A global steering method for nonholonomic systems ☆

Yacine Chitour [a,b,*], Frédéric Jean [c,b], Ruixing Long [d]

[a] *L2S, Université Paris-Sud XI, CNRS and Supélec, Gif-sur-Yvette, France*
[b] *Team GECO, INRIA Saclay, Ile-de-France, France*
[c] *UMA, ENSTA ParisTech, France*
[d] *Department of Chemistry, Princeton University, Princeton, NJ 08544, USA*

A B S T R A C T

In this paper, we present an iterative steering algorithm for nonholonomic systems (also called driftless control-affine systems) and we prove its global convergence under the sole assumption that the Lie Algebraic Rank Condition (LARC) holds true everywhere. That algorithm is an extension of the one introduced in Jean et al. (2005) [21] for regular systems. The first novelty here consists in the explicit algebraic construction, starting from the original control system, of a lifted control system which is regular. The second contribution of the paper is an exact motion planning method for nilpotent systems, which makes use of sinusoidal control laws and which is a generalization of the algorithm described in Murray and Sastry (1993) [29] for chained-form systems.

© 2012 Elsevier Inc. All rights reserved.

Contents

* Corresponding author at: L2S, Université Paris-Sud XI, CNRS and Supélec, Gif-sur-Yvette, France.
   E-mail addresses: yacine.chitour@lss.supelec.fr (Y. Chitour), frederic.jean@ensta-paristech.fr (F. Jean), rlong@princeton.edu (R. Long).

1. Introduction

Nonholonomic systems have been attracting the attention of the scientific community for several years, due to the theoretical challenges they offer and the numerous important applications they cover. From the point of view of control theory, a nonholonomic system is a driftless control-affine system and is written as

$$(\Sigma) \quad \dot{x} = \sum_{i=1}^{m} u_i X_i(x), \quad x \in \Omega, \ u = (u_1, \ldots, u_m) \in \mathbb{R}^m, \tag{1.1}$$

where Ω is an open connected subset of $\mathbb{R}^n$, and $X_1, \ldots, X_m$ are C^∞ vector fields on Ω. Admissible inputs are $\mathbb{R}^m$-valued measurable functions $u(\cdot)$ defined on some interval $[0, T]$ and a trajectory of (Σ), corresponding to some $x_0 \in \Omega$ and to an admissible input $u(\cdot)$, is the (maximal) solution $x(\cdot)$ in Ω of the Cauchy problem defined by $\dot{x}(t) = \sum_{i=1}^{m} u_i(t) X_i(x(t))$, $t \in [0, T]$, and $x(0) = x_0$.

In this paper, we address the *motion planning problem* (MPP for short) for (Σ), namely determine a procedure which associates with every pair of points $(p, q) \in \Omega \times \Omega$ an admissible input $u(\cdot)$ defined on some interval $[0, T]$, such that the corresponding trajectory of (Σ) starting from p at $t = 0$ reaches q at $t = T$. As for the existence of a solution to MPP, this is equivalent to the complete controllability of (Σ). After the works of Chow and Rashevsky in the thirties [14,32], and that of Sussmann and Stefan in the seventies [36,35], the issue of complete controllability for nonholonomic systems is well-understood and it is usually guaranteed by assuming that the Lie Algebraic Rank Condition (also known as the Hörmander condition) holds for (Σ). This easily checkable condition is not only sufficient for complete controllability but also necessary when the vector fields are analytic. From a practical viewpoint, assuming the LARC is, in a sense, the minimal requirement to ensure complete controllability for (Σ) and this is what we will do for all the control systems considered hereafter.

As for the construction of the solutions of the MPP, we present, in this paper, a complete procedure solving the MPP for a nonholonomic system subject to the sole LARC. By "complete procedure", we mean that the following properties must be guaranteed by the proposed procedure.

1. Global character of the algorithm: for every pair of points (p, q) in Ω, the algorithm must produce a steering control. (Note that the core of many algorithms consists in a local procedure and turning the latter into a global one is not always a trivial issue.)
2. Proof of convergence of the algorithm.
3. Regarding numerical implementations, no prohibitive limitation on the state dimension n.
4. Usefulness for practical applications, e.g., robustness with respect to the dynamics, "nice" trajectories produced by the algorithm (no cusps neither large oscillations), and possibility of localizing the algorithm in order to handle obstacles (i.e., reducing the working space Ω to any smaller open and connected subset of $\mathbb{R}^n$).

There exist of course several algorithms addressing the MPP in different contexts but most of them fail to verify all the aforementioned properties.

At first, in the case of specific classes of driftless nonholonomic systems (i.e. where more is known than the sole LARC), effective techniques have been proposed, among which a Lie bracket method for steering nilpotentizable systems (see [24] and [25]), sinusoidal controls for chained-form systems (see [29]), averaging techniques for left-invariant systems defined on a Lie group (see [26,5]), and a trajectory generation method for flat systems (see [15]). Depending on the applications, these methods turn out to be extremely efficient, especially when the system to be steered is shown to be flat with an explicit flat output.

However, the class of systems considered previously is rather restrictive: for 2-input nonholonomic systems (i.e. $m = 2$), under suitable regularity assumptions, a flat system admits a feedback chained-form transformation (cf. [30,28]) and thus is exactly nilpotentizable; on the other hand, when the dimension of the state space is large enough, exact nilpotentizability is clearly a non-generic property among 2-input nonholonomic systems. Moreover, there exist standard nonholonomic systems whose kinematic model does not fall into any of the aforementioned categories. For instance, mobile robots with more than one trailer cannot be transformed in chained-form unless each trailer is hinged to the midpoint of the previous wheel axle, an unusual situation in real vehicles. Another similar example is the rolling-body problem: even the simplest model in this category, the so-called plate–ball system, does not allow any chained-form transformation and is not flat.

Regarding general nonholonomic systems, various steering techniques have been proposed in the literature, and we only mention three of them: the iterative method, the generic loop method, and the continuation method. The first one, introduced in [24] and improved in [25], is an approximation procedure which is exact for nilpotent systems. This method is proved to be convergent with the sole assumption of the LARC and actually meets most of the requirements to be a complete procedure in the sense defined previously. However, either the resulting trajectories in [25] contain a large number of cusps (exponential with respect to the degree of nonholonomy), or the computation of the steering control in [24] requires the inversion of a system of algebraic equations. The latter turns out to be numerically intractable as soon as the dimension of the state is larger than six. Let us also mention a less important limitation for practical use. The iterative method described in [24,25] makes use of several nonlinear changes of coordinates, which must be performed by numerical integration of ODEs at each step of the iterative method, thus leading to spurious on-line computations.

The generic loop method, presented in [34], is based on a local deformation procedure and requires an a priori estimate of some "critical distance" which is, in general, an unknown parameter in practice. That fact translates into a severe drawback for constructing a globally valid algorithm. The continuation method of [39] and [10] belongs to the class of Newton-type methods. See [2] for the application of this method to the rolling-body problem. Proving its convergence amounts to show the global existence for the solution of a nonlinear differential equation, which relies on handling the abnormal extremals associated to the control system. That latter issue turns out to be a hard one, see [11,12] for instance. This is why, in the current state of knowledge, the continuation method can be proved to converge only under restrictive assumptions (see [8,9,13]).

The algorithm considered in the present paper takes as starting point the globally convergent algorithm for steering *regular* nonholonomic systems discussed in [21]. As the iterative method of [24,25], that algorithm can be casted in the realm of nonlinear geometric control and roughly works as follows: one first solves the motion planning problem for a control system which is nilpotent and

"approximates" system (1.1) in a suitable sense; then, one applies the resulting input $\hat{u}$ to (1.1) and iterates the procedure from the current point. If we use $\hat{x}(t, a, \hat{u})$, $t \in [0, T]$, to denote a trajectory of the "approximate" control system starting from a, a local version of this algorithm is summarized below, where d is an appropriate distance (to be defined in the next section) and e is a fixed positive real number.

Algorithm 1 Local steering algorithm.

Require: x_0, x_1, e
    $k = 0$;
    $x^k = x_0$;
    while $d(x^k, x_1) > e$ **do**
        Compute $\hat{u}^k$ such that $x_1 = \hat{x}(T, x^k, \hat{u}^k)$;
        $x^{k+1} = \text{AppSteer}(x^k, x_1) := x(T, x^k, \hat{u}^k)$;
        $k = k + 1$;
    end while

We note that Algorithm 1 converges *locally* provided that the function AppSteer is *locally contractive* with respect to the distance d, i.e., for $x_1 \in \Omega$, there exist $\varepsilon_{x_1} > 0$ and $c_{x_1} < 1$ such that

$$d\big(\text{AppSteer}(x, x_1), x_1\big) \leqslant c_{x_1} d(x_1, x), \tag{1.2}$$

for $x \in \Omega$ and $d(x_1, x) < \varepsilon_{x_1}$.

Assume now that we have a *uniformly locally contractive* function AppSteer on a connected compact set $K \subset \Omega$, i.e. there exist $\varepsilon_K > 0$ and $c_K \in (0, 1) < 1$ such that

$$d\big(\text{AppSteer}(x, x_1), x_1\big) \leqslant c_K d(x_1, x), \tag{1.3}$$

for $x, x_1 \in K$ and $d(x_1, x) < \varepsilon_K$.

Based on the local algorithm, a global approximate steering algorithm on K can be built along the line of the following idea (a similar procedure is proposed in [25]): consider a parameterized path $\gamma \subset K$ connecting x_0 to x_1. Then pick a finite sequence of intermediate goals $\{x_0^d = x_0, x_1^d, \ldots, x_j^d = x_1\}$ on γ such that $d(x_{i-1}^d, x_i^d) < \varepsilon_K/2$, $i = 0, \ldots, j$. One can prove that the iterated application of a uniformly locally contractive $\text{AppSteer}(x^{i-1}, x_i^d)$ from the current state to the next subgoal (having set $x_i^d = x_1$, for $i \geqslant j$) yields a sequence x^i converging to x_1.

To turn the above idea into a practically efficient algorithm, three issues must be successfully addressed:

(P-1) Construct a *uniformly* locally contractive local approximate steering method.
(P-2) The "approximate" control $\hat{u}^k$ must be *exact* for steering the "approximate system" from the current point x^k to the final point x_1. As this computation occurs at each iteration, it must be performed in a reasonable time.
(P-3) Since the knowledge of the "critical distance" ε_K is not available in practice, the algorithm should achieve global convergence without explicit knowledge of ε_K.

Issues (P-1) and (P-3) are solved in [21] under the assumption that the control system is *regular* (cf. Definition 2.9 below). The solution proposed therein relies on the understanding of the geometry defined by the nonholonomic system (cf. [4]). That geometry is a sub-Riemannian one and it endows the working space Ω with a sub-Riemannian metric d for which the aforementioned function AppSteer is contractive. Moreover, the approximation of the original system adapted to the motion planning turns out to be the approximation at the first order with respect to d (cf. [21]). However, the regularity assumption for the control system is rather restrictive: in general, nonholonomic systems do exhibit singularities (cf. [41]). A solution also exists in the case of a non-regular control system [40], but only when the state dimension n is less than or equal to 5. In the present paper, we completely remove the regularity assumption and solve issues (P-1) and (P-3) for every nonholonomic control system. The solution is based on an explicit desingularization procedure: adding new variables

(thus augmenting the dimension of the state space), we construct a "lifted" control system which is regular and whose projection is the original control system. The solution of issue (P-1) described in [21] can thus be applied to the "lifted" control system, as well as the globally convergent motion planning algorithm solving issue (P-3) proposed therein. Note that other desingularization procedures already exist [3,16,20,33], but we insist on the fact that the one we propose here involves only *explicit* polynomial transformations. It numerically translates to the fact that these changes of variables can be performed off-line once each local procedure is identified.

As regards issue (P-2), several algorithms were proposed for computing $\hat{u}$, i.e. for controlling nilpotent systems. In [24], the authors make use of piecewise constant controls and obtain smooth controls by imposing some special parameterization (namely by requiring the control system to stop during the control process). In that case, the smoothness of the inputs is recovered by using a reparameterization of the time, which cannot prevent in general the occurrence of cusps or corners for the corresponding trajectories. However, smoothness of the *trajectories* is generally mandatory for robotic applications. Therefore, the method proposed in [24] is not adapted to such applications. In [25], the proposed controls are polynomial (in time), but an algebraic system must be inverted in order to access to these inputs. The size and the degree of this algebraic system increase exponentially with respect to the dimension of state space, and there does not exist a general efficient exact method to solve it. Even the existence of solutions is a nontrivial issue. Furthermore, the methods [24] and [25] both make use of exponential coordinates which are not explicit and thus require in general numerical integrations of nonlinear differential equations. That prevents the use of these methods in an iterative scheme such as Algorithm 1. Let us also mention the path approximation method by Liu and Sussmann [27], which uses unbounded sequences of sinusoids. Even though this method bears similar theoretical aspects with our method (see especially the argument strategy in order to prove Lemma 5.8, which is borrowed from [27]), it is not adapted from a numerical point of view to the motion planning issue since it relies on a limit process of highly oscillating inputs. In the present paper, we present an *exact* steering algorithm for general nilpotent systems, which uses sinusoidal inputs and which can be applied for controlling the approximate (nilpotent) system used in [21]. Our method generalizes the one proposed in [29] for controlling chained-form systems, which is briefly recalled next: after having brought the system under a "canonical" form, the authors of [29] proceed by controlling component after component by using, for each component, two sinusoids with suitably chosen frequencies. In the present paper, we show for general nilpotent systems that, with more frequencies for each component, one can steer an arbitrary component independently on the other components. We are also able to construct inputs which give rise to C^1-trajectories.

We now describe in a condensed manner the global motion planning strategy developed in this paper. The latter is presented as an algorithmic procedure associated with a given nonholonomic system (Σ) defined on $\Omega \subset \mathbb{R}^n$. The required inputs are initial and final points x^{initial} and x^{final} belonging to Ω, a tolerance $e > 0$, and a compact convex set $K \subset \Omega$ (of appropriate size) equal to the closure of its interior which is a neighborhood of both x^{initial} and x^{final}. For instance, K can be chosen to be a large enough compact tubular neighborhood constructed around a curve joining x^{initial} and x^{final}. The global steering method is summarized in Algorithm 2.

The paper is devoted to the construction of the various steps of this algorithm. We will also show that each of these steps is conceived so that the overall construction is a complete procedure in the sense defined previously. In particular, the convergence issue is addressed in the following theorem.

Theorem 1.1. *Let (Σ) be a nonholonomic system on $\Omega \subset \mathbb{R}^n$ satisfying the LARC. For every $e > 0$, every connected compact set K which is equal to the closure of its interior, and every pair of points $(x^{\text{initial}}, x^{\text{final}})$ in the interior of K, Algorithm 2 steers, in a finite number of steps, the control system (Σ) from x^{initial} to a point $x \in K$ such that $d(x, x^{\text{final}}) < e$.*

Before providing the structure of the paper, we mention possible extensions of our algorithm. The first one concerns the working space Ω. Since it is an arbitrary open connected set of $\mathbb{R}^n$, one can extend the algorithm to the case where the working space is a smooth connected manifold of finite dimension. From a numerical point of view, there would be the additional burden of computing the charts. A second extension deals with the stabilization issue. Indeed, at the heart of the algorithm

---

Algorithm 2 Global approximate steering algorithm: Global(x^{initial}, x^{final}, e, K).

1: Build a decomposition of K into a finite number of compact sets $\mathcal{V}^c_{\mathcal{J}_i}$, with $i = 1, \ldots, M$ (Section 3.3).

2: Construct the connectedness graph $\mathsf{G} := (\mathsf{N}, \mathsf{E})$ associated with this decomposition and choose a simple path $\mathsf{p} := \{j_0, j_1, \ldots, j_{\overline{M}}\}$ in G such that $x^{\text{initial}} \in \mathcal{V}^c_{\mathcal{J}_{j_0}}$ and $x^{\text{final}} \in \mathcal{V}^c_{\mathcal{J}_{j_{\overline{M}}}}$ (Section 3.3).

3: Choose a sequence $(x^i)_{i=1,\ldots,\overline{M}-1}$ such that $x^i \in \mathcal{V}^c_{\mathcal{J}_{j_i}} \cap \mathcal{V}^c_{\mathcal{J}_{j_{i+1}}}$.

4: Set $x := x^{\text{initial}}$.

5: **for** $i = 1, \ldots, \overline{M} - 1$ **do**

6:     Apply the Desingularization Algorithm at $a := x^i$ with $\mathcal{J} := \mathcal{J}_i$ (Section 3.3).
       {the output is an m-tuple of vector fields ξ on $\mathcal{V}_{\mathcal{J}_i} \times \mathbb{R}^{\tilde{n}}$ which is free up to step r.}

7:     Let AppSteer be the LAS method associated to the approximation $\mathcal{A}^\xi$ of ξ on $\mathcal{V}_{\mathcal{J}_i} \times \mathbb{R}^{\tilde{n}}$ defined in Section 4.1 and to its steering law Exact$_{m,r}$ constructed in Section 5.3.

