linear systems, the only invariant under linear conjugacy is thus the ordered list of Kronecker indices. These can be computed from $(B, AB, \ldots, A^{n-1}B)$ as follows : if we put

$$\begin{aligned} r_j \;&=\; \operatorname{Rank}(B, AB, \ldots, A^{j-1}B) \;,\; j \geq 1 \;, \qquad r_0 = 0,\; r_{-1} = -m \;, \\ s_j \;&=\; r_j - r_{j-1} \;,\; j \geq 1 \;, \qquad s_0 = m \;, \end{aligned} \tag{38}$$

then s_j does not increase with j and a moment's thinking will convince the reader that the number of Kronecker indices that are equal to i is $s_i - s_{i+1}$, or equivalently that s_k is the number of κ_j's that are no smaller than k.

To us, it will be more convenient to use as normal form the following permutation of the previous one. Let ρ be the smallest integer such that $s_\rho = 0$, so that

$$0 \;=\; s_\rho \;<\; s_{\rho-1} \;\leq\; s_{\rho-2} \;\leq\; \cdots \;\leq\; s_1 \;\leq\; s_0 \;=\; m \;,$$

with $\sum_{j\geq 1} s_j = n$. From these we define, for $0 \leq i \leq \rho$:

$$\sigma_i \;=\; \sum_{j\geq i} s_j \;=\; n - r_{i-1} \,, \tag{39}$$

so that in particular $\sigma_\rho = 0$, $\sigma_{\rho-1} = s_{\rho-1} > 0$, $\sigma_1 = n$ and $\sigma_0 = n + m$. Note that, from (38), $\sigma_i = n - r_{i-1}$ for $i \geq 1$. We shall write our controllable canonical form as $\dot{z} = A_c z + B_c v$ with

$$A_c=\left(\begin{array}{c|c|c|c|c|c} 0 & J^{s_{\rho-1}}_{s_{\rho-2}} & & & & \\ \hline & 0 & J^{s_{\rho-2}}_{s_{\rho-3}} & & & \\ \hline & & 0 & J^{s_{\rho-3}}_{s_{\rho-4}} & & \\ \hline & & & \ddots & \ddots & \\ \hline & & & & 0 & J^{s_2}_{s_1} \\ \hline & & & & & 0 \end{array}\right), \quad B_c=\left(\begin{array}{c} 0 \\ \hline J^{s_1}_{s_0} \end{array}\right) \tag{40}$$

where for any integers r and s with $s \leq r$, J^s_r is the $s \times r$ matrix

$$J^s_r = \left(\begin{array}{c|c} I_s & 0 \end{array}\right) \tag{41}$$

where I_s is the $s \times s$ identity matrix.

4.2. **Topological classification of linear control systems.** In [29], which is devoted to the topological classification of *linear* control systems and uses the notion of x-conjugacy rather than conjugacy (*cf* Section 3.5), the following result is proved:

Theorem 4.3 (Willems [29])**.** *If two linear control systems $\dot{x} = Ax + Bu$ and $\dot{z} = \widetilde{A}z + \widetilde{B}v$ are topologically x-conjugate, then they have the same list of Kronecker indices, and the non-controllable blocks A_0^c and $\widetilde{A}_0^c$ in their respective canonical forms* (36) *are such that the two linear differential equations $\dot{X}_0 = A_0^c X_0$ and $\dot{Z}_0 = \widetilde{A}_0^c Z_0$ are topologically equivalent.*

As pointed out in Section 3.5, topological conjugacy implies topological x-conjugacy but not conversely. However, for linear control systems having the same number m of inputs, Theorem 4.3 implies that these notions are equivalent. Indeed, if two systems are respectively brought into their canonical form (36) by a linear change of variable on $I\!R^{n+m}$, and if in addition they are x-conjugate, then their non-controllable parts are topologically equivalent while the remaining blocks are identical by equality of the Kronecker indices. Hence, both in the above theorem and in the corollary below, one may use indifferently "x-conjugate" or "conjugate"

Corollary 4.4. *If two linear systems $\dot{x} = Ax + Bu$ and $\dot{z} = \widetilde{A}z + \widetilde{B}v$ are topologically conjugate and one of them is controllable, then the other one is controllable too and they are linearly conjugate.*

Proof. Controllability is preserved, since Kronecker indices are by the theorem. Linear conjugacy follows, as we saw that the list of Kronecker indices is a complete invariant for controllable systems under linear conjugacy. □

In some sense, the results of section 5 can be viewed as a generalization of Corollary 4.4 to a local setting where only one of the two systems is linear.

5. Local linearization for control systems

In this section, we consistently assume that the map f defining system (4) is either smooth or real-analytic.

Definition 5.1. Let $k \in \{\infty, \omega\}$. The system (4) is said to be *locally topologically* (resp. $\mathbf{C}^k$, resp. *quasi*-$\mathbf{C}^k$) *linearizable at* $(\bar{x}, \bar{u}) \in I\!R^{n+m}$ if it is locally topologically (resp. $\mathbf{C}^k$, resp. *quasi*-$\mathbf{C}^k$) conjugate, in the sense of Definition 3.10, to a linear controllable system $\dot{z} = Az + Bv$ (*cf.* Definition 4.2).

This definition of smooth linearizability coincides with linearizability by smooth static feedback as described in the textbooks [10, 19]. In subsection 5.2, we recall classical necessary and sufficient geometric conditions for a system to be smoothly (resp. analytically) linearizable, and we complement them with a characterization of quasi-smooth (resp. quasi-analytic) linearizability.

5.1. **Main result.** If a smooth control system is locally topologically linearizable, then the conjugating homeomorphism has a lot more regularity than required *a priori*. This is in contrast with the Grobman-Hartman theorem for ODE's and constitutes the central result of the paper:

Theorem 5.2. *Let $k \in \{\infty, \omega\}$ and assume that f is of class $\mathbf{C}^k$ on an open set $\Omega \subset I\!R^{n+m}$. Then system (4) is locally topologically linearizable at $(\bar{x}, \bar{u}) \in \Omega$ if, and only if, it is locally quasi-$\mathbf{C}^k$ linearizable at $(\bar{x}, \bar{u})$.*

Proof. See section 6.2. □

Observe from (14), that a quasi-$\mathbf{C}^k$ diffeomorphism χ is a linearizing homeomorphism if and only if it satisfies

$$\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\, f(x,u) \;=\; A\,\chi_{\mathrm{I}}(x) \;+\; B\,\chi_{\mathrm{II}}(x,u)\ . \tag{42}$$

Hence quasi-smooth linearizability is much easier to handle than topological linearizability, that relies on conjugating *solutions* rather than equations.

System (1) of the introduction is topologically, quasi-$\mathbf{C}^\omega$and quasi-$\mathbf{C}^\infty$linearizable at $(0,0)$ but fails to be even $\mathbf{C}^1$ linearizable; hence quasi-$\mathbf{C}^k$ cannot be replaced with $\mathbf{C}^k$ in Theorem 5.2. To study the gap between $\mathbf{C}^k$ and quasi-$\mathbf{C}^k$ linearizability, note that (42) imposes additional regularity on a linearizing quasi-$\mathbf{C}^k$ diffeomorphism :

Proposition 5.3. *Let $k \in \{\infty, \omega\}$ and f in (4) be $\mathbf{C}^k$. If $\chi : \Omega \to \Omega'$ is a quasi-$\mathbf{C}^k$ diffeomorphism that conjugates (4) to the linear system $\dot{z} = Az + Bv$, then :*

1. *the map $B\chi_{\mathrm{II}} : \Omega \to I\!R^m$ is of class $\mathbf{C}^k$,*
2. *for any $(x,u) \in \Omega$ in the neighborhood of which the rank of $\partial f/\partial u$ is constant, one has* $\mathrm{Rank}\, \dfrac{\partial f}{\partial u}(x,u) = \mathrm{Rank}\, B$.
3. *for any open subset O of Ω, one has* $\displaystyle\sup_{(x',u')\in O} \mathrm{Rank}\, \frac{\partial f}{\partial u}(x',u') = \mathrm{Rank}\, B$.

Proof. Point (1) is direct consequence of (42) and the smoothness of χ_{I} and f. To establish (2) and (3), differentiate (42) with respect to u to obtain

$$\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x)\frac{\partial f}{\partial u}(x,u) \;=\; \frac{\partial (B\chi_{\mathrm{II}})}{\partial u}(x,u)\ . \tag{43}$$

Let $\mathcal{V} \subset \Omega$ be open and such that $\mathrm{Rank}\, \partial f/\partial u(x,u) = \rho$ some integer ρ and all $(x,u) \in \mathcal{V}$. Define $\phi : \mathcal{V} \to I\!R^{n+m}$ by $\phi(x,u) = (\chi_{\mathrm{I}}(x), B\chi_{\mathrm{II}}(x,u))$. On the one hand, since χ_{I} is a diffeomorphism, (43) implies that the rank of the Jacobian of ϕ is $n+\rho$, hence, by the constant rank theorem, $\phi(\mathcal{V})$ is a $(n+\rho)$ dimensional immersed sub-manifold of $\mathbb{R}^{n+m}$; on the other hand, since χ is open, $\phi(\mathcal{V})$ is an open subset of the $(n + \operatorname{Rank} B)$-dimensional linear range of $I_n \times B$; hence $\rho = \operatorname{Rank} B$. This proves point 2, and at the same time point 3 because any $O \subset \Omega$ contains an open subset on which the rank of $\partial f/\partial u$ is constant while (43) clearly implies that, for all $(x, u) \in \Omega$, the rank of $\partial f/\partial u(x, u)$ is no larger than $\operatorname{Rank} B$. □

Based on Proposition 5.3, let us divide the points of Ω into three classes.

- A point $(\bar{x}, \bar{u}) \in \Omega$ is called *regular* if it has a neighborhood on which $\partial f/\partial u$ has constant rank. It is easy to see that regular points form an open dense subset of Ω.
- If $(\bar{x}, \bar{u})$ is not regular, it is termed *weakly singular* if each neighborhood $O \subset \Omega$ of this point satisfies
$$\sup_{(x,u)\in O} \operatorname{Rank} \frac{\partial f}{\partial u}(x, u) = m. \tag{44}$$
- A point $(\bar{x}, \bar{u}) \in \Omega$ which is neither regular nor weakly singular is said to be *strongly singular*. This means it has a neighborhood $O \subset \Omega$, such that
$$\sup_{(x,u)\in O} \operatorname{Rank} \frac{\partial f}{\partial u}(x, u) = m' < m\,, \quad \operatorname{Rank} \frac{\partial f}{\partial u}(\bar{x}, \bar{u}) < m'\,. \tag{45}$$

The distinction between topological and smooth linearizability may now be approached *via* the following theorem that complements Theorem 5.2.

Theorem 5.4. *Let $k \in \{\infty, \omega\}$ and f be of class $\mathbf{C}^k$ on an open set $\Omega \subset \mathbb{R}^{n+m}$.*

- *System* (4) *is locally $\mathbf{C}^k$ linearizable at $(\bar{x}, \bar{u}) \in \Omega$ if, and only if it is locally topologically linearizable at $(\bar{x}, \bar{u})$* and *the latter is a regular point.*
- *If* (4) *is locally topologically linearizable at $(\bar{x}, \bar{u})$ and the latter is a weakly singular point, then a linearizing homeomorphism around $(\bar{x}, \bar{u})$ may be chosen to be a map of class $\mathbf{C}^k$, although not necessarily a $\mathbf{C}^k$ diffeomorphism (its inverse may fail to be $\mathbf{C}^k$).*

Proof. The first assertion is a consequence of Theorem 5.2 together with Theorems 5.7 and 5.8 to come, observing that condition (2′) in the latter will automatically hold at a regular point by the constant rank theorem. Next, assume that $\chi : \Omega \to \Omega'$ is a quasi-$\mathbf{C}^k$ diffeomorphism that conjugates the $\mathbf{C}^k$ system (4) to the linear controllable system $\dot{z} = Az + Bv$ at some weakly singular point $(\bar{x}, \bar{u})$. By (3) of Proposition 5.3, the rank of B is m hence it is left invertible; by (1) of the same proposition, χ_{II} is indeed $\mathbf{C}^k$. □

Whether Theorem 5.4 remains true if “weakly singular” gets replaced by “strongly singular” is unknown to the authors. This turns out to be equivalent to the following question in differential topology which is of interest in its own right and seems to have no answer so far.

Open Question 5.5. *Let O be a neighborhood of the origin in $\mathbb{R}^{p+q}$ and $F : O \to \mathbb{R}^p$ a smooth (resp. real-analytic) map. Suppose $G : O \to \mathbb{R}^q$ is a* continuous *map such that $F \times G : O \to \mathbb{R}^p \times \mathbb{R}^q$ is a local homeomorphism at* 0.

Does there exist another neighborhood $O' \subset O$ of the origin and a smooth *(resp.* real-analytic*) map $H : O' \to \mathbb{R}^q$ such that $F \times H : O' \to \mathbb{R}^p \times \mathbb{R}^q$ is still a local homeomorphism at* 0 *?*

If the answer to the open question was yes, then Definitions 5.1 and 3.9 of quasi-smooth (resp. quasi-analytic) linearizability might equivalently require χ to be smooth (resp. analytic) because, assuming the linear system is in normal form

(40)-(41), one could set $F = \pi_{n+s_1} \circ \chi$ and smoothly (resp. analytically) redefine the last $m - s_1$ components of χ.

If the answer to the open question was no, then Definition 5.1 would really be more general than the one obtained by restricting χ to be smooth (resp. analytic). Indeed, if F provides a counterexample to the open question, say, in the $\mathbf{C}^\infty$ case, we may consider on $I\!R^p \times O$ the control system

$$\dot{x} \;=\; F(u) \;\;, \qquad x \in I\!R^p,\; u \in I\!R^{p+q} \tag{46}$$

which is locally quasi-smoothly linearizable at the origin because the local homeomorphism

$$(x,u) \;\mapsto\; (z,v) \;=\; (x, F(u), G(u))$$

conjugates (46) to

$$\dot{z} \;=\; B\,v, \text{ with } B = (\,I_p\,|0). \tag{47}$$

However, no *smooth* homeomorphism

$$\chi \,:\, (x,u) \;\mapsto\; (z,v) \;=\; (\chi_{\mathrm{I}}(x), \chi_{\mathrm{II}}(x,u))$$

exists that quasi-smoothly linearizes (46) at 0: if this was the case, by Corollary 4.4 we may assume up to a linear change of variables that χ conjugates (46) to (47). Then conjugacy would imply

$$\frac{\partial \chi_{\mathrm{I}}}{\partial x}(x) F(u) \;=\; B\,\chi_{\mathrm{II}}(x,u)$$

whence in particular

$$F(u) \;=\; \left(\frac{\partial \chi_{\mathrm{I}}}{\partial x}(0)\right)^{-1} B\,\chi_{\mathrm{II}}(0,u),$$

and the last q components of $\chi_{\mathrm{II}}(0,u)$ would yield a smooth H such that $F \times H$ is a local homeomorphism at 0 in R^{p+q}, contrary to the assumption.

5.2. **Geometric characterization of quasi smooth linearization.** Let $\mathcal{X}$ and $\mathcal{U}$ be two open subsets of $I\!R^n$ and $I\!R^m$ respectively, and assume that f is defined on $\mathcal{X} \times \mathcal{U}$. For each $u \in I\!R^m$, let f_u be the vector field on $\mathcal{X}$ defined by :

$$f_u(x) \;=\; f(x,u). \tag{48}$$

Also, for each $(x,u) \in \mathcal{X} \times \mathcal{U}$, we define below a subspace $D(x,u)$, that coincides with the range of the linear mapping $\partial f/\partial u\,(x,u)$ when its dimension is locally constant. First, we consider the subset $\mathcal{L}_{x,u} \subset I\!R^n$ (not a vector subspace) given by :

$$y \in \mathcal{L}_{x,u} \;\;\Leftrightarrow\;\; \exists (w_n) \in \mathcal{U}^{I\!N}, \;\; \lim_{n\to\infty} w_n = u \tag{49}$$
$$\text{and } \lim_{n\to\infty} \frac{f(x,w_n) - f(x,u)}{\|f(x,w_n) - f(x,u)\|} = y;$$

subsequently we put

$$D(x,u) \;=\; \mathrm{Span}_{I\!R}\, \mathcal{L}_{x,u}\;. \tag{50}$$

In words, $D(x,u)$ is the vector space spanned by all limit directions of straight lines through $f(x,u)$ and $f(x,u')$ as u' approaches u in $I\!R^m$; it is of common use in stratified geometry to generalize the notion of tangent space. Note that the set $\mathcal{L}_{x,u}$ depends on the norm used in (49), but the subspace $D(x,u)$ does not.

Proposition 5.6. *If f is of class* $\mathbf{C}^\infty$ *and if we denote by* $\mathrm{Ran}\, L$ *the range of a linear map L, we have that*

$$D(x,u) \;\supset\; \mathrm{Ran}\, \frac{\partial f}{\partial u}(x,u) \tag{51}$$

and equality holds at every (x,u) *where the rank of* $\partial f/\partial u\,(x,u)$ *is locally constant with respect to* u.

Proof. The inclusion (51) holds because any nonzero element of $\mathrm{Ran}\,\partial f/\partial u(x,u)$ can be written $\partial f/\partial u(x,u).h$ for some h in $I\!R^m$, and one has

$$\frac{\partial f/\partial u\,(x,u).h}{\|\partial f/\partial u\,(x,u).h\|}=\lim_{t\to 0^+}\frac{f(x,u+th)-f(x,u)}{\|f(x,u+th)-f(x,u)\|}\ .$$

Now fix (x,u) and assume that the rank of $\partial f/\partial u$ is locally constant around (x,u), equal to $r\le m$ and use the constant rank-theorem. Up to a permutation of coordinates,

$$(h_1,\dots,h_m)\overset{\lambda}{\mapsto}(f_1(x,u+h)-f_1(x,u),\dots,f_r(x,u+h)-f_r(x,u),h_{r+1},\dots,h_m)$$

is a local diffeomorphism around zero in $I\!R^m$ and, setting $\rho=\lambda^{-1}\circ z\circ\lambda$ with z given by $z(w_1,\dots,w_m)=(w_1,\dots,w_r,0,\dots,0)$, there is a constant c such that

$$\|\rho(h)\|\le c\,\|f(x,u+h)-f(x,u)\|\ \text{ and }\ f(x,u+\rho(h))=f(x,u+h) \tag{52}$$

for all h. Take $y\in\mathcal{L}_{x,u}$; by definition, there is a sequence (h_n) converging to zero and satisfying (49) with $w_n=u+h_n$; from (52), we may re-write it as

$$y=\lim_{n\to\infty}\frac{f(x,u+\rho(h_n))-f(x,u)}{\|\rho(h_n)\|}\,\frac{\|\rho(h_n)\|}{\|f(x,u+h_n)-f(x,u)\|} \tag{53}$$

where both ratios are bounded; extracting a sequence such that both converge, the limit of the first ratio is, by definition of the derivative, $\frac{\partial f}{\partial u}(x,u).h$ with h a limit point of $\rho(h_n)/\|\rho(h_n)\|$; hence $y\in\mathrm{Ran}\,\partial f/\partial u\,(x,u)$. We have proved that $\mathcal{L}_{x,u}\subset\mathrm{Ran}\,\partial f/\partial u\,(x,u)$. From (50), this implies the reverse inclusion of (51) because the left-hand side is a linear subspace. □

We can now characterize smooth (resp. analytic) and quasi-smooth (resp. quasi-analytic) linearizability in parallel. The proofs are given in section 6.1.

