Revisiting the Normal Form of Input-Output Linearization

Manfredi Maggiore, *Senior member, IEEE*

***Abstract*— This paper revisits the normal form arising in the context of input-output feedback linearization for nonlinear control systems possessing well-defined relative degree. The objective is to investigate the validity of the normal form in a neighbourhood of the zero dynamics manifold, as opposed to a neighbourhood of a point on the manifold. The two main results of the paper are necessary and sufficient conditions under which the normal form exists in some neighbourhood of the zero dynamics manifold, or in a given a priori neighbourhood of the manifold in question. A special case is the existence of a global normal form. These results naturally lead to conditions for either local or regional (global, as a special case) asymptotic stabilization of the zero dynamics manifold. To illustrate these contributions, a normal form is derived for the kinematic unicycle model leading to a novel global circular path following controller.**

***Index Terms*— feedback linearization, normal form, nonlinear control, set stabilization, zero dynamics manifold**

I. Introduction

The notion of (vector) relative degree of a nonlinear control system with outputs is one of the fundamental concepts in nonlinear control theory. This notion originated in [11] in the context of invertibility of nonlinear control systems, i.e., the ability to find input signals generating desired output signals, and in that context it was referred to as *relative order*. It is, however, with the work of Byrnes and Isidori that the notion of relative degree came to the fore, and its deep relationships with other fundamental concepts of nonlinear control was established. In [2], the authors defined relative degree with an eye to generalizing the notion of frequency domain zeros of LTI systems to the nonlinear setting, and using this generalization for the stabilization of equilibria for so-called minimum-phase systems. Shortly thereafter, with work in [3], [5], it became clear that the notion of relative degree is intimately connected with the existence of the zero dynamics manifold (ZDM), the maximal controlled invariant subset of the set where the output function is zero. This subject is reviewed in Section III.

A well-defined relative degree is not required in order for the ZDM to be well-defined, and in [15] the authors showed that under certain regularity conditions weaker than relative degree, one can assert the existence of the ZDM in a neighbourhood of a point. Additionally, [15] developed a constructive procedure to find the ZDM reminiscent of Silverman's structure algorithm in [29] and its nonlinear generalization in [11] (see also [30], [6]). This procedure came to be known as the *zero dynamics algorithm* in later work ([4], [5]).

While relative degree is not a required property for the ZDM to exist, there are good reasons for investigating systems with well-defined relative degree. For one, such systems afford a direct characterization of the ZDM as the zero level set of the output function and a number of its Lie derivatives along the drift vector field. Moreover, these systems are feedback equivalent to a cascade connection of a reduced-order nonlinear control system driven by parallel chains of integrators. In the square multi-input multi-output case, this so-called *normal form of input-output linearization* has the structure

$$\begin{aligned}
\dot{z} &= \alpha(z,\xi) + \sum_{i=1}^{m} \beta^i(z,\xi) v_i \\
\dot{\xi}^i_j &= \xi^i_{j+1},\ i \in 1{:}m,\ j \in 1{:}r_i - 1 \\
\dot{\xi}^i_{r_i} &= v_i,\ i \in 1{:}m \\
y_i &= \xi^i_1,\ i \in 1{:}m,
\end{aligned} \tag{1}$$

where the ξ^i_j states constitute m decoupled chains of integrators with inputs v_i resulting from a feedback transformation reviewed in the following, and outputs y_i. In (z,ξ) coordinates, the ZDM is the set $\{(z,\xi) : \xi = 0\}$, and thus z represents the component of the state on the ZDM. The dynamics on the ZDM are represented by the subsystem $\dot{z} = \alpha(z,0)$.

The above normal form, originating in [14], [3], [5], has brought about much insight into the equilibrium stabilization problem. The interested reader may consult the survey paper by Isidori ([13]) for an account of the history of the zero dynamics and research perspectives.

While historically the notion of relative degree and the associated normal form have been developed with the intention of stabilizing equilibria, these tools are in fact important in the broader context of *set* stabilization. Indeed, a wealth of control specifications ultimately involve the asymptotic stabilization of the ZDM itself, not necessarily an equilibrium on it. In this context, the availability of a normal form defined at least *in a neighbourhood of the manifold* is of central interest.

To illustrate this point, suppose the control specification is to design a feedback controller that asymptotically sends a certain function $h(x)$ to zero, where the zero level set $h^{-1}(0)$ has some physical significance (e.g., a path to follow, a constraint to enforce, and so on). It is a well-known consequence of

This research was funded by the Natural Sciences and Engineering Research Council of Canada (NSERC).

M. Maggiore is with the Department of Electrical and Computer Engineering, University of Toronto, 10 King's College Road, Toronto, ON, M5S3G4, Canada (email: maggiore@ece.utoronto.ca).

the Birkhoff theorem on limit sets that any bounded state trajectory $x(t)$ resulting from a controller making $h(x(t))$ converge to zero will, by necessity, converge to a controlled invariant subset of $h^{-1}(0)$. Since the ZDM is the maximal such subset, $x(t)$ will necessarily converge to the ZDM. The punch-line is that if one wants to asymptotically zero out the output of a control system, one should, at a minimum, render the ZDM attractive. From a practical viewpoint, attractivity without stability is an undesirable property, so it is natural to design controllers that *asymptotically stabilize* the ZDM.

The problem of stabilizing the ZDM is central in the literature on bipedal locomotion (e.g., [33], [32]), where a stable walking gait is induced by asymptotically stabilizing the ZDM associated with a suitably defined output function representing a virtual constraint.

Just as the asymptotic stabilization of an equilibrium for a system with relative degree often involves the normal form for input-output linearization in a neighbourhood of that equilibrium, if one wants to investigate the asymptotic stabilization of the ZDM, one would benefit from a normal form valid not just near a point, but at a minimum in a neighbourhood of the ZDM. The literature, however, has only focused on the two extreme cases of normal forms defined in a neighbourhood of a point and one defined globally. This paper fills the gap, as detailed below.

In some cases, the control specification leads directly to a target set to be stabilized that is controlled invariant. The question in this setting is whether such target set could be viewed as the ZDM associated with a certain choice of output function yielding a well-defined relative degree. If this were the case, then the target set could be stabilized using input-output linearization. This question leads to the problem of *transverse feedback linearization*, investigated in [1], [25], [26], and more recently in [7] using tools of exterior differential geometry.

While the focus of this paper is on normal forms for input-output feedback linearization, there are other normal forms that play an important role in nonlinear control theory. For example, in [17], [16], [31] the authors investigate normal forms that are made up of a linear part plus nonlinear terms that are homogeneous of a certain degree. In [21] Menini and Tornambè use the Poincaré-Dulac normal form to determine when there exists an immersion mapping a nonlinear vector field to a linear one.

Contributions of this paper. For square MIMO control-affine systems with well-defined vector relative degree, this paper investigates conditions under which the normal form for input-output linearization is valid on an open subset of the state space containing the ZDM. We present two main results. In Theorem 3, we show that if relative degree is well-defined on the ZDM, there *always* exists a normal form defined in *some* neighbourhood of the ZDM. Here, we expose the role of smooth retractions in determining the coordinate transformation. The second result, in Theorem 6, gives necessary and sufficient conditions for the existence of a *regional* normal form, one valid in a *given a priori* neighbourhood of the zero dynamics manifold. A special case of the theorem gives necessary and sufficient conditions for the existence of a global normal form which slightly generalize those by Byrnes and Isidori in [5] which are only sufficient. Leveraging the results summarized above, we investigate in Propositions 12 and 13 the asymptotic stabilization of the ZDM in the case when this latter is a compact set (we discuss the non-compact case in a remark), giving conditions under which a normal form-based controller either locally asymptotically stabilizes the ZDM, or it does so with a guaranteed basin of attraction.

Finally, we illustrate the construction of Theorem 6 with the derivation of an almost global normal form in the context of the circular path following problem for a kinematic unicycle. A global normal form does not exist for this problem. Using this normal form, we derive a smooth almost global path following controller that does not enforce a direction of traversal of the path. With a hybrid supervisor, this controller is then turned into a global path following controller enforcing a desired direction of traversal.

II. Notation and preliminaries

If $k, l \in \mathbb{N}$ with $k \leq l$, then we let $k\!:\!l$ denote the index set $\{k, k+1, \ldots, l\}$. Given a vector $x \in \mathbb{R}^n$, x_i denotes the i-th element of x. For $v, w \in \mathbb{R}^n$, we denote $\|v\| := (\sum_i v_i^2)^{1/2}$ and $\langle v, w \rangle := v^\top w$. For $v \in \mathbb{R}^k$, $w \in \mathbb{R}^l$, $\operatorname{col}(v, w) \in \mathbb{R}^{k+l}$ is the concatenation of v and w. We use the abbreviations $c_\theta := \cos(\theta)$, $s_\theta := \sin(\theta)$.

We denote by $\mathbb{S}^1$ the set of real numbers modulo 2π, diffeomorphic to the unit circle. If $v \in \mathbb{R}^2$ is nonzero, $\measuredangle v \in \mathbb{S}^1$ denotes the angle of v, i.e., the unique element θ of $\mathbb{S}^1$ such that $v = \|v\| \operatorname{col}(c_\theta, s_\theta)$.

If $\mathcal{X}$ is a smooth manifold and $p \in \mathcal{X}$, $T_p\mathcal{X}$ is the tangent space to $\mathcal{X}$ at p, and $T\mathcal{X} := \{v_p \in T_p\mathcal{X} : p \in \mathcal{X}\}$ is the tangent bundle of $\mathcal{X}$. If $(x_1, \ldots, x_n)$ are local coordinates in a coordinate chart of $\mathcal{X}$, then we denote by $\{\frac{\partial}{\partial x_i}, i \in 1\!:\!n\}$, the coordinate basis for $T_p\mathcal{X}$ at each p in the chart domain. We denote by $\mathfrak{X}^\infty(\mathcal{X})$ the set of smooth vector fields on $\mathcal{X}$, i.e., functions $f : \mathcal{X} \to T\mathcal{X}$ such that $f(p) \in T_p\mathcal{X}$ for each $p \in \mathcal{X}$.

If $F : \mathcal{X} \to \mathcal{Y}$ is a smooth map of manifolds and $p \in \mathcal{X}$, then $dF_p : T_p\mathcal{X} \to T_{F(p)}\mathcal{Y}$ is the differential map at p. In local coordinates, this is the linear function whose matrix representation is the Jacobian matrix of the local coordinate representation of F. Sometimes we will find it useful to use parenthesis around the map to be differentiated, such as $d(\mathsf{H}_{\mathbf{r}})_x$ in what follows.

If $F : \mathcal{X} \to \mathcal{Y}$ is a diffeomorphism of manifolds and $f \in \mathfrak{X}^\infty(\mathcal{X})$, $F_* f$ denotes the pushforward of f, i.e, the unique $g \in \mathfrak{X}^\infty(\mathcal{Y})$ such that $g(F(p)) = dF_p(f(p))$ for all $p \in \mathcal{X}$. We denote by $\Phi_t^f(x_0)$ the flow of $f \in \mathfrak{X}^\infty(\mathcal{X})$, i.e., the solution at time t with initial condition $x(0) = x_0$ of the ODE on $\mathcal{X}$, $\dot{x} = f(x)$. If $f, g \in \mathfrak{X}^\infty(\mathcal{X})$ and $k \in \mathbb{N}$, $ad_f^k g \in \mathfrak{X}^\infty(\mathcal{X})$ denotes the k-th iterated Lie bracket of f and g, defined recursively as $ad_f^0 g := g$ and $ad_f^k g = [f, ad_f^{k-1} g]$. Moreover, if $h : \mathcal{X} \to \mathbb{R}$ is a smooth function, $L_f^k h$ denotes the k-th iterated Lie derivative of h along f, defined recursively as $L_f^0 h := h$ and $L_f^k h = L_f(L_f^{k-1} h)$.

A distribution Δ on $\mathcal{X}$ is the assignment to each point $x \in \mathcal{X}$ of a subspace $\Delta(x) \subset T_x\mathcal{X}$. A distribution is smooth if it is locally spanned by a finite number of smooth vector fields on $\mathcal{X}$, by which is meant the following: in a neighbourhood U of every point of $\mathcal{X}$, there exist an integer k and $f_i \in \mathfrak{X}^\infty(\mathcal{X})$, $i \in 1:k$, such that for each $x \in U$, $\Delta(x) = \operatorname{span}\{f_1(x), \dots, f_k(x)\}$. We denote by $\operatorname{span}\{f_1, \dots, f_k\}$ the distribution spanned by $\{f_1, \dots, f_k\}$ in the manner just described. We denote by $C^\infty(\mathcal{X})$ the ring of smooth real-valued functions on $\mathcal{X}$, and say that the vector fields $\{f_1, \dots, f_k\}$ are $C^\infty(\mathcal{X})$-linearly independent if for any $\alpha_i \in C^\infty(\mathcal{X})$, $i \in 1{:}k$, $\sum_i \alpha_i f_i = 0$ implies $\alpha_i = 0$.

The rank of a distribution Δ at $x \in \mathcal{X}$ is $\dim \Delta(x)$, and Δ is nonsingular if it has constant rank on $\mathcal{X}$. If $f \in \mathfrak{X}^\infty(\mathcal{X})$ and Δ is smooth distribution, we write $f \in \Delta$ to mean that for each $p \in \mathcal{X}$, $f(p) \in \Delta(p)$. A smooth distribution Δ on $\mathcal{X}$ is involutive if for each $\delta_1, \delta_2 \in \Delta$, $[\delta_1, \delta_2] \in \Delta$. Nonsingular and involutive distributions admit maximal integral manifolds by the Frobenius theorem (see, e.g., [19]).

A smooth distribution Δ on $\mathcal{X}$ is invariant under $f \in \mathfrak{X}^\infty(\mathcal{X})$ if for each $\delta \in \Delta$, $[f, \delta] \in \Delta$. We use the shorthand $[f, \Delta] \subset \Delta$ for this property.

III. Review of relative degree and zero dynamics

In this section we review the concept of relative degree for nonlinear control-affine systems as developed in [12], [5] (see also [27]). Let $\mathcal{X}$ be a C^∞ manifold and consider the square multi-input multi-output (MIMO) control-affine system on $\mathcal{X}$

$$\begin{aligned} \dot{x} &= f(x) + \sum_{i=1}^{m} g^i(x) u_i \\ y_i &= h^i(x), \ i \in 1{:}m, \end{aligned} \tag{2}$$

where $f, g^1, \dots g^m \in \mathfrak{X}^\infty(\mathcal{X})$ and $h^i \in C^\infty(\mathcal{X})$. We recall from [5], [12] that system (2) has **vector relative degree** $\{r_1, \dots, r_m\}$ at $x_0 \in \mathcal{X}$ if there exists a neighbourhood of $U \subset \mathcal{X}$ of x_0 such that for each $x \in U$ and $i, j \in 1:m$, $L_{g^i} L_f^k h^j(x) = 0$ for all $k \in 0 : r_j - 2$ and the **decoupling matrix** $A(x_0) \in \mathbb{R}^{m \times m}$ defined as $[A(x_0)]_{ij} := L_{g^j} L_f^{r_i - 1} h^i(x_0)$ is nonsingular. We denote $\mathbf{r} := \{r_1, \dots, r_m\}$ and $r := r_1 + \dots + r_m$, and define a map $\mathsf{H}_\mathbf{r} : \mathcal{X} \to \mathbb{R}^r$ as

$$\mathsf{H}_\mathbf{r} = \operatorname{col}(h^1, \dots, L_f^{r_1-1} h^1, \dots, h^m, \dots, L_f^{r_m - 1} h^m).$$

Since the matrix-valued function $A : \mathcal{X} \to \mathbb{R}^{m \times m}$ is smooth and nonsingular at x_0, by possibly making U smaller one may assume without loss of generality that $A(x)$ is nonsingular for all $x \in U$.