8:     Set $\tilde{x}_0 := (x, 0)$, $\tilde{x}_1 := (x^i, 0)$, and $\mathcal{V}^c := \mathcal{V}^c_{\mathcal{J}_i} \times \overline{B}_R(0)$ with $R > 0$ large enough.

9:     Apply GlobalFree($\tilde{x}_0$, $\tilde{x}_1$, e, $\mathcal{V}^c$, AppSteer) to ξ (Section 4.2).
       {the algorithm stops at a point $\tilde{x}$ which is e-close to $\tilde{x}_1$;}

10:      **return** $x := \pi(\tilde{x})$.
       {$\pi : \mathcal{V}_{\mathcal{J}_i} \times \mathbb{R}^{\tilde{n}} \to \mathcal{V}_{\mathcal{J}_i}$ is the canonical projection.}

11: **end for**

---

lies an iterative procedure such as Algorithm 1, which can be easily adapted for stabilization tasks (cf. [31]). Another possible generalization takes advantage of devising from our algorithm a globally regular input, one can then address the motion planning of dynamical extensions of the nonholonomic control systems considered in the present paper. Finally, let us point out the modular nature of Algorithm 2: one can propose other approaches to obtain uniformly contractive local methods (other desingularization methods or different ways of dealing with singular points), or replace Exact$_{m,r}(\cdot)$ by more efficient control strategies for general nilpotent systems.

The paper is organized as follows. In Section 2, we define properly the notion of first-order approximation. We then propose, in Section 3, a purely polynomial desingularization procedure based on a *lifting* method. In Section 4, we describe in detail the globally convergent steering algorithm given in [21] for regular systems together with a proof of convergence. In Section 5, we present an exact steering procedure for general nilpotent systems using sinusoids, and we gather, in Appendices A–C, the proof of Theorem 1.1 and some additional comments.

2. Notations and definitions

Let n and m be two positive integers. Let Ω and $VF(\Omega)$ be respectively an open connected subset of $\mathbb{R}^n$ and the set of C^∞ vector fields on Ω. Consider m vector fields $X_1, \ldots, X_m$ of $VF(\Omega)$, and the associated driftless control-affine nonholonomic system given by

$$\dot{x} = \sum_{i=1}^m u_i X_i(x), \quad x \in \Omega, \tag{2.1}$$

where $u = (u_1, \ldots, u_m) \in \mathbb{R}^m$ and the input $u(\cdot) = (u_1(\cdot), \ldots, u_m(\cdot))$ is an integrable vector-valued function defined on $[0, T]$, with T a fixed positive real number.

We also assume that (2.1) is *complete*, i.e., for every $a \in \Omega$ and input $u(\cdot)$, the Cauchy problem defined by (2.1) starting from a at $t = 0$ and corresponding to $u(\cdot)$ admits a unique (absolutely continuous) solution $x(\cdot, a, u)$ defined on $[0, T]$ and called the trajectory of (2.1) starting from a at $t = 0$ and corresponding to the input $u(\cdot)$. A point $x \in \Omega$ is said to be *accessible* from a if there exist an input $u : [0, T] \to \mathbb{R}^m$ and a time $t \in [0, T]$ such that $x = x(t, a, u)$. Then, system (2.1) is said to be completely controllable if any two points in Ω are accessible one from each other (see [38]).

We next provide a classical condition ensuring that system (2.1) is controllable. We first need the following definition.

Definition 2.1 (*Lie Algebraic Rank Condition (LARC)*). Let $L(X)$ be the Lie algebra generated by the vector fields $X_1, \ldots, X_m$ and $L(x)$ be the linear subspace of $\mathbb{R}^n$ equal to the evaluation of $L(X)$ at every point

$x \in \Omega$ (see [3]). If $L(x) = \mathbb{R}^n$, we say that the *Lie Algebraic Rank Condition* (LARC for short) is verified at $x \in \Omega$. If this is the case at every point $x \in \Omega$, we say that system (2.1) satisfies the LARC.

Chow's Theorem essentially asserts that, if system (2.1) satisfies the LARC then it is completely controllable (cf. [14]).

Remark 2.1. For the sake of clarity, we assume through this paper that the control set is equal to $\mathbb{R}^m$. However, it is well known that Chow's Theorem only requires that the convex hull of the control set contains a neighborhood of the origin in $\mathbb{R}^m$ (see for instance [23, Chapter 4, Theorem 2]). We will explain later how we can adapt our method to the case with constraints on the control set (see Appendix B). Moreover, it is worth recalling that complete controllability for (Σ) does not imply that LARC holds true for (Σ) if the vector fields $X_1, \ldots, X_m$ are only smooth, but this is the case if $X_1, \ldots, X_m$ are analytic (cf. [1, Chapter 5]).

Throughout this paper, we will only consider driftless control-affine nonholonomic systems of the type (2.1) verifying the LARC, and thus completely controllable. In that context, the *motion planning problem* will be defined as follows: find a procedure which furnishes, for every two points $x_0, x_1 \in \Omega$, an input u steering (2.1) from x_0 to x_1, i.e., $x(T, x_0, u) = x_1$.

Our solution to this problem relies heavily on the underlying geometry, which is a *sub-Riemannian geometry*. We provide in Section 2.1 the useful definitions and refer the reader to [3] for more details. We then introduce in Section 2.2 a notion of approximate steering method related to this geometry.

2.1. Basic facts on sub-Riemannian geometry

2.1.1. Sub-Riemannian distance and nonholonomic order

Definition 2.2 *(Length of an input).* The *length of an input* u is defined by

$$\ell(u) = \int_0^T \sqrt{u_1^2(t) + \cdots + u_m^2(t)} \, dt,$$

and the *length of a trajectory* $x(\cdot, a, u)$ is defined by $\ell(x(\cdot, a, u)) := \ell(u)$.

The appropriate notion of distance associated with the control system (2.1) and closely related to the notion of accessibility is that of *sub-Riemannian distance*, also called *control distance*.

Definition 2.3 *(Sub-Riemannian distance).* The vector fields $X_1, \ldots, X_m$ induce a function d on Ω, defined by $d(x_1, x_2) := \inf_u \ell(x(\cdot, x_1, u))$, for every points x_1, x_2 in Ω, where the infimum is taken over all the inputs u such that $x(\cdot, x_1, u)$ is defined on $[0, T]$ and $x(T, x_1, u) = x_2$. We say that the function d is the *sub-Riemannian distance* associated with $X_1, \ldots, X_m$.

Remark 2.2. The function d defined above is a *distance* in the usual sense, i.e., it verifies (i) $d(x_1, x_2) \geqslant 0$ and $d(x_1, x_2) = 0$ if and only if $x_1 = x_2$; (ii) symmetry: $d(x_1, x_2) = d(x_2, x_1)$; (iii) triangular inequality: $d(x_1, x_3) \leqslant d(x_1, x_2) + d(x_2, x_3)$. Notice that one always has $d(x_1, x_2) < \infty$ since the control system is assumed to be completely controllable.

Definition 2.4 *(Nonholonomic derivatives of a function).* If $f : \Omega \to \mathbb{R}^n$ is a smooth function, the *first-order nonholonomic derivatives* of f are the Lie derivatives $X_i f$ of f along X_i, $i = 1, \ldots, m$. Similarly, $X_i(X_j f)$, $i, j = 1, \ldots, m$, are called the *second-order nonholonomic derivatives* of f, and more generally, $X_{i_1} \cdots X_{i_k} f$, $i_1, \ldots, i_k \in \{1, \ldots, m\}$, are the *kth-order nonholonomic derivatives* of f, where k is any positive integer.

Proposition 2.1. *(See [3, Proposition 4.10, p. 34].) Let s be a non-negative integer. For a smooth function f defined near a ∈ Ω, the following conditions are equivalent:*

(i) $f(x) = O(d^s(x, a))$ *for x in a neighborhood of a;*
(ii) *all the nonholonomic derivatives of order $\leqslant s - 1$ of f vanish at a.*

Definition 2.5 *(Nonholonomic order of a function).* Let s and f be respectively a non-negative integer and a smooth real-valued function defined on Ω. If condition (i) or (ii) of Proposition 2.1 holds, we say that f is of *order $\geqslant s$* at $a \in \Omega$. If f is of order $\geqslant s$ but not of order $\geqslant s + 1$ at a, we say that f is of *order s* at a. The order of f at a will be denoted by $\operatorname{ord}_a(f)$.

Definition 2.6 *(Nonholonomic order of a vector field).* Let q be an integer. A vector field $Y \in VF(\Omega)$ is of *order $\geqslant q$* at $a \in \Omega$ if, for every non-negative integer s and every smooth function f of order s at a, the Lie derivative Yf is of order $\geqslant q + s$ at a. If Y is of order $\geqslant q$ but not $\geqslant q + 1$, it is of *order q* at a. The order of Y at a will be denoted by $\operatorname{ord}_a(Y)$.

Definition 2.7 *(Nonholonomic first-order approximation at a).* An m-tuple $\hat{X}^a := \{\hat{X}_1^a, \ldots, \hat{X}_m^a\}$, defined on $B(a, \rho_a) := \{x \in \Omega,\ d(x, a) \leqslant \rho_a\}$ with $\rho_a > 0$, is said to be a *nonholonomic first-order approximation of* $X := \{X_1, \ldots, X_m\}$ *at* $a \in \Omega$, if the vector fields $X_i - \hat{X}_i^a$, for $i = 1, \ldots, m$, are of order $\geqslant 0$ at a. The positive number ρ_a is called *the approximate radius at a.*

Remark 2.3. As a consequence of Definition 2.7, one gets that the nonholonomic order at a defined by the vector fields $\hat{X}_1^a, \ldots, \hat{X}_m^a$ coincides with the one defined by $X_1, \ldots, X_m$.

2.1.2. Privileged coordinates

The changes of coordinates take an important place in this paper, whether it is to estimate the sub-Riemannian distance, or to compute the order of functions and vector fields, or to transform a control system into a normal form. To avoid heavy notations, we will need some conventions and simplifications that we fix now for the rest of the paper.

A point in $\Omega \subset \mathbb{R}^n$ is denoted by $x = (x_1, \ldots, x_n)$ and the canonical basis of $\mathbb{R}^n$ by $(\partial_{x_1}, \ldots, \partial_{x_n})$. Even though x is a point, we will sometimes refer to $(x_1, \ldots, x_n)$ as the original coordinates. A *system of local coordinates* $y = (y_1, \ldots, y_n)$ at a point $a \in \Omega$ is defined as a diffeomorphism φ between an open neighborhood $N_a \subset \Omega$ of a and an open neighborhood $N_{\varphi(a)} \subset \mathbb{R}^n$ of $\varphi(a)$, $\varphi : x \mapsto y = (y_1, \ldots, y_n)$. If the diffeomorphism φ is defined on Ω, then $y = (y_1, \ldots, y_n)$ is said to be a *system of global coordinates* on Ω. A system of global coordinates is said to be *affine* (resp. *linear*) if the corresponding diffeomorphism φ is affine (resp. linear). If f is a function defined on N_a, the function $f \circ \varphi^{-1}$ defined on $N_{\varphi(a)}$ will be called f *(expressed) in coordinates* $(y_1, \ldots, y_n)$. If $X \in VF(\Omega)$ is a vector field, the push-forward $\varphi_* X = d\varphi \circ X \circ \varphi^{-1} \in VF(N_{\varphi(a)})$ will be called X *(expressed) in coordinates* $(y_1, \ldots, y_n)$.

For the sake of simplicity, we will in general not introduce the notation φ and, with a slight abuse of the notation, replace it by y. Thus we write $y(x)$ or $(y_1(x), \ldots, y_n(x))$ instead of $\varphi(x)$. The function $f \circ \varphi^{-1}$ will be denoted by $f(y)$, and the vector field $\varphi_* X$ by $X(y)$ or $y_* X$. The values at a point $\bar{y} \in N_{\varphi(a)}$ will be denoted respectively by $f(y)_{|y=\bar{y}}$ and $X(y)_{|y=\bar{y}}$.

A special class of coordinates, called *privileged coordinates* and defined below, turns out to be a useful tool to compute the order of functions and vector fields, and to estimate the sub-Riemannian distance d.

We will use $L^s(X)$ to denote the Lie sub-algebra of elements of length (cf. Definitions 3.1 and 3.4) not greater than $s \in \mathbb{N}$. Take $x \in \Omega$ and let $L^s(x)$ be the vector space generated by the values at x of elements belonging to $L^s(X)$. Since system (2.1) verifies the LARC at every point $x \in \Omega$, there exists a smallest integer $r := r(x)$ such that $\dim L^r(x) = n$. This integer is called the *degree of nonholonomy* at x.

Definition 2.8 *(Growth vector).* For $a \in \Omega$, let $n_s(a) := \dim L^s(a)$, $s = 1, \ldots, r$. The sequence $(n_1(a), \ldots, n_r(a))$ is the *growth vector* of X at a.

Definition 2.9 *(Regular and singular points).* A point $a \in \Omega$ is said to be *regular* if the growth vector remains constant in a neighborhood of a and, otherwise, a is said to be *singular*. The nonholonomic system (2.1) (or the m-tuple X) is said to be *regular* if every point in Ω is regular.

Note that regular points form an open and dense set in Ω.

Definition 2.10 *(Weight).* For $a \in \Omega$ and $j = 1, \ldots, n$, let $w_j := w_j(a)$ be the integer defined by setting $w_j := s$ if $n_{s-1} < j \leqslant n_s$, with $n_s := n_s(a)$ and $n_0 := 0$. The integers w_j, for $j = 1, \ldots, n$ are called the *weight at a*.

Remark 2.4. The meaning of Definition 2.10 can be understood in another way. Choose first some vector fields $W_1, \ldots, W_{n_1}$ in $L^1(X)$ such that $W_1(a), \ldots, W_{n_1}(a)$ form a basis of $L^1(a)$. Choose then other vector fields $W_{n_1+1}, \ldots, W_{n_2}$ in $L^2(X)$ such that $W_1(a), \ldots, W_{n_2}(a)$ form a basis of $L^2(a)$ and, for every positive integer s, choose $W_{n_{s-1}+1}, \ldots, W_{n_s}$ in $L^s(X)$ such that $W_1(a), \ldots, W_{n_s}(a)$ form a basis of $L^s(a)$. We obtain in this way a sequence of vector fields $W_1, \ldots, W_n$ such that

$$\begin{cases} W_1(a), \ldots, W_n(a) \text{ is a basis of } \mathbb{R}^n, \\ W_i \in L^{w_i}(X), \quad i = 1, \ldots, n. \end{cases} \tag{2.2}$$

A sequence of vector fields verifying Eq. (2.2) is called an *adapted frame at a*. The word "adapted" means "adapted to the flag $L^1(a) \subset L^2(a) \subset \cdots \subset L^r(a) = \mathbb{R}^n$", since the values at a of an adapted frame contain a basis $W_1(a), \ldots, W_{n_s}(a)$ of every subspace $L^s(a)$ of the flag. The values of $W_1, \ldots, W_n$ at a point b close to a also form a basis of $\mathbb{R}^n$. However, if a is singular, this basis may be not adapted to the flag $L^1(b) \subset L^2(b) \subset \cdots \subset L^{r(b)}(b) = \mathbb{R}^n$.

Definition 2.11 *(Privileged coordinates at a).* A *system of privileged coordinates* at $a \in \Omega$ is a system of local coordinates $(z_1, \ldots, z_n)$ centered at a (the image of a is 0) such that $\mathrm{ord}_a(z_j(x)) = w_j$, for $j = 1, \ldots, n$.

Remark 2.5. For every system of local coordinates $(y_1, \ldots, y_n)$ centered at a, we have, up to a re-ordering, $\mathrm{ord}_a(y_j) \leqslant w_j$ or, without re-ordering, $\sum_{j=1}^n \mathrm{ord}_a(y_j) \leqslant \sum_{j=1}^n w_j$.

The order at $a \in \Omega$ of functions and vector fields expressed in a system of privileged coordinates $(z_1, \ldots, z_n)$ centered at a can be evaluated algebraically as follows:

- the order of the monomial $z_1^{\alpha_1} \cdots z_n^{\alpha_n}$ is equal to its *weighted degree* $w(\alpha) := w_1 \alpha_1 + \cdots + w_n \alpha_n$;
- the order of a function $f(z)$ at $z = 0$ is the least weighted degree of the monomials occurring in the Taylor expansion of $f(z)$ at 0;
- the order of the monomial vector field $z_1^{\alpha_1} \cdots z_n^{\alpha_n} \partial_{z_j}$ is equal to its *weighted degree* $w(\alpha) - w_j$, where one assigns the weight $-w_j$ to ∂_{z_j} at 0;
- the order of a vector field $W(z) = \sum_{j=1}^n W_j(z) \partial_{z_j}$ at $z = 0$ is the least weighted degree of the monomials occurring in the Taylor expansion of W at 0.