Theorem 5.7 (smooth or analytic linearizability)**.** *Let* $k\in\{\infty,\omega\}$ *and* f *be of class* $\mathbf{C}^k$ *on an open set* $\Omega\subset I\!R^{n+m}$. *The control system* (4) *is locally* $\mathbf{C}^k$ *linearizable at* $(\bar x,\bar u)\in\Omega$ *if, and only if there are open neighborhoods* $\mathcal{X}$ *and* $\mathcal{U}$ *of* $\bar x$ *and* $\bar u$ *in* $I\!R^n$ *and* $I\!R^m$, *with* $\mathcal{X}\times\mathcal{U}\subset\Omega$, *such that the following conditions are satisfied.*

(1) $D(x,u)$ *does not depend on* u *for* $(x,u)\in\mathcal{X}\times\mathcal{U}$.

(2) *The rank of* $\dfrac{\partial f}{\partial u}(x,u)$ *is constant in* $\mathcal{X}\times\mathcal{U}$.

(3) *Defining on* $\mathcal{X}$ *the distribution* Δ_0 *by* $\Delta_0(x)=D(x,u)$ *— this is possible if point (1) holds true — and inductively the flag of distributions* (Δ_k) *by :*

$$\Delta_{k+1}\ =\ \Delta_k\ +\ [\,f_{\bar u}\,,\,\Delta_k\,] \tag{54}$$

where $[\ ,\ ]$ *denotes the Lie bracket, then each* Δ_k *for* $0\le k\le n-1$ *is integrable (i.e. has constant dimension over* $I\!R$ *and is closed under Lie bracket) and the rank of* Δ_{n-1} *is* n.

Theorem 5.8 (quasi-smooth or quasi-analytic linearizability)**.** *Let* $k\in\{\infty,\omega\}$ *and* f *be of class* $\mathbf{C}^k$ *on an open set* $\Omega\subset I\!R^{n+m}$. *The control system* (4) *is locally quasi-*$\mathbf{C}^k$ *linearizable at* $(\bar x,\bar u)\in\Omega$ *if, and only if there are open neighborhoods* $\mathcal{X}$ *and* $\mathcal{U}$ *of* $\bar x$ *and* $\bar u$ *in* $I\!R^n$ *and* $I\!R^m$, *with* $\mathcal{X}\times\mathcal{U}\subset\Omega$, *such that conditions (1) and (3) of Theorem 5.7 are met and, instead of condition (2), it holds that*

(2′) *Denoting by* $r_1\le m$ *the constant rank of* Δ_0, *the mapping*

$$\begin{array}{rccc} F\ : & \mathcal{X}\times\mathcal{U} & \to & \mathcal{X}\times I\!R^n\\ & (x,u) & \mapsto & (\,x\,,\,f(x,u)\,)\end{array} \tag{55}$$

restricts to a $\mathbf{C}^0$ *fibration*[4] $\mathcal{W} \to F(\mathcal{W})$ *with fiber* $I\!R^{m-r_1}$ *on some neighborhood* $\mathcal{W}$ *of* $(\bar{x}, \bar{u})$ *in* $\mathcal{X} \times \mathcal{U}$*.*

Theorem 5.7 is of course equivalent to the results in [12, 9, 27], but the conditions are stated here in a slightly different form to parallel Theorem 5.8.

Corollary 5.9. *Assume that* f *is real analytic on some open set* $\Omega \subset I\!R^{n+m}$*. If the control system* (4) *is locally* $\mathbf{C}^\infty$*(resp. quasi-*$\mathbf{C}^\infty$*) linearizable at* $(\bar{x}, \bar{u}) \in \Omega$*, then it is also* $\mathbf{C}^\omega$*(resp. quasi-*$\mathbf{C}^\omega$*) linearizable there.*

Proof. analyticity does not appear in the conditions of the theorems, except for the regularity of f itself. □

5.3. **Linearization versus equivalence to the linear approximation.** For a control system, *smooth* linearizability at an equilibrium implies conjugacy *to its linear approximation*:

Proposition 5.10. *Let* $(\bar{x}, \bar{u})$ *be an equilibrium point of* (4)*, i.e.* $f(\bar{x}, \bar{u}) = 0$*, and let* $\overline{A} = \partial f / \partial x\,(\bar{x}, \bar{u})$*,* $\overline{B} = \partial f / \partial u\,(\bar{x}, \bar{u})$ *so that :*

$$f(x,u) \;=\; \overline{A}\,(x-\bar{x}) \;+\; \overline{B}\,(u-\bar{u}) \;+\; \varepsilon(x-\bar{x}, u-\bar{u})\;, \tag{56}$$

where ε *is little* $\mathrm{o}(\|\mathrm{x}-\bar{\mathrm{x}}\| + \|\mathrm{u}-\bar{\mathrm{u}}\|)$*.*

If system (4) *is locally smoothly linearizable at* $(\bar{x}, \bar{u})$*, then:*

1. *its linear approximation* $(\overline{A}, \overline{B})$ *is controllable (cf. Definition 4.2),*
2. *the system is smoothly conjugate to* $(\overline{A}, \overline{B})$ *at* $(\bar{x}, \bar{u})$*.*

Proof. Let χ be a local diffeomorphism conjugating system (4) to $\dot{z} = Az + Bv$ at $(\bar{x}, \bar{u})$, and observe from (14) in Remark 3.8 that smooth linearizability translates into (42). If we write f as in (56), and if we set $\overline{P} = \frac{\partial \chi_{\mathrm{I}}}{\partial x}(\bar{x})$, $\overline{K} = \frac{\partial \chi_{\mathrm{II}}}{\partial x}(\bar{x}, \bar{u})$, $\overline{Q} = \frac{\partial \chi_{\mathrm{II}}}{\partial u}(\bar{x}, \bar{u})$, we get by differentiating (42) with respect to x and u at $(\bar{x}, \bar{u})$, using the relation $f(\bar{x}, \bar{u}) = 0$, that

$$\overline{P}\,\overline{A} \;=\; A\overline{P} + B\overline{K}\;\;,\qquad \overline{P}\,\overline{B} \;=\; B\overline{Q}\;.$$

Since $\overline{P}$ and $\overline{Q}$ are square invertible matrices by the triangular structure of χ displayed in (13), this implies that the linear systems (A, B) and $(\overline{A}, \overline{B})$ are linearly conjugate, see (35). Since (A, B) is controllable by definition so is $(\overline{A}, \overline{B})$, thereby achieving the proof. □

Proposition 5.10 has no analog if the control system is only topologically linearizable (hence *quasi-smoothly* linearizable according to Theorem 5.2). For example, the system (1) in the introduction is quasi-$\mathbf{C}^\omega$linearizable at $(0,0)$, but its linear approximation $\dot{x} = 0$ is not controllable and it is not topologically equivalent to $\dot{x} = 0$. Apart from such degenerate cases, there also exist systems that are quasi-analytically linearizable at some point with controllable linear approximation there, and still they are not conjugate to this linear approximation. An example when $m = n = 2$ is given by :

$$\dot{x}_1 \;=\; u_1 \;\;,\qquad \dot{x}_2 \;=\; x_1 \;+\; u_2^3\;,$$

This system is quasi-analytically conjugate at $(0,0)$ to

$$\dot{z}_1 \;=\; v_1 \;\;,\qquad \dot{z}_2 \;=\; v_2\;, \tag{57}$$

via $z = x$, $v_1 = u_1$, $v_2 = u_2^3 + x_1$. However, its linear approximation at the origin is $\dot{x}_1 = u_1$, $\dot{x}_2 = x_1$, which is controllable yet not conjugate to (57) (*cf* Theorem 4.3).

[4] A $\mathbf{C}^0$ fibration with fiber $\mathcal{F}$ over $\mathcal{B}$ is a continuous map $g : \mathcal{E} \to \mathcal{B}$ for which every $\xi \in \mathcal{B}$ has a neighborhood $\mathcal{O}$ in $\mathcal{B}$ such that $g^{-1}(\mathcal{O}) \subset \mathcal{E}$ is homeomorphic to $\mathcal{O} \times \mathcal{F}$, the so-called *trivializing* homeomorphism $\psi : g^{-1}(\mathcal{O}) \to \mathcal{O} \times \mathcal{F}$ being such that $\pi \circ \psi = g$ where $\pi : \mathcal{O} \times \mathcal{F} \to \mathcal{O}$ is the natural projection onto the first factor.

5.4. **Non-genericity of linearizability.** Except when $m \geq n$ or $(n,m)=(2,1)$, the conditions of Theorem 5.7 require a certain number of *equalities* (involving f and its partial derivatives) to hold *everywhere*. For example, the integrability of a distribution entails that all Lie brackets be linearly dependent on the original vector fields, *i.e.* certain determinants must be identically zero. This makes smooth (resp. analytic) linearizability of a smooth (resp. analytic) control system highly non-generic in any reasonable sense, because when written in proper jet spaces it is contained in a set of infinite co-dimension. Moreover, small perturbations of a system that does not satisfy these condition will not satisfy them either, while most perturbations of a system which satisfies them will fail to do so. Compare for instance [26] where it is shown that the equivalence class of any system affine in the control has infinite co-dimension in some Whitney topology.

From Theorem 5.4, quasi-smooth or quasi-analytic linearizability, hence also topological linearizability by Theorem 5.2, require the same equalities to hold on an open dense set, although this time some singularities are allowed. This is no more "generic" than smooth linearizability, as opposed to ODE's for which the Grobman-Hartman theorem allows one to linearize around an equilibrium as soon as it is hyperbolic.

6. Proofs

6.1. **Proof of Theorems 5.7 and 5.8.** We begin with a lemma whose cumbersome index arrangement will be rewarded later when constructing the Kronecker indices of the linearized system.

Lemma 6.1. *Let $k \in \{\infty, \omega\}$. Let Δ_0 and $f_{\bar{u}}$ be respectively a distribution and a vector field, both of class $\mathbf{C}^k$ on a connected open neighborhood of $x \in I\!R^n$. Let further Δ_i, $i \geq 0$, be the distributions defined according to* (54)*, and set for convenience $\Delta_{-1}=\{0\}$. Assume they satisfy point (3) of Theorem 5.7 or 5.8. Put*

$$r_i \;=\; \mathrm{Rank}\Delta_{i-1}\,,\;\, i>0, \qquad r_0=0\,,\; r_{-1}=-m\,, \tag{58}$$

so that $r_i=n$ for some $i \leq n-1$; let $\rho \in \{3,\dots,n+1\}$ be the smallest integer such that $r_{\rho-1}=n$. Define also

$$s_i=r_i-r_{i-1}\,,\quad \sigma_i=\sum_{j=i}^{\rho} s_j = n-r_{i-1}\,,\quad 0\leq i\leq \rho \tag{59}$$

(note that $s_\rho=\sigma_\rho=0$).

Then, there exists coordinates $\chi_1,\dots,\chi_n$ of class $\mathbf{C}^k$ on a neighborhood $\mathcal{X}$ of x such that

- *$\chi_1,\dots,\chi_{\sigma_i}$ are independent first integrals of Δ_{i-2} for $i \in \{1,\dots,\rho-1\}$,*
- *$\chi_{\sigma_i+j}=f_{\bar{u}}\chi_{\sigma_{i+1}+j}$ for all integers i,j, $2\leq i\leq \rho-1$, $1\leq j\leq s_i$*

($f_{\bar{u}}\chi_{\sigma_{i+1}+j}$ is the Lie derivative of the function $\chi_{\sigma_{i+1}+j}$ along the vector field $f_{\bar{u}}$).

Proof. Note that when $i=1$, the first point above means that $\chi_1,\dots,\chi_n$ are indeed local coordinates. Now, the Frobenius theorem provides us with $n-r$ independent $\mathbf{C}^k$ first integrals for a $\mathbf{C}^k$ integrable distribution of rank r. This accounts for the regularity of the coordinates if we construct them as follows.

First pick $n-r_{\rho-2}=\sigma_{\rho-1}$ independent first integrals of $\Delta_{\rho-3}$ and call them $\chi_1,\dots,\chi_{\sigma_{\rho-1}}$; define further $\chi_{1+\sigma_{\rho-1}},\dots,\chi_{2\sigma_{\rho-1}}$ by $\chi_{\sigma_{\rho-1}+j}=f_{\bar{u}}\chi_j$ for $1\leq j\leq \sigma_{\rho-1}=s_{\rho-1}$. Clearly, $\chi_1,\dots,\chi_{\sigma_{\rho-1}+s_{\rho-1}}$ satisfy the conditions for $i=\rho-1$. Then proceed inductively : assume that, for some $i_0 \in \{2,\dots,\rho-1\}$, the functions $\chi_1,\dots,\chi_{\sigma_{i_0}+s_{i_0}}$ have been constructed and satisfy the conditions for $i \geq i_0$. *We* *claim* that the differentials $\mathrm{d}\chi_\ell$ are linearly independent at each point of $\mathcal{X}$. Indeed, assume that there is $\bar{x} \in \mathcal{X}$ and real coefficients μ_j and λ_k such that

$$\sum_{j=1}^{\sigma_{i_0}} \mu_j \mathrm{d}\chi_j(\bar{x}) + \sum_{k=1+\sigma_{i_0+1}}^{\sigma_{i_0}} \lambda_k \mathrm{d}\left(f_{\bar{u}}\chi_k\right)(\bar{x}) = 0. \tag{60}$$

Put $\omega_1 = \Sigma\mu_j \mathrm{d}\chi_j$ and $\omega_2 = \Sigma\lambda_k \mathrm{d}\chi_k$. Since d commutes with the Lie derivative, we may rewrite (60) as $\omega_1(\bar{x}) + f_{\bar{u}}\omega_2(\bar{x}) = 0$. In particular, for any $\mathbf{C}^k$-vector field X in Δ_{i_0-2}, we get as $\omega_1(X) \equiv 0$ that $f_{\bar{u}}\omega_2(X)(\bar{x}) = 0$. Now, by virtue of the formula

$$f_{\bar{u}}\left(\omega_2(X)\right) = f_{\bar{u}}\omega_2(X) + \omega_2([f_{\bar{u}}, X]), \tag{61}$$

we obtain since $\omega_2(X) \equiv 0$ that $\omega_2([f_{\bar{u}}, X])(\bar{x}) = 0$, that is, ω_2 annihilates Δ_{i_0-1} at $\bar{x}$. But $\mathrm{d}\chi_1(\bar{x}), \ldots, \mathrm{d}\chi_{\sigma_{i_0+1}}(\bar{x})$ are a basis of the orthogonal space to $\Delta_{i_0-1}(\bar{x})$ by the induction hypothesis, whereas $\omega_2(\bar{x})$ is a linear combination of the $\mathrm{d}\chi_k(\bar{x})$ for $\sigma_{i_0+1} < k \le \sigma_{i_0}$. Therefore, since we know by the induction hypothesis that the $\mathrm{d}\chi_\ell$ are point-wise independent for $1 \le \ell \le \sigma_{i_0}$, we get that the λ_k are zero and then the μ_j are also zero by (60). *This proves the claim.* Next, recall that $\chi_1, \ldots, \chi_{\sigma_{i_0}}$ are first integrals of Δ_{i_0-2}, thus *a fortiori* of Δ_{i_0-3}. For X a $\mathbf{C}^k$-vector field in the latter we deduce from (61), where ω_2 is replaced by $\mathrm{d}\chi_\ell$ with $1 + \sigma_{i_0+1} \le \ell \le \sigma_{i_0}$, that $\chi_{1+\sigma_{i_0}}, \ldots, \chi_{\sigma_{i_0}+s_{i_0}}$ are also first integrals of Δ_{i_0-3}. In case $\sigma_{i_0} + s_{i_0} < \sigma_{i_0-1}$, pick $\chi_{\sigma_{i_0}+s_{i_0}+1}, \ldots, \chi_{\sigma_{i_0-1}}$ so that χ_ℓ for $1 \le \ell \le \sigma_{i_0-1}$ is a complete set of independent integrals of Δ_{i_0-3}. If $i_0 = 2$ we are done, otherwise define $\chi_{\sigma_{i_0-1}+j} = f_{\bar{u}}\chi_{\sigma_{i_0}+j}$ for $1 \le j \le s_{i_0-1}$ in order to complete the induction step. $\square$

Proof of Theorems 5.7 and 5.8. The two proofs run parallel to each other.

We first show necessity, assuming that $k = \infty$ for analyticity does not appear in the conclusions. Assume local (quasi) smooth linearizability, *cf.* Definitions 5.1 and 3.9. Without loss of generality, we assume that $\Omega = \mathcal{X} \times \mathcal{U}$ where $\mathcal{X}$ and $\mathcal{U}$ are open neighborhoods of $\bar{x}$ and $\bar{u}$ in $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Let $\chi : \mathcal{X} \times \mathcal{U} \to \Omega' \subset \mathbb{R}^{n+m}$ be as in (13); recall that χ_{I} is a smooth diffeomorphism $\mathcal{X} \to \chi_{\mathrm{I}}(\mathcal{X})$. We may also assume, after composing χ with a linear invertible map, that the pair (A, B) is in canonical form (40)-(41), but we still write A, B rather than A_c, B_c. Denote by $B_0, \ldots, B_m$ the columns of B and define the vector fields $b_0, \ldots, b_m$ on $\mathbb{R}^n$ by

$$b_i(z) = B_i\,,\ 1 \le i \le m\,, \qquad b_0(z) = Az + B\bar{u} \tag{62}$$

and the distributions Λ_i by

$$\Lambda_0(z) = \mathrm{Span}_{\mathbb{R}}\{b_1(z), \ldots, b_m(z)\} = \mathrm{Ran}B \qquad \Lambda_{i+1} = \Lambda_i + [b_0, \Lambda_i]\,,\ 1 \le i \le m\,. \tag{63}$$

From (42), we have

$$\frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\, f(x,u) \;=\; b_0(\chi_{\mathrm{I}}(x)) \;+\; B\left(\chi_{\mathrm{II}}(x,u) - \chi_{\mathrm{II}}(x,\bar{u})\right). \tag{64}$$

Since χ is a triangular homeomorphism, $\chi_{\mathrm{II}}(x,w) - \chi_{\mathrm{II}}(x,u)$ covers an open neighborhood of 0 in $\mathbb{R}^m$ when w ranges around u in $\mathbb{R}^m$. Thus, in view of (64), $\mathcal{L}_{x,u}$ defined by (49) contains an open set in $\left(\frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\right)^{-1}\mathrm{Ran}B$, and by double inclusion

$$D(x,u) \;=\; \left(\frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\right)^{-1}\mathrm{Ran}B.$$

This proves point 1, and also proves that the distribution Δ_0 in point 3 is the pull-back of Λ_0 by the diffeomorphism χ_{I}, *i.e.* $(\chi_{\mathrm{I}})_*\Delta_0 = \Lambda_0$. Since (64) also implies $(\chi_{\mathrm{I}})_* f_{\bar{u}} = b_0$, we have $(\chi_{\mathrm{I}})_*\Delta_i = \Lambda_i$ for all i. This gives point 3 because it is obviously true with Λ_i instead of Δ_i, and integrability and ranks are preserved by conjugation with the smooth diffeomorphism χ_{I}. In the case of smooth linearizability, point 2 is easily obtained by differentiating (64) with respect to u and using invertibility of $\partial\chi_{\mathrm{II}}/\partial u(x,u)$.