In what follows, we associate with a vector relative degree $\mathbf{r}$ the set of index pairs

$$\mathcal{I}_\mathbf{r} := \{(i,k) : i \in 1{:}m, k \in 1{:}r_i\}.$$

We also use two different notational systems to refer to components of the vector $\xi = \mathsf{H}_\mathbf{r}(x)$. Define the bijection $\mathsf{idx} : \mathcal{I}_\mathbf{r} \to 1 : r$, $\mathsf{idx}(i,k) := \sum_{k=1}^{i-1} r_k + k$. Then, the j-th component of $\xi \in \mathbb{R}^r$ will be denoted either by ξ_j or by ξ_k^i, with $(k,i) = \mathsf{idx}^{-1}(j)$. This way, if $\xi = \mathsf{H}_\mathbf{r}(x)$, then $\xi_k^i = L_f^{k-1} h^i(x)$. We denote by e_k^i the vector whose element in position $\mathsf{idx}(i,k)$ is one and all other elements are zero.

The basic geometric properties of control systems possessing a well-defined relative degree are summarized in the next result.

Theorem 1 ([5], [12]): Suppose there exists an open set $U \subset \mathcal{X}$ and integers $r_i \in 1:n$, $i \in 1:m$, such that for each $x \in U$ and all $i, j \in 1:m$, $L_{g^i} L_f^k h^j(x) = 0$, $k \in 0 : r_j - 2$, and $A(x)$ is nonsingular for all $x \in U$. Then, letting $\mathbf{r} := \{r_1, \dots, r_m\}$ and $r := \sum_i r_i$, we have

(i) The smooth map $\mathsf{H}_\mathbf{r} : U \to \mathbb{R}^r$ is a submersion, i.e., $\operatorname{rank} d(\mathsf{H}_\mathbf{r})_x = r$ for all $x \in U$.

(ii) The distribution $\Delta(x) = \operatorname{Ker}(d(\mathsf{H}_\mathbf{r})_x)$ defined on U is smooth, has constant rank $n - r$, and is involutive.

(iii) The vector fields $\{ad_f^{k-1} g^i : (i,k) \in \mathcal{I}_\mathbf{r}\}$ are $C^\infty(U)$-linearly independent. Moreover, letting G denote the smooth nonsingular distribution on U defined as

$$G(x) := \operatorname{span}\{ad_f^{k-1} g^i : (i,k) \in \mathcal{I}_\mathbf{r}\},$$

we have that

$$(\forall x \in U) \quad \Delta(x) \oplus G(x) = T_x \mathcal{X}.$$

(iv) Letting

$$\begin{aligned} \tilde{f} &:= f - \left[g^1 \cdots g^m\right] A^{-1} \operatorname{col}(L_f^{r_1} h^1, \dots, L_f^{r_m} h^m) \\ \tilde{g}^j &= \sum_i g^i \left[A^{-1}\right]_{ij}. \end{aligned} \tag{3}$$

the following properties hold on U:

$$\begin{aligned} &[\tilde{f}, \Delta] \subset \Delta \\ &[\tilde{g}^j, \Delta] \subset \Delta, \ j \in 1:m. \end{aligned}$$

(v) For each $x_0 \in U$, there exist a neighbourhood $V \subset U$ of x_0 and a smooth function $\mathsf{p} : V \to \mathbb{R}^{n-r}$ such that the map $T : V \to \mathbb{R}^{n-r} \times \mathbb{R}^r$, $x \mapsto (z, \xi) = (\mathsf{p}(x), \mathsf{H}_\mathbf{r}(x))$ is a diffeomorphism onto its image. Moreover, the feedback transformation

$$u = A^{-1}(x) \left(-\operatorname{col}(L_f^{r_1} h^1, \dots, L_f^{r_m} h^m) + v \right), \tag{4}$$

and the coordinate transformation $(z, \xi) = T(x)$ give the **local normal form** (1).

From an input-output viewpoint, part (v) expresses the fact that the feedback transformation (4) gives $d^{r_i} y_i / dt^{r_i} = v_i(t)$, for any output signal $y_i(t) = h^i(x(t))$, where $x(t)$ is a state trajectory contained in V. If $r = n$, the z subsystem is absent and therefore system (2) is locally feedback equivalent to m decoupled chains of integrators.

The combination of a diffeomorphism and a feedback transformation such as (4) is called a **feedback equivalence**.

If system (2) has vector relative degree $\mathbf{r}$ at each $x_0 \in \mathcal{X}$, the system is said to possess a **uniform vector relative degree** $\mathbf{r}$. In this case, the properties (i)-(iv) in Theorem 1 hold with $U = \mathcal{X}$, while property (v) is globalized in the following result by Byrnes and Isidori in [5].

Theorem 2 ([5]): Suppose (2) has uniform vector relative degree $\mathbf{r}$, let $\mathcal{Z} = (\mathsf{H}_\mathbf{r})^{-1}(0)$, and $\tilde{f}, \tilde{g}^i \in \mathfrak{X}^\infty(\mathcal{X})$, $i \in 1{:}m$, be given as in (3). If the vector fields $ad_{\tilde{f}}^{k-1} \tilde{g}^i$, $(i,k) \in \mathcal{I}_\mathbf{r}$, are complete[1], then there exists a smooth function $\mathsf{p} : \mathcal{X} \to \mathcal{Z}$

[1] This means that for each initial condition $x_0 \in \mathcal{X}$, the integral curve of $ad_{\tilde{f}}^{k-1} \tilde{g}^i$ through x_0 is defined for all $t \in \mathbb{R}$.

such that the map $T : \mathcal{X} \to \mathcal{Z} \times \mathbb{R}^r$ given by $x \mapsto (z, \xi) = (\mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x))$ is a diffeomorphism and in (z, ξ) coordinates after the feedback transformation (4), system (2) takes on the normal form (1) with state $(z, \xi) \in \mathcal{Z} \times \mathbb{R}^r$.

When the nonlinear control system (2) is globally feedback equivalent to the normal form (1) with state $(z, \xi) \in \mathcal{Z} \times \mathbb{R}^r$, we say that system (2) admits a **global normal form**.

For a system with uniform vector relative degree, the set

$$\mathcal{Z} = \{x \in \mathcal{X} : L_f^{k-1} h^i(x) = 0,\ (i, k) \in \mathcal{I}_{\mathbf{r}}\}$$

is a closed embedded submanifold of $\mathcal{X}$ of codimension r, and it is called the **zero dynamics manifold (ZDM)** of (2). It is shown in [5] that $\mathcal{Z}$ is the maximal controlled invariant subset of the zero level set of the output, $h^{-1}(0)$, and there is a unique smooth feedback $u^\star : \mathcal{Z} \to \mathbb{R}$ rendering $\mathcal{Z}$ invariant: $u^\star = -A^{-1} \operatorname{col}(L_f^{r_1} h^1, \ldots, L_f^{r_m} h^m)\big|_{\mathcal{Z}}$. The vector field $f^\star \in \mathfrak{X}^\infty(\mathcal{Z})$ defined as $f^\star := (f + g u^\star)\big|_{\mathcal{Z}}$ is called the **zero dynamics vector field**.

IV. Normal form in a neighbourhood of the ZDM

We have reviewed the fact, proved in [5], that when system (2) has a uniform vector relative degree $\mathbf{r}$, the ZDM $\mathcal{Z} = (\mathsf{H}_{\mathbf{r}})^{-1}(0)$ is the maximal controlled invariant subset of $h^{-1}(0)$. For this property to hold, uniform vector relative degree is not required, and the following assumption suffices[2].

Assumption 1: System (2) has vector relative degree $\mathbf{r}$ at each $x_0 \in (\mathsf{H}_{\mathbf{r}})^{-1}(0)$. △

It is natural to ask whether under Assumption 1 one can assert the existence of the normal form (1) valid *in a neighbourhood of $\mathcal{Z}$*. In this section we show that the answer is *yes, always*.

The literature has focused on the two extreme cases reviewed in Section III: the existence of a *local* normal form valid in a neighbourhood of a point $x_0 \in (\mathsf{H}_{\mathbf{r}})^{-1}(0)$, and that of a *global* normal form valid everywhere. The intermediate case of a normal form valid in a neighbourhood of $\mathcal{Z}$ is quite interesting in its own right, as such a normal form plays a role in the local asymptotic stabilization of $\mathcal{Z}$ (see Proposition 12 below).

To construct the normal form, we need the notion of smooth retraction. Let $U \subset \mathcal{X}$ be an open set and $\mathcal{Z} \subset U$ be a closed embedded submanifold of $\mathcal{X}$. A **smooth retraction of U onto $\mathcal{Z}$** is a smooth map $\mathsf{p} : U \to \mathcal{Z}$ satisfying $\mathsf{p}(z) = z$ for all $z \in \mathcal{Z}$. It is a consequence of the tubular neighbourhood theorem (see [10]) that every embedded submanifold $\mathcal{Z} \subset \mathcal{X}$ admits a retraction $\mathsf{p} : U \to \mathcal{Z}$ of a neighbourhood U of $\mathcal{Z}$ onto $\mathcal{Z}$.

Now the main result of this section.

Theorem 3 (Normal form in a neighbourhood of $\mathcal{Z}$): Suppose system (2) satisfies Assumption 1, and let $\mathsf{p} : U \to \mathcal{Z}$ be a smooth retraction of a neighbourhood $U \subset \mathcal{X}$ of $\mathcal{Z}$ onto $\mathcal{Z}$. Let $\sigma : \mathcal{Z} \to \tilde{\mathcal{Z}}$ be a diffeomorphism (one may let $\tilde{\mathcal{Z}} = \mathcal{Z}$ and σ be the identity map $\mathrm{id}_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{Z}$). Then there exists a neighbourhood $V \subset U$ of $\mathcal{Z}$ such that the map

$$T : V \to \tilde{\mathcal{Z}} \times \mathbb{R}^r,\ x \mapsto (z, \xi) = (\sigma \circ \mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x))$$

is a diffeomorphism onto its image and $T(V)$ has the form $T(V) = \tilde{\mathcal{Z}} \times W$, where $W \subset \mathbb{R}^r$ is a neighbourhood of the origin. Moreover, in (z, ξ) coordinates, after the feedback transformation (4), system (2) takes on the normal form (1) with state $(z, \xi) \in \tilde{\mathcal{Z}} \times W$.

The proof of this result is found in Appendix I.

Remark 4: The map σ in the theorem statement allows one to replace $\mathcal{Z}$ by any manifold $\tilde{\mathcal{Z}}$ diffeomorphic to it. To illustrate, when $\mathcal{Z}$ is diffeomorphic to a generalized cylinder $\tilde{\mathcal{Z}} = \mathbb{S}^1 \times \cdots \mathbb{S}^1 \times \mathbb{R} \cdots \times \mathbb{R}$ (this is *always* the case when $\dim \mathcal{Z} = 1$), one gets a global parametrization of $\mathcal{Z}$ by means of the state $z = (\theta_1, \ldots, \theta_s, t_{s+1}, \ldots, t_{n-r}) \in \tilde{\mathcal{Z}}$, with $\theta_i \in \mathbb{S}^1$ and $t_j \in \mathbb{R}$. This parametrization allows one to write the dynamics of the z-subsystem in (1) in a set of global coordinates. For another illustration, if $\mathcal{Z}$ is diffeomorphic to a matrix group, then we can use σ to represent the state z as a matrix in the group. This flexibility will be further illustrated in the unicycle example of Section VII. △

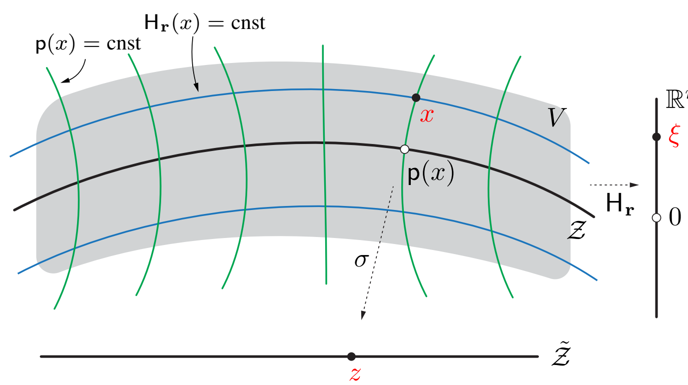


Fig. 1. The coordinate transformation $T : x \mapsto (z, \xi)$ on a neighbourhood V of $\mathcal{Z}$.

Remark 5: The coordinate transformation of Theorem 3 can be understood intuitively as follows (see Figure 1). In a sufficiently small neighbourhood V of $\mathcal{Z}$, we may represent $x \in V$ via a pair (z, ξ). The vector $\xi = \mathsf{H}_{\mathbf{r}}(x) \in \mathbb{R}^r$ determines which level set of $\mathsf{H}_{\mathbf{r}}$ the point x is on, while $z = \sigma \circ \mathsf{p}(x) \in \tilde{\mathcal{Z}}$ indicates where x is situated on the level set of the function $\mathsf{H}_{\mathbf{r}}$. Indeed, the level sets of the function $\sigma \circ \mathsf{p} : V \to \tilde{\mathcal{Z}}$ are embedded submanifolds of $\mathcal{X}$ of dimension r, while the level sets of $\mathsf{H}_{\mathbf{r}} : V \to \mathbb{R}^r$ are embedded submanifolds of $\mathcal{X}$ of complementary dimension $n - r$. The fact that $T : V \to T(V)$ is a diffeomorphism means geometrically that each level set of $\mathsf{H}_{\mathbf{r}} : V \to \mathbb{R}^r$ and each level set of $r : V \to \tilde{\mathcal{Z}}$ intersect at a unique point. Thus, knowing the pair (z, ξ) we can uniquely reconstruct x, and do so smoothly.