Definition 2.12 *(Continuously varying system of privileged coordinates).* A *continuously varying system of privileged coordinates on Ω* is a mapping Φ taking values in $\mathbb{R}^n$, defined and continuous on a neighborhood of the set $\{(x, x), \ x \in \Omega\} \subset \Omega \times \Omega$, and so that the partial mapping $z := \Phi(a, \cdot)$ is a system of privileged coordinates at a. In this case, there exists a continuous function $\bar{\rho} : \Omega \to (0, +\infty)$ such that the coordinates $\Phi(a, \cdot)$ are defined on $B(a, \bar{\rho}(a))$. We call $\bar{\rho}$ an *injectivity radius function* of Φ.

Definition 2.13 *(Pseudo-norm).* Let $a \in \Omega$ and $w_1, \ldots, w_n$ be the weights at a. The application from $\mathbb{R}^n$ to $\mathbb{R}$ defined by $\|z\|_a := |z_1|^{1/w_1} + \cdots + |z_n|^{1/w_n}$, $z = (z_1, \ldots, z_n) \in \mathbb{R}^n$, is called *the pseudo-norm at a*.

2.1.3. Distance and error estimates

Privileged coordinates provide estimates of the sub-Riemannian distance d, according to the following result.

Theorem 2.2 *(Ball–Box Theorem). (See [3].) Consider $(X_1, \ldots, X_m) \in VF(\Omega)^m$, a point $a \in \Omega$, and a system of privileged coordinates z at a. There exist positive constants $C_d(a)$ and $\varepsilon_d(a)$ such that, for every $x \in \Omega$ with $d(a, x) < \varepsilon_d(a)$, one has*

$$\frac{1}{C_d(a)} \big\| z(x) \big\|_a \leqslant d(a, x) \leqslant C_d(a) \big\| z(x) \big\|_a. \tag{2.3}$$

If Ω contains only regular points and if Φ is a continuously varying system of privileged coordinates on Ω, then there exist continuous positive functions $C_d(\cdot)$ and $\varepsilon_d(\cdot)$ on Ω such that Eq. (2.3) holds true with $z = \Phi(a, \cdot)$ at all (x, a) satisfying $d(x, a) < \varepsilon_d(a)$.

Corollary 2.3. *Let K be a compact subset of Ω. Assume that K only contains regular points and there exists a continuously varying system of privileged coordinates Φ on K. Then, there exist positive constants C_K and ε_K such that, for every pair $(a, x) \in K \times K$ verifying $d(a, x) < \varepsilon_K$, one has*

$$\frac{1}{C_K} \big\| \Phi(a, x) \big\|_a \leqslant d(a, x) \leqslant C_K \big\| \Phi(a, x) \big\|_a. \tag{2.4}$$

Privileged coordinates also allow one to measure the error obtained when X is replaced by an approximation $\hat{X}$.

Proposition 2.4. *(See [3, Proposition 7.29].) Consider a point $a \in \Omega$, a system of privileged coordinates z at a, and an approximation $\hat{X}$ of X at a. Then, there exist positive constants $C_e(a)$ and $\varepsilon_e(a)$ such that, for every $x \in \Omega$ with $d(a, x) < \varepsilon_e(a)$ and every integrable input function $u(\cdot)$ with $\ell(u) < \varepsilon_e(a)$, one has*

$$\big\| z\big(x(T, x, u)\big) - z\big(\hat{x}(T, x, u)\big) \big\|_a \leqslant C_e(a) \max\big(\big\| z(x) \big\|_a, \ell(u)\big)\ell(u)^{1/r}, \tag{2.5}$$

where r is the degree of nonholonomy at a, $x(\cdot, x, u)$ and $\hat{x}(\cdot, x, u)$ are respectively the trajectories of $\dot{x} = \sum_{i=1}^m u_i X_i(x)$ and $\dot{x} = \sum_{i=1}^m u_i \hat{X}_i(x)$.

2.2. Approximate steering method

Definition 2.14 *(Nonholonomic first-order approximation on Ω). A nonholonomic first-order approximation of X on Ω is a mapping $\mathcal{A}$ which associates, with every $a \in \Omega$, a nonholonomic first-order approximation of X at a defined on $B(a, \rho_a)$, i.e., $\mathcal{A}(a) := \hat{X}^a$ on $B(a, \rho_a)$. The approximation radius function of $\mathcal{A}$ is the function $\rho : \Omega \to (0, \infty)$ which associates, with every a, its approximate radius ρ_a, i.e., $\rho(a) := \rho_a$.*

In the sequel, *nonholonomic first-order approximations* will simply be called *approximations*. Useful properties of approximations are *continuity* and *nilpotency*.

Definition 2.15 *(Continuity and nilpotency of an approximation). Let $\mathcal{A} : a \mapsto \hat{X}^a$ be an approximation on Ω.*

- We say that $\mathcal{A}$ is *continuous* if
    (i) the mapping $(a, x) \mapsto \mathcal{A}(a)(x)$ is well defined and, for every $a \in \Omega$, is continuous on a neighborhood of $(a, a) \in \Omega \times \Omega$;
    (ii) the approximation radius function ρ of $\mathcal{A}$ is continuous.

● We say that $\mathcal{A}$ is *nilpotent of step* $s \in \mathbb{N}$ if, for every $a \in \Omega$, the Lie algebra generated by $\hat{X}^a$ is nilpotent of step s, i.e. every Lie bracket of length larger than s is equal to zero. (For a definition of the length of a Lie bracket, see Definitions 3.1 and 3.4.)

Consider an m-tuple of vector fields $X = \{X_1, \ldots, X_m\}$ in $VF^m(\Omega)$.

Definition 2.16 (*Steering law of an approximation*). Let $\mathcal{A} : a \mapsto \hat{X}^a$ be an approximation of X on Ω and ρ its approximation radius function. A *steering law* of $\mathcal{A}$ is a mapping which, to every pair $(x, a) \in \Omega \times \Omega$ verifying $d(x, a) < \rho(a)$, associates an integrable input function $\hat{u} : [0, t] \mapsto \mathbb{R}^m$, henceforth called a *steering control*, such that the trajectory $\hat{x}(\cdot, x, \hat{u})$ of the *approximate control system*

$$\dot{x} = \sum_{i=1}^{m} u_i \hat{X}_i^a(x) \tag{2.6}$$

is defined on $[0, T]$ and satisfies $\hat{x}(T, x, \hat{u}) = a$. In other words, $\hat{u}(\cdot)$ steers (2.6) from x to a.

A steering law of an approximation is intended to be used as an approximate steering law for the original system. For that purpose, it is important to have a continuity property of the steering control: the closest are x and a, the smaller is the length of $\hat{u}$. We introduce the stronger notion of *sub-optimality* (which is a sort of Lipschitz continuity of the steering law).

Definition 2.17 (*Sub-optimal steering law*). Let $\mathcal{A}$ be an approximation of X on Ω and, for every $a \in \Omega$, let $\hat{d}_a$ be the sub-Riemannian distance associated to $\mathcal{A}(a)$. We say that a steering law of $\mathcal{A}$ is *sub-optimal* if there exist a constant $C_\ell > 0$ and a continuous positive function $\varepsilon_\ell(\cdot)$ such that, for any $a, x \in \Omega$ with $d(a, x) < \varepsilon_\ell(a)$, the control $\hat{u}(\cdot)$ steering (2.6) from x to a satisfies: $\ell(\hat{u}) \leqslant C_\ell \hat{d}_a(x, a) = C_\ell \hat{d}_a(\hat{x}(0, x, \hat{u}), \hat{x}(T, x, \hat{u}))$.

Due to the definition of the sub-Riemannian distance $\hat{d}_a$, sub-optimal steering laws always exist. Given an approximation $\mathcal{A}$ of X and a steering law for $\mathcal{A}$, we define a *local approximate steering method* for X as follows.

Definition 2.18 (*Local approximate steering*). The *local approximate steering* (LAS for short) method associated to $\mathcal{A}$ and its steering law is the mapping AppSteer$(\cdot, \cdot)$ which associates, with every pair $(x, a) \in \Omega \times \Omega$ verifying $d(x, a) < \rho(a)$, the point $x(T, x, \hat{u})$, i.e.,

$$\text{AppSteer}(x, a) := x(T, x, \hat{u}),$$

where $\hat{u}(\cdot)$ is the steering control of $\mathcal{A}(a)$ associated to (x, a) and ρ is the approximation radius function of $\mathcal{A}$.

Definition 2.19 (*Local contractions and uniform local contractions*). An LAS method is *locally contractive* if, for every $a \in \Omega$, there exist $\varepsilon_a > 0$ and $c_a < 1$ such that one has:

$$d(a, x) < \varepsilon_a \quad \Longrightarrow \quad d\big(a, \text{AppSteer}(x, a)\big) \leqslant c_a d(a, x).$$

An LAS method is *uniformly locally contractive* on a compact set $K \subset \Omega$ if it is locally contractive, and if ε_a and c_a are independent of a, i.e., there exist $\varepsilon_K > 0$ and $c_K < 1$ such that, for every pair $(a, x) \in K \times K$, the following implication holds true:

$$d(a, x) < \varepsilon_K \quad \Longrightarrow \quad d\big(a, \text{AppSteer}(x, a)\big) \leqslant c_K d(a, x).$$

Remark 2.6. We will show that if $\hat{X}$ is an approximation of X at a, the corresponding AppSteer function is locally contractive in a neighborhood of a. By the Fixed Point Theorem, one gets local convergence of Algorithm 1 (LAS). However, in order to obtain a globally convergent algorithm from LAS, one needs AppSteer to be *uniformly* locally contractive. In other words, the mapping $\mathcal{A}$ needs to be *continuous* in the sense of Definition 2.15.

As a direct consequence of Proposition 2.4, we obtain sufficient conditions for an LAS method to be uniformly locally contractive.

Corollary 2.5. *Let K be a compact subset of Ω. Assume that:*

(i) *all points in K are regular;*
(ii) *there exists a continuously varying system of privileged coordinates Φ on K;*
(iii) *there exists a continuous approximation $\mathcal{A}$ of X on K;*
(iv) *$\mathcal{A}$ is provided with a sub-optimal steering law.*

Then, the LAS method AppSteer *associated to $\mathcal{A}$ and its steering law is uniformly locally contractive. Moreover, up to reducing the positive constant ε_K occurring in Corollary 2.3, one has, for every pair $(a, x) \in K \times K$ verifying $d(a, x) < \varepsilon_K$,*

$$d\big(\text{AppSteer}(x, a), a\big) \leqslant \frac{1}{2} d(x, a), \tag{2.7}$$

$$\big\| z\big(\text{AppSteer}(x, a)\big) \big\|_a \leqslant \frac{1}{2} \big\| z(x) \big\|_a. \tag{2.8}$$

Proof. Under the hypotheses (i)–(iv), one immediately extends Proposition 2.4 and obtains that there exist continuous positive functions $C_e(\cdot)$ and $\varepsilon_e(\cdot)$ such that inequality (2.5) holds true, with $z = \Phi(a, \cdot)$ and $\hat{X} = \mathcal{A}(a)$, for every pair $(x, a) \in \Omega \times \Omega$ with $d(x, a) < \varepsilon_e(a)$ and every integrable input function $u(\cdot)$ with $\ell(u) < \varepsilon_e(a)$. The remaining argument is standard and one concludes easily. $\quad\square$

Remark 2.7. Since the growth vector and the weights do not remain constant in any open neighborhood of a singular point, privileged coordinates z cannot vary continuously in any open neighborhood of that singular point. Therefore, around a singular point, the distance estimations provided in Eqs. (2.4) and (2.8) and based on privileged coordinates do not hold true uniformly. In particular, if (a_n) is a sequence of regular points converging to a singular point a (this is possible since regular points are dense in Ω), the sequences $\varepsilon_d(a_n)$ and $\varepsilon_e(a_n)$ tend to zero whereas $\varepsilon_d(a)$ and $\varepsilon_e(a)$ are not equal to zero.

Remark 2.8. A similar discontinuity issue occurs of course for the approximate system. Indeed, if a is a singular point, the growth vector and the weights of the associated privileged coordinates at a change around a, implying a change of the truncation order in the Taylor expansion of the vector fields. Therefore, the approximate vector fields cannot vary continuously in any neighborhood of a singular point.

3. Desingularization by lifting

As it appears in Corollary 2.5, the absence of singular points is one of the key features in order to construct uniformly locally contractive LAS method. As a matter of fact, we will show in Section 4.2 how to construct a globally convergent motion planning algorithm for a regular nonholonomic system (i.e., when all points in Ω are regular).

However, in general, nonholonomic systems do have singular points. For such systems, attempts have been made to construct specific LAS methods (see [40,22]), but additional conditions on the structure of the singularities are required. Our approach here is different: we present in this section a

desingularization procedure of the system, in such a way to replace an MPP for a non-regular system by an MPP for a regular one.

The strategy consists in "*lifting*" the vector fields $\{X_1, \ldots, X_m\} \in VF^m(\Omega)$ defining the control system to some extended domain $\tilde{\Omega} := \Omega \times \mathbb{R}^{\tilde{n}}$, with $\tilde{n} \in \mathbb{N}$ to be defined later. The lifted vector fields $\{\xi_1, \ldots, \xi_m\} \in VF^m(\tilde{\Omega})$ are constructed so that:

(i) for $i = 1, \ldots, m$, ξ_i has the following form,

$$\xi_i(x, y) = X_i(x) + \sum_{j=1}^{\tilde{n}} b_{ij}(x, y) \partial_{y_j}, \quad (x, y) \in \Omega \times \mathbb{R}^{\tilde{n}},$$

where the functions b_{ij}, $j = 1, \ldots, \tilde{n}$, are smooth;
(ii) the Lie algebra generated by $\{\xi_1, \ldots, \xi_m\}$ is free up to step r (see Definition 3.6 below).

Point (ii) guarantees that the nonholonomic system defined by $\{\xi_1, \ldots, \xi_m\}$ is regular, since its growth vector is constant on $\tilde{\Omega}$. Point (i) guarantees that one obtains $X_1, \ldots, X_m$ by *projecting* $\xi_1, \ldots, \xi_m$ on $\mathbb{R}^n$. Indeed, let π be the canonical projector from $\tilde{\Omega}$ to Ω defined by $\pi(\tilde{x}) = x$, where $\tilde{x} = (x, y) \in \tilde{\Omega}$. Then, denoting $d\pi_{\tilde{x}}$ the differential of π at $\tilde{x}$, one has

$$d\pi_{\tilde{x}}\big(\xi_i(\tilde{x})\big) = X_i\big(\pi(\tilde{x})\big).$$

As a consequence, the projection by π of a trajectory $\tilde{x}(\cdot, \tilde{x}_0, u)$ of the control system

$$\dot{\tilde{x}} = \sum_{i=1}^{m} u_i \xi_i(\tilde{x}), \quad \tilde{x} \in \tilde{\Omega}, \tag{3.1}$$

is a trajectory of (2.1) associated to the *same* input, i.e., $\pi(\tilde{x}(\cdot, \tilde{x}_0, u)) = x(\cdot, \pi(\tilde{x}_0), u)$.

Therefore, any control u steering system (3.1) from a point $\tilde{x}_0 := (x_0, 0)$ to a point $\tilde{x}_1 := (x_1, 0)$ also steers system (2.1) from x_0 to x_1. It then suffices to solve the MPP for the regular system (3.1).

Note that distinguished desingularization procedures already exist, cf. [33,3,20]. However, an important property of the desingularization procedure presented here is that all the changes of coordinates and intermediate constructions involved in it are explicit and purely algebraic. Note also that, during the lifting process, we obtain, as a byproduct, a nonholonomic first-order approximation of $\{\xi_1, \ldots, \xi_m\}$ in a "canonical" form, which can be exactly controlled by sinusoids (see Section 5).

We start this section by presenting some general facts on free Lie algebras, namely the *P. Hall basis* in Section 3.1, and the *canonical form* of a nilpotent Lie algebra of step r in Section 3.2. We then give one *desingularization procedure* in Section 3.3. The proofs of the results stated in Section 3.3 will be gathered in Section 3.4.

3.1. P. Hall basis on a free Lie algebra and evaluation map

In this section, we present some general facts on free Lie algebras. The reader is referred to [7] for more details. Consider $\mathcal{I} := \{1, \ldots, m\}$, and the free Lie algebra $\mathcal{L}(\mathcal{I})$ generated by the elements of $\mathcal{I}$. Recall that $\mathcal{L}(\mathcal{I})$ is the $\mathbb{R}$-vector space generated by the elements of $\mathcal{I}$ and their formal brackets [,], together with the relations of skew-symmetry and the Jacobi identity enforced (see [7] for more details). We note that, by construction, for every $I \in \mathcal{L}(\mathcal{I})$, there exists $(I_1, I_2) \in \mathcal{L}(\mathcal{I}) \times \mathcal{L}(\mathcal{I})$ such that $I = [I_1, I_2]$.