To conclude the proof of necessity, let us prove point $2'$ in the case of quasi-smooth linearizability. Let

$$\mathcal{M} = \left\{ \ (x,y) \in \mathcal{X} \times I\!R^n ; \ \frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\, y \ - \ A\chi_{\mathrm{I}}(x) \ \in \ \operatorname{Ran} B \ \right\} .$$

This is a smooth embedded sub-manifold of $\mathcal{X} \times I\!R^n$ of dimension $n + r_1$, where $r_1 = \operatorname{Rank} B \leq m$. If we define F as in (55), it is clear from (42) that

$$F(\mathcal{X} \times \mathcal{U}) \ \subset \ \mathcal{M} \ .$$

Now, take some $(m - r_1) \times m$ matrix C whose rows complement r_1 independent rows of B into a basis of $I\!R^m$. Pick matrices E_1 and E_2 of appropriate sizes such that

$$E_1 B \ + \ E_2 C \ = \ I_m \ .$$

By (42) we get

$$E_1 \left[\ \frac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\, f(x,u) \ - \ A\,\chi_{\mathrm{I}}(x) \right] \ + \ E_2 C \chi_{\mathrm{II}}(x,u) \ = \ \chi_{\mathrm{II}}(x,u) \ . \tag{65}$$

Define

$$\psi : \mathcal{X} \times \mathcal{U} \to \mathcal{M} \times I\!R^{m-r_1}$$

by the formula:

$$\psi(x,u) = (x, f(x,u), C\chi_{\mathrm{II}}(x,u)).$$

From (65), this mapping has an inverse given by

$$\begin{array}{rccl} \psi^{-1} : & \psi(\mathcal{X} \times \mathcal{U}) & \to & \mathcal{X} \times \mathcal{U} \\ & (x,y,z) & \mapsto & \chi^{-1} \left(\ \chi_{\mathrm{I}}(x) \ , \ E_1 \left[\dfrac{\partial\chi_{\mathrm{I}}}{\partial x}(x)\, y \ - \ A\,\chi_{\mathrm{I}}(x) \right] \ + \ E_2\, z \ \right) \end{array}$$

so that ψ defines a homeomorphism from $\mathcal{X} \times \mathcal{U}$ onto its image which is open in $\mathcal{M} \times I\!R^{m-r_1}$ by invariance of the domain. Let $\mathcal{O}$ be a neighborhood of $(\bar{x}, f(\bar{x},\bar{u}))$ in $\mathcal{M}$ and $\mathcal{S}$ an open ball centered at $C\chi_{\mathrm{II}}(\bar{x},\bar{u})$ in $I\!R^{m-r_1}$ such that $\mathcal{O} \times \mathcal{S} \subset \psi(\mathcal{X} \times \mathcal{U})$, and take $\mathcal{W} = \psi^{-1}(\mathcal{O} \times \mathcal{S})$. Then $F : \mathcal{W} \to F(\mathcal{W}) = \mathcal{O}$ is a $\mathbf{C}^0$ fibration with fiber $\mathcal{S}$ and trivializing homeomorphism $\psi : \mathcal{W} \to \mathcal{O} \times \mathcal{S}$. Since $\mathcal{S}$ is homeomorphic to $I\!R^{m-r_1}$, condition $2'$ follows.

We turn to sufficiency. Points 1, 3, and either 2 or $2'$ imply, for all $x \in \mathcal{X}$,

$$\Delta_0(x) \ = \ \operatorname{Span}_{I\!R}\{ f(x,w) - f(x,u), (u,w) \in \mathcal{U} \times \mathcal{U} \} \ . \tag{66}$$

Indeed the right-hand side always contains $D(x,u)$ because it contains all the differences $f(x,w_n) - f(x,u)$ in (49), and point 1 implies the reverse inclusion because $f(x,w) - f(x,u)$ can be computed as the integral on the segment $[u,w] \subset \mathcal{U}$ of a function that, thanks to Proposition 5.6, belongs constantly to $\Delta_0(x)$.

From (66), the distribution Δ_0 is of class $\mathbf{C}^k$. Considering point 3, we may apply Lemma 6.1. We thus obtain some, with r_i, s_i and σ_i the integers defined by (58) and (59), some $\mathbf{C}^k$ coordinates $\chi_1, \ldots, \chi_n$ on a neighborhood of $\bar{x}$ possibly smaller than $\mathcal{X}$ (but that we continue to denote by $\mathcal{X}$), i.e. a diffeomorphism $\chi_{\mathrm{I}} : \mathcal{X} \to \chi_{\mathrm{I}}(\mathcal{X})$, with $\chi_{\mathrm{I}} = (\chi_1, \ldots, \chi_n)$, meeting the conclusions of Lemma 6.1. In particular, $\chi_1, \ldots, \chi_{n-r_1}$ are first integrals of the distribution Δ_0, and from (66), this implies that $\partial\chi_i/\partial x(x)\, f(x,u)$ does not depend on u, and is there fore equal to its value for $u = \bar{u}$:

$$\frac{\partial\chi_i}{\partial x}(x) f(x,u) = f_{\bar{u}}\chi_i\,(x) \,, \quad 1 \leq i \leq n - r_1 \ . \tag{67}$$

For larger i, the left-hand side depends on x and u : define $\lambda : \mathcal{X} \times \mathcal{U} \to I\!R^{m_1}$ by

$$\lambda(x,u) \;=\; (\;\frac{\partial \chi_{n-r_1+1}}{\partial x}(x) f(x,u)\;,\;\ldots\;,\;\frac{\partial \chi_n}{\partial x}(x) f(x,u)\;)\;. \tag{68}$$

Then, defining coordinates $z_1, \ldots, z_n$ by $z = \chi_{\mathrm{I}}(x)$. The equations of system (4) are as follows (the first line gives the derivatives of the $n - r_1$ first coordinates and the second line the last r_1 ones) :

$$\begin{array}{rcll} \dot z_{\sigma_{i+1}+j} & = & z_{\sigma_i + j}\;, & 2 \le i \le \rho - 1\;,\; 1 \le j \le s_i\;, \\ \dot z_\ell & = & \lambda_{n-\ell}(\,\chi_{\mathrm{I}}^{-1}(z)\;,\;u\,)\;, & n - m_1 + 1 \le \ell \le n\;. \end{array} \tag{69}$$

If point 2 is satisfied, the rank of the map $(x,u) \mapsto (\chi_{\mathrm{I}}(x), \frac{\partial \chi_{\mathrm{I}}}{\partial (x)} f(x,u))$ is constant and thus, according to (66), it is equal to $n + r_1$, r_1 being the rank of Δ_0. From (67), the map $(x,u) \mapsto (\chi_{\mathrm{I}}(x), \lambda(x,u))$ has the same constant rank $n + r_1$. Hence there exists $\phi\; : \mathcal{X} \times \mathcal{U} \to I\!R^{m-r_1}$ such that

$$(x,u) \mapsto (\;\chi_{\mathrm{I}}(x)\;,\;\lambda(x,u)\;,\;\phi(x,u)\;) \tag{70}$$

is a diffeomorphism of class $\mathbf{C}^k$. Obviously, defining χ_{II} by $\chi_{\mathrm{II}}(x,u) = (\lambda(x,u), \phi(x,u))$ yields a $\mathbf{C}^k$ diffeomorphism χ that conjugates (4) to a linear controllable system $\dot z = Az + Bu$. This proves sufficiency in Theorem 5.7.

If point $2'$ is satisfied instead, let $\psi : \mathcal{W} \to F(\mathcal{W}) \times I\!R^{n-r_1}$ be the “trivializing” homeomorphism. Recall that, with $\pi : F(\mathcal{W}) \times I\!R^{n-r_1} \to F(\mathcal{W})$ the natural projection, one has $\pi \circ \psi = F$; call $\phi : \mathcal{W} \to I\!R^{n-r_1}$ the map such that $\psi = F \times \phi$ Composing F with $(x,\xi) \mapsto (\chi_{\mathrm{I}}(x), \frac{\partial \chi_{\mathrm{I}}}{\partial (x)} \xi)$, one gets that $(x,u) \mapsto \chi(x,u) = (\chi_{\mathrm{I}}(x), \lambda(x,u), \phi(x,u))$ is a homeomorphism. It clearly conjugates (4) to a linear controllable system $\dot z = Az + Bu$. This proves sufficiency in Theorem 5.8. □

6.2. **Proof of Theorem 5.2.** This theorem for $k = \omega$ is consequence of this theorem for $k = \infty$ and of Corollary 5.9. Hence we only have to prove it for $k = \infty$, *i.e.* we assume that f is infinitely differentiable and we prove that topological linearizability implies quasi-$\mathbf{C}^\infty$linearizability.

Without loss of generality, we suppose that $(\bar x, \bar u) = (0,0)$. Assume there exists a homeomorphism χ from a neighborhood of the origin in $I\!R^{n+m}$ to an open subset of $I\!R^{n+m}$ that conjugates system (4) to the linear controllable system

$$\dot z \;=\; Az \;+\; Bv \tag{71}$$

with $z \in I\!R^n$ and $v \in I\!R^m$. Composing χ with a linear invertible map allows us to suppose that the pair (A,B) is in canonical form (40)-(41), i.e. that (71) can be read

$$\dot z_{\sigma_i + k} \;=\; z_{\sigma_{i-1}+k}\;,\quad 2 \le i \le \rho,\; 1 \le k \le s_{i-1}, \tag{72}$$

where the integers s_i and σ_i were defined in (38) and (39) and where, for notational compactness, we have set :

$$z_{n+k} \;\stackrel{\Delta}{=}\; v_k\;; \tag{73}$$

recall here that $s_0 = m$, and notice that $s_1 < m$ may well occur as it simply means that $\operatorname{Rank} B < m$, in which case some of the controls do not appear in the canonical form. With the aggregate notation :

$$Z_j \;\stackrel{\Delta}{=}\; \begin{pmatrix} z_{\sigma_{j+1}+1} \\ \vdots \\ z_{\sigma_j} \end{pmatrix},\; 1 \le j \le \rho - 1\;,\qquad Z_0 \;\stackrel{\Delta}{=}\; \begin{pmatrix} v_1 \\ \vdots \\ v_m \end{pmatrix}\;, \tag{74}$$

and the matrices J_r^s defined in (41), system (72) can be rewritten as

$$\begin{array}{rcl} \dot Z_{\rho-1} &=& J^{s_{\rho-1}}_{s_{\rho-2}} Z_{\rho-2} \\ \dot Z_{\rho-2} &=& J^{s_{\rho-2}}_{s_{\rho-3}} Z_{\rho-3} \\ &\vdots& \\ \dot Z_2 &=& J^{s_2}_{s_1} Z_1 \\ \dot Z_1 &=& J^{s_1}_{s_0} Z_0 \end{array} \tag{75}$$

and is viewed as a control system with state $(Z_{\rho-1}, \ldots, Z_1)$ and control Z_0. We also make the convention, similar to (73), that

$$x_{n+k} \triangleq u_k , \tag{76}$$

and we use for the controls the aggregate notation :

$$X_0 \triangleq \left(\begin{array}{c} x_{n+1} \\ \vdots \\ x_{n+m} \end{array} \right) = \left(\begin{array}{c} u_1 \\ \vdots \\ u_m \end{array} \right) . \tag{77}$$

Let us now prove that property $\mathcal{P}_\ell$ below is true for $0 \le \ell \le \rho - 1$.

Property $\mathcal{P}_\ell$ ***:*** *there exists a smooth local change of coordinates around* 0 *in* $I\!R^n$*, say*

$$(x_1, \ldots, x_n) \mapsto (\widehat X, X_\ell, \ldots, X_2, X_1),$$

with $\widehat X \in I\!R^{\sigma_{\ell+1}}$ *and* $X_i \in I\!R^{s_i}$ *for* $0 \le i \le \ell$ *(if* $\ell = 0$ *there are no* X_i*'s beyond* X_0 *whereas if* $\ell = \rho - 1$ *there is no* $\widehat X$*), after which system* (4) *reads:*

$$\begin{array}{rcl} \dot{\widehat X} &=& \widehat F(\widehat X, X_\ell) \\ \dot X_\ell &=& F_\ell(\widehat X, X_\ell, X_{\ell-1}) \\ &\vdots& \\ \dot X_2 &=& F_2(\widehat X, X_\ell, \ldots, X_1) \\ \dot X_1 &=& F_1(\widehat X, X_\ell, \ldots, X_1, X_0) \end{array} , \tag{78}$$

and such that (78)*, viewed as a control system with state* $(\widehat X, X_\ell, \ldots, X_1)$ *and control* X_0*, is locally topologically conjugate at* $(0,0)$ *to system* (75) *via a local homeomorphism*

$$(\widehat X, X_\ell, \ldots, X_1, X_0) \mapsto (Z_{\rho-1}, \ldots, Z_0)$$

which is, together with its inverse, of the block triangular form :

$$\begin{array}{lll} (Z_{\rho-1}, \ldots, Z_{\ell+1}) = \widehat\Phi(\widehat X) & \qquad & \widehat X = \widehat\Psi(Z_{\rho-1}, \ldots, Z_{\ell+1}) \\ Z_\ell = \Phi_\ell(\widehat X, X_\ell) & & X_\ell = \Psi_\ell(Z_{\rho-1}, \ldots, Z_\ell) \\ \vdots & & \vdots \\ Z_1 = \Phi_1(\widehat X, X_\ell, \ldots, X_1) & & X_1 = \Psi_1(Z_{\rho-1}, \ldots, Z_1) \\ Z_0 = \Phi_0(\widehat X, X_\ell, \ldots, X_1, X_0) & & X_0 = \Psi_0(Z_{\rho-1}, \ldots, Z_1, Z_0) \end{array}$$

where Φ_i *and* Ψ_i *are, for* $1 \le i \le \ell$*, continuously differentiable with respect to* X_i *and* Z_i *respectively, have an invertible derivative, and satisfy for* $1 \le i \le \ell$ *the relation :*

$$\begin{array}{l} F_i(\widehat X, X_\ell, \ldots, X_i, X_{i-1}) = F_i(\widehat X, X_\ell, \ldots, X_i, 0) \\ \quad + \left(\frac{\partial \Phi_i}{\partial X_i}(\widehat X, X_\ell, \ldots, X_i) \right)^{-1} J^{s_i}_{s_{i-1}} \Big(\\ \qquad \Phi_{i-1}(\widehat X, X_\ell, \ldots, X_i, X_{i-1}) - \Phi_{i-1}(\widehat X, X_\ell, \ldots, X_i, 0) \Big) \ ; \end{array} \tag{79}$$

furthermore, the partial homeomorphism

$$(\widehat{X}, X_\ell) \quad \mapsto \quad (Z_{\rho-1}, \ldots, Z_\ell) \tag{80}$$

locally topologically conjugates, at $(0,0) \in I\!R^{\sigma_{\ell+1}+s_\ell}$*, the reduced control system*

$$\dot{\widehat{X}} \quad = \quad \widehat{F}(\widehat{X}, X_\ell), \tag{81}$$

with state $\widehat{X}$ *and input* X_ℓ*, to the reduced linear control system*

$$\begin{array}{rcl} \dot{Z}_{\rho-1} & = & J_{s_{\rho-2}}^{s_{\rho-1}} Z_{\rho-2} \ , \\ & \vdots & \\ \dot{Z}_{\ell+1} & = & J_{s_\ell}^{s_{\ell+1}} Z_\ell \end{array} \tag{82}$$

wit state $(Z_{\rho-1}, \ldots, Z_{\ell+1})$ *and input* Z_ℓ*.*

Indeed, $\mathcal{P}_0$ is merely the original assumption on local topological conjugacy of systems (4) and (75), where the triangular structure (13) of the conjugating homeomorphism was taken into account; note that, in $\mathcal{P}_0$, (79) is empty and that the reduced system (81) is the original system. Next, supposing that $\mathcal{P}_\ell$ holds for some $\ell \geq 0$, we apply Lemmas 6.2 and 6.3 (see below) to the reduced systems (81), (82), and to the partial homeomorphism (80), with

$$d = \sigma_{\ell+1}, \ r = s_\ell, \ s = s_{\ell+1}, \ U = X_\ell, \ (x_1, \ldots, x_d) = \widehat{X},$$

$$Z^1 = (Z_{\rho-1}, \ldots, Z_{\ell+2}), \ Z^2 = Z_{\ell+1}, \text{ and } V = Z_\ell,$$

and then, upon renaming $\widetilde{X}^2$ as $X_{\ell+1}$, $\widetilde{f}^2$ as $F_{\ell+1}$, and choosing $\widetilde{X}^1$ to be the new $\widehat{X}$, we get $\mathcal{P}_{\ell+1}$.

Now, $\mathcal{P}_{\rho-1}$, where we specialize (79) to $i = 1$, provides us with a *smooth* change of variables around 0 in $I\!R^n$:

$$(x_1, \ldots, x_n) \quad \mapsto \quad (X_{\rho-1}, \ldots, X_2, X_1)$$

with $X_i \in I\!R^{s_i}$ such that, in the new coordinates, system (4) reads

$$\begin{array}{rcl} \dot{X}_{\rho-1} & = & F_{\rho-1}(X_{\rho-1}, X_{\rho-2}) \\ \dot{X}_{\rho-2} & = & F_{\rho-2}(X_{\rho-1}, X_{\rho-2}, X_{\rho-3}) \\ & \vdots & \\ \dot{X}_2 & = & F_2(X_{\rho-1}, \ldots, X_1) \\ \dot{X}_1 & = & F_1(X_{\rho-1}, \ldots, X_1, X_0), \end{array} \tag{83}$$

and also such that the local homeomorphism Φ that topologically conjugates system (83) to system (75) at $(0,0)$ is, together with its inverse Ψ, of the triangular form :

$$\begin{array}{rclcrcl} Z_{\rho-1} & = & \Phi_{\rho-1}(X_{\rho-1}) & \quad & X_{\rho-1} & = & \Psi_{\rho-1}(Z_{\rho-1}) \\ Z_{\rho-2} & = & \Phi_{\rho-2}(X_{\rho-1}, X_{\rho-2}) & & X_{\rho-2} & = & \Psi_{\rho-2}(Z_{\rho-1}, Z_{\rho-2}) \\ & \vdots & & & & \vdots & \\ Z_1 & = & \Phi_1(X_{\rho-1}, \ldots, X_1) & & X_1 & = & \Psi_1(Z_{\rho-1}, \ldots, Z_1) \\ Z_0 & = & \Phi_0(X_{\rho-1}, \ldots, X_1, X_0) & & X_0 & = & \Psi_0(Z_{\rho-1}, \ldots, Z_1, Z_0), \end{array} \tag{84}$$

where the following three properties hold :

(1) Each Φ_k and Ψ_k for $k \geq 1$ is continuously differentiable with respect to X_k and Z_k respectively; in particular, $\partial\Phi_k/\partial X_k$ is invertible throughout the considered neighborhood.

(2) For $k \geq 2$, $\text{Rank}\dfrac{\partial F_k}{\partial X_{k-1}}(0, \ldots, 0) \ = \ s_k$, i.e. this rank is maximum, equal to the number of rows.