One can think of the retraction $\mathsf{p} : V \to \mathcal{Z}$ as a nonlinear projection of V onto $\mathcal{Z}$. In fact, if $\mathcal{X}$ is a Riemannian manifold, one can always define p to be the orthogonal projection onto $\mathcal{Z}$, and the tubular neighbourhood theorem guarantees that if V is a sufficiently small neighbourhood of $\mathcal{Z}$ such orthogonal projection is well-defined and smooth. In some cases, however, the orthogonal projection may not be the most convenient choice of retraction. We also remark that p can *always* be chosen to take the form (8) reviewed in the next section. △

[2] Recall from the introduction that a well-defined vector relative degree is not required for the ZDM to exist and have the mentioned properties.

V. Regional normal form

In this section we investigate necessary and sufficient conditions for the existence of a *regional* normal form, one valid in a given a priori neighbourhood of the ZDM. A special case is when the neighbourhood is the entire state space $\mathcal{X}$, in which case the problem has been studied by Byrnes and Isidori in [5] and its solution was reviewed in Theorem 2. We thus begin our treatment by reviewing the ideas of [5]. We will build upon them to get the main result.

We assume that system (2) has uniform vector relative degree $\mathbf{r}$ and the vector fields $ad_{\tilde{f}}^{k-1}\tilde{g}^i$, $(i,k)\in\mathfrak{I}_{\mathbf{r}}$, are complete.

Let $\lambda_l^j:\mathcal{X}\to\mathbb{R}$ be the differentiated output

$$\lambda_l^j := L_f^{l-1}h^j,\ (j,l)\in\mathfrak{I}_{\mathbf{r}}. \tag{5}$$

Byrnes and Isidori in [5] show that there exist vector fields $\tau_k^i\in\mathfrak{X}^\infty(\mathcal{X})$, $(i,k)\in\mathfrak{I}_{\mathbf{r}}$, such that

$$L_{\tau_k^i}\lambda_l^j=\begin{cases}1 & i=j,\ k+l=r_i+1\\ 0 & \text{otherwise.}\end{cases} \tag{6}$$

In other words, the output of the dynamical system

$$\begin{aligned}\dot{x}&=\tau_k^i(x)\\ s&=\lambda_l^j(x),\end{aligned}$$

has the property that its time derivative along solutions satisfies $\dot{s}=1$ if $i=j$ and $k+l=r_i+1$, and $\dot{s}=0$ otherwise. In terms of the flow of the vector field τ_k^i, this property can be rewritten as

$$\lambda_l^j\circ\Phi_t^{\tau_k^i}(x)=\begin{cases}\lambda_l^j(x)+t & i=j,\ k+l=r_i+1\\ \lambda_l^j(x) & \text{otherwise,}\end{cases} \tag{7}$$

for all $x\in\mathcal{X}$ and all $t\in\mathbb{R}$ for which the flow is defined. In [5] it is shown that the choice

$$\tau_k^i=(-1)^{k-1}ad_{\tilde{f}}^{k-1}\,\tilde{g}^i,\ (i,k)\in\mathfrak{I}_{\mathbf{r}},$$

satisfies (6), and the assumption that $ad_{\tilde{f}}^{k-1}\tilde{g}^i$ is complete ensures that the identities in (7) hold for all $t\in\mathbb{R}$.

The authors in [5] define[3] a map $\mathsf{p}:\mathcal{X}\to\mathcal{Z}$ as

$$\begin{aligned}\mathsf{p}(x) :=& \Phi^{\tau_1^m}_{-\lambda^m_{r_m}(x)}\circ\cdots\circ\Phi^{\tau^m_{r_m}}_{-\lambda^m_1(x)}\circ\cdots\\ &\cdots\circ\Phi^{\tau_1^1}_{-\lambda^1_{r_1}(x)}\circ\cdots\circ\Phi^{\tau^1_{r_1}}_{-\lambda^1_1(x)}(x).\end{aligned} \tag{8}$$

The completeness assumption ensures that p is a well-defined smooth map on $\mathcal{X}$, and the identities (7) ensure that $h^1(\mathsf{p}(x))=0$. Indeed, recalling that $h^1(x)=\lambda_1^1(x)$ and using (7) recursively r times, we get

$$\begin{aligned}h^1(\mathsf{p}(x))&=\lambda_1^1\circ\Phi^{\tau_1^m}_{-\lambda^m_{r_m}(x)}\circ\cdots\circ\Phi^{\tau^1_{r_1}}_{-\lambda^1_1(x)}(x)\\ &=\lambda_1^1\circ\Phi^{\tau_2^m}_{-\lambda^m_{r_m-1}(x)}\circ\cdots\circ\Phi^{\tau^1_{r_1}}_{-\lambda^1_1(x)}\\ &\cdots\\ &=\lambda_1^1\circ\Phi^{\tau^1_{r_1}}_{-\lambda^1_1(x)}(x)\\ &=\lambda_1^1(x)-\lambda_1^1(x)=0.\end{aligned}$$

[3]The order of composition of flows in [5] is different than in (8) but the order is immaterial, as discussed below.

In a similar manner, one shows that $L_f^{k-1}h^i(\mathsf{p}(x))=0$, $(i,k)\in\mathfrak{I}_{\mathbf{r}}$, which implies that $\mathsf{p}(x)\in\mathcal{Z}$ for all $x\in\mathcal{X}$. Note moreover that if $x\in\mathcal{Z}$, then $\lambda_l^j(x)=0$, $(j,l)\in\mathfrak{I}_{\mathbf{r}}$, and therefore $\mathsf{p}(x)=x$, meaning that the restriction of p to $\mathcal{Z}$ is the identity map. Thus the function $\mathsf{p}(x)$ is a smooth retraction of $\mathcal{X}$ onto $\mathcal{Z}$.

One can see that the map $T:\mathcal{X}\to\mathcal{Z}\times\mathbb{R}^r$, $x\mapsto(\mathsf{p}(x),\mathsf{H}_{\mathbf{r}}(x))$ is a diffeomorphism with inverse

$$T^{-1}(z,\xi)=\Phi^{\tau^1_{r_1}}_{\xi^1_1}\circ\cdots\circ\Phi^{\tau^1_1}_{\xi^1_{r_1}}\circ\cdots\circ\Phi^{\tau^m_{r_m}}_{\xi^m_1}\circ\cdots\circ\Phi^{\tau^m_1}_{\xi^m_{r_m}}(z). \tag{9}$$

This is the main idea in the proof of Theorem 2.

The order of the composition of flows in the definition of p does not affect the analysis, even though the vector fields τ_k^i generally do not commute. A permutation of the vector fields in p will give a different retraction onto $\mathcal{Z}$, and therefore a different diffeomorphism T, but both of these are equally viable choices.

The idea presented above can be generalized in two directions. First, the vector fields $(-1)^{k-1}ad_{\tilde{f}}^{k-1}\tilde{g}^i$ are not the only ones yielding the identity (6). If we add to $\tilde{f}$ and $\tilde{g}^i$, $i\in 1\!:\!m$ any vector fields in the distribution Δ of Theorem 1, the identity (6) is still satisfied (this is proved in Lemma 18 in the appendix). As long as there exist complete vector fields τ_k^i satisfying (6), Byrnes and Isidori's proof goes through unchanged and gives a global normal form, even if the vector fields $ad_{\tilde{f}}^{k-1}\tilde{g}^i$ are not complete.

The second, more important, generalization concerns the domain of validity of the diffeomorphism T. If the control system has a well-defined vector relative degree in a certain neighbourhood of the ZDM, it is still possible to employ the retraction p in (8), or a variation of it, as long as the flows of the vector fields used to construct p are well defined. We next make this idea precise.

We begin with an assumption on the properties of the set $\mathcal{U}\subset\mathcal{X}$ over which relative degree is well-defined. First we need some notation. Recall the map $\mathsf{idx}:\mathfrak{I}_{\mathbf{r}}\to 1\!:\!r$ defined in Section III. Given a bijection $\pi:\mathfrak{I}_{\mathbf{r}}\to\mathfrak{I}_{\mathbf{r}}$ (called a **permutation** in what follows), for $j\in 1:r$ define indices $(i_j,k_j):=\pi\circ\mathsf{idx}^{-1}(j)$ and $l_j:=\mathsf{idx}(i_j,k_j)$. Then, π induces the isomorphism $\mathsf{P}_\pi:\mathbb{R}^r\to\mathbb{R}^r$, $\mathsf{P}_\pi(\xi):=\mathrm{col}(\xi_{l_1},\ldots,\xi_{l_r})$.

In what follows, we will use π and P_π to change the ordering of flows in the retraction (8).

Assumption 2: System (2) has vector relative degree $\mathbf{r}$ at each $x_0\in\mathcal{U}:=(\mathsf{H}_{\mathbf{r}})^{-1}(\mathcal{C})$, where $\mathcal{C}\subset\mathbb{R}^r$ is a convex open set containing the origin and enjoying the following property. There exists a permutation $\pi:\mathfrak{I}_{\mathbf{r}}\to\mathfrak{I}_{\mathbf{r}}$ such that for each $\xi\in\mathcal{C}$ and each $j\in 1\!:\!r-1$, $\mathrm{col}\big(0,\ldots,0,\xi_{l_{j+1}},\ldots,\xi_{l_r}\big)\in\mathsf{P}_\pi(\mathcal{C})$. △

The assumption is illustrated in Figure 2.

Theorem 6 (Normal form on $\mathcal{U}$): Suppose system (2) satisfies Assumption 2. Let $\tilde{f}$ and $\tilde{g}$ be the vector fields defined in (3) and $\sigma:\mathcal{Z}\to\tilde{\mathcal{Z}}$ be a diffeomorphism. If, and only if, there exist vector fields $\delta_1,\delta_2^i\in\Delta$, $i\in 1:m$, such that, letting $\hat{f}:=\tilde{f}+\delta_1$, $\hat{g}^i:=\tilde{g}^i+\delta_2^i$, the vector fields $\tau_k^i:=(-1)^{k-1}ad_{\hat{f}}^{k-1}\,\hat{g}^i$, $(i,k)\in\mathfrak{I}_{\mathbf{r}}$, enjoy the following properties for each $x\in\mathcal{U}$ and each $(i,k)\in\mathfrak{I}_{\mathbf{r}}$:

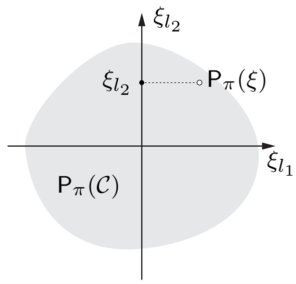


Fig. 2. Illustration of Assumption 2 in the case $m = 1, r = 2$.

(i) the flow $t \mapsto \Phi_t^{\tau_k^i}(x)$ is defined on $\{t \in \mathbb{R} : \mathsf{H}_\mathbf{r}(x) + te^i_{r_i-k+1} \in \mathcal{C}\}$, and
(ii) the vector fields $\{\tau_k^i\}_{(i,k)\in\mathcal{I}_\mathbf{r}}$ commute, i.e., $[\tau_k^i, \tau_l^j] = 0$ for all $(i,k),(j,l) \in \mathcal{I}_\mathbf{r}$,

then there exists a diffeomorphism $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$ giving the normal form (1) with state $(z,\xi) \in \tilde{\mathcal{Z}} \times \mathcal{C}$ after the feedback transformation (4). Moreover, if the diffeomorphism $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$ exists, it can be chosen as $T : x \mapsto (z,\xi) = (\sigma \circ \mathsf{p}(x), \mathsf{H}_\mathbf{r}(x))$, with $\mathsf{p} : \mathcal{X} \to \mathcal{Z}$ defined as

$$\mathsf{p}(x) := \Phi^{\bar{\tau}_r}_{-\bar{\lambda}_r(x)} \circ \cdots \circ \Phi^{\bar{\tau}_1}_{-\bar{\lambda}_1(x)}(x), \tag{10}$$

where

$$\begin{aligned} \bar{\lambda}_j &:= L_f^{k_j-1} h^{i_j}, \\ \bar{\tau}_j &:= \tau^{i_j}_{r_{i_j}-k_j+1} = (-1)^{r_{i_j}-k_j} ad_{\hat{f}}^{r_{i_j}-k_j} \hat{g}^{i_j}, \end{aligned} \tag{11}$$

with $j \in 1\!:\!r$.

The proof of the theorem is found in Appendix II.

Remark 7: The commutativity property (ii) in Theorem 6 is not needed for a global normal form to exist. In other words, the existence of $\delta_1, \delta_2^i \in \Delta$, $i \in 1\!:\!m$, inducing property (i) is sufficient for the existence of a global normal form (this will be evident in the proof, where commutativity is not used in the construction of the diffeomorphism). However, if there exist vector fields δ_1, δ_2^i inducing property (i), then these can always be chosen to induce property (ii). △

Remark 8: A special case of Theorem 6 is when the set $\mathcal{C}$ in Assumption 2 is the entire $\mathbb{R}^r$, in which case the diffeomorphism T is a map $\mathcal{X} \to \tilde{\mathcal{Z}} \times \mathbb{R}^r$ and we get a global normal form. The result in this case is a small generalization of Theorem 2 in that the conditions for existence of a global normal form are both necessary and sufficient. The necessary and sufficient conditions in this special case are a uniform vector relative degree and the existence of smooth vector fields δ_1, δ_2^i, $i \in 1\!:\!m$, such that the vector fields $ad_{\hat{f}}^{k-1}\hat{g}^i$, $(i,k) \in \mathcal{I}_\mathbf{r}$ are complete, where $\hat{f} = \tilde{f} + \delta_1$ and $\hat{g}^i = g^i + \delta_2^i$. Theorem 2 corresponds to the choice $\delta_1, \delta_2^i = 0$. The completeness requirement just mentioned is quite strong and restrictive, and this is the reflection of the fact that the existence of a global normal form is rather exceptional. This observation further highlights the importance of Theorem 6 in seeking *regional*, as opposed to *global*, normal forms. △

Remark 9: Recall from part (iii) of Theorem 1 that for each $x \in \mathcal{X}$, $\Delta(x) \oplus G(x) = T_x\mathcal{X}$. Property (ii) in Theorem 6 implies that there exists a nonsingular involutive distribution $G' : \mathcal{X} \to T\mathcal{X}$ such that $\Delta(x) \oplus G'(x) = T_x\mathcal{X}$. This distribution is spanned by the vector fields $ad_{\hat{f}}^{k-1}\hat{g}^i$, $(i,k) \in \mathcal{I}_\mathbf{r}$, and its maximal integral manifolds are the connected components of the level sets of the smooth retraction $\mathsf{p} : \mathcal{X} \to \mathcal{Z}$. This observation is linked to the discussion in Remark 5. △

Example 10: We illustrate the theorem with an elementary example. Later, in Section VII, we investigate a nontrivial example. The single integrator system with $\mathcal{X} = \mathbb{R}$,

$$\begin{aligned} \dot{x} &= u \\ y &= \arctan(x) \end{aligned}$$

has uniform relative degree 1 and trivial ZDM $\mathcal{Z} = \{0\}$. However, the control system does not admit a global normal form because the image of the output function is the open interval $(-\pi/2, \pi/2)$ and therefore the map T in Theorem 2 cannot be a diffeomorphism onto $\{0\} \times \mathbb{R}$. This observation is confirmed by the fact that the assumption of Theorem 2 does not hold. Indeed, the feedback transformation $u = (1+x^2)v$ gives $\dot{y} = v$ and $\tilde{f}(x) = 0$, $\tilde{g}(x) = (1+x^2)\frac{\partial}{\partial x}$. The flow of $\tilde{g}$ is

$$\Phi_t^{\tilde{g}}(x) = \tan(\arctan(x) + t),$$

and we see that there are finite escape times when $t = -\arctan(x) \pm \pi/2$ so this vector field is not complete, and Theorem 2 is not applicable.