Definition 3.1 *(Length of the elements of $\mathcal{L}(\mathcal{I})$).* The *length* of an element I of a free Lie algebra $\mathcal{L}(\mathcal{I})$, denoted by $|I|$, is defined inductively by

$$|I| := 1, \quad \text{for } I = 1, \dots, m; \tag{3.2}$$

$$|I| := |I_1| + |I_2|, \quad \text{for } I = [I_1, I_2], \text{ with } I_1, I_2 \in L(X). \tag{3.3}$$

We use $\mathcal{L}^s(\mathcal{I})$ to denote the subspace generated by elements of $\mathcal{L}(\mathcal{I})$ of length not greater than s. Let $\tilde{n}_s$ be the dimension of $\mathcal{L}^s(\mathcal{I})$.

The *P. Hall basis* of $\mathcal{L}(\mathcal{I})$ is a totally ordered subset of $\mathcal{L}(\mathcal{I})$ defined as follows.

Definition 3.2 (P. Hall basis). A subset $\mathcal{H} = \{I_j\}_{j \in \mathbb{N}}$ of $\mathcal{L}(\mathcal{I})$ is the *P. Hall basis* of $\mathcal{L}(\mathcal{I})$ if (H1), (H2), (H3), and (H4) are verified.

(H1) If $|I_i| < |I_j|$, then $I_i \prec I_j$;
(H2) $\{1, \dots, m\} \subset \mathcal{H}$, and we impose that $1 \prec 2 \prec \cdots \prec m$;
(H3) every element of length 2 in $\mathcal{H}$ is in the form $[I_i, I_j]$ with $(I_i, I_j) \in \mathcal{I} \times \mathcal{I}$ and $I_i \prec I_j$;
(H4) an element $I_k \in \mathcal{L}(\mathcal{I})$ of length greater than 3 belongs to $\mathcal{H}$ if $I_k = [I_{k_1}, [I_{k_2}, I_{k_3}]]$ with I_{k_1}, I_{k_2}, I_{k_3}, and $[I_{k_2}, I_{k_3}]$ belonging to $\mathcal{H}$, $I_{k_2} \prec I_{k_3}$, $I_{k_2} \prec I_{k_1}$ or $I_{k_2} = I_{k_1}$, and $I_{k_1} \prec [I_{k_2}, I_{k_3}]$.

The elements of $\mathcal{H}$ form a basis of $\mathcal{L}(\mathcal{I})$, and "$\prec$" defines a strict and total order over the set $\mathcal{H}$. In the sequel, we use I_k to denote the kth element of $\mathcal{H}$ with respect to the order "$\prec$". Let $\mathcal{H}^s$ be the subset of $\mathcal{H}$ of all the elements of length not greater than s. The elements of $\mathcal{H}^s$ form a basis of $\mathcal{L}^s(\mathcal{I})$ and $\text{Card}(\mathcal{H}^s) = \tilde{n}_s$. The set $\mathcal{G}^s := \mathcal{H}^s \setminus \mathcal{H}^{s-1}$ contains the elements in $\mathcal{H}$ of length equal to s. Its cardinal will be denoted by $\tilde{k}_s = \text{Card}(\mathcal{G}^s)$.

By (H1)–(H4), every element $I_j \in \mathcal{H}$ can be expanded in a unique way as

$$I_j = \left[I_{k_1}, \left[I_{k_2}, \dots, [I_{k_i}, I_k] \dots \right] \right], \tag{3.4}$$

with $k_1 \geqslant \cdots \geqslant k_i$, $k_i < k$, and $k \in \{1, \dots, \tilde{n}_1\}$. In that case, the element I_j is said to be a *direct descendent* of I_k, and we write $\phi(j) := k$. For $I_j \in \mathcal{H}^r$, the expansion (3.4) also associates with $I_j \in \mathcal{H}$ a sequence $\alpha_j = (\alpha_j^1, \dots, \alpha_j^{\tilde{n}_r})$ in $\mathbb{Z}^{\tilde{n}_r}$ defined by

$$\alpha_j^\ell := \text{Card}\{s \in \{1, \dots, i\}, \ k_s = \ell\}.$$

By construction, one has $\alpha_j^\ell = 0$ for $\ell \geqslant j$, and $\alpha_j = (0, \dots, 0)$ for $1 \leqslant j \leqslant \tilde{n}_1$.

Consider now a family of m vector fields $X = \{X_1, \dots, X_m\}$ and the Lie algebra $L(X)$ they generate. The P. Hall basis $\mathcal{H}$ induces, via the *evaluation map*, a family of vector fields spanning $L(X)$ as a linear space.

Definition 3.3 (Evaluation map). The *evaluation map* E_X defined on $\mathcal{L}(\mathcal{I})$, with values in $L(X)$, assigns to every $I \in \mathcal{L}(\mathcal{I})$ the vector field $X_I = E_X(I)$ obtained by plugging in X_i, $i = 1, \dots, m$, for the corresponding letter i.

Definition 3.4 (Length of the elements of $L(X)$). With the notations of Definition 3.3, if $X_I = E_X(I)$, the length of X_I, denoted by $\Delta(X_I)$, is set to be equal to $|I|$.

Definition 3.5 (P. Hall family). The *P. Hall family* H_X associated with the vector fields $X = \{X_1, \dots, X_m\}$ is defined by $H_X := \{E_X(I), \ I \in \mathcal{H}\}$, where E_X is the evaluation map and $\mathcal{H}$ is the P. Hall basis of the free Lie algebra $\mathcal{L}(\mathcal{I})$. The family H_X inherits the ordering and the numbering of the elements in $\mathcal{H}$ induced by (H1)–(H4).

Note that H_X is a spanning set of $L(X)$, but not always a basis.

Definition 3.6 *(Free up to step s).* Let s be a positive integer such that $1 \leqslant s \leqslant r$. A family of vector fields $\xi = \{\xi_1, \dots, \xi_m\}$ defined on a subset $\tilde{\Omega}$ of $\mathbb{R}^{\tilde{n}_r}$ is said to be *free up to step s* if, for every $\tilde{x} \in \tilde{\Omega}$, the growth vector $(n_1(\tilde{x}), \dots, n_s(\tilde{x}))$ is equal to $(\tilde{n}_1, \dots, \tilde{n}_s)$.

Remark 3.1. If ξ defined on $\tilde{\Omega} \subset \mathbb{R}^{\tilde{n}_r}$ is free up to step r, then every point of $\tilde{\Omega}$ is *regular*.

Definition 3.7 *(Free weights).* Let $\xi = \{\xi_1, \dots, \xi_m\}$ be free up to step r on $\tilde{\Omega} \subset \mathbb{R}^{\tilde{n}_r}$. The integers $\tilde{w}_1, \dots, \tilde{w}_{\tilde{n}_r}$, where $\tilde{w}_j = s$ if $n_{s-1}(\tilde{x}) < j \leqslant n_s(\tilde{x})$ for every $\tilde{x} \in \tilde{\Omega}$, are called the *free weights* of step r.

3.2. Canonical form

We present in this section the construction of a canonical form for nilpotent systems proposed by Grayson and Grossman in [17] and [18]. Similar results were also obtained by Sussmann in [37].

The construction takes place in $\mathbb{R}^{\tilde{n}_r}$, where r is a positive integer and $\tilde{n}_r$ the dimension of $\mathcal{L}^r(\mathcal{I})$. We denote by $v = (v_1, \dots, v_{\tilde{n}_r})$ the points in $\mathbb{R}^{\tilde{n}_r}$, and by $(\partial_{v_1}, \dots, \partial_{v_{\tilde{n}_r}})$ the canonical basis of $\mathbb{R}^{\tilde{n}_r}$. For $j = 1, \dots, \tilde{n}_r$, we assign to the coordinate function v_j the weight $\tilde{w}_j$ and to the vector ∂_{v_j} the weight $-\tilde{w}_j$. The *weighted degree* of a monomial $v_1^{\alpha_1} \cdots v_{\tilde{n}_r}^{\alpha_{\tilde{n}_r}}$ is then defined as $\tilde{w}(\alpha) := \tilde{w}_1 \alpha_1 + \cdots + \tilde{w}_{\tilde{n}_r} \alpha_{\tilde{n}_r}$, where $\alpha = (\alpha_1, \dots, \alpha_{\tilde{n}_r})$, and the weighted degree of a monomial vector field $v_1^{\alpha_1} \cdots v_{\tilde{n}_r}^{\alpha_{\tilde{n}_r}} \partial_{v_j}$ is defined as $\tilde{w}(\alpha) - \tilde{w}_j$.

For every $I_j \in \mathcal{H}^r$, let α_j be the sequence associated with I_j (see Section 3.1). Define the monomial $P_j(v)$ associated with I_j by

$$P_j(v) := \frac{v^{\alpha_j}}{\alpha_j!}, \tag{3.5}$$

where $v^{\alpha_j} := \prod_\ell v_\ell^{\alpha_j^\ell}$ and $\alpha_j! := \prod_\ell \alpha_j^\ell!$. Note that P_j satisfies the following inductive formulas:

$$P_j(v) = 1 \quad \text{if } I_j \in \mathcal{H}^1;$$
$$P_j(v) = \frac{v_{j_1}}{\alpha_{j_2}^{j_1} + 1} P_{j_2}(v) \quad \text{if } I_j = [I_{j_1}, I_{j_2}]. \tag{3.6}$$

Theorem 3.1. *(See [17,18].)* We define the vector fields $D_1, \dots, D_m$ on $\mathbb{R}^{\tilde{n}_r}$ as follows:

$$D_1 = \partial_{v_1},$$
$$D_2 = \partial_{v_2} + \sum_{\substack{2 \leqslant |I_j| \leqslant r \\ \phi(j)=2}} P_j(v) \partial_{v_j},$$

$$\vdots$$

$$D_m = \partial_{v_m} + \sum_{\substack{2 \leqslant |I_j| \leqslant r \\ \phi(j)=m}} P_j(v) \partial_{v_j}.$$

Then, the Lie algebra $L(D)$ generated by $(D_1, \dots, D_m)$ is free up to step r, and one has

$$D_{I_j}(0) = \partial_{v_j}, \quad \text{for } I_j \in \mathcal{H}^r,$$

where $D_{I_j} := E_D(I_j)$ is defined through the evaluation map E_D with values in $L(D)$.

The proof of Theorem 3.1 goes by induction on the length of elements in $L(D)$. The reader is referred to [18] for a complete development.

Corollary 3.2. *For all* $I_k \in \mathcal{H}^r$, D_{I_k} *has the following form*

$$D_{I_k} = \partial_{v_k} + \sum_{I_j \in \mathcal{H}^r, \, |I_j| > |I_k|} P_j^k(v) \partial_{v_j}, \tag{3.7}$$

where every non-zero polynomial P_j^k *is homogeneous of weighted degree equal to* $|I_j| - |I_k|$.

Remark 3.2. The explicit expression of the polynomials $P_j^k(v)$ as functions of the monomials $P_j(v)$ is obtained through an induction formula.

Corollary 3.3. *For* $i = 1, \ldots, m$, *we define* m *vector fields* $\check{D}_i$ *as follows*:

$$\check{D}_i := \partial_{v_i} + \sum_{\substack{2 \leqslant |I_k| \leqslant \mathcal{H}^{r-1} \\ \phi(k) = i}} P_k(v) \partial_{v_k} + \sum_{\substack{I_j \in S \\ \phi(j) = i}} P_j(v) \partial_{v_j},$$

where S *is an arbitrary non-empty subset of* $\mathcal{G}^r$. *Then,*

- *if* $I_k \in \mathcal{H}^{r-1} \cup S$, *we have* $\check{D}_{I_k} = \partial_{v_k} + \sum_{I_j \in \mathcal{H}^{r-1} \cup S, \, |I_j| > |I_k|} P_j^k(v) \partial_{v_j}$;
- *if* $I_k \in \mathcal{G}^r \setminus S$, *we have* $\check{D}_{I_k} = 0$.

Definition 3.8 *(Canonical form).* Let $X_1, \ldots, X_m$ be m vector fields on an open subset Ω of $\mathbb{R}^{\tilde{n}_r}$ and v a local system of coordinates on Ω. The control system associated to $\{X_1, \ldots, X_m\}$ is said to be in *canonical form in the coordinates* v if one has $v_* X_i = D_i$, for $i = 1, \ldots, m$, where we use $v_* X_i$ to denote the push-forward of X_i by v.

Consider now the control system given by

$$\dot{v} = \sum_{i=1}^m u_i D_i(v), \quad v \in \mathbb{R}^{\tilde{n}_r}. \tag{3.8}$$

Writing (3.8) component by component, one has, for $j = 1, \ldots, \tilde{n}_r$,

$$\dot{v}_j = P_j(v_1, \ldots, v_{j-1}) u_i, \quad \text{where } i = \phi(j), \tag{3.9}$$

or inductively,

$$\dot{v}_j = \frac{v_{j_1}}{\alpha_{j_2}^{j_1} + 1} \dot{v}_{j_2}, \quad \text{where } I_j = [I_{j_1}, I_{j_2}]. \tag{3.10}$$

More explicitly, one has

$$\dot{v}_j = \frac{1}{k!} v_{j_1}^k \dot{v}_{j_2}, \quad \text{if } X_{I_j} = \mathrm{ad}_{X_{I_{j_1}}}^k X_{I_{j_2}}, \tag{3.11}$$

where $\mathrm{ad}^k_{X_{I_{j_1}}} X_{I_{j_2}} := \underbrace{[X_{I_{j_1}}, [X_{I_{j_1}}, \ldots, [X_{I_{j_1}}, X_{I_{j_2}}]]]}_{k \text{ times}}$, with $X_{I_{j_2}} = [X_{I_{j_3}}, X_{I_{j_4}}]$ and $I_{j_3} \prec I_{j_1}$. The inductive formula (3.11) will be used in Section 5.

A particular system of coordinates, called *canonical coordinates* (a terminology arising from Lie group theory), allows one to obtain canonical forms. Consider m vector fields $X_1, \ldots, X_m$ on $\Omega \subset \mathbb{R}^{\tilde{n}_r}$, let $v \in \Omega$, and $W = \{W_1, \ldots, W_n\}$ be a set of vector fields in $L(X)$ such that $W_1(v), \ldots, W_n(v)$ is a basis of $\mathbb{R}^{\tilde{n}_r}$. The *canonical coordinates of the second kind* at v associated with W are the system of local coordinates at v defined as the inverse of the local diffeomorphism

$$(z_1, \ldots, z_{\tilde{n}_r}) \mapsto e^{z_{\tilde{n}_r} W_{\tilde{n}_r}} \circ \cdots \circ e^{z_1 W_1}(v), \tag{3.12}$$

where we use e^{zW_i} to denote the flow of W_i. When the system $(X_1, \ldots, X_m)$ is nilpotent, the above diffeomorphism defines global coordinates on Ω for every $v \in \Omega$.

Theorem 3.4. *(See [37].) Assume that the vector fields $(X_1, \ldots, X_m)$ generate a Lie algebra which is both nilpotent of step r and free up to step r. Then, in the canonical coordinates of the second kind associated with the P. Hall basis H^r_X, the control system defined by $(X_1, \ldots, X_m)$ is in canonical form.*

Remark 3.3. The canonical coordinates of the second kind require to determine the flow of the control vector fields, i.e., to integrate some differential equations. In general, there does not exist algebraic change of coordinates between an arbitrary system of coordinates and the canonical coordinates of the second kind. For more details about the construction of nilpotent approximation based on canonical coordinates of the second kind, see the review paper by Hermes [19].

3.3. Desingularization Algorithm

Let $X = \{X_1, \ldots, X_m\} \subset VF(\Omega)$ be a family of m vector fields on $\Omega \subset \mathbb{R}^n$, and K be a compact subset of Ω. We assume that the LARC is satisfied at every point of K. Therefore, the degree of nonholonomy of X is *bounded* on K and we denote by r its maximal value.

Recall that $\mathcal{H}^r$ denote the elements of the P. Hall basis of length smaller or equal to r. For every n-tuple $\mathcal{J} = (I_1, \ldots, I_n)$ of elements of $\mathcal{H}^r$, we define the domain $\mathcal{V}_{\mathcal{J}} \subset \Omega$ by

$$\mathcal{V}_{\mathcal{J}} := \big\{ p \in \Omega \text{ such that } \det\big(X_{I_1}(p), \ldots, X_{I_n}(p)\big) \neq 0 \big\}, \tag{3.13}$$

where $X_{I_j} = E_X(I_j)$. Such a set $\mathcal{V}_{\mathcal{J}}$ is open in Ω (possibly empty) and for every $p \in \mathcal{V}_{\mathcal{J}}$, the vectors $X_{I_1}(p), \ldots, X_{I_n}(p)$ form a basis of $\mathbb{R}^n$.