(3) F_1 satisfies

$$\begin{aligned} F_1(X_{\rho-1},\ldots,X_1,X_0) &= F_1(X_{\rho-1},\ldots,X_1,0) \\ &\quad + \left(\frac{\partial\Phi_1}{\partial X_1}(X_{\rho-1},\ldots,X_1)\right)^{-1} J_m^{s_1} \Big(\\ &\quad \Phi_0(X_{\rho-1},\ldots,X_1,X_0) - \Phi_0(X_{\rho-1},\ldots,X_1,0)\Big) \,. \end{aligned} \tag{85}$$

From the maximum rank assumption on $\partial F_{\rho-1}/\partial X_{\rho-2}$, it is possible to define $Y_{\rho-2}$ whose first $s_{\rho-1}$ entries are those of $F_{\rho-1}(X_{\rho-1},X_{\rho-2})$ and whose remaining $s_{\rho-2}-s_{\rho-1}$ entries are suitable components of $X_{\rho-2}$, in such a way that

$$(X_{\rho-1},\ldots,X_1) \;\mapsto\; (X_{\rho-1},Y_{\rho-2},X_{\rho-3}\ldots,X_1)$$

is a local *smooth* change of coordinates around 0 in $I\!R^n$. After performing this change of coordinates and setting $Y_{\rho-1}=X_{\rho-1}$ for notational homogeneity, system (83) reads

$$\begin{aligned} \dot Y_{\rho-1} &= J_{s_{\rho-2}}^{s_{\rho-1}}\,Y_{\rho-2} \\ \dot Y_{\rho-2} &= \widetilde F_{\rho-2}(Y_{\rho-1},Y_{\rho-2},X_{\rho-3}) \\ &\;\vdots \\ \dot X_2 &= \widetilde F_2(Y_{\rho-1},Y_{\rho-2},X_{\rho-3},\ldots,X_1) \\ \dot X_1 &= \widetilde F_1(Y_{\rho-1},Y_{\rho-2},X_{\rho-3},\ldots,X_1,X_0) \end{aligned}$$

where the $\widetilde F$'s enjoy the same properties than the F's, in particular the maximality of $\operatorname{Rank}\partial\widetilde F_k/\partial X_{k-1}(0,\ldots,0)$ for $\rho-2\geq k\geq 2$. One may iterate this procedure, limited only by the fact that the maximum rank property mentioned above only holds for $k\geq 2$ but not necessarily for $k=1$. Altogether, this yields a *smooth* local change of coordinates around 0 in $I\!R^n$:

$$(X_{\rho-1},\ldots,X_1) \;\mapsto\; (Y_{\rho-1},\ldots,Y_1),$$

after which system (83) is of the form

$$\begin{aligned} \dot Y_{\rho-1} &= J_{s_{\rho-2}}^{s_{\rho-1}}\,Y_{\rho-2} \\ &\;\vdots \\ \dot Y_2 &= J_{s_1}^{s_2}\,Y_1 \\ \dot Y_1 &= F_1(Y_{\rho-1},,\ldots,Y_1,X_0)\,, \end{aligned} \tag{86}$$

where we abuse the notation F_1 for simplicity because, although it needs not be the same as in (83), this new F_1 enjoys the same property (85) for some suitably redefined Φ_1 and Φ_0. Now, we may rewrite (85) as

$$F_1(Y_{\rho-1},,\ldots,Y_1,X_0) \;=\; J_m^{s_1}\,H(Y_{\rho-1},,\ldots,Y_1,X_0) \tag{87}$$

where H, in the aggregate notation $Y=(Y_{\rho-1},,\ldots,Y_1)$, is defined by

$$H(Y,X_0) \;=\; \begin{pmatrix} F_1(Y,0) \\ 0 \end{pmatrix} + \begin{pmatrix} \frac{\partial\Phi_1}{\partial Y_1}(Y)^{-1} & 0 \\ 0 & I_{m-s_1} \end{pmatrix} \big(\Phi_0(Y,X_0)-\Phi_0(Y,0)\big).$$

Since Φ has the triangular structure displayed in (84), the map $X_0\mapsto\Phi_0(Y,X_0)$ is injective for fixed $Y=(Y_{\rho-1},\ldots,Y_1)$ in the neighborhood of 0 where it is defined in $I\!R^m$. Consequently, $(Y,X_0)\mapsto(Y,H(Y,X_0))$ is also injective in the neighborhood of 0 where it is defined in $I\!R^{n+m}$; since it is continuous, it is a local homeomorphism of $I\!R^{n+m}$ at $(0,0)$ by invariance of the domain, and then (86), (87) make it clear that system (83) is locally quasi-smoothly linearizable at this point.

Since (83) is smoothly conjugate to the original system (4), this proves local quasi-smooth linearizability of the latter hence the theorem.

Two lemmas. The following two lemmas are applied recursively in the above proof of Theorem 5.2 to obtain the forms (83), (75), and (84). Although these lemmas team up into a single result in the above-mentioned proof, they have been stated here separately for the sake of clarity.

We will consider two control systems with state in $\mathbb{R}^d$ and control in $\mathbb{R}^r$. Expanded in coordinates, the first system reads

$$\begin{array}{rcl} \dot{x}_1 & = & f_1(x_1,\ldots,x_d,x_{d+1},\ldots,x_{d+r}) \\ & \vdots & \\ \dot{x}_d & = & f_d(x_1,\ldots,x_d,x_{d+1},\ldots,x_{d+r})\ , \end{array} \tag{88}$$

with state variable $(x_1,\ldots,x_d)$ and control variable $(x_{d+1},\ldots,x_{d+r}) \in \mathbb{R}^r$, the functions $f_1,\cdots,f_d$ being smooth $\mathbb{R}^{d+r} \to \mathbb{R}$. The second system has state variable $(z_1,\ldots,z_d)$ and control variable $(z_{d+1},\ldots,z_{d+r}) \in \mathbb{R}^r$, and it assumes the special form :

$$\begin{array}{rcl} \dot{z}_1 & = & g_1(z_1,\ldots,z_d) \\ & \vdots & \\ \dot{z}_{d-s} & = & g_{d-s}(z_1,\ldots,z_d) \\ \dot{z}_{d-s+1} & = & z_{d+1} \\ & \vdots & \\ \dot{z}_d & = & z_{d+s}\ , \end{array} \tag{89}$$

where $0 < s \leq d$ and $s \leq r$ while $g_1,\cdots,g_{d-s}$ are again smooth $\mathbb{R}^d \to \mathbb{R}$. Nothing prevents us here from having $s < r$, in which case some of the controls do not enter the equation. It will be convenient to use the aggregate notations

$$\begin{array}{ll} X \triangleq (x_1,\ldots,x_d)\,, & U \triangleq (x_{d+1},\ldots,x_{d+r})\,, \\ Z \triangleq (z_1,\ldots,z_d)\,, & V \triangleq (z_{d+1},\ldots,z_{d+r})\,, \end{array}$$

and to further split Z into (Z^1,Z^2) with

$$Z^1 \triangleq (z_1,\ldots,z_{d-s})\,,\quad Z^2 \triangleq (z_{d-s+1},\ldots,z_d)\,, \tag{90}$$

so as to write (88) in the form

$$\dot{X} \;=\; f(X,U) \tag{91}$$

and (89) as

$$\begin{array}{rcl} \dot{Z}^1 & = & g^1(Z^1,Z^2) \\ \dot{Z}^2 & = & J_r^s\,V\ , \end{array} \tag{92}$$

with J_r^s the $s\times r$ matrix, defined in (41), that selects the first s entries of a vector.

Lemma 6.2. *Let d, r and s be strictly positive integers with $s \leq d$ and $s \leq r$. Suppose, for some $\varepsilon > 0$, that*

$$\varphi : (-\varepsilon,\varepsilon)^{d+r} \to \mathbb{R}^{d+r}$$

is a homeomorphism onto its image, with inverse ψ, that conjugates system (91) *to system* (92)*. Then, there exists* $0 < \varepsilon' < \varepsilon$ *and a* smooth *local change of coordinates around* $0 \in \mathbb{R}^d$ *:*

$$\theta :\ (-\varepsilon',\varepsilon')^d \ \ \to\ \ \theta\big((-\varepsilon',\varepsilon')^d\big) \subset (-\varepsilon,\varepsilon)^d$$

that fixes the origin and is such that, in the new coordinates $\widetilde{X} = \theta^{-1}(X)$, both the system (91) *and the conjugating homeomorphism $\widetilde{\varphi} = \varphi \circ (\theta \times \mathrm{id})$ assume a block triangular structure with respect to the partition $\widetilde{X} = (\widetilde{X}^1,\widetilde{X}^2)$, where $\widetilde{X}^1 \triangleq (\widetilde{x}_1,\ldots,\widetilde{x}_{d-s})$ and $\widetilde{X}^2 \triangleq (\widetilde{x}_{d-s+1},\ldots,\widetilde{x}_d)$; that is to say, on $(-\varepsilon',\varepsilon')^{d+r}$, we have that*

- *system* (88) *reads :*

$$\begin{array}{rcl} \dot{\widetilde{X}}^1 & = & \widetilde{f}^1(\widetilde{X}^1, \widetilde{X}^2) \\ \dot{\widetilde{X}}^2 & = & \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, U), \end{array} \tag{93}$$

- *On their respective domains of definition, the homeomorphism* $\widetilde{\varphi}$ *and its inverse* $\widetilde{\psi} = (\theta^{-1} \times \mathrm{id}) \circ \psi$ *read :*

$$\begin{array}{rclcrcl} Z^1 & = & \widetilde{\varphi}^1(\widetilde{X}^1) & & \widetilde{X}^1 & = & \widetilde{\psi}_1(Z^1) \\ Z^2 & = & \widetilde{\varphi}^2(\widetilde{X}^1, \widetilde{X}^2) & & \widetilde{X}^2 & = & \widetilde{\psi}_2(Z^1, Z^2) \\ V & = & \widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, U) & & U & = & \widetilde{\psi}_3(Z^1, Z^2, V) \,. \end{array} \tag{94}$$

Lemma 6.3. *Let*

$$\widetilde{\varphi} : (-\varepsilon', \varepsilon')^{d+r} \to I\!\!R^{d+r}$$

be a homeomorphism onto its image, having the block triangular structure displayed in (94)*, and assume that it conjugates the smooth system* (93) *to the smooth system* (92)*. Necessarily then,* $\widetilde{\varphi}$ *has the following properties :*

(1) *The map* $\widetilde{\varphi}^2$ *is continuously differentiable with respect to its second argument* $\widetilde{X}^2$*, and* $\dfrac{\partial \widetilde{\varphi}^2}{\partial \widetilde{X}^2}(0,0)$ *is invertible.*

(2) *On some neighborhood of* $0 \in I\!\!R^{d+r}$ *included in* $(-\varepsilon', \varepsilon')^{d+r}$*, one has :*

$$\begin{array}{rcl} \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, U) & = & \\ \widetilde{f}^2(\widetilde{X}^1, \widetilde{X}^2, 0) & + & \left(\dfrac{\partial \widetilde{\varphi}^2}{\partial \widetilde{X}^2}(\widetilde{X}^1, \widetilde{X}^2) \right)^{-1} J_r^s \left(\widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, U) - \widetilde{\varphi}^3(\widetilde{X}^1, \widetilde{X}^2, 0) \right) \end{array} \tag{95}$$

(3) *On some neighborhood of* $0 \in I\!\!R^{d}$ *included in* $(-\varepsilon', \varepsilon')^{d}$*, the partial homeomorphism*

$$(\, \widetilde{X}^1 \,,\, \widetilde{X}^2 \,) \;\mapsto\; (\, \widetilde{\varphi}^1(\widetilde{X}^1) \,,\, \widetilde{\varphi}^2(\widetilde{X}^1, \widetilde{X}^2) \,) \tag{96}$$

conjugates the control system

$$\dot{\widetilde{X}}^1 = \widetilde{f}^1(\widetilde{X}^1, \widetilde{X}^2), \tag{97}$$

with state $\widetilde{X}^1$ *and control* $\widetilde{X}^2$*, to the control system*

$$\dot{Z}^1 = g^1(Z^1, Z^2) \tag{98}$$

with state Z^1 *and input* Z^2*.*

Note that (97) and (98) are reduced systems from (93) and (92).

Proof of Lemma 6.2. Since the homeomorphism φ conjugates (91) to (92), we know, by Proposition 3.6, that φ and ψ split component-wise into :

$$\begin{array}{rclcrcl} Z & = & \varphi_{\mathrm{I}}(X) & & X & = & \psi_{\mathrm{I}}(Z) \\ V & = & \varphi_{\mathrm{II}}(X, U) & & U & = & \psi_{\mathrm{II}}(Z, V) \;. \end{array} \tag{99}$$

Consider the map $f : (-\varepsilon, \varepsilon)^{d+r} \to I\!\!R^d$ given in (91), and let us define $g : \varphi((-\varepsilon, \varepsilon)^{d+r}) \to I\!\!R^d$ analogously from (92), namely g is the concatenated map whose first $d - s$ components are given by $g^1(Z)$ and whose last s components are given by $J_r^s V$. Define two families of continuous vector fields $\mathcal{F}'$ and $\mathcal{G}'$, on $(-\varepsilon, \varepsilon)^d$ and $\varphi_{\mathrm{I}}((-\varepsilon, \varepsilon)^d)$ respectively, by the following formulas (compare (142)) :

$$\mathcal{F}' \;=\; \{\, \delta f_{\alpha_1, \alpha_2} \,;\, \alpha_1, \alpha_2 \text{ feedbacks on } (-\varepsilon, \varepsilon)^{d+r} \,\} \;, \tag{100}$$

$$\mathcal{G}' \;=\; \{\, \delta g_{\beta_1, \beta_2} \,;\, \beta_1, \beta_2 \text{ feedbacks on } \varphi\big((-\varepsilon, \varepsilon)^{d+r}\big) \,\} \;. \tag{101}$$

Applying Proposition 3.13 twice, first to $\chi = \varphi$ and then to $\chi = \psi$, we see that each integral curve of a vector field in $\mathcal{F}'$ is mapped by φ_{I} to some integral curve of a vector field in $\mathcal{G}'$ and *vice-versa* upon replacing φ_{I} by ψ_{I}. This shows in particular that uniqueness of solutions to the Cauchy problem associated to vector fields is preserved, i.e. if we define the families of vector fields (compare (143)) :

$$\mathcal{F}'' = \{\, Y \in \mathcal{F}' \,,\ Y \text{ has a flow} \,\} \,, \tag{102}$$

$$\mathcal{G}'' = \{\, Y \in \mathcal{G}' \,,\ Y \text{ has a flow} \,\} \,, \tag{103}$$

we also have that each integral curve of a vector field in $\mathcal{F}''$ is mapped by φ_{I} to an integral curve of a vector field in $\mathcal{G}''$ and *vice-versa* upon replacing φ_{I} by ψ_{I}. By concatenation, using Proposition B.5, it follows that

$$\left.\begin{array}{l} \textit{for any } X \in (-\varepsilon,\varepsilon)^d,\ \varphi_{\mathrm{I}} \textit{ defines a homeomorphism,} \\ \textit{for the orbit topologies, from the orbit of } \mathcal{F}'' \textit{ through } \mathrm{X} \\ \textit{onto the orbit of } \mathcal{G}'' \textit{ through } \varphi_{\mathrm{I}}(X), \end{array}\right\} \tag{104}$$

where the orbit topology as described in Proposition B.5 (by definition the restriction of φ_{I} is bi-continuous for the topologies induced by the ambient space; bi-continuity for the orbit topologies requires the description of these topologies as given in Proposition B.5).

Now, the vector fields $\delta g_{\beta_1,\beta_2}$ appearing in (101) inherit from the structure of g, displayed in (92), the following particular form :

$$\delta g_{\beta_1,\beta_2}(Z) = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \beta_{1,1}(Z) - \beta_{2,1}(Z) \\ \vdots \\ \beta_{1,s}(Z) - \beta_{2,s}(Z) \end{pmatrix} , \tag{105}$$

where $\beta_{i,1},\ldots,\beta_{i,s}$ designate, for $i = 1, 2$, the first s component of the feedback β_i. This will allow for us to describe explicitly the orbits of $\mathcal{G}''$, namely :

$$\left.\begin{array}{l} \textit{the orbit of } \mathcal{G}'' \textit{ through } Z_0 = (c_1,\ldots,c_d) \\ \textit{is the connected component containing } Z_0 \textit{ of the set} \\ \{Z \in \varphi_{\mathrm{I}}\left((-\varepsilon,\varepsilon)^d\right) \,,\ z_1 = c_1,\ldots,z_{d-s} = c_{d-s}\}. \end{array}\right\} \tag{106}$$

Indeed, the orbit in question is contained in this set, because it is connected, and because all the vector fields in $\mathcal{G}''$ have their first $d - s$ components equal to zero by (105).