There does however exist a *regional* normal form. For this, we use Theorem 6 with $\mathcal{U} = \mathbb{R}$ and $\mathcal{C} = h(\mathcal{U}) = (-\pi/2, \pi/2)$. The set $\mathcal{C}$ satisfies Assumption 2 with π the trivial permutation. We check the necessary and sufficient condition of Theorem 6. In light of Remark 7, we only need to check condition (i) in the theorem statement.

For each $x \in \mathbb{R}$, the flow of $\tau_1^1 = \tilde{g}$ is defined on $\{t \in \mathbb{R} : \arctan(x) + t \in (-\pi/2, \pi/2)\}$. Is it true that for all t in this interval, $\mathsf{H}_\mathbf{r}(x) + t$ is contained in $\mathcal{C}$? Yes, because if $t \in (-\arctan(x) - \pi/2, -\arctan(x) + \pi/2)$ then $\mathsf{H}_\mathbf{r}(x) + t \in (-\pi/2, \pi/2) = \mathcal{C}$.

We construct the diffeomorphism T. First, the map $\mathsf{p} : \mathbb{R} \to \mathcal{Z}$ is

$$\begin{aligned} \mathsf{p}(x) &= \Phi^{\tilde{g}}_{-h(x)}(x) = \tan(\arctan(x) - \arctan(x)) \\ &= \tan(0) = 0. \end{aligned}$$

This is natural, as $\mathcal{Z}$ is the point $\{0\}$ so p above is the only possible retraction $\mathcal{U} \to \mathcal{Z}$. Now Theorem 6 states that the map $T : \mathbb{R} \to \{0\} \times (-\pi/2, \pi/2)$ given by $T(x) := (0, \arctan(x))$ is a diffeomorphism and it gives the normal form

$$\dot{\xi} = v, \quad y = \xi,$$

with state space $\mathcal{C} = (-\pi/2, \pi/2)$. △

Remark 11: We conclude this section by pointing out that the existence of a global normal form imposes strong topological and geometric properties on $\mathcal{Z}$ and $\mathcal{X}$. Specifically, it is easy to show that the following are necessary conditions for the existence of a global normal form:

(a) $\mathcal{X}$ and $\mathcal{Z}$ are homotopy equivalent.
(b) $\mathcal{X}$ can be given the structure of a trivial smooth vector bundle over $\mathcal{Z}$.

In the (z,ξ) coordinates of the normal form, the homotopy equivalence in part (a) is established by the strong deformation retract $H : \mathcal{X} \times [0,1] \to \mathcal{X}$ defined as

$$H(x,\lambda) = T^{-1}\left(\mathrm{pr}_z(T(x)), (1-\lambda)\,\mathrm{pr}_\xi(T(x))\right),$$

where pr_z and pr_z denote the projections $(z,\xi) \mapsto z$ and $(z,\xi) \mapsto \xi$. One can see that H is smooth, that for every $x \in \mathcal{X}$, $H(x,0) = x$ and $H(x,1) \in \mathcal{Z}$. Moreover, $H(x,\lambda) = x$ for all $x \in \mathcal{Z}$ and all $\lambda \in [0,1]$. We will use this remark to rule out the existence of a normal form in the unicycle example of Section VII. △

VI. Stabilization of the ZDM

In this section we investigate the problem of asymptotic stabilization of the ZDM for systems with well-defined vector relative degree, by which is meant the design of a smooth feedback controller rendering the ZDM an asymptotically stable set (see, e.g., [18] for definitions of set stability). We assume that $\mathcal{X}$ is a connected geodesically complete Riemannian manifold so that, by the Hopf-Rinow theorem [20], it is endowed with a distance function $\|\cdot\| : \mathcal{X} \to \mathbb{R}$, and we can define the point-to-set distance $\|x\|_{\mathcal{Z}} := \inf_{y\in\mathcal{Z}}(\|x-y\|)$. The next result concerns the local asymptotic stabilization of the ZDM.

Proposition 12: Suppose system (2) satisfies Assumption 1 and the ZDM $\mathcal{Z}$ is compact.

Consider the coordinate transformation $T : V \to \tilde{\mathcal{Z}} \times W$, $x \mapsto (z,\xi) = (\sigma \circ \mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x))$ arising from Theorem 3, where $W \subset \mathbb{R}^r$ is a neighbourhood of the origin in $\mathbb{R}^r$, and the resulting normal form (1) valid on $\tilde{\mathcal{Z}} \times W$. If $v = \bar{v}(\xi)$ is a smooth feedback controller asymptotically stabilizing the origin of the ξ-subsystem in (1), then the controller given by the feedback transformation (4) with $v = \bar{v} \circ \mathsf{H}_{\mathbf{r}}(x)$ asymptotically stabilizes the ZDM.

Proof: Since $\mathcal{Z} = \mathsf{H}_{\mathbf{r}}^{-1}(0)$, the function $\|\mathsf{H}_{\mathbf{r}}(\cdot)\| : \mathcal{X} \to \mathbb{R}$ is continuous and positive definite with respect to $\mathcal{Z}$. Since $\mathcal{Z}$ is compact, there exist two class-$\mathcal{K}$ functions $\alpha_1, \alpha_2 : [0,a) \to \mathbb{R}$ such that

$$\alpha_1(\|x\|_{\mathcal{Z}}) \le \|\mathsf{H}_{\mathbf{r}}(x)\| \le \alpha_2(\|x\|_{\mathcal{Z}}), \tag{12}$$

for every $x \in \{x \in \mathcal{X} : \|x\|_{\mathcal{Z}} < a\}$. Consider the normal form (1) with feedback $v = \bar{v}(\xi)$. Since $\xi = 0$ is asymptotically stable, there exists a neighbourhood $B_{\varepsilon_1}(0) := \{\xi \in \mathbb{R}^r : \|\xi\| < \varepsilon_1\}$ contained in W such that for each $\xi(0) \in B_{\varepsilon_1}(0)$, the solution $\xi(t)$ is contained in $W \cap B_{\alpha_1(a/2)}(0)$ and converges to 0. In particular, it is bounded. We may assume that ε_1 is in the image of α_2.

For each $(z(0),\xi(0)) \in \tilde{\mathcal{Z}} \times W$, the solution $(z(t),\xi(t))$ is bounded because $\xi(t)$ is bounded and $z(t) \in \tilde{\mathcal{Z}}$, this latter a compact set since $\mathcal{Z}$ is compact by assumption. Since the solution $(z(t),\xi(t))$ is bounded, it is defined for all $t \ge 0$. Moreover, $x(t) := T^{-1}(z(t),\xi(t))$ is the solution of the closed-loop system (2) with the controller given by the feedback transformation (4) with $v = \bar{v} \circ \mathsf{H}_{\mathbf{r}}(x)$ and initial condition $x(0) = T^{-1}(z(0),\xi(0))$.

For each initial condition $x(0) \in \{x \in \mathcal{X} : \|x\|_{\mathcal{Z}} < \alpha_2^{-1}(\varepsilon_1)\}$ we have $\|\mathsf{H}_{\mathbf{r}}(x(0)\| \le \alpha_2(\|x(0)\|_{\mathcal{Z}}) < \varepsilon_1$, implying that the signal $\mathsf{H}_{\mathbf{r}}(x(t)))$ converges to zero and is contained in $W \cap B_{\alpha_1(a/2)}(0)$, which implies by (12) that $x(t) \in \{x \in \mathcal{X} : \|x\|_{\mathcal{Z}} < a\}$. Since $\alpha_1(\|x(t)\|_{\mathcal{Z}}) \le \|\mathsf{H}_{\mathbf{r}}(x(t))\|$, $\|x(t)\|_{\mathcal{Z}} \to 0$. This proves that $\mathcal{Z}$ is attractive.

For stability, let $\varepsilon_2 \in (0,a)$ be arbitrary. Since the origin $\xi = 0$ is stable, there exists $\delta_1 > 0$ such that for each $\xi(0) \in B_{\delta_1}(0)$, the solution $\xi(t)$ is contained in W and $\xi(t) \in B_{\alpha_1(\varepsilon_2)}(0)$. Let $\delta_2 := \alpha_2^{-1}(\delta_1)$. For each $\|x(0)\|_{\mathcal{Z}} < \delta_2$, we have $\|\mathsf{H}_{\mathbf{r}}(x(0))\| < \alpha_2(\|x(0)\|_{\mathcal{Z}}) < \delta_1$. Thus, $\|\mathsf{H}_{\mathbf{r}}(x(t))\| < \alpha_1(\varepsilon_2)$ for all $t \ge 0$, implying by (12) that $\|x(t)\|_{\mathcal{Z}} < \varepsilon_2$ for all $t \ge 0$. This proves stability of $\mathcal{Z}$. ■

Next, we turn to the asymptotic stabilization of the ZDM with a *guaranteed* basin of attraction. A special case is the global asymptotic stabilization of the ZDM.

Proposition 13: Suppose there exists a diffeomorphism $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$ transforming system (2) to the the normal form (1) after the feedback transformation (4), where $\tilde{\mathcal{Z}}$ is diffeomorphic to $\mathcal{Z}$ and $\mathcal{Z}$ is compact. Let $v = \bar{v}(\xi)$ be a smooth feedback controller for (1) meeting the following specifications:

(i) The set $\mathcal{C} \subset \mathbb{R}^r$ is positively invariant for the ξ-subsystem.
(ii) The origin $\xi = 0$ is asymptotically stable with basin of attraction containing $\mathcal{C}$.

Then the feedback controller given by the feedback transformation (4) with $v = \bar{v} \circ \mathsf{H}_{\mathbf{r}}(x)$ asymptotically stabilizes the ZDM with basin of attraction containing $\mathcal{U}$.

Proof: The ZDM $\mathcal{Z}$ is asymptotically stable by Proposition 12. We only need to show that $\mathcal{U}$ is contained in the basin of attraction of $\mathcal{Z}$. Let $x(0) \in \mathcal{U}$ be arbitrary, and let $(z(0),\xi(0)) := T(x(0))$. The corresponding solution $(z(t),\xi(t))$ of the normal form (1) with feedback $v = \bar{v}(\xi)$ is contained in $\tilde{\mathcal{Z}} \times \mathcal{C}$ by assumption (i), and therefore the solution of system (2) with controller given by the feedback transformation (4) with $v = \bar{v} \circ \mathsf{H}_{\mathbf{r}}(x)$ is given by $x(t) = T^{-1}(z(t),\xi(t))$. By assumption (ii), $\xi(t) \to 0$, implying also that $\xi(t)$ is bounded. Since $\tilde{\mathcal{Z}}$ is compact, $z(t)$ is also bounded. This implies that $(z(t),\xi(t))$, and therefore $x(t)$, is defined for all $t \ge 0$. Consider the class-$\mathcal{K}$ functions $\alpha_i : [0,a) \to \mathbb{R}$ defined in the proof of Proposition 12, and for each $\varepsilon \in (0,a)$, let $t_1 \ge 0$ be such that $\|\xi(t)\| < \alpha_1(\varepsilon)$ for all $t \ge t_1$. For each $t \ge t_1$, we have

$$\alpha_1(\|x(t)\|_{\mathcal{Z}}) \le \|\mathsf{H}_{\mathbf{r}}(x(t))\| = \|\xi(t)\| < \alpha_1(\varepsilon),$$

from which it follows that $\|x(t)\|_{\mathcal{Z}} < \varepsilon$. This proves that $x(0)$ is contained in the basin of attraction of $\mathcal{Z}$. ■

Remark 14: The compactness hypothesis on $\mathcal{Z}$ in Propositions 12 and 13 was used in two ways: to assert the existence of class-$\mathcal{K}$ functions α_1, α_2 satisfying the inequalities in (12), and to rule out finite escape times arising from the z-subsystem in the normal form. If one *assumes* these two properties, one may remove the compactness assumption. Thus the results in the propositions hold true when $\mathcal{Z}$ is unbounded provided $\mathsf{H}_{\mathbf{r}}$ satisfies inequalities (12) for some class-$\mathcal{K}$ functions α_i, $i \in 1:2$, and provided that solutions of the normal form are defined for all $t \ge 0$. While the first assumption is relatively mild, the second one must be verified case-by-case by a dedicated analysis, typically of Lyapunov nature, or by dedicated control design, leading to a feedback $\bar{v}(z,\xi)$ which may depend on z, rather than just ξ. △

Remark 15: If the assumptions of Proposition 13 hold with $\mathcal{U} = \mathcal{X}$, then the feedback in the proposition globally asymptotically stabilizes the ZDM even if a global normal form does not exist. To illustrate, consider the single-integrator system in Example 10 which does not admit a global normal form, but can be transformed in normal form via a diffeomorphism $T : \mathcal{X} \to \{0\} \times (-\pi/2, \pi/2)$. The feedback $v = -\xi$ obviously meets the requirements of Proposition 13, and therefore it globally asymptotically stabilizes the ZDM, which in this example is the origin. In x coordinates, the feedback is $u = -(1+x^2)\arctan(x)$, and one can verify that this feedback does indeed globally asymptotically stabilize the origin. △

VII. Application: Circular path following for kinematic unicycle.

In this section we revisit an example investigated in [24], bringing new light to the problem by means of the construction in Theorem 6.

Consider a kinematic unicycle with position $\operatorname{col}(x_1, x_2) \in \mathbb{R}^2$ and heading angle $x_3 \in \mathbb{S}^1$:

$$\begin{aligned}\dot{x}_1 &= \cos(x_3)\\ \dot{x}_2 &= \sin(x_3)\\ \dot{x}_3 &= u.\end{aligned}$$

The state space is $\mathcal{X} = \mathbb{R}^2 \times \mathbb{S}^1$. The control objective is circular path following: make the unicycle converge to and traverse the unit circle in an unspecified direction. A variation of this problem where the desired direction of traversal is specified a priori can be solved globally my means of a smooth feedback (see [8]) without using feedback linearization, but here we focus on using input-output linearization in the spirit of [23]. We note the work by Samson [28] on unicycle path following (see also the review in [22]), which proposes to represent the unicycle dynamics in a Frenet-Serret frame attached to the path one wants to follow, and use this representation to solve the path following problem locally for a large class of paths. This idea has been used extensively in the motion control literature.