Since K is compact, there exists a finite family of n-tuples $\mathcal{J}_1, \ldots, \mathcal{J}_M$ of elements of $\mathcal{H}^r$ such that

$$K \subset \bigcup_{i=1}^{M} \mathcal{V}_{\mathcal{J}_i}. \tag{3.14}$$

One deduces from (3.14) a compact covering of K in the form

$$K \subset \bigcup_{i=1}^{M} \mathcal{V}^c_{\mathcal{J}_i}, \tag{3.15}$$

where, for $i = 1, \ldots, M$, the set $\mathcal{V}^c_{\mathcal{J}_i} \subset \mathcal{V}_{\mathcal{J}_i}$ is compact.

Definition 3.9. Let $(S_i)_{i \in I}$ be a finite set of subsets of Ω. The *connectedness graph* $\mathsf{G} := (\mathsf{N}, \mathsf{E})$ associated with $(S_i)_{i \in I}$ is defined as follows:

– the set of nodes $\mathsf{N} := I$;
– a pair (i, j) with i and j in N belongs to the set of edges E if $S_i \cap S_j \neq \emptyset$.

A *simple path* on G is a subset $\mathsf{p} := \{i_1, \ldots, i_L\}$ of two by two distinct elements of N such that, for $j = 1, \ldots, L - 1$, the pair (i_j, i_{j+1}) belongs to E.

Remark 3.4. With the notations of Definition 3.9, if we assume that all the sets S_i are open or, all of them are closed, and the set $S := \bigcup_{i \in I} S_i$ is connected, then, for every $(x_0, x_1) \in S \times S$, there exists a simple path on G denoted by $\mathsf{p} := \{i_1, \ldots, i_L\}$ such that $x_0 \in S_{i_1}$ and $x_1 \in S_{i_L}$.

Take $\mathcal{J} = (I_1, \ldots, I_n)$ among $\mathcal{J}_1, \ldots, \mathcal{J}_M$, and pick a point a in $\mathcal{V}_{\mathcal{J}}$. In the sequel, we construct, by induction on the length of elements in a free Lie algebra, a family of m vector fields $\xi = \{\xi_1, \ldots, \xi_m\}$ defined on $\mathcal{V}_{\mathcal{J}} \times \mathbb{R}^{\tilde{n}_r - n}$, which is free up to step r and has its projection on $\mathcal{V}_{\mathcal{J}}$ equal to X. At the same time, we give an approximation of ξ at $\tilde{a} := (a, 0) \in \mathcal{V}_{\mathcal{J}} \times \mathbb{R}^{\tilde{n}_r - n}$ in canonical form.

We define $\mathcal{J}^s := \{I_j \in \mathcal{J}, \text{ with } |I_j| = s\}$, for $s \geqslant 1$, and $\mathcal{G}^s := \mathcal{H}^s \setminus \mathcal{H}^{s-1}$, for $s \geqslant 2$. We denote by k_s the cardinal of $\mathcal{J}^s$, and by $\tilde{k}_s$ the cardinal of $\mathcal{G}^s$. We are now ready to describe in detail our Desingularization Algorithm.

Desingularization Algorithm (DA)

- Step 1:
  (1-1) Set $\mathcal{V}^1 := \mathcal{V}_{\mathcal{J}} \times \mathbb{R}^{\tilde{k}_1 - k_1}$, $a^1 := (a, 0) \in \mathcal{V}^1$, $\mathcal{K}^1 := \mathcal{H}^1 \cup (\mathcal{J} \setminus \mathcal{J}^1)$.
  (1-2) Define $\{\xi_1^1, \ldots, \xi_m^1\}$ on $\mathcal{V}^1$ as follows:

$$\forall (x, v^1) \in \mathcal{V}^1, \qquad \xi_i^1(x, v^1) := X_i(x) + \begin{cases} 0, & \text{for } i \in \mathcal{J}^1, \\ \partial_{v_i^1}, & \text{for } i \in \mathcal{G}^1 \setminus \mathcal{J}^1. \end{cases}$$

  (1-3) Compute the coordinates y^1 on $\mathcal{V}^1$ defined as the unique affine system of coordinates on $\mathcal{V}^1$ such that $\partial_{y_j^1} = \xi_{I_j}^1(a^1)$, for $I_j \in \mathcal{K}^1$, and $y^1(a^1) = 0$.
  (1-4) Construct the system of global coordinates z^1 on $\mathcal{V}^1$ by

$$z_j^1 := y_j^1, \quad \text{for } j \in \mathcal{H}^1,$$

$$z_j^1 := y_j^1 - \sum_{k=1}^{\tilde{n}_1} (\xi_k^1 \cdot y_k^1)(y^1)_{|y^1=0} y_k^1, \quad \text{for } I_j \in \mathcal{K}^1 \setminus \mathcal{H}^1,$$

  where I_j denotes the jth element in $\mathcal{K}^1$.
- Step s, $2 \leqslant s \leqslant r$:
  (s-1) Set $\mathcal{V}^s := \mathcal{V}^{s-1} \times \mathbb{R}^{\tilde{k}_s - k_s}$, $a^s := (a, 0) \in \mathcal{V}^s$, and $\mathcal{K}^s := \mathcal{K}^{s-1} \cup (\mathcal{G}^s \setminus \mathcal{J}^s)$. Denote by v^s the points in $\mathbb{R}^{\tilde{k}_s - k_s}$.
  (s-2) Define $\{\xi_1^s, \ldots, \xi_m^s\}$ as the vector fields on $\mathcal{V}^s$ which write in coordinates (z^{s-1}, v^s) as:

$$\xi_i^s(z^{s-1}, v^s) = \xi_i^{s-1}(z^{s-1}) + \sum_{\substack{I_k \in \mathcal{G}^s \setminus \mathcal{J}^s \\ \phi(k) = i}} P_k(z^{s-1}) \partial_{v_k^s}.$$

  (s-3) Compute the system of global coordinates y^s on $\mathcal{V}^s$ as the unique isomorphism $(z^{s-1}, v^s) \mapsto y^s$ such that $\partial_{y_{\phi(I)}^s} = \xi_I^s(a^s)$ for every $I \in \mathcal{K}^s$.
  (s-4) Construct the system of global coordinates $\tilde{z}^s$ on $\mathcal{V}^s$ by the following recursive formulas:

(s-4)-(a) for $I_j \in \mathcal{H}^s$,

$$\tilde{z}_j^s := y_j^s + \sum_{k=2}^{|I_j|-1} r_k(y_1^s, \ldots, y_{j-1}^s), \tag{3.16}$$

where, for $k = 2, \ldots, |I_j| - 1$,

$$r_k(y_1^s, \ldots, y_{j-1}^s) = - \sum_{\substack{|\beta|=k \\ \omega(\beta)<|I_j|}} \left[(\xi_{I_1}^s)^{\beta_1} \cdots (\xi_{I_{j-1}}^s)^{\beta_{j-1}} \cdot \left(y_j^s + \sum_{q=2}^{k-1} r_q \right) \right] (y^s)_{|y^s=0}$$

$$\times \frac{(y_1^s)^{\beta_1}}{\beta_1!} \cdots \frac{(y_{j-1}^s)^{\beta_{j-1}}}{\beta_{j-1}!};$$

(s-4)-(b) for $I_j \in \mathcal{K}^s \setminus \mathcal{H}^s$,

$$\tilde{z}_j^s := y_j^s + \sum_{k=2}^{s} r_k(y_1^s, \ldots, y_{\tilde{n}_s}^s), \tag{3.17}$$

where, for $k = 2, \ldots, s$,

$$r_k(y_1^s, \ldots, y_{\tilde{n}_s}^s) = - \sum_{\substack{|\beta|=k \\ \omega(\beta)\leqslant s}} \left[(\xi_{I_1}^s)^{\beta_1} \cdots (\xi_{I_{\tilde{n}_s}}^s)^{\beta_{\tilde{n}_s}} \cdot \left(y_j^s + \sum_{q=2}^{s} r_q \right) \right] (y^s)_{|y^s=0}$$

$$\times \frac{(y_1^s)^{\beta_1}}{\beta_1!} \cdots \frac{(y_{j-1}^s)^{\beta_{\tilde{n}_s}}}{\beta_{\tilde{n}_s}!}.$$

(s-5) Construct the system of global coordinates z^s as follows:

(s-5)-(a) for $j > \tilde{n}_s$, set $z_j^s := \tilde{z}_j^s$;

(s-5)-(b) for $j = 1, \ldots, \tilde{n}_s$, set $z_j^s := \Psi_j^s(\tilde{z}_1^s, \ldots, \tilde{z}_{j-1}^s)$, where all Ψ_j^s are polynomials such that the two following conditions are satisfied:

– if we impose the weight of z_j^s to be $\tilde{w}_j$ for $j = 1, \ldots, \tilde{n}_s$, then every Ψ_j^s is homogeneous of weighted degree equal to $\tilde{w}_j$;

– denote by $\mathrm{ord}_{a^s}^s(\cdot)$ the nonholonomic order defined by $(\xi_1^s, \ldots, \xi_m^s)$ at a^s, and by $\xi_{i,j}^s(z^s)$ the jth component of $\xi_i^s(z^s)$; then one has

$$\xi_{i,j}^s(z^s) = \delta_{i,\phi(j)} P_j(z_1^s, \ldots, z_{j-1}^s) + R_{i,j}(z^s), \quad j = 1, \ldots, \tilde{n}_s, \tag{3.18}$$

where $\mathrm{ord}_{a^s}^s(R_{i,j}) \geqslant \tilde{w}_j$ ($\delta_{i,k}$ denotes the Kronecker symbol). Note that $\mathrm{ord}_{a^s}^s(P_j) = \tilde{w}_j - 1$.

Theorem 3.5. *Let $\xi_i := \xi_i^r$, for $i = 1, \ldots, m$, and $z := z^r$, where ξ_i^r and z^r are given by the Desingularization Algorithm. Then, the family of vector fields $\xi = \{\xi_1, \ldots, \xi_m\}$ defined on $\Omega \times \mathbb{R}^{\tilde{n}_r - n}$ is free up to step r. Moreover,*

the system of coordinates $z = (z_1, \ldots, z_{\tilde{n}_r})$ is a system of privileged coordinates at $\tilde{a}$ for ξ, and the family of vector fields $\hat{\xi} = \{\hat{\xi}_1, \ldots, \hat{\xi}_m\}$ defined in the coordinates z by the canonical form:

$$\hat{\xi}_i = \partial_{z_i} + \sum_{\substack{2 \leqslant |I_j| \leqslant \tilde{n}_r \\ \phi(j)=i}} P_j(z_1, \ldots, z_{j-1}) \partial_{z_j}, \quad for\ i = 1, \ldots, m, \tag{3.19}$$

is a nonholonomic first-order approximation of ξ at $\tilde{a}$.

Remark 3.5. We note that the desingularization procedure does not a priori require that

(a) the coordinates z are privileged coordinates;
(b) the system $\hat{\xi}$ is a first-order approximation of ξ at a.

However, (a) and (b) can be used directly at the first step of the motion planning algorithm presented in Section 4.

Remark 3.6. If we assume that the original system $X = \{X_1, \ldots, X_m\}$ is nilpotent, then, by adapting the proof of Theorem 3.5 presented in Section 3.4, one can show that the corresponding "lifted" system $\xi = \{\xi_1, \ldots, \xi_m\}$ given by the Desingularization Algorithm proposed in this section remains *nilpotent* with the same order of nilpotency. Moreover, when expressed in the privileged coordinates z, the system ξ is equal to its own first-order approximation in the canonical form. In other words, for any nilpotent system of step k, the Desingularization Algorithm constructs a nilpotent system of step k and free up to step k which is in the canonical form in coordinates z.

3.4. Proof of Theorem 3.5

The proof of Theorem 3.5 is based on the following proposition.

Proposition 3.6. *The Desingularization Algorithm is feasible from $s = 1$ to $s = r$. At each step s of the construction $(s = 1, \ldots, r)$, the following properties hold true:*

(A1) *the vectors $\{\xi_I^s(a^s): I \in \mathcal{K}^s\}$ are linearly independent;*
(A2) *if $|I_j| \leqslant s$, then $\mathrm{ord}_{a^s}^s(\tilde{z}_j^s) = |I_j|$ and $\mathrm{ord}_{a^s}^s(z_j^s) = |I_j|$;*
(A3) *if $|I_j| > s$, then $\mathrm{ord}_{a^s}^s(z_j^s) > s$;*
(A4) *the change of coordinates $(\Psi_j^s)_{j=1,\ldots,\tilde{n}_s}$ is well defined;*
(A5) *for $I_k \in \mathcal{K}^s$, the vector fields $\xi_{I_k}^s$ have the following form in coordinates z^s,*

$$\xi_{I_k}^s(z^s) = \sum_{I_j \in \mathcal{H}^s} \left(P_j^k(z^s) + R_j^k(z^s) \right) \partial_{z_j^s} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{\,k}(z^s) \partial_{z_\ell^s}, \tag{3.20}$$

with $\mathrm{ord}_{a^s}^s(R_j^k) > |I_j| - |I_k|$, $\mathrm{ord}_{a^s}^s(Q_\ell^{\,k}) > s - |I_k|$, and P_j^k given by Eq. (3.7).
More precisely, if one defines $\check{\xi}_i^s := \sum_{I_j \in \mathcal{H}^s,\ \phi(j)=i} P_j(z^s) \partial_{z_j^s}$, then, one has $\check{\xi}_{I_k}^s = \sum_{I_j \in \mathcal{H}^s} P_j^k(z^s) \partial_{z_j^s}$,
where the polynomials P_j^k verify the following properties:
- *if $I_k \in \mathcal{H}^s$, then*
  - *for $|I_j| < |I_k|$, $P_j^k = 0$;*
  - *for $|I_j| = |I_k|$, $P_j^j = 1$, and $P_j^k = 0$ if $k \neq j$;*
  - *for $|I_j| > |I_k|$, $\mathrm{ord}_{a^s}^s(P_j^k) = |I_j| - |I_k|$;*
- *if $I_k \in \mathcal{K}^s \setminus \mathcal{H}^s$, $P_j^k = 0$ for all $j = 1, \ldots, \tilde{n}_s$.*

Remark 3.7. Property (A1) implies that Step (s-3) is feasible, which, in turn, guarantees that Steps (s-4)-(a) and (s-4)-(b) are well defined, and $\tilde{z}^s$ is a system of coordinates because the differential of the application $y^s \mapsto \tilde{z}^s$ at 0 is equal to the identity map. Property (A4) guarantees that Step (s-5)-(b) is feasible. Property (A2) ensures that, at the end of the algorithm, the system of coordinates z^r is a system of privileged coordinates. Property (A5) finally ensures that for $s = r$, the approximation $\hat{\xi}$ of ξ is in canonical form.

By Remark 3.7, Theorem 3.5 is a consequence of Proposition 3.6 by induction on s.

Proof of Proposition 3.6. We begin by showing that properties (A1)–(A5) hold true for $s = 1$.

Claim 1. *The family of vectors $\{\xi_I^1(a^1)\}_{I \in \mathcal{K}^1}$ is linearly independent, i.e., property* (A1) *holds true for $s = 1$.*

Proof. By construction, for every $I \in \mathcal{J}$, one has $\xi_I^1(a^1) = X_I(a)$, which belongs to $\mathbb{R}^n \times \{0\}$. For $i \in \mathcal{G}^1 \setminus \mathcal{J}^1$, the vector $\xi_i^1(a^1)$ belongs to $\mathbb{R}^n \times \mathbb{R}^{\tilde{k}_1 - k_1}$, and the family of vectors $\{\xi_i^1(a^1)\}_{i \in \mathcal{G}^1 \setminus \mathcal{J}^1}$ is linearly independent. Therefore, the family of vectors $\{\xi_I^1(a^1)\}_{I \in \mathcal{K}^1}$ is linearly independent and Claim 1 holds true. □

Claim 2. *For $j = 1, \ldots, \tilde{n}_1$, one has $\mathrm{ord}_{a^1}^1(z_j^1) = 1$, i.e., property* (A2) *holds true for $s = 1$.*

Proof. For $j = 1, \ldots, \tilde{n}_1$, one has by construction $\xi_j^1 \cdot z_j^1(a^1) = 1$. Thus, one has $\mathrm{ord}_{a^1}^1(z_j^1) \leqslant 1$. Since z^1 is a system of coordinates centered at a^1, one has $z_j^1(a^1) = 0$, which implies that $\mathrm{ord}_{a^1}^1(z_j^1) > 0$. Therefore, one gets $\mathrm{ord}_{a^1}^1(z_j^1) = 1$ and Claim 2 holds true. □

Claim 3. *For $I_j \in \mathcal{K}^1$ with $|I_j| > 1$, one has $\mathrm{ord}_{a^1}^1(z_j^1) > 1$, i.e., property* (A3) *holds true for $s = 1$.*

Proof. For $|I_j| \geqslant 2$, i.e. $I_j \in \mathcal{K}^1 \setminus \mathcal{J}^1$, one computes $\xi_k^1 \cdot z_j^1$ at a^1 for every $k \in \{1, \ldots, \tilde{n}_1\}$:

$$\xi_k^1 \cdot z_j^1(a^1) = \xi_k^1 \cdot y_j^1(a^1) - \sum_{i=1}^{\tilde{n}_1} (\xi_i^1 \cdot y_j^1)(a^1)(\xi_k^1 \cdot y_i^1)(a^1)$$

$$= \xi_k^1 \cdot y_j^1(a^1) - \xi_k^1 \cdot y_j^1(a^1) = 0.$$

Then, by definition, one has $\mathrm{ord}_{a^1}^1(z_j^1) > 1$ for $|I_j| > 1$ and Claim 3 holds true. □

Claim 4. *For $i = 1, \ldots, m$, and $j = 1, \ldots, \tilde{n}_1$, the jth component of ξ_i^1 in coordinates z^1 is equal to 1 if $i = j$, and equal to 0 otherwise. In other words, for $i = 1, \ldots, m$, the $\tilde{n}_1$ first components of ξ_i^1 verify Eq.* (3.18). *Properties* (A4) *and* (A5) *hold true for $s = 1$.*

Proof. By Claim 1, $\xi_1^1(a^1), \ldots, \xi_{\tilde{n}_1}^1(a^1)$ is a basis of $\mathbb{R}^{\tilde{n}_1}$, and thus the linear change of coordinates y^1 exists. As $\partial_{y_j^1} = \xi_j^1(a^1)$, and $z_j^1 = y_j^1$ for $j = 1, \ldots, \tilde{n}_1$, Claim 4 holds true. □

Therefore, properties (A1)–(A5) hold true for $s = 1$. Let $1 \leqslant s \leqslant r$. Let us now assume that properties (A1)–(A5) hold true for $1 \leqslant s' \leqslant s$. We will show that they still hold true for $s + 1$.