To prove the reverse inclusion, it is enough to show that the orbit of $\mathcal{G}''$ through Z_0, denoted hereafter by $\mathcal{O}_{\mathcal{G}'',Z_0}$, contains all the points sufficiently close to Z_0 having the same first $d - s$ coordinates as Z_0. Indeed, since Z_0 was arbitrary, this will imply that the connected component defined by (106) splits into a disjoint union of open orbits hence consists of a single one by connectedness. That is to say, putting $Z_0 = (Z_0^1, Z_0^2)$ according to (90), 106 will follow from the existence of a $\rho > 0$ such that

$$\{Z_0^1\} \times B(Z_0^2,\rho) = B(Z_0,\rho) \cap \mathcal{O}_{\mathcal{G}'',Z_0}. \tag{107}$$

Now, it follows from Remark B.3 that, for sufficiently small ρ, each connected component of $B(Z_0,\rho) \cap \mathcal{O}_{\mathcal{G}'',Z_0}$ is an embedded sub-manifold of $B(Z_0,\rho)$. Then, the connected component of $B(Z_0,\rho) \cap \mathcal{O}_{\mathcal{G}'',Z_0}$ containing Z_0 is, by inclusion, an embedded sub-manifold of the linear manifold $\{Z_0^1\} \times B(Z_0^2,\rho)$. In particular, since no strict sub-manifold can be densely embedded in a given manifold, we see that (107) will hold is only we can prove that

$$\begin{array}{l} \textit{The connected component containing } Z_0 \textit{ of } B(Z_0,\rho) \cap \mathcal{O}_{\mathcal{G}'',Z_0} \\ \textit{is dense in } \{Z_0^1\} \times B(Z_0^2,\rho) \textit{ for the Euclidean topology.} \end{array} \tag{108}$$

To prove (108), pick V_0 such that $(Z_0,V_0)\in\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$ and observe, since the latter is an open set, that shrinking ρ further, if necessary, allows us to assume $\overline{B}(Z_0,\rho)\times\overline{B}(V_0,\rho)\subset\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$. *We claim that any continuous map* $\overline{B}(Z_0,\rho)\to\overline{B}(V_0,\rho)$ *extends to a feedback on* $\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$. Indeed, in view of the one-to-one correspondence $\beta\to\psi\,\blacksquare\,\beta$ between feedbacks on $\varphi\big((-\varepsilon,\varepsilon)^{d+r}\big)$ and feedbacks on $(-\varepsilon,\varepsilon)^{d+r}$ (*cf* the discussion leading to (23)-(24)), it is enough to prove that every continuous map $\psi_{\mathrm{I}}\big(\overline{B}(Z_0,\rho)\big)\to(-\varepsilon,\varepsilon)^r$ extends to a continuous map $(-\varepsilon,\varepsilon)^d\to(-\varepsilon,\varepsilon)^r$, and this in turn follows from the Tietze extension theorem since $\psi_{\mathrm{I}}\big(\overline{B}(Z_0,\rho)\big)$ is closed in $(-\varepsilon,\varepsilon)^d$ and since $(-\varepsilon,\varepsilon)^r$ is a poly-interval. *This proves the claim.*

From the claim, it follows that the restriction to $\overline{B}(Z_0,\rho)$ of the $I\!R^s$-valued vector field $J^s_r(\beta_1(Z)-\beta_2(Z))$, accounting for the lower half of the right-hand side in (105), can be assigned *arbitrarily*, by choosing adequately the feedbacks β_1 and β_2, among *continuous* vector fields $\overline{B}(Z_0,\rho)\to\overline{B}(0,\rho)$ (take β_2 to extend the constant map V_0 on $\overline{B}(Z_0,\rho)$). Of course, the corresponding vector field $\delta g_{\beta_1,\beta_2}$ in (105) belongs to $\mathcal{G}'$ but not necessarily to $\mathcal{G}''$ since continuous vector fields need not have a flow. However, since $\delta g_{\beta_1,\beta_2}$ has a flow at least when β_1 and β_2 are smooth, we deduce from Proposition 3.4 that the restriction to $\overline{B}(Z_0,\rho)$ of the vector fields in $\mathcal{G}''$ are of the form $\{0\}\times Y$, where Y ranges over a uniformly dense subset Υ of all $I\!R^s$-valued continuous maps $\overline{B}(Z_0,\rho)\to\overline{B}(0,\rho)$. Now, every point in $B(Z_0^2,\rho)$ can be attained from Z_0^2 upon integrating, *within* $B(Z_0^2,\rho)$, a constant vector field of arbitrary small norm. By Lemma A.2 applied with $\mathcal{U}=B(Z_0^2,\rho)$ and $K=\{Z_0^2\}$, the corresponding trajectory can be approximated uniformly by integral curves *that remain in* $B(Z_0^2,\rho)$ of vector fields in Υ. Therefore, every point in $\{z_0^1\}\times B(Z_0^2,\rho)$ is the limit of endpoints of integral curves of $\mathcal{G}''$ *that remain in* $\{z_0^1\}\times B(Z_0^2,\rho)$, which proves (108) and thus (106). In particular, the orbits of $\mathcal{G}''$ are *embedded* sub-manifolds in $\varphi_{\mathrm{I}}\big((-\varepsilon,\varepsilon)^d\big)$.

Next, we turn to the orbits of $\mathcal{F}''$, and we designate by $\mathcal{O}_{\mathcal{F}'',p}$ the orbit of $\mathcal{F}''$ in $]-\varepsilon,\varepsilon[^d$ through the point p. On the one hand, Proposition B.5 and Theorem B.2 show that $\mathcal{O}_{\mathcal{F}'',p}$ is a smooth immersed sub-manifold of $]-\varepsilon,\varepsilon[^d$. On the other hand, by (104), this immersed sub-manifold is sent homeomorphically by φ_{I}, both for the orbit topology and the ambient topology, onto $\mathcal{O}_{\mathcal{G}'',\varphi_{\mathrm{I}}(p)}$ which is a smooth *embedded* s-dimensional sub-manifold of $\varphi_{\mathrm{I}}\big((-\varepsilon,\varepsilon)^d\big)$, as we saw from (106). This entails that all orbits of $\mathcal{F}''$ in $]-\varepsilon,\varepsilon[^d$ are *embedded* sub-manifolds of dimension s. Consequently, still from Proposition B.5 and Theorem B.2, there are coordinates $(\xi_1,\ldots,\xi_d)$ defined on an open neighborhood W_0 of the origin in $]-\varepsilon,\varepsilon[^d$ —this neighborhood may be assumed to be of the form $\{(\xi_1,\ldots,\xi_d),\,|\xi_i|<\varepsilon'\}$ — such that, in these coordinates,

$$W_0\cap\mathcal{O}_{\mathcal{F}'',0}\;=\;\{\,(\xi_1,\ldots,\xi_d),\text{ with }(\xi_{s+1},\ldots,\xi_d)\in T\,\}\;,$$

with T a subset of $]-\varepsilon',\varepsilon'[^{d-s}$ containing $(0,\ldots,0)$, the tangent space to $W_0\cap\mathcal{O}_{\mathcal{F}'',0}$ at each of its points being *spanned* by $\partial/\partial\xi_1,\ldots,\partial/\partial\xi_s$, while at any point $p\in W_0$ the vector fields $\partial/\partial\xi_1,\ldots,\partial/\partial\xi_s$ *belong* to the tangent space of $\mathcal{O}_{\mathcal{F}'',p}$. But since we saw that *all* orbits are smooth sub-manifolds of dimension s, these vector fields actually *span* the tangent space to the orbit at every point. Hence all the vector fields $\delta f_{\alpha_1,\alpha_2}$ in $\mathcal{F}''$ have their last $d-s$ components equal to zero on W_0 in the ξ coordinates, and this holds in particular when α_1, α_2 range over all constant feedbacks $(-\varepsilon,\varepsilon)^d\to(-\varepsilon,\varepsilon)^r$. This implies, by the very definition of $\delta f_{\alpha_1,\alpha_2}$, that $(\dot{\xi}_{s+1},\ldots,\dot{\xi}_d)$ — as computed from (91) upon performing the change of variable $X\mapsto(\xi_1,\ldots,\xi_d)$ — does not depend on the control variable U. Choose for $\widetilde{X}$ the ξ coordinates arranged in reverse order, and let $\widetilde{f}$ be the analog of f in the new coordinates $(\widetilde{X}, U)$. Then the first $d-s$ components of $\widetilde{f}$ do not depend on U so that (93) holds. Moreover, if $\widetilde{\varphi}$ denotes the new homeomorphism that conjugates (93) to (92) over $(-\varepsilon,\varepsilon)^{d+r}$, $\widetilde{\varphi}((-\varepsilon,\varepsilon)^{d+r})$, and if $\widetilde{\psi}$ denotes its inverse, it follows from (104) and the above characterization of the orbits that $\widetilde{\varphi}_{\mathrm{I}}$ maps the sets where $\tilde{x}_1,\dots,\tilde{x}_{d-s}$ are constant to those where $z_1,\dots,z_{d-s}$ are constant, thus the functions $\widetilde{\varphi}_1,\dots,\widetilde{\varphi}_{d-s}$ and $\widetilde{\psi}_1,\dots,\widetilde{\psi}_{d-s}$ depend only on their $d-s$ first arguments whence (94) follows. □

Proof of Lemma 6.3. We use again the concatenated notation $\widetilde{\varphi}_{\mathrm{I}} = (\widetilde{\varphi}^1,\widetilde{\varphi}^2)$, $\widetilde{\psi}_{\mathrm{I}} = (\widetilde{\psi}^1,\widetilde{\psi}^2)$, these partial homeomorphisms being inverse of each other. Let $(Z_0,V_0)\in\widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$ and ε'' be so small that the product neighborhood $(Z_0,V_0)+(-\varepsilon'',\varepsilon'')^{d+r}$ lies entirely within $\widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$. The restriction to $(Z_0,V_0)+(-\varepsilon'',\varepsilon'')^{d+r}$ of $\widetilde{\psi}$ conjugates (92) to (93). Consequently, for any $\overline{V}\in(-\varepsilon'',\varepsilon'')^r$, we may apply Proposition 3.13 to this restriction and to the constant feedbacks $\alpha_1(Z)=V_0+\overline{V}$ and $\alpha_2(Z)=V_0$; this yields that $\widetilde{\psi}_{\mathrm{I}}$, given by

$$(Z^1,Z^2) \;\mapsto\; (\widetilde{X}^1,\widetilde{X}^2) \;=\; (\widetilde{\psi}^1(Z^1),\widetilde{\psi}^2(Z^1,Z^2)),$$

maps every solution of

$$\dot{Z}^1 \;=\; 0 \;\;,\quad \dot{Z}^2 \;=\; J_r^s\overline{V} \tag{109}$$

that remains in $Z_0+(-\varepsilon'',\varepsilon'')^d$ to a solution of

$$\begin{aligned}\dot{\widetilde{X}}^1 = 0\,,\quad \dot{\widetilde{X}}^2 &= \widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),V_0+\overline{V}))\\ &\quad -\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),V_0))\end{aligned} \tag{110}$$

that remains in $\widetilde{\psi}_{\mathrm{I}}(Z_0+(-\varepsilon'',\varepsilon'')^d)$, and *vice versa* upon applying Proposition 3.13 in the other direction.

Integrating (109) explicitly with initial condition $Z(0)=Z_0$, we get that

$$t \;\;\mapsto\;\; \begin{pmatrix}\widetilde{\psi}^1(Z_0^1)\\ \widetilde{\psi}^2(Z_0^1,Z_0^2\,+\,tJ_r^s\overline{V})\end{pmatrix}$$

solves (110) for sufficiently small t, hence $\widetilde{\psi}^2(Z^1,Z^2)$ is differentiable at Z_0 with respect to its second argument in the direction $J_r^s\overline{V}$, with directional derivative

$$\begin{aligned}\frac{\partial\widetilde{\psi}^2}{\partial Z^2}(Z_0^1,Z_0^2)\,J_r^s\overline{V} \;\;&=\;\; \widetilde{f}^2(\widetilde{\psi}^1(Z_0^1),\widetilde{\psi}^2(Z_0^1,Z_0^2),\widetilde{\psi}^3(Z_0^1,Z_0^2,V_0+\overline{V}))\\ &\quad -\widetilde{f}^2(\widetilde{\psi}^1(Z_0^1),\widetilde{\psi}^2(Z_0^1,Z_0^2),\widetilde{\psi}^3(Z_0^1,Z_0^2,V_0))\;.\end{aligned} \tag{111}$$

In particular, since Z_0 can be any member of $\widetilde{\varphi}_{\mathrm{I}}((-\varepsilon',\varepsilon')^d)$ while $J_r^s\overline{V}$ can be assigned arbitrarily in $(-\varepsilon'',\varepsilon'')^s$, we conclude that $\partial\widetilde{\psi}^2/\partial Z^2(Z^1,Z^2)$ exists and is continuous since this holds for the partial derivatives. Next we prove that $\partial\widetilde{\psi}^2/\partial Z^2$ is invertible at every point by showing that its kernel reduces to zero. In fact, if the left-hand side of (111) vanishes, so does the right-hand side which is also the value of the right-hand side of (110) for $\widetilde{X}=\widetilde{\psi}_{\mathrm{I}}(Z_0)$. Therefore the constant map $t\mapsto\widetilde{\psi}_{\mathrm{I}}(Z_0)$ is a solution to (110) over a suitable time interval, and by conjugation the constant map $t\mapsto Z_0$ is a solution to (109) over that time interval which clearly entails $J_r^s\overline{V}=0$, as desired. Now, since $\partial\widetilde{\psi}^2/\partial Z^2$ is invertible at every $(Z^1,Z^2)\in\widetilde{\varphi}_{\mathrm{I}}((-\varepsilon',\varepsilon')^d)$, the triangular structure of (94) and the inverse function theorem together imply that

$$\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2) = \left(\frac{\partial\widetilde{\psi}^2}{\partial Z^2}(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2))\right)^{-1} \tag{112}$$

continuously exists and is invertible for $(\widetilde{X}^1,\widetilde{X}^2)\in(-\varepsilon',\varepsilon')^d$. This proves point 1.

Let us turn to point 2. Select an open neighborhood $\mathcal{W}$ of 0 having compact closure in $(-\varepsilon',\varepsilon')^d$, so there is $\eta>0$ such that $\widetilde{\varphi}(\widetilde{X},0)+(-\eta,\eta)^{d+r}$ is included in $\widetilde{\varphi}((-\varepsilon',\varepsilon')^{d+r})$ whenever $\widetilde{X}\in\mathcal{W}$. If $\overline{V}\in(-\eta,\eta)^r$, we can apply (111) to $(Z_0,V_0)=\widetilde{\varphi}(\widetilde{X},0)$ with $\widetilde{X}\in\mathcal{W}$, and we obtain in view of (112) :

$$\begin{aligned}\left(\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2)\right)^{-1} J_r^s\overline{V} \;&=\; -\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,0) \\ &+\widetilde{f}^2\left(\widetilde{X}^1,\widetilde{X}^2,\widetilde{\psi}^3\big(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),\widetilde{\varphi}^3(\widetilde{X}^1,\widetilde{X}^2,0)+\overline{V}\big)\right).\end{aligned} \tag{113}$$

Set

$$U=\widetilde{\psi}^3(\widetilde{\varphi}^1(\widetilde{X}^1),\widetilde{\varphi}^2(\widetilde{X}^1,\widetilde{X}^2),\widetilde{\varphi}^3(\widetilde{X}^1,\widetilde{X}^2,0)+\overline{V}) \tag{114}$$

and observe that $(\widetilde{X},\overline{V})\mapsto(\widetilde{X},U)=\widetilde{\psi}(\widetilde{\varphi}(\widetilde{X},0)+(0,\overline{V}))$ defines a continuous map $h:\mathcal{W}\times(-\eta,\eta)^r\to(-\varepsilon',\varepsilon')^{d+r}$, such that $h(0)=0$, which is injective. By invariance of the domain, h is a homeomorphism onto some open neighborhood of 0, say $\mathcal{N}\subset(-\varepsilon',\varepsilon')^{d+r}$. For $(\widetilde{X},U)\in\mathcal{N}$, (114) can be inverted as

$$\overline{V}=\widetilde{\varphi}^3(\widetilde{X},U)-\widetilde{\varphi}^3(\widetilde{X},0), \tag{115}$$

and substituting (114) and (115) in (113) yields (95).

Finally we prove point 3, keeping in mind the previous definitions and properties of h, $\mathcal{W}$, η and $\mathcal{N}$. For $\widetilde{X}=(\widetilde{X}^1,\widetilde{X}^2)\in(-\varepsilon',\varepsilon')^d$, define $\overline{V}(\widetilde{X})\in\mathbb{R}^s\times\{0\}\subset\mathbb{R}^r$ by the formula :

$$J_r^s\overline{V}(\widetilde{X})=\frac{\partial\widetilde{\varphi}^2}{\partial\widetilde{X}^2}(\widetilde{X}^1,\widetilde{X}^2)\big(\widetilde{f}^2(0,0,0)-\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,0)\big). \tag{116}$$

Clearly $\overline{V}:(-\varepsilon',\varepsilon')^d\to\mathbb{R}^r$ is continuous and $\overline{V}(0)=0$, so there exists an open neighborhood $\mathcal{V}\subset\mathcal{W}$ of 0 in $\mathbb{R}^d$ such that $\overline{V}(\widetilde{X})\in(-\eta,\eta)^r$ as soon as $\widetilde{X}\in\mathcal{V}$; then, if we set $h(\widetilde{X},\overline{V}(\widetilde{X}))=(\widetilde{X},U(\widetilde{X}))\in\mathcal{N}$, it follows from (116), (115), and (95) that

$$\widetilde{f}^2(\widetilde{X}^1,\widetilde{X}^2,U(\widetilde{X}))=\widetilde{f}^2(0,0,0),\qquad \widetilde{X}\in\mathcal{V}. \tag{117}$$

We will show, using Proposition 3.12*, that the restriction of* $\widetilde{\varphi}_{\mathrm{I}}$ *to any relatively compact open subset* $\mathcal{X}$ *of* $\mathcal{V}$ *conjugates* (97) *and* (98) *over* $\mathcal{X}$, $\widetilde{\varphi}(\mathcal{X})$*, and this will achieve the proof.* To this effect, let $\mathcal{C}$ to be the collection of all piecewise affine maps $\mathbb{R}\to\mathbb{R}^s$ with constant slope $\widetilde{f}^2(0,0,0)$ (*cf* the discussion before Proposition 3.12) and note that, for any open set $\mathcal{O}\subset\mathbb{R}^s$ and any compact interval $J\subset\mathbb{R}$, the restriction of $\mathcal{C}$ to J contains, in its uniform closure, the set all piecewise continuous maps $J\to\mathcal{O}$. Now, consider a solution $\gamma:I\to\mathcal{V}$ of the control system :

$$\dot{\widetilde{X}}^1 \;=\; \widetilde{f}^1(\widetilde{X}^1,\Upsilon) \tag{118}$$

with state $\widetilde{X}^1$ and control Υ; hereafter, $\mathcal{V}_{\mathrm{I}}\subset\mathbb{R}^{d-s}$ and $\mathcal{V}_{\mathrm{II}}\subset\mathbb{R}^s$ will indicate the projections of $\mathcal{V}$ onto the first $d-s$ and the last s components respectively, and similarly for any other open set in $\mathbb{R}^d$. Assume that the control function $\gamma_{\mathrm{II}}:I\to\mathcal{V}_{\mathrm{II}}$ is the restriction to I of some member of $\mathcal{C}$. By definition, if a,b are the endpoints of I (that may belong to I or not), there are time instants $a=t_0<t_1<\cdots<t_N=b$, and vectors $\bar{\xi}_1,\dots,\bar{\xi}_N\in\mathbb{R}^s$ such that, for $1\le j<N$, one has

$$t_{j-1}<t<t_j \;\Rightarrow\; \gamma_{\mathrm{II}}(t) \;=\; \bar{\xi}_j+t\widetilde{f}^2(0,0,0), \tag{119}$$

while at the points t_j themselves γ_{II} is either right or left continuous when $1<j<N$. *We claim* that $\widetilde{\varphi}_{\mathrm{I}}(\gamma(t))$ is a solution that remains in $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ of the control system :

$$\dot{Z}^1 \;=\; g^1(Z^1,\Gamma) \tag{120}$$

with state Z^1 and control Γ. In fact, since γ_{I} is continuous by definition of a solution, so is $\widetilde{\varphi}^1(\gamma_{\mathrm{I}})$ and therefore, as $\widetilde{\varphi}_{\mathrm{I}}(\gamma(t))$ lies in $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ for all $t \in I$ by construction, it is enough to check that

$$\widetilde{\varphi}^1(\gamma_{\mathrm{I}}(T_2)) - \widetilde{\varphi}^1(\gamma_{\mathrm{I}}(T_1)) = \int_{T_1}^{T_2} g^1\big(\widetilde{\varphi}^1(\gamma_{\mathrm{I}}(t)), \widetilde{\varphi}^2(\gamma_{\mathrm{I}}(t), \gamma_{\mathrm{II}}(t))\big)\,\mathrm{d}t \tag{121}$$

whenever $t_{j-1} < T_1 < T_2 < t_j$ for some $j > 1$. However, the restriction of $\gamma(t)$ to (t_{j-1}, t_j) is a solution that remains in $\mathcal{V}$ of the differential equation :

$$\begin{array}{rcl} \dot{\gamma}_{\mathrm{I}} & = & \widetilde{f}^1(\gamma_{\mathrm{I}}, \gamma_{\mathrm{II}}) \\ \dot{\gamma}_{\mathrm{II}} & = & \widetilde{f}^2(0,0,0), \end{array}$$

hence $(\gamma(t), U(\gamma(t))$ is, by (117), a solution of (93) that remains in $\mathcal{N}$, and therefore (121) follows from the triangular structure (94) of $\widetilde{\varphi}$ and the fact that it conjugates system (93) to system (92). *This proves the claim.*