We define the output $y = h(x) := x_1^2 + x_2^2 - 1$, a smooth function on $\mathcal{X}$. We will denote by $\mathbf{n}(x) := \operatorname{col}(x_1, x_2)/(x_1^2 + x_2^2)^{1/2}$ the unit normal vector to the circle with radius $(x_1^2 + x_2^2)^{1/2}$ centred at 0, and $\mathbf{t}(x) = \operatorname{col}(x_2, -x_1)/(x_1^2+x_2^2)^{1/2}$ the unit tangent vector to the same circle in the clockwise direction. We will further denote by $\mathbf{h}(x) := \operatorname{col}(c_{x_3}, s_{x_3})$ the unicycle's unit heading vector. Letting

$$\mu(x) := x_1 s_{x_3} - x_2 c_{x_3},$$

we have $\ddot{y} = 2 - 2\mu(x)u$, and therefore the system has relative degree $r = 2$ on the set

$$\mathcal{U} := \{x \in \mathcal{X} : \mu(x) \neq 0\},$$

which can be partitioned into sets

$$\begin{aligned}\mathcal{U}^+ &= \{x \in \mathcal{X} : \mu(x) > 0\}\\ \mathcal{U}^- &= \{x \in \mathcal{X} : \mu(x) < 0\}.\end{aligned}$$

Noting that $\mu(x) = 0$ if either $(x_1, x_2) = (0,0)$ or $\langle \mathbf{t}(x), \mathbf{h}(x)\rangle = 0$, relative degree fails when either the unicycle is at the origin or its heading vector is perpendicular to the target path. We have

$$\mathsf{H}_{\mathbf{r}}(x) = \begin{bmatrix} x_1^2 + x_2^2 - 1 \\ 2(x_1 c_{x_3} + x_2 s_{x_3}) \end{bmatrix},$$

and $\mathcal{Z} = (\mathsf{H}_{\mathbf{r}})^{-1}(0)$. If $x \in (\mathsf{H}_{\mathbf{r}})^{-1}(0)$, then $(x_1, x_2) \neq (0,0)$, and $\langle \mathbf{n}(x), \mathbf{h}(x)\rangle = 0$, and this implies that $\mu(x) \neq 0$. Thus the system has well-defined relative degree on $\mathcal{Z}$, and $\mathcal{Z}$ is the ZDM.

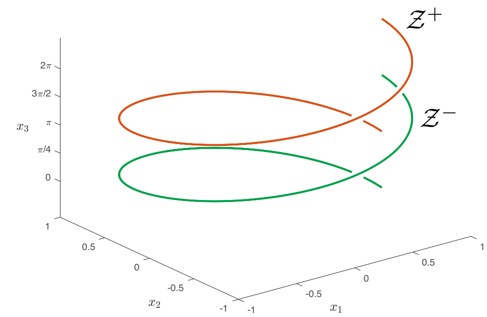


Fig. 3. The ZDM for unicycle path following. Since $x_3 \in \mathbb{S}^1$, the endpoints of the two curves are identified.

We note that $\mathsf{H}_{\mathbf{r}}(x) = 0$ when (x_1, x_2) is on the unit circle and $\langle \mathbf{n}(x), \mathbf{h}(x)\rangle = 0$. Geometrically then, $\mathcal{Z}$ is the set where the unicycle is on the unit circle with heading $\mathbf{h}(x)$ tangent to it with either clockwise or counterclockwise orientation. Accordingly, $\mathcal{Z}$ is the union of two disjoint *closed* curves in $\mathcal{X}$ which we denote $\mathcal{Z}^+$ and $\mathcal{Z}^-$,

$$\mathcal{Z} = \mathcal{Z}^+ \sqcup \mathcal{Z}^-,$$

with plus indicating counterclockwise orientation and minus indicating clockwise orientation; see Figure 3. We note that

$$\mathcal{Z}^+ \subset \mathcal{U}^+, \ \mathcal{Z}^- \subset \mathcal{U}^-.$$

Since $\mathcal{Z}$ is not connected while the state space $\mathcal{X}$ *is* connected, $\mathcal{X}$ and $\mathcal{Z}$ are not homotopy equivalent, and therefore by the discussion in Remark 11, a global normal form does not exist.

Since the curves making up $\mathcal{Z}$ are closed, each connected component of $\mathcal{Z}$ is diffeomorphic to $\mathbb{S}^1$. Accordingly, let $\tilde{\mathcal{Z}}$ denote the disjoint union of two copies of $\mathbb{S}^1$, denoted $\mathbb{S}^+$ and $\mathbb{S}^-$,

$$\tilde{\mathcal{Z}} := \mathbb{S}^+ \sqcup \mathbb{S}^-,$$

and define parametrizations

$$\begin{aligned}\phi^+ &: \mathbb{S}^+ \to \mathcal{Z}^+, \ z \mapsto (c_z, s_z, z + \pi/2)\\ \phi^- &: \mathbb{S}^- \to \mathcal{Z}^-, \ z \mapsto (c_z, s_z, z - \pi/2).\end{aligned}$$

Now let $\phi : \tilde{\mathcal{Z}} \to \mathcal{Z}$ be the map defined by $\phi\big|_{\mathbb{S}^+} := \phi^+$ and $\phi\big|_{\mathbb{S}^-} := \phi^-$. By [19, Corollary 2.8]), this map is well-defined and smooth. Moreover, its inverse is the smooth map

$$\sigma : \mathcal{Z} \to \tilde{\mathcal{Z}}, \ x \mapsto z = \measuredangle \operatorname{col}(x_1, x_2),$$

so σ is a diffeomorphism.

Having determined the ZDM and characterized its geometry, we turn to checking the hypotheses of Theorem 6. The set $\mathcal{C} = \mathsf{H}_{\mathbf{r}}(\mathcal{U})$ is given by

$$\mathcal{C} = \left\{\xi \in \mathbb{R}^2 : \xi_1 > (\xi_2/2)^2 - 1\right\},$$

and we claim that $\mathcal{C}$ satisfies Assumption 2. For this problem we have one output with relative degree 2 so $\mathfrak{I}_{\mathbf{r}} = \{(1,1),(1,2)\}$. Consider the permutation $\pi : \mathfrak{I}_{\mathbf{r}} \to \mathfrak{I}_{\mathbf{r}}$, $\pi(1,1) = (1,2), \pi(1,2) = (2,1)$, and associated indices $l_1 = \mathsf{idx}(1,2) = 2$ and $l_2 = \mathsf{idx}(1,1) = 1$. The isomorphism P_π is given by $\mathsf{P}_\pi(\xi) = \mathrm{col}(\xi_2, \xi_1)$. It is readily seen that $\mathcal{C}$ is open, convex, and it contains the origin. Moreover, the condition

$$\xi \in \mathcal{C} \implies (0, \xi_{l_2}) \in \mathsf{P}_\pi(\mathcal{C}),$$

is equivalent to

$$\xi \in \mathcal{C} \implies (\xi_1, 0) \in \mathcal{C},$$

and one can see that this latter condition is verified. Thus Assumption 2 holds with $l_1 = 2, l_2 = 1$.

Next, we check the assumptions of Theorem 6. The feedback transformation

$$u = (2\mu(x))^{-1}(2 - v) \tag{13}$$

gives $\ddot{y} = v$ and

$$\begin{aligned}
\tilde{f} &= c_{x_3}\frac{\partial}{\partial x_1} + s_{x_3}\frac{\partial}{\partial x_2} + \frac{1}{\mu(x)}\frac{\partial}{\partial x_3}\\
\tilde{g} &= \frac{-1}{2\mu(x)}\frac{\partial}{\partial x_3}\\
ad_{\tilde{f}}\,\tilde{g} &= \frac{-1}{2\mu(x)}\left(s_{x_3}\frac{\partial}{\partial x_1} - c_{x_3}\frac{\partial}{\partial x_2}\right).
\end{aligned}$$

We won't need to use vector fields δ_1, δ_2 in this example, so we set $\tau_1^1 = \tilde{g}$ and $\tau_2^1 = -ad_{\tilde{f}}\,\tilde{g}$.

We compute the retraction p in (10),

$$\mathsf{p}(x) = \Phi^{\bar{\tau}_2}_{-\bar{\lambda}_2(x)} \circ \Phi^{\bar{\tau}_1}_{-\bar{\lambda}_1(x)}(x) = \Phi^{\tau_2^1}_{-h(x)} \circ \Phi^{\tau_1^1}_{-L_f h(x)}(x). \tag{14}$$

To find the flow of τ_1^1, we use separation of variables to integrate the associated differential equation and employ basic trigonometric identities. By so doing, we get

$$\Phi_t^{\tau_1^1}(x) = \begin{bmatrix} x_1 \\ x_2 \\ \measuredangle\mathbf{n}(x) \pm \cos^{-1}\left(\frac{(t/2) + x_1 c_{x_3} + x_2 s_{x_3}}{\sqrt{x_1^2 + x_2^2}}\right)\end{bmatrix},$$

with plus sign if $x \in \mathcal{U}^+$, and minus sign if $x \in \mathcal{U}^-$. We note that this flow is well-defined for all $x \in \mathcal{U}$ and all $t \in \mathbb{R}$ such that the argument of the function $\cos^{-1}(\cdot)$ is in the interval $(-1, 1)$, or using the expressions for h and $L_f h$,

$$t \in \left(-L_f h(x) - 2\sqrt{h(x)+1}, -L_f h(x) + 2\sqrt{h(x)+1}\right).$$

For $x \in \mathcal{U}$ and t in the open interval above, we have

$$\mathsf{H}_{\mathbf{r}}(x) + t e_2^1 = \mathrm{col}(h(x), L_f h(x) + t),$$

and $L_f h(x) + t \in \left(-2\sqrt{h(x)+1}, 2\sqrt{h(x)+1}\right)$. Since $x \in \mathcal{U}$, $h(x) + 1 > 0$, and so

$$\mathsf{H}_{\mathbf{r}}(x) + t e_2^1 \in \left\{\xi \in \mathbb{R}^2 : \xi_1 > (\xi_2/2)^2 - 1\right\}.$$

This latter set is $\mathcal{C}$ and therefore τ_1^1 satisfies condition (i) of Theorem 6. In light of Remark 7, we do not need to check condition (ii).

Now we turn to the flow $\Phi_t^{\tau_2^1}(x)$. We notice that this flow preserves x_3, and therefore its orbits on the (x_1, x_2) plane are straight lines with tangent vector $\mathrm{col}(-s_{x_3}, c_{x_3})$. The flow, therefore, has the form

$$\Phi_t^{\tau_2^1}(x) = x + \lambda(t,x)\begin{bmatrix} -s_{x_3} \\ c_{x_3} \\ 0\end{bmatrix}, \tag{15}$$

where λ is a real-valued function to be determined such that $\lambda(0,x) = 0$. Differentiating the above expression with respect to t and imposing that the result be equal to $\tau_2^1(\Phi_t^{\tau_2^1}(x))$ we obtain an expression for $\dot{\lambda}$,

$$\dot{\lambda} = \frac{1}{-2\mu(x) + 2\lambda}.$$

Using separation of variables and imposing $\lambda|_{t=0} = 0$, we integrate to get

$$\lambda^2 - 2\mu(x)\lambda = t. \tag{16}$$

The flow is defined for all t such that the above polynomial in the variable λ has real roots, or $t > -\mu(x)^2$ which can be expressed as $t > -h(x) - 1 + (L_f h(x))^2/4$. Now we check the flow condition of Theorem 6. Letting $x \in \mathcal{U}$, we have

$$\mathsf{H}_{\mathbf{r}}(x) + t e_1^1 = \begin{bmatrix} h(x) + t \\ L_f h(x)\end{bmatrix}.$$

If $t > -h(x) - 1 + (L_f h(x))^2/4$, then $h(x) + t > -1 + (L_f h(x))^2/4$, so the vector $\mathsf{H}_{\mathbf{r}}(x) + t e_1^1$ is in the set

$$\left\{\xi \in \mathbb{R}^2 : \xi_1 > (\xi_2/2)^2 - 1\right\},$$

which coincides with $\mathcal{C}$. Therefore τ_2^1 satisfies condition (30). To obtain the flow of τ_2^1, we solve the polynomial (16), discarding the root incompatible with the condition $\lambda(0,x) = 0$, thus obtaining

$$\lambda(t,x) = \begin{cases} \mu(x) - \sqrt{\mu(x)^2 + t} & x \in \mathcal{U}^+ \\ \mu(x) + \sqrt{\mu(x)^2 + t} & x \in \mathcal{U}^-. \end{cases} \tag{17}$$

The flow of τ_2^1 is given by (15) and (17).

Using the flows of τ_1^1, τ_2^1, we may now compute the retraction $\mathsf{p} : \mathcal{U} \to \mathcal{Z}$ in (14). Using the fact that $L_f h(x) = 2x_1 c_{x_3} + 2x_3 s_{x_3}$, we have

$$\Phi^{\tau_1^1}_{-L_f h(x)}(x) = \begin{bmatrix} x_1 \\ x_2 \\ \measuredangle\mathbf{n}(x) \pm \pi/2\end{bmatrix},$$

with plus sign when $x \in \mathcal{U}^+$ and minus sign when $x \in \mathcal{U}^-$. Further composing with $\Phi^{\tau_2^1}_{-h(x)}$ we get

$$\mathsf{p}(x) = \begin{bmatrix} \mathbf{n}(x) \\ \measuredangle\mathbf{n}(x) \pm \pi/2\end{bmatrix}.$$

Note that p projects the position of the unicycle (x_1, x_2) orthogonally on the unit circle, and it projects the heading angle x_3 onto the tangent line to the circle, in the counterclockwise

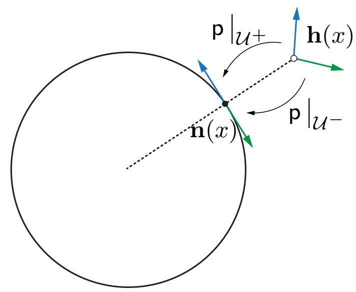


Fig. 4. An illustration of the smooth retraction $\mathsf{p} : \mathcal{U} \to \mathcal{Z}$.

or clockwise direction depending on whether, respectively, x is in $\mathcal{U}^+$ or $\mathcal{U}^-$. We note that

$$\mathsf{p}\,\big|_{\mathcal{U}^+} : \mathcal{U}^+ \to \mathcal{Z}^+$$
$$\mathsf{p}\,\big|_{\mathcal{U}^-} : \mathcal{U}^- \to \mathcal{Z}^-.$$

This retraction is illustrated in Figure 4.

Next, we compute $\sigma \circ \mathsf{p} : \mathcal{U} \to \tilde{\mathcal{Z}}$ as $\sigma \circ \mathsf{p}(x) = \measuredangle \mathbf{n}(x)$. By Theorem 6, the map $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$,

$$\begin{aligned}(z, \xi) &= (\sigma \circ \mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x)) \\ &= \big(\measuredangle \mathbf{n}(x), \mathrm{col}(x_1^2 + x_2^2 - 1, 2x_1 c_{x_3} + 2x_2 s_{x_3})\big)\end{aligned}$$

is a diffeomorphism. Note that $\mathcal{U}$ is a set of full measure in $\mathcal{X}$, and therefore we are about to produce an almost global feedback equivalence to a normal form (recall that a global normal form does not exist for this problem).