Claim 5. *The vector fields $\{\xi_i^{s+1}\}_{i=1,\ldots,m}$ are well defined. Moreover, one has $\mathrm{ord}_{a^{s+1}}^s(P_k) = s$.*

Proof. Consider $I_k \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}$, then one has $I_k = [I_{k_1}, I_{k_2}]$. By Eq. (3.6), one has

$$P_k(z^s) = \frac{z_{k_1}^s}{\alpha_{k_2}^{k_1} + 1} P_{k_2}(z^s).$$

Since $|I_{k_1}| \leqslant s$ and $|I_{k_2}| \leqslant s$, we have $k_1 \leqslant \tilde{n}_s$ and $k_2 \leqslant \tilde{n}_s$, thus the right-hand side of the above equation is well defined in coordinates $z^s = (z_1^s, \ldots, z_{\tilde{n}_s}^s)$. Therefore, the new vector fields $\{\xi_i^{s+1}\}_{i=1,\ldots,m}$ are also well defined. Since $\mathrm{ord}_{a^{s+1}}^s(z_{k_1}^s P_{k_2}) = \mathrm{ord}_{a^{s+1}}^s(z_{k_1}^s) + \mathrm{ord}_{a^{s+1}}^s(P_{k_2})$, and by inductive hypothesis (namely (A2) holds true at step s), one has $\mathrm{ord}_{a^{s+1}}^s(z_{k_1}^s) = |I_{k_1}|$, and $\mathrm{ord}_{a^{s+1}}^s(P_{k_2}) = |I_{k_2}| - 1$. Therefore, one has $\mathrm{ord}_{a^{s+1}}^s(P_k) = |I_{k_1}| + |I_{k_2}| - 1 = s$. $\quad\square$

Claim 6. *For $I_k \in \mathcal{K}^{s+1}$ with $|I_k| \leqslant s + 1$, one has*

$$\xi_{I_k}^{s+1}(z^s, v^{s+1}) = \xi_{I_k}^s(z^s) + \sum_{I_j \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} \tilde{P}_j^k(z^s) \partial_{v_j^{s+1}}, \tag{3.21}$$

where $\tilde{P}_j^k(z^s) = P_j^k(z_1^s, \ldots, z_{\tilde{n}_s}^s) + \tilde{R}_j^k(z^s)$, with $\mathrm{ord}_{a^{s+1}}^s(P_j^k) = |I_j| - |I_k|$ and $\mathrm{ord}_{a^{s+1}}^s(\tilde{R}_j^k) > |I_j| - |I_k|$.

Proof. The proof goes by induction on the length $|I_k|$. For $|I_k| = 1$, one has

$$\xi_k^{s+1}(z^s, v^{s+1}) = \xi_k^s(z^s) + \sum_{\substack{I_j \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1} \\ \phi(j) = k}} P_j(z^s) \partial_{v_j^{s+1}}.$$

By Claim 5, if $\phi(j) = k$, then $\mathrm{ord}_{a^{s+1}}^s(P_j) = s = |I_j| - |I_k|$. Claim 6 holds true for $|I_k| = 1$.

Assume that Claim 6 holds true for every $I \in \mathcal{K}^{s+1}$ of length less than or equal to s_1. Consider $I_k \in \mathcal{K}^{s+1}$ with $|I_k| = s_1 + 1$. In coordinates (z^s, v^{s+1}), one has

$$\xi_{I_k}^{s+1} = [\xi_{I_{k_1}}^{s+1}, \xi_{I_{k_2}}^{s+1}] = \left[\xi_{I_{k_1}}^s + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} (P_i^{k_1} + \tilde{R}_i^{k_1}) \partial_{v_i^{s+1}}, \xi_{I_{k_2}}^s + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} (P_i^{k_2} + \tilde{R}_i^{k_2}) \partial_{v_i^{s+1}} \right]$$

$$= [\xi_{I_{k_1}}^s, \xi_{I_{k_2}}^s] + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} \{ \xi_{I_{k_2}}^s \cdot (P_i^{k_1} + \tilde{R}_i^{k_1}) - \xi_{I_{k_1}}^s \cdot (P_i^{k_2} + \tilde{R}_i^{k_2}) \} \partial_{v_i^{s+1}}.$$

Since (A5) holds true up to step s, one has

$$\xi_{I_{k_1}}^s \cdot (P_i^{k_2} + \tilde{R}_i^{k_2})$$

$$= \left[\sum_{I_j \in \mathcal{H}^s} (P_j^{k_1} + R_j^{k_1}) \partial_{z_j^s} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{k_1} \partial_{z_\ell^s} \right] \cdot (P_i^{k_2} + \tilde{R}_i^{k_2})$$

$$= \sum_{I_j \in \mathcal{H}^s} (P_j^{k_1} + R_j^{k_1}) \partial_{z_j^s} P_i^{k_2} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{k_1} \partial_{z_\ell^s} P_i^{k_2} + \sum_{I_j \in \mathcal{H}^s} (P_j^{k_1} + R_j^{k_1}) \partial_{z_j^s} \tilde{R}_i^{k_2} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{k_1} \partial_{z_\ell^s} \tilde{R}_i^{k_2}$$

$$= \sum_{I_j \in \mathcal{H}^s} P_j^{k_1} \partial_{z_j^s} P_i^{k_2} + \left[\sum_{I_j \in \mathcal{H}^s} R_j^{k_1} \partial_{z_j^s} P_i^{k_2} + \sum_{I_j \in \mathcal{H}^s} (P_j^{k_1} + R_j^{k_1}) \partial_{z_j^s} \tilde{R}_i^{k_2} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{k_1} \partial_{z_\ell^s} \tilde{R}_i^{k_2} \right]$$

$$:= \sum_{I_j \in \mathcal{H}^s} P_j^{k_1} \partial_{z_j^s} P_i^{k_2} + \mathcal{R}_{i,1}.$$

We first show that every term in $\mathcal{R}_{i,1}$ has, at a^{s+1}, an order strictly greater than $s+1-|I_k|$.

Indeed, for $I_j \in \mathcal{H}^s$, since $\mathrm{ord}_{a^{s+1}}^s(z_j) = |I_j|$, $\mathrm{ord}_{a^{s+1}}^s(P_i^{k_2}) = |I_i| - |I_{k_2}|$, and $\mathrm{ord}_{a^{s+1}}^s(R_j^{k_1}) > |I_j| - |I_{k_1}|$, one has $\mathrm{ord}_{a^{s+1}}^s(R_j^{k_1} \partial_{z_j^s} P_i^{k_2}) > |I_j| - |I_{k_1}| + (|I_i| - |I_{k_2}|) - |I_j| = |I_i| - |I_k|$, with $|I_i| = s+1$. Note that $\mathrm{ord}_{a^{s+1}}^s((P_j^{k_1} + R_j^{k_1})\partial_{z_j^s} \tilde{R}_i^{k_2}) = \mathrm{ord}_{a^{s+1}}^s(P_j^{k_1} \partial_{z_j^s} \tilde{R}_i^{k_2})$. Since $\mathrm{ord}_{a^{s+1}}^s(P_j^{k_1}) = |I_j| - |I_{k_1}|$, and $\mathrm{ord}_{a^{s+1}}^s(\tilde{R}_i^{k_2}) > |I_i| - |I_{k_1}|$, then, one has $\mathrm{ord}_{a^{s+1}}^s(P_j^{k_1} \partial_{z_j^s} \tilde{R}_i^{k_2}) > |I_j| - |I_{k_2}| + |I_i| - |I_{k_1}| - |I_j| = |I_i| - |I_k|$. Recall that, by definition, all the functions have positive order. Therefore, one gets $\mathrm{ord}_{a^{s+1}}^s(Q_\ell^{k_1} \partial_{z_\ell^s} \tilde{R}_i^{k_2}) \geqslant \mathrm{ord}_{a^{s+1}}^s(Q_\ell^{k_1}) > s - |I_{k_1}| \geqslant s+1 - |I_k|$. In conclusion, one gets $\mathrm{ord}_{a^{s+1}}^s(\mathcal{R}_{i,1}) > s+1 - |I_k|$.

A similar computation shows that

$$
\xi_{I_{k_2}}^s \cdot \left(P_i^{k_1} + \tilde{R}_i^{k_1} \right)
$$
$$
= \sum_{I_j \in \mathcal{H}^s} P_j^{k_2} \partial_{z_j^s} P_i^{k_1} + \left[\sum_{I_j \in \mathcal{H}^s} R_j^{k_2} \partial_{z_j^s} P_i^{k_1} + \sum_{I_j \in \mathcal{H}^s} (P_j^{k_2} + R_j^{k_2}) \partial_{z_j^s} \tilde{R}_i^{k_1} + \sum_{I_\ell \in \mathcal{K}^s \setminus \mathcal{H}^s} Q_\ell^{k_2} \partial_{z_\ell^s} \tilde{R}_i^{k_1} \right]
$$
$$
:= \sum_{I_j \in \mathcal{H}^s} P_j^{k_2} \partial_{z_j^s} P_i^{k_1} + \mathcal{R}_{i,2}, \quad \text{with } \mathrm{ord}_{a^{s+1}}^s(\mathcal{R}_{i,2}) > s+1 - |I_k|.
$$

Therefore, one gets

$$
\xi_{I_k}^{s+1} = \xi_{I_k}^s + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} \left\{ \sum_{I_j \in \mathcal{H}^s} \left(P_j^{k_1} \partial_{z_j^s} P_i^{k_2} - P_j^{k_2} \partial_{z_j^s} P_i^{k_1} \right) \right\} \partial_{v_i} + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} (\mathcal{R}_{i,1} + \mathcal{R}_{i,2}) \partial_{v_i},
$$

with $\mathrm{ord}_{a^{s+1}}^s(\mathcal{R}_{i,1} + \mathcal{R}_{i,2}) \geqslant \min(\mathrm{ord}_{a^{s+1}}^s(\mathcal{R}_{i,1}), \mathrm{ord}_{a^{s+1}}^s(\mathcal{R}_{i,2})) > s+1 - |I_k|$.

Since Corollary 3.3 implies that $\sum_{I_j \in \mathcal{H}^s} (P_j^{k_1} \partial_{z_j^s} P_i^{k_2} - P_j^{k_2} \partial_{z_j^s} P_i^{k_1}) = P_i^k$, and $\mathrm{ord}_{a^{s+1}}^s(P_i^k) = |I_i| - |I_k|$, one gets

$$
\xi_{I_k}^{s+1}(z^s, v^{s+1}) = \xi_{I_k}^s(z^s) + \sum_{I_i \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}} (P_i^k(z^s) + \tilde{R}_i^k(z^s)) \partial_{v_i},
$$

with $\mathrm{ord}_{a^{s+1}}^s(P_i^k) = s+1 - |I_k|$ and $\mathrm{ord}_{a^{s+1}}^s(\tilde{R}_i^k) > s+1 - |I_k|$. Therefore, Claim 6 still holds true for $I_k \in \mathcal{K}^{s+1}$ with $|I_k| = s_1 + 1$. This terminates the induction, and Claim 6 is now proved. □

Claim 7. *The family of vectors $\{\xi_{I_k}^{s+1}(a^{s+1})\}_{I_k \in \mathcal{K}^{s+1}}$ is linearly independent, i.e., property* (A1) *holds true at step $s+1$.*

Proof. Claim 6 implies that for every $I_k \in \mathcal{K}^s$, one has $\xi_{I_k}^{s+1}(a^{s+1}) = \xi_{I_k}^s(a^s) \in \mathbb{R}^{\tilde{n}_s} \times \{0\}$. Corollary 3.3 implies that for every $I_k \in \mathcal{G}^{s+1} \setminus \mathcal{J}^{s+1}$, one has

$$
\xi_{I_k}^{s+1}(a^{s+1}) = \xi_{I_k}^s(a^s) + \partial_{v_k} \in \mathbb{R}^{\tilde{n}_s} \times \mathbb{R}^{\tilde{k}_{s+1} - k_{s+1}}.
$$

Therefore, by (A1) at step s, the vectors $\{\xi_{I_k}^{s+1}(a^{s+1})\}_{I_k \in \mathcal{K}^{s+1}}$ are linearly independent. □

Claim 8. *After performing* (s+1)-4-(a) *and* (s+1)-4-(b) *in the Desingularization Algorithm, one has, for every $I_j \in \mathcal{H}^{s+1}$, $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) = |I_j|$, and for every $I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}$, $\mathrm{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) > s+1$.*

The proof of Claim 8 is based on the following result due to Bellaïche [3, Lemma 4.12].

Lemma 3.7. *Let* $\{X_1, \ldots, X_m\}$ *be a family vector fields defined on* Ω. *Consider* $\{W_1, \ldots, W_n\}$ *a frame adapted to the flag* $L^1(a) \subset \cdots \subset L^r(a) = \mathbb{R}^n$ *at* $a \in \Omega$ *(Remark* 2.4*). A function* f *is of order strictly greater than* s *at* a *if and only if* $(W_1^{\alpha_1} \cdots W_n^{\alpha_n} f)(a) = 0$, *for all* $\alpha = (\alpha_1, \ldots, \alpha_n)$ *such that* $w(\alpha) \leqslant s$.

Proof of Claim 8. Claim 7 guarantees that $\{\xi_{I_k}^{s+1}\}_{I_k \in \mathcal{H}^{s+1}}$ is a basis adapted to the flag $L^1(a^{s+1}) \subset \cdots \subset L^{s+1}(a^{s+1})$. Complete $\{\xi_{I_k}^{s+1}\}_{I_k \in \mathcal{H}^{s+1}}$ by other elements of the Lie algebra generated by $\{\xi_i^{s+1}\}_{i=1,\ldots,m}$ in order to get a basis adapted to the flag $L^1(a^{s+1}) \subset \cdots \subset L^{s+1}(a^{s+1}) \subset \cdots \subset L^r(a^{s+1})$.

For $I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}$, formula (s+1)-4-(b) ensures that $((\xi_{I_1}^{s+1})^{\beta_1} \cdots (\xi_{I_{\tilde{n}_{s+1}}}^{s+1})^{\beta_{\tilde{n}_{s+1}}} \cdot \tilde{z}_j^{s+1})(a^{s+1}) = 0$, for all $\beta = (\beta_1, \ldots, \beta_{\tilde{n}_{s+1}})$ such that $w(\beta) \leqslant s + 1$. By Lemma 3.7, one has $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) > s + 1$, for $I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}$.