In the other direction, we observe since it is included in $\mathcal{W}$ that $\mathcal{V}$ has compact closure in $(-\varepsilon', \varepsilon')^d$, and therefore that $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ in turn has compact closure in $\widetilde{\varphi}_{\mathrm{I}}\big((-\varepsilon', \varepsilon')^d\big)$. Pick $\eta' > 0$ such that $\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V}) \times (-\eta', \eta')^r \subset \widetilde{\varphi}((-\varepsilon', \varepsilon')^{d+r})$, and let $\mathcal{C}'$ denote the collection of all piecewise smooth maps $I\!R \to I\!R^s$ whose derivative is strictly bounded by η' component-wise. The restriction of $\mathcal{C}'$ to any compact real interval J is uniformly dense in the set all piecewise continuous maps $J \to \mathcal{O}$, for any open set $\mathcal{O} \subset I\!R^s$. Clearly, any solution $\gamma' : I \to \widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})$ of system (120), whose control function $\gamma'_{\mathrm{II}} : I \to \big(\widetilde{\varphi}_{\mathrm{I}}(\mathcal{V})\big)_{\mathrm{II}}$ is the restriction to I of some member of $\mathcal{C}'$, satisfies the differential equation

$$\begin{array}{rcl} \dot{\gamma}'_{\mathrm{I}} & = & g^1(\gamma'_{\mathrm{I}}, \gamma'_{\mathrm{II}}) \\ \dot{\gamma}'_{\mathrm{II}} & = & J^s_r\,(d\gamma'_{\mathrm{II}}/dt\,,\,0) \end{array}$$

on every interval where it is smooth. By the very definition of η' and $\mathcal{C}'$, it follows that $\big(\gamma'(t), (d\gamma'_{\mathrm{II}}(t)/dt, 0)\big)$ is, on such intervals, a solution to (92) that remains in $\widetilde{\varphi}((-\varepsilon', \varepsilon'))^{d+r}$ and, since $\widetilde{\psi}$ conjugates system (92) to system (93), we argue as before to the effect that $\widetilde{\psi}_{\mathrm{I}}(\gamma')$ is a solution to system (118) that remains in $\mathcal{V}$. Appealing to Proposition 3.12, we conclude that $\widetilde{\varphi}_{\mathrm{I}}$ conjugates system (118) to system (120) on relatively compact open subsets of $\mathcal{V}$, as desired. □

Appendix A. Four lemmas on ODEs

Throughout this section, we let $\mathcal{U}$ be an open subset of $I\!R^d$. We say that a continuous vector field $X : \mathcal{U} \to I\!R^d$ *has a flow* if the Cauchy problem $\dot{x}(t) = X(x(t))$ with initial condition $x(0) = x_0$ has a unique solution, defined for $t \in (-\varepsilon, \varepsilon)$ with $\varepsilon = \varepsilon(x_0) > 0$. The flow of X at time t is denoted by X_t, in other words we have with the preceding notations that $X_t(x_0) = x(t)$. It is easy to see that the domain of definition of $(t, x) \mapsto X(t, x)$ is open in $I\!R \times \mathcal{U}$.

Lemma A.1. *If $X : \mathcal{U} \to I\!R^d$ is a continuous vector field that has a flow, the map $(t, x) \mapsto X_t(x)$ is continuous on the open subset of $I\!R \times \mathcal{U}$ where it is defined.*

Proof. This is an easy consequence of the Ascoli-Arzela theorem, and actually a special case of [8, chap. V, Theorem 2.1]. □

Lemma A.2. *Assume that the sequence of continuous vector fields $X^k : \mathcal{U} \to I\!R^d$ converges to X, uniformly on compact subsets of $\mathcal{U}$, and that all the X^k as well as X itself have a flow. Suppose that $X_t(x)$ is defined for all $(t, x) \in [0, T] \times K$ with $T > 0$ and $K \subset \mathcal{U}$ compact. Then $X^k_t(x)$ is also defined on $[0, T] \times K$ for k large enough, and the sequence of mappings $(t, x) \mapsto X^k_t(x)$ converges to $(t, x) \mapsto X_t(x)$, uniformly on $[0, T] \times K$.*

Proof. By assumption,

$$K_1 = \{X_t(x);\ (t,x) \in [0,T] \times K\}$$

is a well-defined subset of $\mathcal{U}$ that contains K, and it is compact by Lemma A.1. Let K_0 be another compact subset of $\mathcal{U}$ whose interior contains K_1, and put $d(K_1, \mathcal{U} \setminus K_0) = \eta > 0$ where $d(E_1, E_2)$ indicates the distance between two sets E_1, E_2. From the hypothesis there is $M > 0$ such that $\|X^k\| \leq M$ on K_0 for all k, hence the maximal solution to $\dot{x}(t) = X^k(x(t))$ with initial condition $x(0) = x_0 \in K$ remains in K_0 as long as $t \leq \eta/2M$. Consequently the flow $(t,x) \mapsto X_t^k(x)$ is defined on $[0, \eta/2M] \times K$ for all k, with values in K_0. *We claim* that it is a bounded equicontinuous sequence of functions there. Boundedness is clear since these functions are K_0-valued, so we must show that, to every $(t,x) \in [0,\eta/2M] \times K$ and every $\varepsilon > 0$, there is $\alpha > 0$ such that $\|X^k(t',x') - X^k(t,x)\| < \varepsilon$ for all k as soon as $|t - t'| + \|x - x'\| < \alpha$. By the mean-value theorem and the uniform majorization $\|X^k(X_t^k(x))\| \leq M$, it is sufficient to prove this when $t = t'$. Arguing by contradiction, assume for some subsequence k_l and some sequence x_l converging to x in K that

$$\|X_t^{k_l}(x) - X_t^{k_l}(x_l)\| \geq \varepsilon \text{ for all } l \in \mathbb{N}. \tag{122}$$

Then, by Lemma A.1, the index k_l tends to infinity with l. Next consider the sequence of maps $F_l : [0, \eta/2M] \to K_0$ defined by $F_l(t) = X_t^{k_l}(x_l)$. Again, by the mean value theorem, it is a bounded equicontinuous family of functions and, by the Ascoli-Arzela theorem, it is relatively compact in the topology of uniform convergence (compare [8, chap. II, Theorem 3.2]). But if $\Phi : [0, \eta/2M] \to K_0$ is the uniform limit of some subsequence F_{l_j}, and since $X^{k_{l_j}}$ converges uniformly to X on K_0 as $j \to \infty$, taking limits in the relation

$$X_t^{k_{l_j}}(x_{l_j}) = x_{l_j} + \int_0^t X^{k_{l_j}}(X_s^{k_{l_j}}(x_{l_j}))\, ds$$

gives us

$$\Phi(t) = x + \int_0^t X(\Phi(s))\, ds$$

so that $\Phi(t) = X_t(x)$ since X has a flow. Altogether $F_l(t)$ converges uniformly to $X_t(x)$ on $[0, \eta/2M]$ because this is the only accumulation point, and then (122) becomes absurd. *This proves the claim.* From the claim it follows, using the Ascoli-Arzela theorem again, that the family of functions $(t,x) \mapsto X_t^k(x)$ is relatively compact for the topology of uniform convergence $[0, \eta/2M] \times K \to K_0$, and in fact it converges to $(t,x) \mapsto X_t(x)$ because, by the same limiting argument as was used to prove the claim, every accumulation point $\Phi(t,x)$ must be a solution to

$$\Phi(t,x) = x + \int_0^t X(s, \Phi(s,x))\, ds$$

hence for fixed x is an integral curve of X with initial condition x. In particular, by definition of K_1, we shall have that $d(X_t^k(x), K_1) < \eta/2$ for all $(t,x) \in [0, \eta/2M] \times K$ as soon as k is large enough. For such k the flow $(t,x) \mapsto X_t^k(x)$ will be defined on $[0, \eta/M] \times K$ with values in K_0, and we can repeat the whole argument again to the effect that $X_t^k(x)$ converges uniformly to $X_t(x)$ there. Proceeding inductively, we obtain after $[2TM/\eta] + 1$ steps at most that $(t,x) \mapsto X_t^k(x)$ is defined on $[0,T] \times K$ with values in K_0 for k large enough, and converges uniformly to $(t,x) \mapsto X_t(x)$ there, as was to be shown. □

The next lemma stands analogous to Lemma A.2 for time-dependent vector fields, assuming that the convergence holds boundedly almost everywhere in time.

The assumption that the vector fields have a flow is replaced here by a local Lipschitz condition that we now comment upon.

By definition, a time-dependent vector field $X : [t_1, t_2] \times \mathcal{U} \to I\!R^d$ is locally Lipschitz with respect to the second variable if every $(t_0, x_0) \in [t_1, t_2] \times \mathcal{U}$ has a neighborhood there such that $\|X(t, x') - X(t, x)\| < c\|x' - x\|$, for some constant c, whenever (t, x) and (t, x') belong to that neighborhood. This of course entails that X is bounded on compact subsets of $[t_1, t_2] \times \mathcal{U}$. Next, by the compactness of $[t_1, t_2]$, the local Lipschitz character of X strengthens to the effect that each $x_0 \in \mathcal{U}$ has a neighborhood $\mathcal{N}_{x_0}$ such that $\|X(t, x') - X(t, x)\| < c_{x_0}\|x' - x\|$, for some constant c_{x_0}, whenever $x, x' \in \mathcal{N}_{x_0}$ and $t \in [t_1, t_2]$. If now $\mathcal{K} \subset \mathcal{U}$ is compact, we can cover it by finitely many $\mathcal{N}_{x_{0,k}}$ as above and find $\varepsilon > 0$ such that $x, x' \in \mathcal{K}$ and $\|x - x'\| < \varepsilon$ is impossible unless x, x' lie in some common $\mathcal{N}_{x_0}$. Consequently there is $c_\mathcal{K} > 0$ such that $\|X(t, x') - X(t, x)\| < c_\mathcal{K}\|x' - x\|$ whenever $x, x' \in \mathcal{K}$ and $t \in [t_1, t_2]$, because if $\|x - x'\| < \varepsilon$ we can take $c_\mathcal{K} \geq \max_k c_{x_{0,k}}$, whereas if $\|x - x'\| \geq \varepsilon$ it is enough to take $c_\mathcal{K} > 2M/\varepsilon$ where M is a bound for $\|X\|$ on $[t_1, t_2] \times \mathcal{K}$. Finally, if $X(t, x)$ happens to vanish identically for x outside some compact $\mathcal{K}' \subset \mathcal{U}$, we can choose $\mathcal{K}$ such that

$$\mathcal{K}' \ \subset\ \overset{\circ}{\mathcal{K}} \ \subset\ \mathcal{K} \ \subset\ \mathcal{U}$$

and construct $c_\mathcal{K}$ as before except that we also pick $\varepsilon > 0$ so small that $\|x - x'\| < \varepsilon$ is impossible for $x \in \mathcal{K}'$ and $x' \notin \mathcal{K}$. Then it holds that $\|X(t, x') - X(t, x)\| < c_\mathcal{K}\|x' - x\|$ for all $x, x' \in \mathcal{U}$ and all $t \in [t_1, t_2]$, that is to say $X(t, x)$ becomes globally Lipschitz with respect to x. These remarks will be used in the proof to come.

Lemma A.3. *Let $t_1 < t_2$ be two real numbers and $X^k : [t_1, t_2] \times \mathcal{U} \to I\!R^d$ a sequence of time-dependent vector fields, measurable with respect to t, locally Lipschitz continuous with respect to $x \in \mathcal{U}$, and bounded on compact subsets of $[t_1, t_2] \times \mathcal{U}$ independently of k. Let $X : [t_1, t_2] \times \mathcal{U} \to I\!R^d$ be another time-dependent vector field, measurable with respect to t, locally Lipschitz continuous with respect to $x \in \mathcal{U}$, and assume that, to each compact $\mathcal{K} \subset \mathcal{U}$, there is $E_\mathcal{K} \subset [t_1, t_2]$ of zero measure such that, whenever $t \notin E_\mathcal{K}$, the sequence $X^k(t, x)$ converges to $X(t, x)$ as $k \to \infty$, uniformly with respect to $x \in \mathcal{K}$. Suppose finally that $\gamma : [t_1, t_2] \to \mathcal{U}$ is, for some $(t_0, x_0) \in [t_1, t_2] \times \mathcal{U}$, a solution to the Cauchy problem*

$$\dot{\gamma}(t) \ =\ X(t, \gamma(t)) \ , \qquad \gamma(t_0) \ =\ x_0. \tag{123}$$

Then, for k large enough, there is a unique solution $\gamma_k : [t_1, t_2] \to \mathcal{U}$ to the Cauchy problem

$$\dot{\gamma}_k(t) \ =\ X^k(t, \gamma_k(t)) \ , \qquad \gamma_k(t_0) \ =\ x_0, \tag{124}$$

and the sequence (γ_k) converges to γ, uniformly on $[t_1, t_2]$.

Proof. Upon multiplying $X^k(t, x)$ and $X(t, x)$ by a smooth function $\varphi(x)$ which is compactly supported $\mathcal{U} \to I\!R$ and identically 1 on a neighborhood of $\gamma([t_1, t_2])$, we may assume in view of the discussion preceding the lemma that $X(t, x)$ and $X^k(t, x)$ are defined and bounded $[t_1, t_2] \times I\!R^d \to I\!R^d$ independently of k, measurable with respect to t, and (globally) Lipschitz continuous with respect to x.

Then, by classical results [23, Proposition C 3.8., Theorem 54], the solution to (124), say γ_k uniquely exists $[t_1, t_2] \to I\!R^d$ for each k :

$$\gamma_k(t) = x_0 + \int_{t_0}^{t} X^k(s, \gamma_k(s))\, ds, \ t \in [t_1, t_2]. \tag{125}$$

From the boundedness of X^k, it is clear that γ_k is an equicontinuous and bounded family of functions, hence it is relatively compact in the topology of uniform convergence on $[t_1, t_2]$. All we have to prove then is that every accumulation point of γ_k coincides with γ. Extracting a subsequence if necessary, let us assume that γ_k converges to some $\bar\gamma$, uniformly on $[t_1,t_2]$. Let $\mathcal{K}\subset I\!\!R^d$ be a compact set containing $\gamma_k([t_1,t_2])$ for all k; such a set exists by the boundedness of γ_k. If we let $E_{\mathcal{K}}\subset[t_1,t_2]$ be the set of zero measure granted by the hypothesis, there exists to each $s\in[t_1,t_2]\backslash E_{\mathcal{K}}$ and each $\varepsilon>0$ an integer $k_{s,\varepsilon}$ such that $\|X^k(s,x)-X(s,x)\|<\varepsilon$ as soon as $x\in\mathcal{K}$ and $k>k_{s,\varepsilon}$. In another connection, the Lipschitz character of X with respect to the second argument and the uniform convergence of γ_k to $\bar\gamma$ shows that that $\|X(s,\gamma_k(s))-X(s,\bar\gamma(s))\|<\varepsilon$ for k large enough. Altogether, by a 2-ε majorization , we find that

$$\lim_{k\to\infty}\|X^k(s,\gamma_k(s))-X(s,\bar\gamma(s))\|=0,$$

that is to say the integrand in the right-hand side of (125) converges point-wise almost everywhere to $X(s,\bar\gamma(s))$. Since X^k is bounded we can apply the dominated convergence theorem and, taking limits on both sides of (125) as $k\to\infty$, we find that $\bar\gamma$ is a solution to (123) whereas the latter is unique. Hence $\bar\gamma-\gamma$ as desired. □

The following averaging lemma for continuous vector fields is less classical than in the locally Lipschitz case, where the Cauchy problem has a unique solution.

Lemma A.4. *Let $t_1<t_2$ be real numbers and $(X^{1,\ell})_{\ell\in I\!\!N}$, $(X^{2,\ell})_{\ell\in I\!\!N}$, be two sequences of continuous time-dependent vector fields $[t_1,t_2]\times I\!\!R^d\to I\!\!R^d$, uniformly bounded with respect to ℓ, that converge uniformly on compact subsets of $[t_1,t_2]\times I\!\!R^d$ to some vector fields X^1 and X^2 respectively. Denoting by $L=t_2-t_1$ the length of the time interval, define, for each $\ell\in I\!\!N$, the "average" vector field $G_\ell:[t_1,t_2]\times I\!\!R^d\to I\!\!R^d$ by :*

$$\begin{array}{rcl} t\in[t_1+\frac{j}{\ell}L\,,\,t_1+\frac{2j+1}{2\ell}L)\ \Rightarrow\ G_\ell(t,x) &=& X^{1,\ell}(t,x)\ ,\\ t\in[t_1+\frac{2j+1}{2\ell}L\,,\,t_1+\frac{j+1}{\ell}L)\ \Rightarrow\ G_\ell(t,x) &=& X^{2,\ell}(t,x)\ \ , \end{array}\tag{126}$$

for $j\in\{0,\dots,\ell-1\}$ and, say, $G_\ell(t_2,x)=X^{2,\ell}(t_2,x)$ for definiteness.

Let $\gamma_\ell:[t_1,t_2]\to I\!\!R^d$ be a solution to

$$\gamma_\ell(t)\ -\ \bar x\ \ =\ \int_{t_1}^t G_\ell(\tau,\gamma_\ell(\tau))\mathrm{d}\tau\ .\tag{127}$$

Then the sequence (γ_ℓ) is compact in $\mathbf{C}^0([t_1,t_2],I\!\!R^d)$, and every accumulation point γ_∞ is a solution to

$$\gamma_\infty(t)\ -\ \bar x\ \ =\ \ \frac12\int_{t_1}^t\left(X^1(\tau,\gamma_\infty(\tau))+X^2(\tau,\gamma_\infty(\tau))\right)\mathrm{d}\tau\ .\tag{128}$$

Proof. Let

$$M\ \ =\ \sup_{t,x,i,\ell}\|X^{i,\ell}(t,x)\|\ .\tag{129}$$

From (126)-(127), it is clear that M is a Lipschitz constant for γ_ℓ, regardless of ℓ. In particular $\gamma_\ell(t)$ stays in a fixed compact ball B of radius ML, and the family (γ_ℓ) is equicontinuous. From Ascoli-Arzela's theorem this implies compactness of the sequence (γ_ℓ) in the uniform topology on $[t_1,t_2]$.

Rewrite (127) as

$$\begin{aligned}\gamma_\ell(t)\ -\ \bar x\ \ =&\ \int_{t_1}^t\left(G_\ell(\tau,\gamma_\ell(\tau))\ -\ \frac{X^{1,\ell}(\tau,\gamma_\ell(\tau))+X^{2,\ell}(\tau,\gamma_\ell(\tau))}{2}\right)\mathrm{d}\tau\\ &+\int_{t_1}^t\left(\frac{X^{1,\ell}(\tau,\gamma_\ell(\tau))+X^{2,\ell}(\tau,\gamma_\ell(\tau))}{2}\ -\ \frac{X^{1}(\tau,\gamma_\ell(\tau))+X^{2}(\tau,\gamma_\ell(\tau))}{2}\right)\mathrm{d}\tau\\ &+\int_{t_1}^t\frac{X^{1}(\tau,\gamma_\ell(\tau))+X^{2}(\tau,\gamma_\ell(\tau))}{2}\,\mathrm{d}\tau\ .\end{aligned}\tag{130}$$

By the uniform convergence of $X^{i,\ell}$ to X^i, it will clearly follow that any accumulation point γ_∞ of (γ_ℓ) satisfies (128) if only we can show that the first integral in the right-hand side of (130) converges to zero as $\ell \to \infty$.