Using the identities

$$\begin{aligned}\frac{d}{dt}\measuredangle \mathbf{n}(x) &= \frac{x_1\dot{x}_2 - x_2\dot{x}_1}{x_1^2 + x_2^2} \\ \mathrm{col}(x_1, x_2) &= (\xi_1 + 1)^{1/2}\,\mathrm{col}(c_z, s_z) \\ \mu(x) &= (\xi_1 + 1)^{1/2} s_{x_3 - z} \\ c_{x_3 - z} &= \frac{\xi_2}{2(\xi_1 + 1)^{1/2}},\end{aligned}$$

and the feedback transformation (13), we arrive at the normal form

$$\begin{aligned}\dot{z} &= \pm\frac{\left(4(\xi_1 + 1) - \xi_2^2\right)^{1/2}}{2(\xi_1 + 1)} \\ \dot{\xi}_1 &= \xi_2 \\ \dot{\xi}_2 &= v,\end{aligned} \tag{18}$$

with plus sign if $z \in \mathbb{S}^+$, and minus sign if $z \in \mathbb{S}^-$. The state space of (18) is $\mathcal{C} = \{(z, \xi_1, \xi_2) \in \tilde{\mathcal{Z}} : \xi_1 > (\xi_2/2)^2 - 1\}$.

Having established an almost global feedback equivalence between the unicycle model and system (18), we now asymptotically stabilize the ZDM. For the double-integrator with state ξ in (18), we wish to design a smooth feedback $v = \bar{v}(\xi)$ meeting the control specifications (i) and (ii) listed in Proposition 13. We remark that $\mathcal{Z}$ is compact so the proposition is indeed applicable, and it implies for each feedback $v = \bar{v}(\xi)$ meeting the aforementioned specifications, the feedback $u = (2\mu(x))^{-1}(2 - \bar{v}(\mathsf{H}_{\mathbf{r}}(x)))$ will render $\mathcal{Z}$ asymptotically stable with basin of attraction containing the positively invariant set $\mathcal{U}$. Thus the controller in question will render $\mathcal{Z}$ almost globally asymptotically stable.

Recall that $\mathcal{U}$ and $\mathcal{Z}$ are unions of two disjoint sets, $\mathcal{U} = \mathcal{U}^+ \cup \mathcal{U}^-$, and $\mathcal{Z} = \mathcal{Z}^+ \cup \mathcal{Z}^-$, with $\mathcal{Z}^+ \subset \mathcal{U}^+$ and $\mathcal{Z}^- \subset \mathcal{U}^-$. Then, $\mathcal{Z}^+$ and $\mathcal{Z}^-$ will each be asymptotically stable with basin of attraction given by the positively invariant sets $\mathcal{U}^+$ and $\mathcal{U}^-$, respectively. More concretely, if $\mu(x(0)) > 0$ the unicycle will stay in $\mathcal{U}^+$, converge to the circle and follow it counterclockwise, while if $\mu(x(0)) < 0$ the unicycle will stay in $\mathcal{U}^-$ and follow the circle in the clockwise direction.

Proposition 16: For each $K_d > 1$, the feedback

$$\bar{v}(\xi) = -2\xi_1 + \frac{\xi_2^2}{2} - K_d\xi_2$$

meets specifications (i)-(ii) in Proposition 13, therefore inducing almost global circular path following.

Proof: The ξ subsystem with feedback $\bar{v}$ is

$$\dot{\xi}_1 = \xi_2, \;\; \dot{\xi}_2 = -2\xi_1 + \frac{\xi_2^2}{2} - K_d\xi_2. \tag{19}$$

Letting $W(\xi) = (\xi_2)^2/4 - (\xi_1 + 1)$, the set $\mathcal{C}$ can be written as $\mathcal{C} = \{\xi \in \mathbb{R}^2 : W(\xi) < 0\}$. The time derivative of W along solutions of (19) is $\dot{W} = -(K_d/2)\xi_2^2 + \xi_2 W(\xi)$. Since $\dot{W}|_{W=0} \leq 0, \mathcal{C}$ is positively invariant for (19) and specification (i) is met. For specification (ii), we will use the reduction theorem for asymptotic stability of compact sets in [9]. Let

$$\Gamma := \{\xi \in \mathcal{C} : \xi_2 \leq 2/K_d\}.$$

For each initial condition in $\mathcal{C}$, the solution $\xi(t)$ satisfies $W(\xi(t)) < 0$ implying that

$$\dot{\xi}_2 = -2\xi_1 + \frac{\xi_2^2}{2} - K_d\xi_2 = -K_d\xi_2 + 2W(\xi) + 2 \leq -K_d\xi_2 + 2.$$

By the comparison lemma (e.g., [18]) we deduce that Γ is globally asymptotically stable relative to $\mathcal{C}$ (i.e., restricting initial conditions to be contained in $\mathcal{C}$). Now consider solutions $\xi(t)$ initialized in Γ, for which we have $\xi_2(t) \leq 2/K_d$. The derivative of the Lyapunov function $V(\xi) = \xi_1^2 + (1/2)\xi_2^2$ along solutions of (19) is

$$\dot{V} = -\xi_2^2(K_d - \xi_2/2) \leq -\xi_2^2\frac{K_d^2 - 1}{K_d}.$$

Since $K_d > 1$, $\dot{V} \leq 0$ and the LaSalle-Krasovskii invariance principle implies that the origin $\xi = 0$ is globally asymptotically stable relative to Γ. Now [9, Corollary 11] implies that the origin $\xi = 0$ is globally asymptotically stable relative to $\mathcal{C}$. ■

Figure 5 shows simulations results for 20 random initial conditions. The green curves represent solutions with initial condition in $\mathcal{U}^+$, while red curves represent initial conditions in $\mathcal{U}^-$. As predicted by the theory, green curves approach and follow the circle in the counterclockwise direction, while red curves do so in the clockwise direction.

The proposed controller can be enhanced with a hybrid supervisor yielding global path following in the clockwise or counterclockwise direction (i.e., global asymptotic stability of $\mathcal{Z}^+$ or $\mathcal{Z}^-$). Suppose we want to globally asymptotically stabilize $\mathcal{Z}^+$ for counterclockwise path following. For initial conditions in $\mathcal{U}^+$, the supervisor does nothing. For every other initial condition, the supervisor applies a constant control

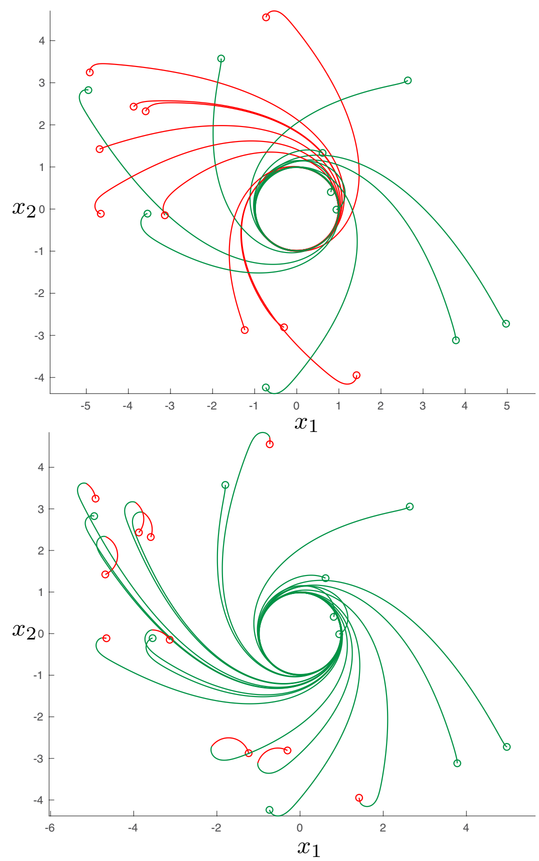


Fig. 5. Simulation results for unicycle path following. The first figure shows the behaviour of the controller in Proposition 16 almost globally stabilizing $\mathcal{Z}$. The second figure shows the results of adding a hybrid supervisor that globally asymptotically stabilizes $\mathcal{Z}^+$. Green arcs represent states in $\mathcal{U}^+$, while red arcs correspond to states in $\mathcal{U}^-$.

input $u = \bar{u} > 0$ until the unicycle enters[4] $\mathcal{U}^+$, at which point it resumes normal operation with the controller proposed above. The positive invariance of $\mathcal{U}^+$ is a key feature because it ensures that once the unicycle enters $\mathcal{U}^+$ it stays there. Simulation results for this hybrid supervisor with the same initial conditions used previously and with $\bar{u} = 2$ are displayed in Figure 5.

VIII. Conclusions

For nonlinear control systems with well-defined relative degree, we have investigated the existence of normal forms for input-output linearization valid in some neighbourhood of the ZDM or in a given a priori neighbourhood of it. Finding the coordinate transformation involves finding a smooth retraction onto the ZDM. While the work of Byrnes and Isidori [5] provides a formula for such a retraction, one that we have relied on in our proof of Theorem 6, its computation requires flows of vector fields that are generally unavailable in closed form. It might generally be easier to seek a smooth retraction yielding the required diffeomorphism. In the unicycle example, the retraction we arrived at using Byrnes and Isidori's formula is the obvious choice. The analysis of the unicycle example might possibly be extended in two directions: the path following of arbitrary embedded curves in the plane for the unicycle, and the extension to other nonholonomic models. Finally, we mention that the results of this paper can be easily extended to handle control systems with more inputs than outputs.

[4]Setting $\bar{u} > 0$, one has $\ddot{\mu} = \bar{u} - \bar{u}^2\mu$, and therefore for any initial condition, $\mu(x(t))$ is a sinusoidal signal with positive average $\bar{u}$, ensuring that there exists a time $\bar{t} > 0$ such that $\mu(x(\bar{t})) > 0$, and thus $x(\bar{t}) \in \mathcal{U}^+$.

Appendix I
Proof of Theorem 3

Before proving the theorem, we recall that if $U \subset \mathcal{X}$ is open, $H : U \to \mathbb{R}^r$ is a submersion, and $\mathcal{Z} = H^{-1}(c)$ is a nonempty level set of H, then $\mathcal{Z}$ is an embedded submanifold of $\mathcal{X}$ of dimension $n - r$, and for each $x \in \mathcal{Z}$, $T_x\mathcal{Z} = \operatorname{Ker} dH_x$. We also need the Generalized Inverse Function Theorem.

Theorem 17 ([10]): Let $\mathcal{Z}$ be an embedded submanifold of $\mathcal{X}$, and $F : \mathcal{X} \to \mathcal{Y}$ be a smooth map of manifolds enjoying the following properties:

(i) For each $z \in \mathcal{Z}$, the differential $dF_z : T_z\mathcal{X} \to T_{F(z)}\mathcal{Y}$ is an isomorphism.
(ii) The restriction $F|_{\mathcal{Z}} : \mathcal{Z} \to F(\mathcal{Z}) \subset \mathcal{Y}$ is a diffeomorphism.

Then, there exists a neighbourhood $V \subset \mathcal{X}$ of $\mathcal{Z}$ such that $F|_V : V \to F(V)$ is a diffeomorphism.
In the above, note that $F(V)$ is a neighbourhood of $F(\mathcal{Z})$.

Proof of Theorem 3. The theorem is trivially true if $r = n$, as in this case $\mathcal{Z}$ is a zero dimensional manifold and hence a set of isolated points $\{x_i\}_{i\in\mathcal{I}} \subset \mathcal{X}$. Applying Theorem 1 with $r = n$ at each of these points we obtain diffeomorphisms T_i on open sets $U_i \subset \mathcal{X}$, and we may assume that $U_i \cap U_j = \emptyset$ for all $i \neq j$ since the points in $\{x_i\}_{i\in\mathcal{I}}$ are isolated. Letting $U = \bigcup_{i\in\mathcal{I}} U_i$, there is a unique smooth map $T : U \to \{x_i\}_{i\in\mathcal{I}} \times \mathbb{R}^r$ such that $T|_{U_i} = T_i$ (see [19, Corollary 2.8]), and this map is a diffeomorphism onto its image.

Now consider the case $r < n$. In order to show that the map $x \in \mathcal{X} \mapsto (z, \xi) = (\sigma \circ \mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x))$ is a diffeomorphism of a neighbourhood of $\mathcal{Z}$ onto its image, we use Theorem 17. Fix $x \in \mathcal{Z}$. Regarding property (i) in Theorem 17, we have

$$dT_x : T_x\mathcal{X} \to T_{(z,\xi)}(\tilde{\mathcal{Z}} \times \mathbb{R}^r) \simeq T_z\tilde{\mathcal{Z}} \times \mathbb{R}^r,$$

where $(z, \xi) = T(x)$. Since the domain and codomain of dT_x are vector spaces of equal dimension, to show that dT_x is an isomorphism it suffices to show that $\operatorname{Ker} dT_x = \{0\}$. Let $v_x \in \operatorname{Ker} dT_x$. Then,

$$d(\sigma \circ \mathsf{p})_x(v_x) = 0 \tag{20a}$$
$$d(\mathsf{H}_{\mathbf{r}})_x(v_x) = 0. \tag{20b}$$

By part (i) of Theorem 1, $\operatorname{rank} d(\mathsf{H}_{\mathbf{r}})_x = r$ for all $x \in (\mathsf{H}_{\mathbf{r}})^{-1}(0)$. Then, since $\mathcal{Z} = (\mathsf{H}_{\mathbf{r}})^{-1}(0)$ we have that $T_x\mathcal{Z} = \operatorname{Ker} d(\mathsf{H}_{\mathbf{r}})_x$, and identity (20b) implies that

$$v_x \in T_x\mathcal{Z}. \tag{21}$$

Now consider identity (20a). Since $x \in \mathcal{Z}$ and p is a retraction onto $\mathcal{Z}$, we have $\mathsf{p}(x) = x$. Using the chain rule, we get

$$d\sigma_x \circ d\mathsf{p}_x(v_x) = 0. \tag{22}$$

Since $\sigma : \mathcal{Z} \to \tilde{\mathcal{Z}}$ is a diffeomorphism, the map $d\sigma_x : T_x\mathcal{Z} \to T_z\tilde{\mathcal{Z}}$, with $z = \sigma(x)$, is an isomorphism. Then, noting that $d\mathsf{p}_x(v_x) \in T_x\mathcal{Z}$, (22) implies that

$$d\mathsf{p}_x(v_x) = 0. \tag{23}$$

By (21), there exists a smooth curve $\gamma : (-\varepsilon, \varepsilon) \to \mathcal{Z}$ such that $\gamma(0) = x$ and $\dot{\gamma}(0) = v_x$. Since $\mathsf{p}|_{\mathcal{Z}}$ is the identity map, we have

$$\mathsf{p}(\gamma(t)) \equiv \gamma(t).$$

Differentiating with respect to t, using the chain rule, and evaluating the result at $t = 0$ we get

$$d\mathsf{p}_x(v_x) = v_x.$$

But this, by (23), implies that $v_x = 0$. We have thus proved that $\operatorname{Ker} dT_x = \{0\}$, and hence dT_x is an isomorphism for any $x \in \mathcal{Z}$.