For $I_j \in \mathcal{H}^{s+1}$, formula (s+1)-4-(a) implies that $((\xi_{I_1}^{s+1})^{\beta_1} \cdots (\xi_{I_{j-1}}^{s+1})^{\beta_{j-1}} \cdot \tilde{z}_j^{s+1})(a^{s+1}) = 0$, for all $\beta = (\beta_1, \ldots, \beta_{j-1})$ such that $w(\beta) \leqslant |I_j| - 1$. Using again Lemma 3.7, one has $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) > |I_j| - 1$, for $I_j \in \mathcal{H}^{s+1}$. By construction, one already has that $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) \leqslant \tilde{w}_j = |I_j|$. Therefore, one finally gets $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) = |I_j|$, for $I_j \in \mathcal{H}^{s+1}$. Claim 8 is now proved. $\square$

Claim 9. *The change of coordinates* $(\Psi_j^{s+1})_{j=1,\ldots,\tilde{n}_{s+1}}$ *is well defined, i.e., property* (A4) *holds true.*

Proof. After performing Steps (s+1)-4-(a) and (s+1)-4-(b), one obtains a new system of coordinates $\tilde{z}^{s+1}$. In this system of coordinates, one can write ξ_i^{s+1} as

$$\xi_i^{s+1}(\tilde{z}^{s+1}) = \partial_{\tilde{z}_i^{s+1}} + \sum_{\substack{I_j \in \mathcal{H}^{s+1} \\ |I_j| \geqslant 2}} \left(\tilde{P}_{i,j}(\tilde{z}^{s+1}) + \tilde{R}_{i,j}(\tilde{z}^{s+1}) \right) \partial_{\tilde{z}_j^{s+1}} + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} \tilde{Q}_{i,\ell}(\tilde{z}^{s+1}) \partial_{\tilde{z}_\ell^{s+1}},$$

where $\tilde{P}_{i,j}$, $\tilde{R}_{i,j}$, and $\tilde{Q}_{i,\ell}$ are polynomials with $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{P}_{i,j}) = \tilde{w}_j - 1$, $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{R}_{i,j}) \geqslant \tilde{w}_j$, and $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{Q}_{i,\ell}) > s$. Since $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) = \tilde{w}_j$, for $I_j \in \mathcal{H}^{s+1}$, and $\mathrm{ord}_{a^{s+1}}^{s+1}(\tilde{z}_j^{s+1}) > s + 1$, for $I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}$, the polynomials $\tilde{P}_{i,j}$ contain only variables $\tilde{z}_k^{s+1}$ with $\tilde{w}_k \leqslant \tilde{w}_j - 1$.

Let us now show that there exists a change of coordinates Ψ^{s+1} which transforms coordinates $\tilde{z}^{s+1}$ into new coordinates z^{s+1} such that

$$\mathrm{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) = \tilde{w}_j, \quad \text{for } I_j \in \mathcal{H}^{s+1},$$

$$\mathrm{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) > s + 1, \quad \text{for } I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1},$$

and in the new coordinates, the $\tilde{n}_{s+1}$ first components $\xi_{i,j}^{s+1}(z^{s+1})$ of $\xi_i^{s+1}(z^{s+1})$ are in the form

$$\xi_{i,j}^{s+1}(z^{s+1}) = \delta_{i,\phi(j)} P_j(z_1^{s+1}, \ldots, z_{j-1}^{s+1}) + R_{i,j}(z^{s+1}), \quad j = 1, \ldots, \tilde{n}_{s+1},$$

with $\mathrm{ord}_{a^{s+1}}^{s+1}(R_{i,j}) \geqslant \tilde{w}_j$.

We first note that, once one has $\mathrm{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) = \tilde{w}_j$ for $I_j \in \mathcal{H}^{s+1}$, and $\mathrm{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) > s + 1$ for $I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}$, then, the order of $P_{i,j}$ will be equal to its weighted degree, and thus automatically equal to $\tilde{w}_j - 1$ by construction of these polynomials.

Consider now $\check{\xi}_i^{s+1}$ defined in coordinates $\tilde{z}^{s+1}$ by

$$\check{\xi}_i^{s+1}(\tilde{z}^{s+1}) = \partial_{\tilde{z}_i^{s+1}} + \sum_{I_j \in \mathcal{H}^{s+1}} \tilde{P}_{i,j}(\tilde{z}^{s+1}) \partial_{\tilde{z}_j^{s+1}}.$$

Recall that, by construction, the vector fields $\{\check{\xi}_i\}_{i=1,\dots,m}$ generate a free nilpotent Lie algebra of step $s+1$. Moreover, in the canonical coordinates of the second kind $(z_1^{s+1}, \dots, z_{\tilde{n}_{s+1}}^{s+1})$ associated with $\{\check{\xi}_{I_k}^{s+1}\}_{I_k \in \mathcal{H}^{s+1}}$, the vector fields $\check{\xi}_i^{s+1}$ are in the canonical form, i.e.

$$\check{\xi}_i^{s+1} = \partial_{z_i^{s+1}} + \sum_{\substack{I_j \in \mathcal{H}^{s+1} \\ \phi(j)=i}} P_j(z^{s+1}) \partial_{z_j^{s+1}}.$$

By definition of a system of coordinates, there exist $\tilde{n}_{s+1}$ smooth functions $(\Psi_1^{s+1}, \dots, \Psi_{\tilde{n}_{s+1}}^{s+1})$ such that, for $j = 1, \dots, \tilde{n}_{s+1}$, one has

$$z_j^{s+1} = \Psi_j^{s+1}(\tilde{z}_1^{s+1}, \dots, \tilde{z}_{\tilde{n}_{s+1}}^{s+1}).$$

Expand now Ψ_j^{s+1} in Taylor series. Since $\operatorname{ord}_{a^{s+1}}^{s+1}(z^{s+1}) = \tilde{w}_j$, the Taylor expansion of Ψ_j^{s+1} is a *polynomial* of weighted degree equal to $\tilde{w}_j$. Claim 9 is now proved. $\square$

Remark 3.8. The change of coordinates $(\Psi_j^{s+1})_{j=1,\dots,\tilde{n}_{s+1}}$ is computed by identification. Indeed, since $\operatorname{ord}_{a^{s+1}}^{s+1}(z_j^{s+1}) = \tilde{w}_j$, and the nonholonomic order does not depend on any system of coordinates, then Ψ_j^{s+1} is a function of order $\tilde{w}_j$ at a^{s+1}, i.e., the Taylor expansion of Ψ_j^{s+1} at a^{s+1} contains only monomials of weighted degree equal to $\tilde{w}_j$, and there is a finite number of such monomials. Therefore, the function Ψ_j^{s+1} is necessarily in the following form

$$\Psi_j^{s+1}(\tilde{z}^{s+1}) = \sum_{w(\alpha)=\tilde{w}_j} \beta_j^\alpha (\tilde{z}_1^{s+1})^{\alpha_1} \cdots (\tilde{z}_{\tilde{n}_{s+1}}^{s+1})^{\alpha_{\tilde{n}_{s+1}}}, \tag{3.22}$$

where β_j^α are real numbers. Eq. (3.22) is a finite sum and therefore the scalar coefficients (φ_j^α) can be obtained by identification. Claim 9 guarantees that such a set of real numbers (φ_j^α) exists. Note also that, due to the constraint on the weight, Eq. (3.22) only involves variables $\tilde{z}_k^{s+1}$ of weight less than $\tilde{w}_j$, implying that the change of coordinates $(\Psi_j^{s+1})_{j=1,\dots,\tilde{n}_{s+1}}$ is naturally triangular.

Remark 3.9. Let us now illustrate Remark 3.8 with a simple example. Consider here a nilpotent system of step 2 generated by two vector fields (ξ_1, ξ_2). We have $\xi_{I_1} = \xi_1$, $\xi_{I_2} = \xi_2$ and $\xi_{I_3} = [\xi_1, \xi_2]$. In coordinates $\tilde{z} = (\tilde{z}_1, \tilde{z}_2, \tilde{z}_3)$, ξ_1 and ξ_2 are necessarily in the form $\xi_1 = (1, 0, \alpha_1 \tilde{z}_1 + \alpha_2 \tilde{z}_2)$, and $\xi_2 = (0, 1, \beta_1 \tilde{z}_1 + \beta_2 \tilde{z}_2)$, where α_1, α_2, β_1 and β_2 are real numbers verifying $\beta_1 - \alpha_2 = 1$. As mentioned in Remark 3.8, in the change of coordinates (Ψ_1, Ψ_2, Ψ_3), every Ψ_j is a homogeneous polynomial of weighted degree equal to $\tilde{w}_j$. Set $z = (\Psi_1(\tilde{z}), \Psi_2(\tilde{z}), \Psi_3(\tilde{z})) =: (\tilde{z}_1, \tilde{z}_2, \tilde{z}_3 + a\tilde{z}_1\tilde{z}_2 + b\tilde{z}_1^2 + c\tilde{z}_2^2)$, with a, b, and c to be determined. One imposes that $\xi_2(z) = (0, 1, z_1)$. After computation, one gets

$$(\alpha_1 + 2b)\tilde{z}_1 + (\alpha_2 + a)\tilde{z}_2 = 0,$$

$$(\beta_1 + a)\tilde{z}_1 + (\beta_2 + 2c)\tilde{z}_2 = z_1 = \tilde{z}_1.$$

By identification, one gets $a = -\alpha_2$, $b = -\frac{\alpha_1}{2}$, $c = -\frac{\beta_2}{2}$, and in that case, $\beta_1 + a = \beta_1 - \alpha_2 = 1$ is automatically verified. Then, the triangular change of coordinates

$$(z_1, z_2, z_3) = \left(\tilde{z}_1, \tilde{z}_2, \tilde{z}_3 - \alpha_2 \tilde{z}_1 \tilde{z}_2 - \frac{\alpha_1}{2}\tilde{z}_1^2 - \frac{\beta_2}{2}\tilde{z}^2\right)$$

puts ξ_1 and ξ_2 into the canonical form.

Claim 10. *Property* (A5) *holds true at step* $s + 1$.

Proof. The proof goes by induction on the length of $I_k \in \mathcal{K}^{s+1}$. It is similar to the one of Claim 6. For $|I_k| = 1$, one has

$$\xi_i^{s+1}(z^{s+1}) = \sum_{\substack{I_j \in \mathcal{H}^{s+1} \\ \phi(j)=i}} \left(P_j(z^{s+1}) + R_{i,j}(z^{s+1})\right)\partial_{z_j^{s+1}} + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_{i,\ell}(z^{s+1})\partial_{z_\ell^{s+1}},$$

with $\mathrm{ord}_{a^{s+1}}^{s+1}(P_j) = |I_j| - 1$, $\mathrm{ord}_{a^{s+1}}^{s+1}(R_{i,j}) > |I_j| - 1$, $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_{i,\ell}) > s$. Claim 10 holds true for $|I_k| = 1$.

Assume that Claim 10 holds for brackets of length less than s_1. We show that it still holds true for brackets of length $s_1 + 1$. Consider $I_k \in \mathcal{K}^{s+1}$ with $|I_k| = s_1 + 1$. Then, one has

$$
\begin{aligned}
\xi_{I_k}^{s+1} &= \left[\xi_{I_{k_1}}^{s+1}, \xi_{I_{k_2}}^{s+1}\right] \\
&= \left[\sum_{I_j \in \mathcal{H}^{s+1}} \left(P_j^{k_1} + R_j^{k_1}\right)\partial_{z_j^{s+1}} + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_\ell^{k_1}\partial_{z_\ell^{s+1}}, \sum_{I_j \in \mathcal{H}^{s+1}} \left(P_j^{k_2} + R_j^{k_2}\right)\partial_{z_j^{s+1}} \right. \\
&\qquad \left. + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_\ell^{k_2}\partial_{z_\ell^{s+1}}\right] \\
&= \sum_{I_j \in \mathcal{H}^{s+1}} \left[\sum_{I_i \in \mathcal{H}^{s+1}} P_i^{k_1}\partial_{z_i^{s+1}}P_j^{k_2} - P_i^{k_2}\partial_{z_i^{s+1}}P_j^{k_1}\right]\partial_{z_j^{s+1}} \\
&\quad + \sum_{I_j \in \mathcal{H}^{s+1}}\left[\sum_{I_i \in \mathcal{H}^{s+1}} \left\{R_i^{k_1}\partial_{z_i^{s+1}}\left(P_j^{k_2} + R_j^{k_2}\right) - R_i^{k_2}\partial_{z_i^{s+1}}\left(P_j^{k_1} + R_j^{k_1}\right)\right\}\right. \\
&\quad + \left\{P_i^{k_1}\partial_{z_i^{s+1}}R_j^{k_2} - P_i^{k_2}\partial_{z_i^{s+1}}R_j^{k_1}\right\} \\
&\quad \left. + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_\ell^{k_1}\partial_{z_\ell^{s+1}}\left(P_j^{k_2} + R_j^{k_2}\right) - Q_\ell^{k_2}\partial_{z_\ell^{s+1}}\left(P_j^{k_1} + R_j^{k_1}\right)\right]\partial_{z_j^{s+1}} \\
&\quad + \sum_{I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}}\left[\sum_{I_i \in \mathcal{H}^{s+1}} \left(P_i^{k_1} + R_i^{k_1}\right)\partial_{z_i^{s+1}}Q_j^{k_2} - \left(P_i^{k_2} + R_i^{k_2}\right)\partial_{z_i^{s+1}}Q_j^{k_1}\right. \\
&\quad \left. + \sum_{I_\ell \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_\ell^{k_1}\partial_{z_\ell^{s+1}}Q_j^{k_2} - Q_\ell^{k_2}\partial_{z_\ell^{s+1}}Q_j^{k_1}\right]\partial_{z_j^{s+1}}.
\end{aligned}
$$

By the inductive hypothesis, one can proceed as follows.

- Taking into account the relation

$$\mathrm{ord}_{a^{s+1}}^{s+1}\left(R_i^{k_1}\right) > |I_i| - |I_{k_1}|, \quad \text{and} \quad \mathrm{ord}_{a^{s+1}}^{s+1}\left(\partial_{z_i^{s+1}}\left(P_j^{k_2} + R_j^{k_2}\right)\right) \geqslant |I_j| - |I_{k_2}| - |I_i|,$$

one deduces that

$$\mathrm{ord}_{a^{s+1}}^{s+1} R_i^{k_1}\partial_{z_i^{s+1}}\left(P_j^{k_2} + R_j^{k_2}\right) > |I_j| - |I_k|.$$

By a similar argument, $\mathrm{ord}_{a^{s+1}}^{s+1} R_i^{k_2} \partial_{z_i^{s+1}}(P_j^{k_1} + R_j^{k_1}) > |I_j| - |I_k|$. Therefore, one gets

$$\mathrm{ord}_{a^{s+1}}^{s+1}\left(R_i^{k_1} \partial_{z_i^{s+1}}(P_j^{k_2} + R_j^{k_2}) - R_i^{k_2} \partial_{z_i^{s+1}}(P_j^{k_1} + R_j^{k_1})\right) > |I_j| - |I_k|.$$

- Since $\mathrm{ord}_{a^{s+1}}^{s+1}(P_i^{k_1}) = |I_i| - |I_{k_1}|$ and $\mathrm{ord}_{a^{s+1}}^{s+1}(\partial_{z_i^{s+1}} R_j^{k_2}) > |I_j| - |I_{k_2}| - |I_i|$, then $\mathrm{ord}_{a^{s+1}}^{s+1}(P_i^{k_1} \partial_{z_i^{s+1}} R_j^{k_2}) > |I_j| - |I_k|$. By a similar argument, $\mathrm{ord}_{a^{s+1}}^{s+1}(P_i^{k_2} \partial_{z_i^{s+1}} R_j^{k_1}) > |I_j| - |I_k|$. One thus obtains

$$\mathrm{ord}_{a^{s+1}}^{s+1}\left(P_i^{k_1} \partial_{z_i^{s+1}} R_j^{k_2} - P_i^{k_2} \partial_{z_i^{s+1}} R_j^{k_1}\right) > |I_j| - |I_k|.$$

- Using the fact that $\mathrm{ord}_{a^{s+1}}^{s+1}(\partial_{z_\ell^{s+1}}(P_j^{k_2} + R_j^{k_2})) > |I_j| - |I_{k_2}| - (s+1)$ and $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1}) > s + 1 - |I_{k_1}|$, then $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1} \partial_{z_\ell^{s+1}}(P_j^{k_2} + R_j^{k_2})) > |I_j| - |I_k|$. By a similar argument, $\mathrm{ord}_{a^{s+1}}^{s+1} Q_\ell^{k_2} \partial_{z_\ell^{s+1}}(P_j^{k_1} + R_j^{k_1}) > |I_j| - |I_k|$. One deduces $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1} \partial_{z_\ell^{s+1}}(P_j^{k_2} + R_j^{k_2}) - Q_\ell^{k_2} \partial_{z_\ell^{s+1}}(P_j^{k_1} + R_j^{k_1})) > |I_j| - |I_k|$.
- Recall that $\mathrm{ord}_{a^{s+1}}^{s+1}(P_i^{k_1} + R_i^{k_1}) = |I_i| - |I_{k_1}|$, and $\mathrm{ord}_{a^{s+1}}^{s+1}(\partial_{z_i^{s+1}} Q_j^{k_2}) > s + 1 - |I_{k_2}| - |I_i|$, then $\mathrm{ord}_{a^{s+1}}^{s+1}((P_i^{k_1} + R_i^{k_1}) \partial_{z_i^{s+1}} Q_j^{k_2}) > s + 1 - |I_k|$. By a similar argument, $\mathrm{ord}_{a^{s+1}}^{s+1}((P_i^{k_2} + R_i^{k_2}) \partial_{z_i^{s+1}} Q_j^{k_1}) > s + 1 - |I_k|$. Therefore, it yields $\mathrm{ord}_{a^{s+1}}^{s+1}((P_i^{k_1} + R_i^{k_1}) \partial_{z_i^{s+1}} Q_j^{k_2} - (P_i^{k_2} + R_i^{k_2}) \partial_{z_i^{s+1}} Q_j^{k_1}) > s + 1 - |I_k|$.
- Since $\partial_{z_\ell^{s+1}} Q_j^{k_2}$ is a function, one knows by definition that $\mathrm{ord}_{a^{s+1}}^{s+1}(\partial_{z_\ell^{s+1}} Q_j^{k_2}) \geqslant 0$. As $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1}) > s + 1 - |I_{k_1}|$, one has $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1} \partial_{z_\ell^{s+1}} Q_j^{k_2}) > s + 1 - |I_{k_1}| = s + 1 - (|I_k| - |I_{k_2}|) > s + 1 - |I_k|$. By a similar argument, $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_2} \partial_{z_\ell^{s+1}} Q_j^{k_1}) > s + 1 - |I_k|$. One hence derives $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^{k_1} \partial_{z_\ell^{s+1}} Q_j^{k_2} - Q_\ell^{k_2} \partial_{z_\ell^{s+1}} Q_j^{k_1}) > s + 1 - |I_k|$.