To prove this, we compute, from the definition of G_ℓ :

$$
\begin{aligned}
&\int_{t_1+\frac{j}{\ell}L}^{t_1+\frac{j+1}{\ell}L} \left(G_\ell(\tau,\gamma_\ell(\tau)) \;-\; \frac{X^{1,\ell}(\tau,\gamma_\ell(\tau)) + X^{2,\ell}(\tau,\gamma_\ell(\tau))}{2} \right) \mathrm{d}\tau \\
&= \int_{t_1+\frac{j}{\ell}L}^{t_1+\frac{2j+1}{2\ell}L} \frac{X^{1,\ell}(\tau,\gamma_\ell(\tau)) - X^{2,\ell}(\tau,\gamma_\ell(\tau))}{2}\,\mathrm{d}\tau \\
&\qquad - \int_{t_1+\frac{2j+1}{2\ell}L}^{t_1+\frac{j+1}{\ell}L} \frac{X^{1,\ell}(\tau,\gamma_\ell(\tau)) - X^{2,\ell}(\tau,\gamma_\ell(\tau))}{2}\,\mathrm{d}\tau \\
&= \int_{t_1+\frac{j}{\ell}L}^{t_1+\frac{2j+1}{2\ell}L} \left(\Delta_\ell(\tau,\gamma_\ell(\tau)) - \Delta_\ell(\tau+\tfrac{L}{2\ell},\gamma_\ell(\tau+\tfrac{L}{2\ell}))\right)\mathrm{d}\tau
\end{aligned}
\tag{131}
$$

with $\Delta_\ell = \frac{1}{2}(X^{1,\ell} - X^{2,\ell})$. On the compact set $[t_1,t_2]\times B$, the vector field Δ_ℓ is uniformly continuous with a modulus of continuity that does not depend on ℓ ; consequently, by the uniform Lipschitz property of γ_ℓ, we see for arbitrary $\varepsilon > 0$ that the norm of the last integral is less that $\varepsilon/2\ell$ as soon as ℓ is large enough, independently of j.

Now, the first integral in (130) can be decomposed into a sum of at most ℓ integrals like these we just studied plus an integral over an interval of length smaller that $1/\ell$. Since the norm of the integrand is bounded by $2M$, the norm of the last term is less than $2M/\ell$. Summing over j, the above estimates tell us that, for $t\in[t_1,t_2]$ and for ℓ is large enough,

$$
\int_{t_1}^{t} \left(G_\ell(\tau,\gamma_\ell(\tau)) \;-\; \frac{X^1(\tau,\gamma_\ell(\tau)) + X^2(\tau,\gamma_\ell(\tau))}{2} \right)\mathrm{d}\tau \;\;\le\;\; \frac{\varepsilon}{2} + \frac{2M}{\ell}\,.
$$

This achieves the proof since $\varepsilon > 0$ was arbitrary. □

Appendix B. Orbits of families of vector fields

In the proof of lemma 6.2 we need results from [25] on orbits[5] of families of smooth vector fields, that were recently exposed in the textbook [14, chapter II]. We recall them below, in a slightly expanded form.

Let $\mathcal{F}$ be a family of smooth vector fields defined on an open subset U of $\mathbb{R}^d$. For any positive integer N and vector fields $X^1,\ldots,X^N$ belonging to $\mathcal{F}$, given $m\in U$, consider the map F given by

$$
(t_1,\ldots,t_N) \;\;\mapsto\;\; X^1_{t_1}(X^2_{t_2}(\cdots(X^N_{t_N}(m))\cdots)) \tag{132}
$$

where the standard notation $X_t(x)$ indicates the flow of X from x at time t; of course, F depends on the choice of the vector fields X^j and of the point m. This map is defined on some open connected neighborhood of the origin, hereafter denoted by $\mathrm{dom}(F)$, and takes values in U. In fact, $(t_1,\ldots,t_N)\in\mathrm{dom}(F)$ if, and only if, for every $j\in\{1,\ldots,N\}$, the solution $x(\tau)$ to $\dot x = X^j(x)$, with initial condition $x(0) = X^{j-1}_{t_{j-1}}(\cdots(X^1_{t_1}(m))\cdots)$, exists in U for all $\tau\in[0,t_j]$ (or $[t_j,0]$ if $t_j<0$).

---

[5] One of the motivations in [25] was to generalize the notion of integral manifolds to vector fields that are smooth but not real analytic. Note that the orbits of a family of *real analytic* vector fields actually coincide with the maximal integral manifolds of the closure of this family under Lie brackets [25, 16, 18]. However, even if we assume the control system (4) to be real analytic, integral manifolds are of no help to us because topological conjugacy does not preserve tangency nor Lie brackets. Using orbits of families of vector fields instead is much more efficient, because topological conjugacy does preserve integral curves.

The *orbit* of the family $\mathcal{F}$ through a point $m \in U$ is the set of all points that lie in the image of F for at least one choice of the vector fields $X^1, \ldots, X^N$. In words, the orbit of the family $\mathcal{F}$ through m is the set of points that may be linked to m in U upon concatenating finitely many integral curves of vector fields in the family. We shall denote by $\mathcal{O}_{\mathcal{F},p}$ the orbit of $\mathcal{F}$ through m.

Note that the definition depends on U in a slightly subtle manner : if $\mathcal{F}$ defines by restriction a family of vector fields $\mathcal{F}|_V$ on a smaller open set $V \subset U$ and if $m \in V$, then

$$V \cap \mathcal{O}_{\mathcal{F},m} \quad \supset \quad \mathcal{O}_{\mathcal{F}|_V,m}, \tag{133}$$

but the inclusion is generally strict because of the requirement that the integral curves used to construct $\mathcal{O}_{\mathcal{F}|_V,m}$ should lie entirely in V.

We turn to topological considerations. The topology of U is the usual Euclidean topology. The topology of $\mathcal{O}_{\mathcal{F},m}$ *as an orbit* is the finest that makes all the maps F, arising from (132), continuous on their respective domains of definition, the latter being endowed with the Euclidean topology. The classical smoothness of the flow implies that each F is continuous $\mathrm{dom}(F) \to \mathbb{R}^d$, hence the topology of $\mathcal{O}_{\mathcal{F},m}$ as an orbit is finer than the Euclidean topology induced by the ambient space U. *It can be strictly finer*, and this is why we speak of the *orbit topology*, as opposed to the *induced topology.*

Starting from $\mathcal{F}$, one defines a larger family of vector fields $P_{\mathcal{F}}$, consisting of all the push-forwards[6] of vector fields in $\mathcal{F}$ through all local diffeomorphisms of the form $X^1_{t_1} \circ X^2_{t_2} \circ \cdots \circ X^N_{t_N}$ where $X^1, \ldots, X^N$ belong to $\mathcal{F}$. That is to say, vector fields in $P_{\mathcal{F}}$ are of the form

$$\left(X^1_{t_1} \circ \cdots \circ X^N_{t_N}\right)_\star X^0 \tag{134}$$

where $X^0, X^1, \ldots, X^N$ belong to $\mathcal{F}$.

Remark B.1. Note that a member of $P_{\mathcal{F}}$ is defined on an open set which is generally a strict subset of U, whereas members of $\mathcal{F}$ are defined over the whole of U, and it is understood that a curve $\gamma : I \to U$, where I is a real interval, will be called an integral curve of $Y \in P_{\mathcal{F}}$ only when $\gamma(I)$ is included in the domain of definition of Y.

For $x \in U$, we denote by $P_{\mathcal{F}}(x)$ the subspace of $\mathbb{R}^d$ spanned by all the vectors $Y(x)$, where $Y \in P_{\mathcal{F}}(x)$ is defined in a neighborhood of x.

Theorem B.2 below, which is the central result in this appendix, describes the topological nature of the orbits. To interpret the statement correctly, it is necessary to recall (see for instance [24]) that an *immersed* sub-manifold of a manifold is a subset of the latter which is a manifold in its own right, and is such that the inclusion map is an immersion. This allows one to naturally identify the tangent space to an immersed sub-manifold at a given point with a linear subspace of the tangent space to the ambient manifold at the same point. The topology of an immersed sub-manifold is in general finer than the one induced by the ambient manifold; when these two topologies coincide, the sub-manifold is called *embedded.*

Theorem B.2 (Orbit Theorem, Sussmann [25]). *Let $\mathcal{F}$ be a family of smooth vector fields defined on an open set $U \subset \mathbb{R}^d$, and m be a point in U. If $\mathcal{O}_{\mathcal{F},m}$ denotes the orbit of $\mathcal{F}$ through m, then:*

[6] Recall that the push-forward of a vector field $X : V \to \mathbb{R}^d$ through a diffeomorphism $\varphi : V \to \varphi(V)$ is the vector field $\varphi_\star X$ on $\varphi(V)$ whose flow at each time is the conjugate of the flow of X under the diffeomorphism φ; it can be defined as $\varphi_\star X(\varphi(x)) = D\varphi(x)X(x)$, where $D\varphi(x)$ is the derivative of φ at $x \in V$.

(i) *Endowed with the orbit topology, $\mathcal{O}_{\mathcal{F},m}$ has a unique differential structure that makes it a smooth connected immersed sub-manifold of U, for which the maps* (132) *are smooth.*
(ii) *The tangent space to $\mathcal{O}_{\mathcal{F},m}$ at $x \in \mathcal{O}_{\mathcal{F},m}$ is $P_{\mathcal{F}}(x)$.*
(iii) *There exists an open neighborhood W of m in U, and smooth local coordinates $\xi : W \to (-\eta, \eta)^d \subset I\!R^d$, with $\xi(m) = 0$, such that*
  (a) *in these coordinates, $W \cap \mathcal{O}_{\mathcal{F},m}$ is a product :*

$$W \cap \mathcal{O}_{\mathcal{F},m} = (-\eta, \eta)^q \times T \tag{135}$$

  where $\eta > 0$, q is the dimension of $\mathcal{O}_{\mathcal{F},m}$, and T is some subset of $(-\eta, \eta)^{d-q}$ containing the origin. The orbit topology of $\mathcal{O}_{\mathcal{F},m}$ induces on $W \cap \mathcal{O}_{\mathcal{F},m}$ the product topology where $(-\eta, \eta)^q$ is endowed with the usual Euclidean topology and T with the discrete topology.
  (b) *if $\gamma : [t_1, t_2] \to W \cap \mathcal{O}_{\mathcal{F},m}$ is an integral curve of a vector field $Y \in P_{\mathcal{F}}$ (see remark B.1), then $t \mapsto \xi_i(\gamma(t))$, $q + 1 \le i \le d$, are constant mappings,*
  (c) *the tangent space to $\mathcal{O}_{\mathcal{F},m}$ at each point $p \in W \cap \mathcal{O}_{\mathcal{F},m}$ is spanned by the vector fields $\partial/\partial\xi_1, \ldots, \partial/\partial\xi_q$,*
  (d) *at any point $p \in W$, the vector fields $\partial/\partial\xi_1, \ldots, \partial/\partial\xi_q$ belong to the tangent space to the orbit of $\mathcal{F}$ through p.*

Remark B.3. Another description of the product topology in point $(iii) - (a)$ is as follows. The connected components of $W \cap \mathcal{O}_{\mathcal{F},m}$ are the sets

$$S_{W,a} = (-\eta, \eta)^q \times \{a\} \tag{136}$$

for $a \in T$, and the topology on each of these connected components is the topology induced by the ambient Euclidean topology. In particular each $S_{W,a}$ is an *embedded* sub-manifold of U.

Proof of Theorem B.2. Assertion (i) is the standard form of the orbit theorem (cf *e.g.* [14, Chapter 2, Theorem 1]), while assertion (ii) is a rephrasing of [25, Theorem 4.1, point (b)]. Assertion (iii) apparently cannot be referenced exactly in this form, but we shall deduce it from the previous ones as follows.

By point (ii), the tangent space to $\mathcal{O}_{\mathcal{F},m}$ at $m \in S$ is the linear span over $I\!R$ of $Y^1(m), \ldots, Y^q(m)$, where $Y^1, \ldots, Y^q$ are q vector fields belonging to $P_{\mathcal{F}}$, defined on some neighborhood of m, and such that $Y^1(m), \ldots, Y^q(m)$ are linearly independent (recall that q is the dimension of $\mathcal{O}_{\mathcal{F},m}$). Let us write

$$Y^j = \left(X^{j,1}_{t_{j,1}} \circ \cdots \circ X^{j,N_j}_{t_{j,N_j}}\right)_\star X^{j,0}, \ 1 \le j \le q,$$

where $X^{j,k} \in \mathcal{F}$ for $0 \le k \le N_j$, and where the $t_{j,k}$'s are real numbers for which the concatenated flow exists, locally around m (compare (134)).

Since $Y^1(m), \ldots, Y^q(m)$ are linearly independent, one may complement them into a basis of $I\!R^d$ by adjunction of $d - q$ independent vectors that may, without loss of generality, be regarded as values at m of $d-q$ smooth vector fields in U, say $Y^{q+1}, \ldots, Y^d$. Then, the smooth map

$$L(\xi_1, \ldots, \xi_d) = \left(Y^1_{\xi_1} \circ \cdots \circ Y^q_{\xi_q} \circ Y^{q+1}_{\xi_{q+1}} \circ \cdots \circ Y^d_{\xi_d}\right)(m) \tag{137}$$

defines a diffeomorphism from some poly-interval $\mathcal{I}_\eta = \{(\xi_1, \ldots, \xi_d)\,,\ |\xi_i| < \eta\}$ onto an open neighborhood W of m in U, simply because the derivative of L is invertible at the origin as $Y^1(m), \ldots, Y^d(m)$ are linearly independent by construction. Let $\xi : W \to \mathcal{I}_\eta$ denote its inverse.

By the characteristic property of push-forwards, we locally have, for $1 \leq j \leq q$, that

$$Y^j_{\xi_j} = X^{j,1}_{t_{j,1}} \circ \cdots \circ X^{j,N_j}_{t_{j,N_j}} \circ X^{j,0}_{\xi_j} \circ X^{j,N_j}_{-t_{j,N_j}} \circ \cdots \circ X^{j,1}_{-t_{j,1}} \, . \tag{138}$$

This implies that, in (137), the images under L of those d-tuples sharing a common value of $\xi_{q+1}, \ldots, \xi_d$ all lie in the same orbit $\mathcal{O}_{\mathcal{F},L(0,\ldots,0,\xi_{q+1},\ldots,\xi_d)}$. In particular, the map

$$\tau_1, \ldots, \tau_q \quad \mapsto \quad \left(Y^1_{\tau_1+\xi_1} \circ \cdots \circ Y^q_{\tau_q+\xi_q} \circ Y^{q+1}_{\xi_{q+1}} \circ \cdots \circ Y^d_{\xi_d}\right)(m)$$

is defined $\Pi^q_{j=1}(-\eta - \xi_j, \eta - \xi_j) \to W \cap \mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$, and this map is smooth from the Euclidean to the orbit topology by (138) and point (i). If we compose it with the immersive injection $J_W : W \cap \mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)} \to W$ (keeping in mind that $W \cap \mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$ is open in $\mathcal{O}_{\mathcal{F},L(\xi_1,\ldots,\xi_d)}$ since the orbit topology is finer than the Euclidean one), and if we subsequently apply ξ, we get the affine map

$$\tau_1, \ldots, \tau_q \quad \mapsto \quad (\tau_1 + \xi_1, \cdots, \tau_q + \xi_q, \xi_{q+1}, \cdots, \xi_d). \tag{139}$$

Thus the derivative of (139) factors through the derivative of $\xi \circ J_W$ at $L(\xi_1, \ldots, \xi_d)$, which implies (d); from this (c) follows, because q is the dimension of the orbit through m. If $Y \in P_{\mathcal{F}}$ is defined over an open subset of W, and if we write in the ξ coordinates $Y(\xi) = \sum_i a_i(\xi)\partial/\partial\xi_i$, then, since $Y(\xi)$ is tangent to $\mathcal{O}_{\mathcal{F},\xi}$ by (ii), we deduce from (c), that the functions $a_{q+1}, \ldots, a_d$ vanish on $\mathcal{O}_{\mathcal{F},m}$, whence (b) holds.

We finally prove (a). Considering (137) and (138), a moment's thinking will convince the reader that $W \cap \mathcal{O}_{\mathcal{F},m}$ consists exactly, in the ξ coordinates, of those $(\xi_1, \ldots, \xi_d)$ such that

$$\left(Y^{q+1}_{\xi_{q+1}} \circ \cdots \circ Y^d_{\xi_d}\right)(m) \in \mathcal{O}_{\mathcal{F},m}, \tag{140}$$

which accounts for (135) where T is the set of $(d-q)$-tuples $(\xi_{q+1}, \ldots, \xi_d)$ such that (140) holds. To prove that the orbit topology is the product topology on $(-\eta, \eta)^q \times T$ where T is discrete, consider a map F as in (132), and pick $\bar{t} = (\bar{t}_1, \ldots, \bar{t}_N) \in \mathrm{dom}(F)$ such that $F(\bar{t}) \in W$ (hence $F(\bar{t}) \in W \cap \mathcal{O}_{\mathcal{F},m}$) ; then F is continuous at $\bar{t}$ for the product topology because, for t close enough to $\bar{t}$, the values $\xi_{q+1}(F(t)), \ldots, \xi_d(F(t))$ do not depend on t by (b) (moving t_i means following the flow of a vector field in $P_{\mathcal{F}}$, namely the push-forward of X^i through $X^1_{t_1} \circ \cdots \circ X^{i-1}_{t_{i-1}}$) while $\xi_1(F(t)), \ldots, \xi_q(F(t))$ vary continuously with t according to the continuous dependence on time and initial conditions of solutions to differential equations. Since this is true for all maps F, the orbit topology on $W \cap \mathcal{O}_{\mathcal{F},m}$ is finer than the product topology. To show that it cannot be strictly finer, it is enough to prove that the orbit topology coincides with the Euclidean topology on each set $S_{W,a}$ defined in (136), a basis of which consists of the sets $O \times \{a\}$ where O is open in $(-\eta, \eta)^q$. Being open for the product topology, these sets are open the orbit topology as well by what precedes and, since $\mathcal{O}_{\mathcal{F},m}$ is a manifold by (i), each point $(y, a) \in O \times \{a\}$ has, in the orbit topology, a neighborhood $\mathcal{N}_y \subset O \times \{a\}$ which is homeomorphic to an open ball of $\mathbb{R}^q$ *via* some coordinate map. When viewed in these coordinates, the injection $\mathcal{N}_y \to O \times \{a\}$ from the orbit topology to the Euclidean topology is a continuous injective map from an open ball in $\mathbb{R}^q$ into $\mathbb{R}^q$, and therefore it is a homeomorphism onto its image by invariance of the domain. As (y, a) was arbitrary in $O \times \{a\}$, this shows the latter is a union of open sets for the orbit topology, as desired. □

Consider now the control system :

$$\dot{x} \quad = \quad f(x, u), \tag{141}$$

with state $x \in \mathbb{R}^d$ and control $u \in \mathbb{R}^r$, the function f being smooth on $\mathbb{R}^d \times \mathbb{R}^r$. Let Ω be an open subset of $\mathbb{R}^d \times \mathbb{R}^r$ and, following the notation introduced in section 3, put $\Omega_{\mathbb{R}^d}$ to denote its projection onto the first factor. In the proof of Theorem 5.2, we shall be concerned with the following family of vector fields on $\Omega_{\mathbb{R}^d}$:

$$\mathcal{F}' \;=\; \{\, \delta f_{\alpha_1,\alpha_2}\,,\ \alpha_1,\alpha_2 \text{ feedbacks on } \Omega\,\}\;, \tag{142}$$

where feedbacks on Ω were introduced in Definition 3.3 and the notation $\delta f_{\alpha_1,\alpha_2}$ was fixed in (25), (26).