Now we turn to property (ii) of Theorem 17. The restriction $T|_{\mathcal{Z}} : \mathcal{Z} \to \mathcal{X}$ is the map $x \mapsto (\sigma \circ \mathsf{p}(x), 0) = (\sigma(x), 0) \in \tilde{\mathcal{Z}} \times \mathbb{R}^r$. Since $\sigma : \mathcal{Z} \to \tilde{\mathcal{Z}}$ is a diffeomorphism, the map $T|_{\mathcal{Z}} : \mathcal{Z} \to \tilde{\mathcal{Z}} \times \{0\}$ is a diffeomorphism as well.

By Theorem 17, there exists a neighbourhood $V \subset \mathcal{X}$ of $\mathcal{Z}$ such that $T|_V : V \to T(V)$ is a diffeomorphism. Since $T(\mathcal{Z}) = \tilde{\mathcal{Z}} \times \{0\}$, $T(V)$ is a neighbourhood of $\tilde{\mathcal{Z}} \times \{0\}$, and thus it has the form $T(V) = \tilde{\mathcal{Z}} \times W$, with $W \subset \mathbb{R}^r$ a neighbourhood of the origin. The rest of the theorem follows from part (v) of Theorem 1. ∎

Appendix II
Proof of Theorem 6

Lemma 18: Let $\tilde{f}, \tilde{g}^i$ be the vector fields defined in (3), and let $\lambda_l^j := L_{\tilde{f}}^{l-1} h^j$, $(j,l) \in \mathcal{I}_{\mathbf{r}}$. For any smooth vector fields $\delta_1, \delta_2^i \in \Delta$, $i \in 1\!:\!m$, letting $\hat{f} := \tilde{f} + \delta_1$, $\hat{g}^i := \tilde{g}^i + \delta_2^i$, the vector fields $\tau_k^i := ad_{\hat{f}}^{k-1} \hat{g}^i$ satisfy the identity

$$L_{\tau_k^i} \lambda_l^j = \begin{cases} 1 & i = j,\ k + l = r_i + 1 \\ 0 & \text{otherwise.} \end{cases} \tag{24}$$

Proof: We claim that

$$(\forall k \in \mathbb{N})(\forall i \in 1\!:\!m)\ [ad_{\tilde{f}}^k \tilde{g}^i, \Delta] \subset \Delta. \tag{25}$$

Identity (25) is true for $k = 0$ because $[\tilde{g}^i, \Delta] \subset \Delta$ for all $i \in 1\!:\!m$ by part (iv) of Theorem 1. Using induction, assume for $k \geq 1$ that $[ad_{\tilde{f}}^{k-1} \tilde{g}^i, \Delta] \subset \Delta$ for all $i \in 1 : m$, and let $\delta \in \Delta$ be arbitrary. For each $i \in 1 : m$, using the Jacobi identity we have

$$\begin{aligned}[ad_{\tilde{f}}^k \tilde{g}^i, \delta] &= -[\delta, [\tilde{f}, ad_{\tilde{f}}^{k-1} \tilde{g}^i]] \\ &= [\tilde{f}, [ad_{\tilde{f}}^{k-1} \tilde{g}^i, \delta]] + [ad_{\tilde{f}}^{k-1} \tilde{g}^i, [\delta, \tilde{f}]].\end{aligned}$$

We have $[ad_{\tilde{f}}^{k-1} \tilde{g}^i, \delta] \in \Delta$ by the induction hypothesis and $[\delta, \tilde{f}] \in \Delta$ by part (iv) of Theorem 1. Thus the inner Lie brackets in the sum above give vector fields in Δ. Using again the induction hypothesis and part (iv) of Theorem 1, we have that $[\tilde{f}, [ad_{\tilde{f}}^{k-1} \tilde{g}^i, \delta]] \in \Delta$ and $[ad_{\tilde{f}}^{k-1} \tilde{g}^i, [\delta, \tilde{f}]] \in \Delta$. This proves that (25) holds.

Next, we show that

$$(\forall k \in \mathbb{N})(\forall i \in 1\!:\!m)\ ad_{\hat{f}}^k \hat{g}^i - ad_{\tilde{f}}^k \tilde{g}^i \in \Delta. \tag{26}$$

The identity is true for $k = 0$ and all $i \in 1\!:\!m$, since $\hat{g}^i - \tilde{g}^i = \delta_2^i \in \Delta$. By induction on k, suppose $ad_{\hat{f}}^{k-1} \hat{g}^i - ad_{\tilde{f}}^{k-1} \tilde{g}^i = \delta \in \Delta$. Then,

$$\begin{aligned} ad_{\hat{f}}^k \hat{g}^i - ad_{\tilde{f}}^k \tilde{g}^i &= [\hat{f}, ad_{\hat{f}}^{k-1} \hat{g}^i] - [\tilde{f}, ad_{\tilde{f}}^{k-1} \tilde{g}^i] \\ &= [\tilde{f}, ad_{\hat{f}}^{k-1} \hat{g}^i] + [\delta_1, ad_{\hat{f}}^{k-1} \hat{g}^i] - [\tilde{f}, ad_{\tilde{f}}^{k-1} \tilde{g}^i] \\ &= [\tilde{f}, ad_{\hat{f}}^{k-1} \hat{g}^i - ad_{\tilde{f}}^{k-1} \tilde{g}^i] + [\delta_1, \delta + ad_{\tilde{f}}^{k-1} \tilde{g}^i] \\ &= [\tilde{f}, \delta] + [\delta_1, \delta] + [\delta_1, ad_{\tilde{f}}^{k-1} \tilde{g}^i]. \end{aligned}$$

Using (25), the fact that Δ is involutive, and the fact that $[\tilde{f}, \Delta] \subset \Delta$ we conclude that three summands are in Δ, and thus identity (26) holds.

By identity (26), for each $(i,k) \in \mathcal{I}_{\mathbf{r}}$ there exists a smooth vector field $\delta_k^i \in \Delta$ such that

$$\tau_k^i = (-1)^{k-1} ad_{\hat{f}}^{k-1}\ \hat{g}^i = (-1)^{k-1} ad_{\tilde{f}}^{k-1}\ \tilde{g}^i + \delta_k^i.$$

For each $(i,k), (j,l) \in \mathcal{I}_{\mathbf{r}}$, we have

$$L_{\delta_k^i} \lambda_l^j(x) = d(L_{\tilde{f}}^{l-1} h^j)_x \delta_k^i(x) = 0,$$

since $\delta_k^i(x) \in \Delta(x) = \operatorname{Ker} d(\mathsf{H}_{\mathbf{r}})_x \subset \operatorname{Ker} d(L_{\tilde{f}}^{l-1} h^j)_x$. Then,

$$\begin{aligned} L_{\tau_k^i} \lambda_l^j &= (-1)^{k-1} L_{ad_{\tilde{f}}^{k-1} \tilde{g}^i} \lambda_l^j + L_{\delta_k^i} \lambda_l^j \\ &= (-1)^{k-1} L_{ad_{\tilde{f}}^{k-1} \tilde{g}^i} \lambda_l^j \\ &= \begin{cases} 1 & i = j,\ k + l = r_i + 1 \\ 0 & \text{otherwise.} \end{cases} \end{aligned}$$

The last identity was proved by [5], [12], as discussed in Section V. ∎

Proof of Theorem 6. (Necessity). Suppose a diffeomorphism $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$ exists giving the normal form (1) after the feedback transformation (4). In (z, ξ) coordinates, and in any set of local coordinates $(z_1, \ldots, z_{n-r})$ for $\mathcal{Z}$, the vector fields $\tilde{f}$ and $\tilde{g}^i$ are given by

$$\begin{aligned} T_* \tilde{f} &= \left(\sum_{j=1}^{n-r} \alpha_j(z,\xi) \frac{\partial}{\partial z_j} \right) + \sum_{(i,k) \in \mathcal{I}_{\mathbf{r}}} \xi_{k+1}^i \frac{\partial}{\partial \xi_k^i} \\ T_* \tilde{g}^i &= \left(\sum_{j=1}^{n-r} \beta_j^i(z,\xi) \frac{\partial}{\partial z_j} \right) + \frac{\partial}{\partial \xi_{r_i}^i}, \end{aligned}$$

where α_j, β_j^i are the components of the local coordinate representation of α, β^i in (1). Let δ_1 be the pullback via T of the vector field $-\sum_j \alpha_j(z,\xi) \frac{\partial}{\partial z_j}$ and δ_2^i be the pullback of $-\sum_j \beta_j^i(z,\xi) \frac{\partial}{\partial z_j}$. Note that defining δ_1 and δ_2^i in every coordinate chart defines them globally on $\mathcal{X}$. Setting $\hat{f} = \tilde{f} + \delta_1$, $\hat{g}^i = \tilde{g} + \delta_2^i$, we get

$$T_* \hat{f} = \sum_{(i,k) \in \mathcal{I}_{\mathbf{r}}} \xi_{k+1}^i \frac{\partial}{\partial \xi_k^i},\ T_* \hat{g}^i = \frac{\partial}{\partial \xi_{r_i}^i}. \tag{27}$$

We see that $T_* \hat{f}$ is linear and $T_* \hat{g}^i$ is constant. By the naturality of the Lie bracket ([19, Prop.8.30]), we have $T_*[\hat{f}, \hat{g}^i] = [T_* \hat{f}, T_* \hat{g}^i]$, using which we get

$$T_* \tau_k^i = (-1)^{k-1} T_* ad_{\hat{f}}^{k-1}\ \hat{g}^i = \frac{\partial}{\partial \xi_{r_i - k + 1}^i},\ (i,k) \in \mathcal{I}_{\mathbf{r}}. \tag{28}$$

The flow of $T_*\tau_k^i$ from a point $(z,\xi) \in \mathcal{C}$ is $t \mapsto (z, \xi + te^i_{r_i-k+1})$, and it is well defined as long as $\xi + te^i_{r_i-k+1} \in \mathcal{C}$. This proves that property (i) holds. Moreover,

$$0 = \left[\frac{\partial}{\partial \xi^i_{r_i-k+1}}, \frac{\partial}{\partial \xi^j_{r_j-l+1}}\right] = [T_* ad_{\hat{f}}^{k-1}\, \hat{g}^i, T_* ad_{\hat{f}}^{l-1}\, \hat{g}^j]$$
$$= T_*[\tau_k^i\,,\, \tau_l^j],$$

proving that property (ii) holds.

(Sufficiency). Claim 1. The map p in (10) is a smooth retraction $\mathcal{U} \to \mathcal{Z}$.

Proof of Claim 1. Defining recursively, for $j \in 2\!:\!r$,

$$\begin{aligned} \mathsf{p}_1(x) &:= \Phi^{\bar{\tau}_1}_{-\bar{\lambda}_1(x)}(x) \\ \mathsf{p}_j(x) &:= \Phi^{\bar{\tau}_j}_{-\bar{\lambda}_j(x)} \circ \mathsf{p}_{j-1}(x), \end{aligned} \tag{29}$$

we have that $\mathsf{p}(x) = \mathsf{p}_r(x)$.

By assumption (i), we have the following property

$$(\forall x \in \mathcal{U}) \Phi^{\bar{\tau}_j}_t(x) \text{ is defined on } \{t \in \mathbb{R} : \mathsf{H}_{\mathbf{r}}(x) + te^{i_j}_{k_j} \in \mathcal{C}\}. \tag{30}$$

We begin by using (30) with $t = -\bar{\lambda}_1(x)$ to show that p_1 is well-defined. For each $x \in \mathcal{U}$, $\mathsf{H}_{\mathbf{r}}(x) \in \mathcal{C}$, and we have

$$\begin{aligned} &\mathsf{P}_\pi\left(\mathsf{H}_{\mathbf{r}}(x) + (-\bar{\lambda}_1(x))e^{i_1}_{k_1}\right) \\ &\quad = \mathsf{P}_\pi\left(\mathsf{H}_{\mathbf{r}}(x) - L_f^{k_1-1}h^{i_1}(x)e^{i_1}_{k_1}\right) \\ &\quad = \mathrm{col}(0, L_f^{k_2-1}h^{i_2}(x), \cdots, L_f^{k_r-1}h^{i_r}(x)). \end{aligned}$$

Since $\mathsf{H}_{\mathbf{r}}(x) \in \mathcal{C}$, by Assumption 2 the vector above is in $\mathsf{P}_\pi(\mathcal{C})$, implying that $\mathsf{H}_{\mathbf{r}}(x) + (-\bar{\lambda}_1(x))e^{i_1}_{k_1} \in \mathcal{C}$. By (30), we conclude that the map p_1 is well-defined and smooth on $\mathcal{U}$. We next show that the image of p_1 is a subset of $\mathcal{U}$. By Lemma 18, the vector fields τ_k^i satisfy the identity (6) and thus also identity (7), which we rewrite as:

$$L_f^{l-1}h^i \circ \Phi^{\bar{\tau}_j}_t(x) = \begin{cases} L_f^{l-1}h^i(x) + t & i = i_j,\ l = k_j \\ L_f^{l-1}h^i(x) & \text{otherwise.} \end{cases} \tag{31}$$

Applying identity (31) to each of the components of $\mathsf{H}_{\mathbf{r}}$ and using the definition of $\bar{\lambda}_1$ in (11), we get

$$\mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_1(x) = \mathsf{H}_{\mathbf{r}}(x) - e^{i_1}_{k_1} L_f^{k_1-1}h^{i_1}(x),$$

from which it follows that

$$\mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_1(x) = \mathrm{col}(0, L_f^{k_2-1}h^{i_2}(x), \ldots, L_f^{k_r-1}h^{i_r}(x)).$$

Since $\mathsf{H}_{\mathbf{r}}(x) \in \mathcal{C}$, using Assumption 2 we deduce from the identity above that $\mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_1(x) \in \mathsf{P}_\pi(\mathcal{C})$, or equivalently, $\mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_1(x) \in \mathcal{C}$, which implies that $\mathsf{p}_1(x) \in \mathcal{U}$. This proves that $\mathsf{p}_1 : \mathcal{U} \to \mathcal{U}$ is well-defined and smooth.

Now for the induction, suppose that for some $j \in 2\!:\!r-1$, the map $\mathsf{p}_{j-1} : \mathcal{U} \to \mathcal{U}$ is well-defined and smooth, and that letting $\xi = \mathsf{H}_{\mathbf{r}}(x)$, it holds that

$$\mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_{j-1}(x) = \mathrm{col}(0, \ldots, 0, \xi_{l_j}, \ldots, \xi_{l_r}). \tag{32}$$

We want to show that $\mathsf{p}_j : \mathcal{U} \to \mathcal{U}$ is well-defined and smooth and $\mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_j(x) = \mathrm{col}(0, \ldots, 0, \xi_{l_{j+1}}, \ldots, \xi_{l_r})$. Let $w := \mathsf{p}_{j-1}(x) \in \mathcal{U}$, then p_j in (29) can be rewritten as

$$\mathsf{p}_j(x) = \Phi^{\bar{\tau}_j}_{-L_f^{k_j-1}h^{i_j}(x)}(w),$$

and property (32) implies that

$$L_f^{k_j-1}h^{i_j}(w) = \xi_{l_j} = L_f^{k_j-1}h^{i_j}(x), \tag{33}$$

where we used the fact that $l_j = \mathsf{idx}(i_j, k_j)$ and that, according to the notation in Section III, $\xi_{l_j} = \xi^{i_j}_{k_j}$. Using (33), we may rewrite $\mathsf{p}_j(x)$ as

$$\mathsf{p}_j(x) = \Phi^{\bar{\tau}_j}_{-L_f^{k_j-1}h^{i_j}(w)}(w).$$

Using (30) with $t = -L_f^{k_j-1}h^{i_j}(w)$, we now show that p_j is well-defined. By (32) and the first identity in (33) we have

$$\mathsf{P}_\pi\big(\mathsf{H}_{\mathbf{r}}(w) - L_f^{k_j-1}h^{i_j}(w)e^{i_j}_{k_j}\big) = \mathrm{col}(0, \ldots, \xi_{l_{j+1}}, \ldots, \xi_{l_r}),$$

which is in $\mathsf{P}_\pi(\mathcal{C})$ by Assumption 2. By (30), the map p_j is well-defined and smooth on $\mathcal{U}$. Now we show that its image is a subset of $\mathcal{U}$. Indeed, applying again identity (31) to each component of $\mathsf{H}_{\mathbf{r}}$ and using Assumption 2, we get

$$\mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_j(x) = \mathrm{col}(0, \ldots, 0, \xi_{l_{j+1}}, \ldots, \xi_{l_r}) \in \mathsf{P}_\pi(\mathcal{C}),$$

implying that for each $x \in \mathcal{U}$, $\mathsf{H}_{\mathbf{r}} \circ \mathsf{p}_j(x) \in \mathcal{C}$ or, what is the same, $\mathsf{p}_j(x) \in \mathcal{U}$. Thus the image of p_j is contained in $\mathcal{U}$ and $\mathsf{p}_j : \mathcal{U} \to \mathcal{U}$ is well-defined and smooth.

By induction, we conclude that $\mathsf{p} : \mathcal{U} \to \mathcal{U}$ is smooth. By the arguments in [5], [12] reviewed in Section V, the image of p is $\mathcal{Z}$, hence the map $\mathsf{p} : \mathcal{U} \to \mathcal{Z}$ is well-defined and smooth. Finally, for each $x \in \mathcal{Z}$ we have $\bar{\lambda}_j(x) = 0$, $j \in 1\!:\!r$, implying that $\mathsf{p}(x) = x$. Thus $\mathsf{p} : \mathcal{U} \to \mathcal{Z}$ is a smooth retraction, as claimed. △

Claim 2. The map $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$, $x \mapsto (\sigma \circ \mathsf{p}(x), \mathsf{H}_{\mathbf{r}}(x))$ is a diffeomorphism.

Proof of Claim 2. We will show that the map $S : \tilde{\mathcal{Z}} \times \mathcal{C} \to \mathcal{U}$ defined as

$$S(z, \xi) = \Phi^{\bar{\tau}_1}_{\xi_{l_1}} \circ \cdots \circ \Phi^{\bar{\tau}_r}_{\xi_{l_r}} \circ \sigma^{-1}(z)$$

is the smooth inverse of T. Assume for a moment that S is well-defined. Then it is straightforward to check that $T \circ S = \mathrm{id}_{\tilde{\mathcal{Z}} \times \mathcal{C}}$ and $S \circ T = \mathrm{id}_{\mathcal{U}}$. Therefore, we only need to show that S is well-defined and smooth.

For $j \in 2\!:\!r$, define recursively

$$\begin{aligned} S_1(z, \xi) &:= \Phi^{\bar{\tau}_r}_{\xi_{l_r}} \circ \sigma^{-1}(z) \\ S_{j+1}(z, \xi) &:= \Phi^{\bar{\tau}_{r-j}}_{\xi_{l_{r-j}}} \circ S_j(z, \xi), \end{aligned} \tag{34}$$

so that $S(z, \xi) = S_r(z, \xi)$.

Since $\sigma^{-1}(z) \in \mathcal{Z} \subset \mathcal{U}$, $\mathsf{H}_{\mathbf{r}}(\sigma^{-1}(z)) = 0$, and property (30) guarantees that $S_1(z, \xi)$ is well-defined for all ξ_{l_r} such that $\xi_{l_r} e^{i_r}_{k_r} \in \mathcal{C}$. By Assumption 2, for each $\xi \in \mathcal{C}$, $\xi_{l_r} e^{i_r}_{k_r} \in \mathcal{C}$ as well. Thus by (30) the map S_1 is indeed well-defined and smooth on $\tilde{\mathcal{Z}} \times \mathcal{C}$. We show that its image is contained in $\mathcal{U}$. Using (31) r times and the fact that $\sigma^{-1}(z) \in \mathcal{Z} = (\mathsf{H}_{\mathbf{r}})^{-1}(0)$, we get

$$\begin{aligned} &L_f^{k_r-1}h^{i_r} \circ S_1(z, \xi) = \xi_{l_r}, \\ &L_f^{k-1}h^i \circ S_1(z, \xi) = 0,\ (i, k) \in \mathfrak{I}_{\mathbf{r}}, (i, k) \neq (i_r, k_r). \end{aligned}$$

Therefore, for each $(z, \xi) \in \tilde{\mathcal{Z}} \times \mathcal{C}$ we have

$$S_1(z, \xi) \in \{x \in \mathcal{X} : \mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}}(x) = \mathrm{col}(0, \ldots, 0, \xi_{l_r})\} \subset \mathcal{U}.$$

The last inclusion follows from Assumption 2. Thus $S_1 : \tilde{\mathcal{Z}} \times \mathcal{C} \to \mathcal{U}$ is well-defined and smooth.

Now inductively suppose that for some $j \in 1 : r-1$, the map $S_j : \tilde{\mathcal{Z}} \times \mathcal{C} \to \mathcal{U}$ is well-defined and smooth, and that for each $(z, \xi) \in \tilde{\mathcal{Z}} \times \mathcal{C}$,

$$S_j(z,\xi) \in \{x \in \mathcal{X} : \mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}}(x) = \\ = \mathrm{col}(0, \dots, 0, \xi_{l_{r-j+1}}, \dots, \xi_{l_r})\}. \tag{35}$$

Using (30) with $t = \xi_{l_{r-j}}$, the map S_{j+1} in (34) is well-defined for all (z, ξ) such that

$$\mathsf{H}_{\mathbf{r}} \circ S_j(z,\xi) + \xi_{l_{r-j}} e^{i_{r-j}}_{k_{r-j}} \in \mathcal{C}.$$

Using Assumption 2 and (35), we have

$$\mathsf{P}_\pi(\mathsf{H}_{\mathbf{r}} \circ S_j(z,\xi) + \xi_{l_{r-j}} e^{i_{r-j}}_{k_{r-j}}) \\ = \mathrm{col}(0, \dots, 0, \xi_{l_{r-j}}, \dots, \xi_{l_r}) \in \mathsf{P}_\pi(\mathcal{C}),$$

and thus property (30) implies that the map S_{j+1} is indeed well-defined and smooth on $\tilde{\mathcal{Z}} \times \mathcal{C}$.

Using (31) in a manner similar to what we have done for S_1, we get that

$$S_{j+1}(z,\xi) \in \{x \in \mathcal{X} : \mathsf{P}_\pi \circ \mathsf{H}_{\mathbf{r}}(x) \\ = \mathrm{col}(0, \dots, 0, \xi_{l_{r-j}}, \dots, \xi_{l_r})\},$$

and thus by Assumption 2 the image of S_{j+1} is contained in $\mathcal{U}$. By induction, the map $S : \tilde{\mathcal{Z}} \times \mathcal{C} \to \mathcal{U}$ is well-defined and smooth. This concludes the proof of Claim 3. △

Since $T : \mathcal{U} \to \tilde{\mathcal{Z}} \times \mathcal{C}$ is a diffeomorphism, the rest of the theorem follows from part (v) of Theorem 1. ■

References

[1] A. Banaszuk and J. Hauser, "Feedback linearization of transverse dynamics for periodic orbits," *Systems and Control Letters*, vol. 26, no. 2, pp. 95–105, Sept. 1995.

[2] C. I. Byrnes and A. Isidori, "A frequency domain philosophy for nonlinear systems, with applications to stabilization and to adaptive control," in *The 23rd IEEE Conference on Decision and Control*. IEEE, 1984, pp. 1569–1573.

[3] ——, "Global feedback stabilization of nonlinear systems," in *24th IEEE Conference on Decision and Control*. IEEE, 1985, pp. 1031–1037.

[4] ——, "Local stabilization of minimum-phase nonlinear systems," *Systems & Control Letters*, vol. 11, no. 1, pp. 9–17, 1988.

[5] ——, "Asymptotic stabilization of minimum phase nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 36, no. 10, pp. 1122–1137, 1991.

[6] M. D. Di Benedetto, J. Grizzle, and C. Moog, "Rank invariants of nonlinear systems," *SIAM journal on control and optimization*, vol. 27, no. 3, pp. 658–672, 1989.

[7] R. S. D'Souza and C. Nielsen, "An exterior differential characterization of single-input local transverse feedback linearization," *Automatica*, vol. 127, 2021.

[8] M. El-Hawwary and M. Maggiore, "Global path following for the unicycle and other results," in *American Control Conference*, Seattle, USA, June 2008.

[9] ——, "Reduction theorems for stability of closed sets with application to backstepping control design," *Automatica*, vol. 49, pp. 214–222, 2013.

[10] V. Guillemin and A. Pollack, *Differential Topology*. New Jersey: Prentice Hall, 1974.

[11] R. M. Hirschorn, "Invertibility of nonlinear control systems," *SIAM Journal on Control and Optimization*, vol. 17, no. 2, pp. 289–297, 1979.

[12] A. Isidori, *Nonlinear Control Systems, 3rd Edition*. Springer Verlag, 1995.

[13] ——, "The zero dynamics of a nonlinear system: From the origin to the latest progresses of a long successful story," *European Journal of Control*, vol. 19, no. 5, pp. 369–378, 2013.

[14] A. Isidori, A. Krener, C. Gori-Giorgi, and S. Monaco, "Nonlinear decoupling via feedback: a differential geometric approach," *IEEE transactions on automatic control*, vol. 26, no. 2, pp. 331–345, 1981.

[15] A. Isidori and C. H. Moog, "On the nonlinear equivalent of the notion of transmission zeros," in *Modelling and Adaptive Control*. Springer, 1988, pp. 146–158.

[16] W. Kang, "Extended controller form and invariants of nonlinear control systems with a single input," *Journal of Mathematical Systems Estimation and control*, vol. 6, pp. 27–52, 1996.

[17] W. Kang and A. J. Krener, "Extended quadratic controller normal form and dynamic state feedback linearization of nonlinear systems," *SIAM Journal on Control and Optimization*, vol. 30, no. 6, pp. 1319–1337, 1992.

[18] H. K. Khalil, *Nonlinear Systems*, 3rd ed. Prentice Hall, 2002.

[19] J. Lee, *Introduction to Smooth Manifolds*, ser. Graduate Texts in Mathematics. Springer, 2012.

[20] ——, *Introduction to Riemannian manifolds*, ser. Graduate Texts in Mathematics. Springer, 2018, vol. 2.

[21] L. Menini and A. Tornambè, "Linearization through state immersion of nonlinear systems admitting Lie symmetries," *Automatica*, vol. 45, no. 8, pp. 1873–1878, 2009.

[22] P. Morin and C. Samson, "Motion control of wheeled mobile robots." *Springer handbook of robotics*, vol. 1, pp. 799–826, 2008.

[23] C. Nielsen, C. Fulford, and M. Maggiore, "Path following using transverse feedback linearization: Application to a maglev positioning system," *Automatica*, vol. 46, no. 3, pp. 585–590, 2010.

[24] C. Nielsen and M. Maggiore, "Maneuver regulation via transverse feedback linearization: Theory and examples," in *Symposium on Nonlinear Control Systems (NOLCOS)*, Stuttgart, Germany, September 2004.

[25] ——, "Output stabilization and maneuver regulation: A geometric approach," *Systems & Control Letters*, vol. 55, no. 5, pp. 418–427, 2006.

[26] ——, "On local transverse feedback linearization," *SIAM Journal on Control and Optimization*, vol. 47, no. 5, pp. 2227–2250, 2008.

[27] H. Nijmeijer and A. V. der Schaft, *Nonlinear dynamical control systems*. Springer, 1990, vol. 175.

[28] C. Samson, "Path following and time-varying feedback stabilization of a wheeled mobile robot," in *International Conference on Control, Automation, Robotics and Vision*, 1992.

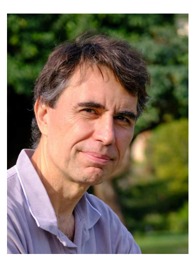
[29] L. Silverman, "Inversion of multivariable linear systems," *IEEE Transactions on automatic control*, vol. 14, no. 3, pp. 270–276, 1969.

[30] S. Singh, "A modified algorithm for invertibility in nonlinear systems," *IEEE Transactions on Automatic Control*, vol. 26, no. 2, pp. 595–598, 1981.

[31] I. A. Tall and W. Respondek, "Feedback classification of nonlinear single-input control systems with controllable linearization: normal forms, canonical forms, and invariants," *SIAM Journal on Control and Optimization*, vol. 41, no. 5, pp. 1498–1531, 2003.

[32] E. Westervelt, J. Grizzle, C. Chevallereau, J. Choi, and B. Morris, *Feedback Control of Dynamic Bipedal Robot Locomotion*. Florida, United States: CRC Press, 2007.

[33] E. Westervelt, J. Grizzle, and D. Koditschek, "Hybrid zero dynamics of planar biped walkers," *IEEE Transactions on Automatic Control*, vol. 48, no. 1, pp. 42–56, January 2003.



Manfredi Maggiore was born in Genoa, Italy. He received the "Laurea" degree in Electronic Engineering in 1996 from the University of Genoa and the PhD degree in Electrical Engineering from the Ohio State University, USA, in 2000. Since 2000 he has been with the Edward S. Rogers Sr. Department of Electrical and Computer Engineering, University of Toronto, Canada, where he is currently Professor. He has been a visiting Professor at the University of Bologna (2007-2008), and the Laboratoire des Signaux et Systèmes, Ecole CentraleSupélec (2015-2016). His research focuses on mathematical nonlinear control and geometric mechanics.