Since feedbacks are only required to be *continuous*, $\mathcal{F}'$ is a family of continuous *but not necessarily differentiable* vector fields on $\Omega_{\mathbb{R}^d}$ and, though the existence of solutions to differential equations with continuous right-hand side makes it still possible to define the orbit as the collection of endpoints of all concatenated integrations like (132), Theorem B.2 does not apply in this case.

To overcome this difficulty, we will consider instead of $\mathcal{F}'$ the smaller family :

$$\mathcal{F}'' \;=\; \{\, X\in\mathcal{F}'\,,\ X \text{ has a flow}\,\}\;, \tag{143}$$

where the sentence "X has a flow" means, as in appendix A, that the Cauchy problem $\dot{x}(t) = X(x(t))$, $x(0) = x_0$, has a unique solution, defined for $|t| < \varepsilon_0$ where ε_0 may depend on x_0, whenever x_0 lies in the domain of definition of X. Let us consider the orbit $\mathcal{O}_{\mathcal{F}'',m}$ of $\mathcal{F}''$ through $m\in\Omega_{\mathbb{R}^d}$, which is still defined as the union of images of all maps (132) where $X^j\in\mathcal{F}''$, the domain of each such map F being again a connected open neighborhood $\mathrm{dom}(F)$ of the origin in $\mathbb{R}^N$ by repeated application of Lemma A.1. As before, we define the orbit topology on $\mathcal{O}_{\mathcal{F}'',m}$ to be the finest that makes all the maps (132) continuous, and since uniqueness of solutions implies continuous dependence on initial conditions (see Lemma A.1), the orbit topology is again finer than the Euclidean topology. *A priori*, we know very little about $\mathcal{O}_{\mathcal{F}'',m}$ and its orbit topology as Theorem B.2 does not apply. However, Proposition B.5 below will establish that these notions coincide with those arising from the family $\mathcal{F}$ of *smooth* vector fields obtained by setting :

$$\mathcal{F} \;=\; \{\, \delta f_{\alpha_1,\alpha_2}\,,\ \alpha_1,\alpha_2 \ \underline{\text{smooth}} \text{ feedbacks on } \Omega\,\}. \tag{144}$$

Note that, from the definitions (142), (143) and (144), we obviously have

$$\mathcal{F}\;\subset\;\mathcal{F}''\;\subset\;\mathcal{F}'\;, \tag{145}$$

hence the orbits of these families through a given point obey the same inclusions.

Remark B.4. It may of course happen that the family $\mathcal{F}'$ is empty because Ω admits no feedback at all. However, if $\mathcal{F}'$ is not empty, then $\mathcal{F}$ is not empty either by Proposition 3.4.

Proposition B.5. *Suppose that $f:\mathbb{R}^d\times\mathbb{R}^r\to\mathbb{R}^d$ is smooth, and let Ω be an open subset of $\mathbb{R}^d\times\mathbb{R}^r$. Let $\mathcal{F}''$ be defined by* (142)-(143)*.*

For any $m\in\Omega_{\mathbb{R}^d}$, the orbit $\mathcal{O}_{\mathcal{F}'',m}$ of $\mathcal{F}''$ through m coincides with the orbit through m of the family $\mathcal{F}$ of smooth *vector fields defined by* (144)*, and the topology of $\mathcal{O}_{\mathcal{F}'',m}$, as an orbit of $\mathcal{F}$, coincides with its topology as an orbit of $\mathcal{F}''$. In particular, the conclusions of Theorem B.2 hold if we replace $\mathcal{F}$ by $\mathcal{F}''$ and U by $\Omega_{\mathbb{R}^d}$.*

Remark B.6. With a limited amount of extra-work, it is possible to show that the orbits of $\mathcal{F}'$ also coincide with those of $\mathcal{F}$. Hence they turn out to be manifolds despite the possible non-uniqueness of solutions to the Cauchy problem. However, (132) is no longer convenient to define the orbit topology in this case because the maps F may be multiply-valued when $X^j\in\mathcal{F}'$, and it is simpler to work with the family $\mathcal{F}''$ anyway.

The proof of the proposition is based on the following lemma.

Lemma B.7. *For $m \in \Omega_{\mathbb{R}^d}$ and $X^1, \ldots, X^N \in \mathcal{F}''$, let $F : \mathrm{dom}(F) \to \Omega_{\mathbb{R}^d}$ be defined by* (132)*. Fix $\bar{t} = (\bar{t}_1, \ldots, \bar{t}_N) \in \mathrm{dom}(F)$ and set $\overline{m} = F(\bar{t})$.*

Then, there is a neighborhood $\mathcal{T}$ of $\bar{t}$ in $\mathrm{dom}(F)$, with $F(\mathcal{T}) \subset \mathcal{O}_{\mathcal{F},\overline{m}}$, such that $F : \mathcal{T} \to \mathcal{O}_{\mathcal{F},\overline{m}}$ is continuous from the Euclidean topology to the orbit topology.

Assuming the lemma for a while, we first prove the proposition.

Proof of Proposition B.5. We noticed already from (145) that the orbit of $\mathcal{F}''$ through m contains the orbit of $\mathcal{F}$ through m. To get the reverse inclusion, consider the map F defined by (132) for some vector fields $X^1, \ldots, X^N$ belonging to $\mathcal{F}''$. Then, observe from Lemma B.7 that F takes values in a disjoint union of orbits of $\mathcal{F}$, and that it is continuous if each orbit in this union is endowed with the orbit topology. Since $\mathrm{dom}(F)$ is connected, F takes values in a single orbit, which can be none but $\mathcal{O}_{\mathcal{F},m}$. As F was arbitrary, we conclude that $\mathcal{O}_{\mathcal{F}'',m} \subset \mathcal{O}_{\mathcal{F},m}$ and therefore the two orbits agree as sets. Moreover, since each map F was continuous $\mathrm{dom}(F) \to \mathcal{O}_{\mathcal{F},m}$, the orbit topology of $\mathcal{O}_{\mathcal{F}'',m}$ is by definition finer than the orbit topology of $\mathcal{O}_{\mathcal{F},m}$; but since it is also coarser, by definition of the orbit topology on $\mathcal{O}_{\mathcal{F},m}$, because $\mathcal{F} \subset \mathcal{F}''$, the two topologies in turn agree as desired. □

Proof of Lemma B.7. Theorem B.2 applied to the family $\mathcal{F}$, at the point $\overline{m} = F(\bar{t})$, yields an open neighborhood W of $\overline{m}$ in $\Omega_{\mathbb{R}^d}$ and smooth local coordinates $(\xi_1, \ldots, \xi_d) : W \to (-\eta, \eta)^d$ satisfying properties $(iii) - (a)$ to $(iii) - (d)$ of that theorem. For $\varepsilon > 0$ denote by $\mathcal{T}_\varepsilon$ the compact poly-interval :

$$\mathcal{T}_\varepsilon \;=\; \{t = (t_1, \ldots, t_N) \in \mathbb{R}^N,\ |t_i - \bar{t}_i| \le \varepsilon\}\ .$$

By Lemma A.1, F is continuous $\mathrm{dom}(F) \to \Omega_{\mathbb{R}^d}$ and, since $\mathrm{dom}(F)$ is an open neighborhood of $\bar{t}$ in $\mathbb{R}^N$, we can pick $\varepsilon > 0$ such that

$$\mathcal{T}_\varepsilon \subset \mathrm{dom}(F) \quad \text{and} \quad F(\mathcal{T}_\varepsilon) \subset W\ .$$

As $X^1, \ldots, X^N$ belong to $\mathcal{F}'' \subset \mathcal{F}'$, we can write

$$X^\ell \;=\; \delta f_{\alpha_1^\ell, \alpha_2^\ell}\ , \quad 1 \le \ell \le N$$

for some collection of feedbacks α_1^ℓ, α_2^ℓ on Ω. From Proposition 3.4, there exists for each $(\ell, l) \in \{1, \ldots, N\} \times \{1, 2\}$ a sequence of *smooth* feedbacks on Ω, say $(\beta_l^{\ell,k})_{k \in \mathbb{N}}$, converging to α_l^ℓ uniformly on $\Omega_{\mathbb{R}^d}$. Subsequently, we let $Y^{\ell,k}$ denote, for $1 \le \ell \le N$ and $k \in \mathbb{N}$, the smooth vector field on $\Omega_{\mathbb{R}^d}$

$$Y^{\ell,k} \;=\; \delta f_{\beta_1^{\ell,k}, \beta_2^{\ell,k}}\ .$$

Clearly $Y^{\ell,k} \in \mathcal{F}$ and, for each ℓ, we have that $Y^{\ell,k}$ converges to X^ℓ as $k \to \infty$, uniformly on compact subsets of $\Omega_{\mathbb{R}^d}$.

Now, pick $j \in \{1, \ldots, N\}$ and consider a N-tuple $t^{(j)} \in \mathcal{T}_\varepsilon$ of the form :

$$t^{(j)} = (\bar{t}_1, \ldots, \bar{t}_{j-1}, t_j, \ldots, t_N)\,,\ |t_\ell - \bar{t}_\ell| \le \varepsilon \ \text{ for } j \le \ell \le N.$$

Let also $\mathbf{1}_j$ designate, for simplicity, the N-tuple $(0, \ldots, 1, \ldots, 0)$ with zero entries except for the j-th one which is 1. Then, for $|\lambda| \le \varepsilon$, we have that

$$t^{(j)} + \lambda \mathbf{1}_j = (\bar{t}_1, \ldots, \bar{t}_{j-1}, \bar{t}_j + \lambda, t_{j+1}, \ldots, t_N)\ \in \mathcal{T}_\varepsilon,$$

and a simple computation allows us to rewrite $F(t + \lambda \mathbf{1}_j)$ as :

$$F(t^{(j)} + \lambda \mathbf{1}_j) \;=\; X^1_{\bar{t}_1} \circ \cdots \circ X^{j-1}_{\bar{t}_{j-1}} \circ X^j_\lambda \circ X^{j-1}_{-\bar{t}_{j-1}} \circ \cdots \circ X^1_{-\bar{t}_1}(F(t)).$$

Let us set

$$A_k(\lambda) \;=\; Y^{1,k}_{\bar{t}_1} \circ \cdots \circ Y^{j-1,k}_{\bar{t}_{j-1}} \circ Y^{j,k}_\lambda \circ Y^{j-1,k}_{-\bar{t}_{j-1}} \circ \cdots \circ Y^{1,k}_{-\bar{t}_1}(F(t)).$$

Repeated applications of Lemmas A.1 and A.2 show that, for fixed j and $t^{(j)}$, the map $\lambda \mapsto A_k(\lambda)$ is well-defined $[-\varepsilon, \varepsilon] \to W$ as soon as the integer k is sufficiently large, and moreover that $A_k(\lambda)$ converges to $F(t^{(j)}+\lambda\mathbf{1}_j)$ as $k\to+\infty$, uniformly with respect to $\lambda\in[-\varepsilon,\varepsilon]$. Now, by the characteristic property push forwards, $\lambda\mapsto A_k(\lambda)$ is an integral curve of the smooth vector field

$$Z^k \;=\; \left(Y^{1,k}_{\bar t_1}\circ\cdots\circ Y^{j-1,k}_{\bar t_{j-1}}\right)_\star Y^{j,k}\ ,$$

which is defined on a neighborhood of $\{F(t^{(j)}+\lambda\mathbf{1}_j);\ |\lambda|\le\varepsilon\}$ in W. Since $Z^k\in P_{\mathcal F}$ (*cf* equation (134)), it follows from point $(iii)-(b)$ of Theorem B.2 that, for k large enough,

$$\xi_i\circ A_k(\lambda)\;=\;\xi_i\circ A_k(0)\ ,\quad \forall\lambda\in[-\varepsilon,\varepsilon],\ i\in\{q+1,\dots,d\}.$$

It is clear from the definition that $A_k(0)=F(t^{(j)})$; hence, using the continuity of ξ_i and taking, in the above equation, the limit as $k\to+\infty$, we get

$$\xi_i\circ F(t^{(j)}+\lambda\mathbf{1}_j)=\xi_i\circ F(t^{(j)}),\ \forall\lambda\in[-\varepsilon,\varepsilon],\ i\in\{q+1,\dots,d\}. \tag{146}$$

Since $\xi_{q+1}\circ F(\bar t)=\cdots=\xi_d\circ F(\bar t)=0$ by definition of W, successive applications of (146) for $j=N,\dots,1$ lead us to the conclusion that

$$\xi_{q+1}\circ F(t)=\cdots=\xi_d\circ F(t)=0,\quad \forall t\in\mathcal{T}_\varepsilon\ . \tag{147}$$

Equation (147) means that, in the ξ-coordinates, $F(\mathcal{T}_\varepsilon)\subset(-\eta,\eta)^q\times\{0\}$. Hence, from the local description of the orbits in (135) (where m is to be replaced by $\overline{m}$), we deduce that $F(\mathcal{T}_\varepsilon)\subset\mathcal{O}_{\mathcal{F},\overline{m}}$. Actually, with the notations of (136), we even get the stronger conclusion that

$$F(\mathcal{T}_\varepsilon)\;\subset\;S_{W,0}$$

which achieves the proof of the lemma, with $\mathcal{T}=\mathcal{T}_\varepsilon$, because the orbit topology on $S_{W,0}$ is the Euclidean topology by Remark B.3. □

Acknowledgments. The authors wish to acknowledge fruitful discussions with Prof. M. Chyba, from University of Hawaii (USA). Thanks are also due to Prof. C.T.C. Wall from the University of Liverpool for his comments on the open question of section 5.1.

References

[1] V. I. Arnold. *Equations différentielles ordinaires.* MIR, 3rd edition, 1974.

[2] V. I. Arnold. *Chapitres supplémentaires de la théorie des équations différentielles ordinaires.* MIR, 1980.

[3] L. Baratchart, M. Chyba, and J.-B. Pomet. A Grobman-Hartman theorem for control systems. *J. Dyn. Differ. Equ.*, 19:75-107, 2007.
URL: `http://dx.doi.org/10.1007/s10884-006-9014-5`.

[4] P. Brunovský. A classification of linear controllable systems. *Kybernetika*, 6:176–188, 1970.

[5] S. Celikovský. Topological equivalence and topological linearization of controlled dynamical systems. *Kybernetika*, 31(2):141–150, 1995.

[6] E. A. Coddington and N. Levinson. *Theory of ordinary differential equations.* McGraw-Hill Book Company, Inc., New York-Toronto-London, 1955.

[7] F. Colonius and W. Kliemann. Some aspects of control systems as dynamical systems. *J. Dyn. Differ. Equations*, 5:469–494, 1993.

[8] P. Hartman. *Ordinary Differential Equations.* Birkhäuser, 2 edition, 1982.

[9] L. R. Hunt, R. Su, and G. Meyer. Design for multi-input nonlinear systems. In R. Brockett, editor, *Differential Geometric Control Theory*, pages 258–298. Birkhäuser, 1983.

[10] A. Isidori. *Nonlinear Control Systems.* Comm. in Control Engineering. Springer-Verlag, 3rd edition, 1995.

[11] B. Jakubczyk. Equivalence and invariants of nonlinear control systems. In *Nonlinear controllability and optimal control*, pages 177–218. Dekker, New York, 1990.

[12] B. Jakubczyk and W. Respondek. On linearization of control systems. *Bull. Acad. Polon. Sci. Sér. Sci. Math.*, 28(9-10):517–522, 1980.

[13] A. Juditsky and H. Hjalmarsson and A. Benveniste and B. Delyon and L. Ljung and J. Sjöberg and Q. Zhang. Nonlinear black-box models in system identification: mathematical foundations. *Automatica*, 31(12):1725–1750, 1995.

[14] V. Jurdjevic. *Geometric Control Theory*, volume 51 of *Cambridge Studies in Advanced Mathematics*. Cambridge Univ. Press, 1997.

[15] V. Jurdjevic and H. J. Sussmann. Controllability of nonlinear systems. *J. Differential Equations*, 12:95–116, 1972.

[16] C. Lobry. Contrôlabilité des systèmes non linéaires. *SIAM J. on Control*, 8:573–605, 1970.

[17] J. R. Munkres. *Elements of Algebraic Topology*. Addison-Wesley, 1984.

[18] T. Nagano. Linear differential systems with singularities and an application to transitive lie algebras. *J. Math. Soc. Jap.*, 18:398–404, 1966.

[19] H. Nijmeijer and A. J. van der Schaft. *Nonlinear Dynamical Control Systems.* Springer-Verlag, New York, 1990.

[20] R. Roussarie. *Modèles locaux de champs et de formes*, volume 30 of *Astérisque*. Soc. Math. de France, 1975.

[21] W. Rudin. *Analyse réelle et complexe.* Masson, Paris, 1975.

[22] J. Sjöberg and Q. Zhang and L. Ljung and A. Benveniste and B. Delyon and P.Y. Glorennec and H. Hjalmarsson and A. Juditsky. Nonlinear black-box modeling in system identification: a unified overview. *Automatica*, 31(12):1691–1724, 1995.

[23] E. D. Sontag. *Mathematical Control Theory.* Springer-Verlag, New York, second edition, 1998. Deterministic finite-dimensional systems.

[24] M. Spivak. *A Comprehensive Introduction to Differential Geometry*, volume 1. Publish or Perish, Houston (U.S.A.), second edition, 1979.

[25] H. J. Sussmann. Orbits of families of vector fields and integrability of distributions. *Transactions of the AMS*, 317:171–188, 1973.

[26] K. Tchoń. The only stable normal forms of affine systems under feedback are linear. *Syst. & Control Lett.*, 8:359–365, 1987.

[27] A. J. van der Schaft. Linearization and input-output decoupling for general nonlinear systems. *Syst. & Control Lett.*, 5:27–33, 1984.

[28] S. van Strien. smooth linearization of hyperbolic fixed points without resonance conditions. *J. Differential Equations*, 85(1):66–90, 1990.

[29] J. C. Willems. Topological classification and structural stability of linear systems. *J. Differential Equations*, 35:306–318, 1980.

INRIA, B.P. 93, 06902 Sophia Antipolis cedex, France
E-mail address: Laurent.Baratchart@sophia.inria.fr

INRIA, B.P. 93, 06902 Sophia Antipolis cedex, France
E-mail address: Jean-Baptiste.Pomet@sophia.inria.fr