Summing up the above terms, one gets, for $I_k \in \mathcal{K}^{s+1}$ of length $s_1 + 1$, that the bracket $\xi_{I_k}^{s+1}$ can be written in the form

$$\xi_{I_k}^{s+1}(z^{s+1}) = \sum_{I_j \in \mathcal{H}^{s+1}} (P_j^k(z^{s+1}) + R_j^k(z^{s+1})) \partial_{z_j^{s+1}} + \sum_{I_j \in \mathcal{K}^{s+1} \setminus \mathcal{H}^{s+1}} Q_j^k(z^{s+1}) \partial_{z_j^{s+1}},$$

with $\mathrm{ord}_{a^{s+1}}^{s+1}(P_j^k) = |I_j| - |I_k|$, $\mathrm{ord}_{a^{s+1}}^{s+1}(R_j^k) > |I_j| - |I_k|$, and $\mathrm{ord}_{a^{s+1}}^{s+1}(Q_\ell^k) > s + 1 - |I_k|$. Claim 10 is now proved. $\square$

In conclusion, properties (A1)–(A5) still hold true at step $s + 1$ in the Desingularization Algorithm. The induction step is established, which terminates the proof of Proposition 3.6. $\square$

4. Global steering method for regular systems

By taking into account the Desingularization Algorithm presented in Section 3, we assume in this section and without loss of generality that the family of vector fields $X = \{X_1, \dots, X_m\}$ is free up to step r (cf. Definition 3.6). Recall that, in that case, every point $x \in \Omega$ is regular and the growth vector is constant on Ω. We present in Section 4.1 an algebraic construction of privileged coordinates and a nonholonomic first-order approximation of X under canonical form. For regular systems, this construction also provides a continuously varying system of privileged coordinates. We then propose in Section 4.2 a global motion planning algorithm for regular systems.

4.1. Construction of the approximate system $\mathcal{A}^X$

For every point a in Ω, we construct the first-order approximate system $\mathcal{A}^X(a)$ of the system X at a (cf. Definition 2.14) as follows:

Step (1) Take $\{X_{I_j}\}_{I_j \in \mathcal{H}^r}$. Set $w_j = \tilde{w}_j$ for $j = 1, \ldots, n$.

Step (2) Construct the linear system of coordinates $y = (y_1, \ldots, y_n)$ such that $\partial_{y_j} = X_{I_j}(a)$.

Step (3) Build the system of privileged coordinates $\tilde{z} = (\tilde{z}_1, \ldots, \tilde{z}_n)$ by the following iterative formula: for $j = 1, \ldots, n$,

$$\tilde{z}_j := y_j + \sum_{k=2}^{w_j - 1} h_k(y_1, \ldots, y_{j-1}), \tag{4.1}$$

where, for $k = 2, \ldots, w_j - 1$,

$$h_k(y_1, \ldots, y_{j-1}) = - \sum_{\substack{|\alpha| = k \\ w(\alpha) < w_j}} X_{I_1}^{\alpha_1} \cdots X_{I_{j-1}}^{\alpha_{j-1}} \cdot \left(y_j + \sum_{q=2}^{k-1} h_q \right)(y)_{|y=0} \frac{y_1^{\alpha_1}}{\alpha_1!} \cdots \frac{y_{j-1}^{\alpha_{j-1}}}{\alpha_{j-1}!},$$

with $|\alpha| := \alpha_1 + \cdots + \alpha_n$.

Step (4) For $i = 1, \ldots, m$, compute the Taylor expansion of $X_i(\tilde{z})$ at 0, and express every vector field as a sum of vector fields which are homogeneous with respect to the weighted degree defined by the sequence $(w_j)_{j=1,\ldots,n}$:

$$X_i(\tilde{z}) = X_i^{(-1)}(\tilde{z}) + X_i^{(0)}(\tilde{z}) + \cdots,$$

where we use $X_i^{(k)}(\tilde{z})$ to denote the sum of all the terms of weighted degree equal to k. Set $\hat{X}_i^a(\tilde{z}) := X_i^{(-1)}(\tilde{z})$.

Step (5) For $j = 1, \ldots, n$, identify homogeneous polynomials Ψ_j of weighted degree equal to w_j such that, in the system of privileged coordinates $z := (z_1, \ldots, z_n)$ defined by

$$z_j := \Psi_j(\tilde{z}_1, \ldots, \tilde{z}_{j-1}), \quad \text{for } j = 1, \ldots, n,$$

the approximate system

$$\hat{X}^a(z) = \left\{ z_* \hat{X}_1^a(\tilde{z}), \ldots, z_* \hat{X}_m^a(\tilde{z}) \right\}$$

is in the canonical form.

Step (6) Set $\mathcal{A}^X(a) := \hat{X}^a$ and $\Phi^X(a, \cdot) :=$ the mapping $x \mapsto z$.

Remark 4.1. Steps (1)–(3) construct a system of privileged coordinates $\tilde{z}$. The proof that $\tilde{z}$ is a system of privileged coordinates is essentially based on Lemma 3.7. Roughly speaking, the idea to obtain $\tilde{z}_j$ from y_j goes as follows: for every $\alpha = (\alpha_1, \ldots, \alpha_n)$ with $w(\alpha) < w_j$ (so $\alpha_j = \cdots = \alpha_n = 0$), compute $X_{I_1}^{\alpha_1} \cdots X_{I_{j-1}}^{\alpha_{j-1}} \cdot y_j(y)_{|y=0}$. If it is not equal to zero, then replace y_j by

$$y_j - \left(X_{I_1}^{\alpha_1} \cdots X_{I_{j-1}}^{\alpha_{j-1}} \cdot y_j \right)(y)_{|y=0} \frac{y_1^{\alpha_1}}{\alpha_1!} \cdots \frac{y_{j-1}^{\alpha_{j-1}}}{\alpha_{j-1}!}.$$

With that new value of y_j, one gets $X_{I_1}^{\alpha_1} \cdots X_{I_{j-1}}^{\alpha_{j-1}} \cdot y_j(y)_{|_{y=0}} = 0$. Therefore, by Lemma 3.7, one has $\mathrm{ord}_a(\tilde{z}_j) \geqslant w_j$ for $j = 1, \ldots, n$. On the other hand, since Step (3) of the construction does not modify the linear part, the system of coordinates $\tilde{z}$ remains adapted. By Remark 2.5, one also has $\mathrm{ord}_a(\tilde{z}_j) \leqslant w_j$, and therefore, $\mathrm{ord}_a(\tilde{z}_j) = w_j$.

Remark 4.2. The existence of Ψ_j in Step (5) is guaranteed by a simple modification of Claim 9, page 1926, see also Remarks 3.8 and 3.9. The key point is, in the current case, the exponential coordinates are *algebraic*.

Remark 4.3. We will propose in Section 5 an effective and exact method for steering general nilpotent systems given in the canonical form.

It results from [3] that, for *regular systems*, the mapping $\Phi^X : (a, x) \to z$ is a continuously varying system of privileged coordinates on Ω. Note also that the coordinates z are obtained from y by expressions of the form

$$z_1 = y_1,$$

$$z_2 = y_2 + \mathrm{pol}_2(y_1),$$

$$\vdots$$

$$z_n = y_n + \mathrm{pol}_n(y_1, \ldots, y_{n-1}),$$

where, for $j = 1, \ldots, n$, the function $\mathrm{pol}_j(\cdot)$ is a polynomial which does not contain constant nor linear terms. Due to the triangular form of this change of coordinates, the inverse change of coordinates from z to y bears exactly the same form. Therefore, the mapping $z = \Phi^X(a, \cdot)$ is defined on the whole Ω, i.e., Φ^X has an infinite injectivity radius. We also note that, by construction, $\mathcal{A}^X$ is a nonholonomic first-order approximation (cf. Definition 2.14) and its continuity results from the continuity of the mapping $\Phi^X : (a, x) \mapsto z$. In summary, we have the following proposition.

Proposition 4.1. *The mapping Φ^X is a continuously varying system of privileged coordinates on Ω and the mapping $\mathcal{A}^X$ is a continuous approximation of X on Ω.*

The following theorem is a consequence of Proposition 4.1 and Corollary 2.5.

Theorem 4.2. *Let $\mathcal{V}^c$ be a compact subset of Ω. If $\mathcal{A}^X$ is provided with a sub-optimal steering law (cf. Definitions 2.16 and 2.17), then the LAS method* AppSteer *associated with $\mathcal{A}^X$ and its steering law (cf. Definition 2.18) is uniformly locally contractive on $\mathcal{V}^c$.*

Remark 4.4. Due to Step (5) in the construction procedure, the approximate system $\mathcal{A}^X(a)$ is under canonical form in a system of privileged coordinates z. Therefore, $\mathcal{A}^X(a)$ has always the same form, regardless of the control system X or the approximate point $a \in \Omega$. The specificity of each system or each approximate point is hidden in the change of coordinates Φ^X.

Remark 4.5. It is important to notice that the approximate system used in the LAS method is a nonholonomic first-order approximation at the *goal* point a (cf. Definition 2.18). Therefore, the steering control always displaces $\mathcal{A}^X(a)$ from some position (which is the image by $\Phi^X(a, \cdot)$ of the current point of the original system) to 0 (which is $\Phi^X(a, a)$ by construction) in coordinates z. The latter fact plays a crucial role in getting the *sub-optimality* for the steering law (see Section 5.3 for more details).

4.2. Approximate steering algorithm

Let $\mathcal{V}^c \subset \Omega$ be a connected compact set equal to the closure of its interior and $(x^{\text{initial}}, x^{\text{final}}) \in \mathcal{V}^c \times \mathcal{V}^c$. We devise, under the assumptions of Theorem 4.2, an algorithm (Algorithm 3 below) which steers system (2.1) from x^{initial} to x^{final}. That algorithm does not require any a priori knowledge on the critical distance $\varepsilon_{\mathcal{V}^c}$. Note that this algorithm bears similarities with *trust-region* methods in optimization (see [6] for more details).

Recall first that the family of vector fields $X = \{X_1, \ldots, X_m\}$ is assumed to be free up to step r. As a consequence the weights $(w_1, \ldots, w_n)$ are equal at every point $a \in \mathcal{V}^c$ to $(\tilde{w}_1, \ldots, \tilde{w}_n)$, the free weights of step r. Hence the pseudo-norm $\| \cdot \|_a$ (see Definition 2.13) does not depend on $a \in \mathcal{V}^c$ and will be denoted as $\| \cdot \|_r$.

The parameterized path $t \mapsto \delta_t(x)$ is defined by

$$\delta_t(x) := \left(t^{w_1} z_1(x), \ldots, t^{w_n} z_n(x) \right), \quad \text{for } x \in \Omega,$$

where $z := \Phi^X(x^{\text{final}}, \cdot)$. Note that δ_t is the (weighted) *dilatation* in privileged coordinates at x^{final} with parameter t. In particular, $\|z(\delta_t(x))\|_r = |t| \|z(x)\|_r$. We also define the function Subgoal as follows:

Subgoal($\bar{x}, \eta, j$)
1. $t_j := \max(0, 1 - \frac{j\eta}{\|z(\bar{x})\|_r})$;
2. Subgoal($\bar{x}, \eta, j$) := $\delta_{t_j}(\bar{x})$

We note that the formula for generating t_j guarantees that

$$\left\| z\big(\text{Subgoal}(\bar{x}, \eta, j)\big) - z\big(\text{Subgoal}(\bar{x}, \eta, j-1)\big) \right\|_r \leqslant \eta,$$

and that $x^d = x^{\text{final}}$ for j large enough.

Algorithm 3 GlobalFree($x^{\text{initial}}, x^{\text{final}}, e, \mathcal{V}^c$, AppSteer).

1: $i := 0$; $j := 1$;
2: $x_i := x^{\text{initial}}$; $\bar{x} := x^{\text{initial}}$;
3: $\eta := \|z(x^{\text{initial}})\|_r$;   {initial choice of the maximum step size;}
4: **while** $\|z(x_i)\|_r > e$ **do**
5:     $x^d := $ Subgoal $(\bar{x}, \eta, j)$;
6:     $x := $ AppSteer (x_i, x^d);
7:     **if** $\|\Phi^X(x^d, x)\|_r > \frac{1}{2} \|\Phi^X(x^d, x_i)\|_r$ **then** {if the system is not approaching the subgoal,}
8:         $\eta := \frac{\eta}{2}$;   {reduce the maximum step size,}
9:         $\bar{x} := x_i$; $j := 1$;   {change the path $\delta_{0,t}(\bar{x})$.}
10:     **else**
11:         $i := i + 1$; $j := j + 1$;
12:         $x_i := x$; $x_i^d := x^d$;
13:     **end if**
14: **end while**
15: **return** x_i.

The global convergence of Algorithm 3 is established in the following theorem. For the sake of clarity, we first assume that the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant 0}$ constructed by Algorithm 3 both stay within $\mathcal{V}^c$. This assumption being of a purely numerical nature, we explain at the end of this section how we can remove it by adding suitable intermediate steps to Algorithm 3.

Theorem 4.3. *Let $\mathcal{V}^c \subset \Omega$ be a connected compact set equal to the closure of its interior. Assume that*

(i) *the approximate system $\mathcal{A}^X$ is provided with a sub-optimal steering law;*
(ii) *the LAS method* AppSteer *is associated with $\mathcal{A}^X$ and its steering law.*

Then, $\forall (x^{\text{initial}}, x^{\text{final}}) \in \mathcal{V}^c \times \mathcal{V}^c$, Algorithm 3 *terminates in a finite number of steps for any choice of the tolerance $e > 0$ provided that the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant 0}$ both belong to $\mathcal{V}^c$.*

Proof. Note first that, if the conditional statement of Line 7 is not true for every i greater than some i_0, then $x_i^d = x^{\text{final}}$ after a finite number of iterations. In this case, the error $\|z(x_i)\|_r$ is reduced at each iteration and the algorithm stops when it becomes smaller than the given tolerance e. This happens in particular if $d(x_i, x^d) < \varepsilon_{\mathcal{V}^c}$ for all i greater than i_0 because condition (2.8) is verified. Another preliminary remark is that, due to the continuity of the control distance and of the function $\|z(\cdot)\|_r$, there exists $\bar{\eta} > 0$ such that, for every pair $(x_1, x_2) \in \mathcal{V}^c \times \mathcal{V}^c$, one has

$$\left\| z(x_1) - z(x_2) \right\|_r < \bar{\eta} \quad \Longrightarrow \quad d(x_1, x_2) < \frac{\varepsilon_{\mathcal{V}^c}}{2}. \tag{4.2}$$

In the following, we will prove by induction that if, at some step i_0, one has $\eta < \bar{\eta}$, then, for all $i > i_0$, one has $d(x_{i-1}, x_i^d) < (1/2 + \cdots + (1/2)^{i-i_0})\varepsilon_{\mathcal{V}^c} < \varepsilon_{\mathcal{V}^c}$.

We assume without loss of generality that $i_0 = 0$ and $\bar{x} = x_0$. For $i = 1$, by construction, $x^d = \text{Subgoal}(x_0, \eta, 1)$ and $\|z(x_0) - z(x^d)\|_r \leqslant \eta < \bar{\eta}$.

In view of (4.2), one then has $d(x_0, x^d) < \varepsilon_{\mathcal{V}^c}/2$. In view of (2.8), the conditional statement of Line 7 is not true, therefore $x_1^d = x^d$ and one has $d(x_0, x_1^d) < \varepsilon_{\mathcal{V}^c}/2$.

Assume now that for $i > 1$ one has:

$$d\left(x_{i-2}, x_{i-1}^d\right) < \left(1/2 + \cdots + (1/2)^{i-1}\right)\varepsilon_{\mathcal{V}^c}. \tag{4.3}$$

The subgoal x_{i-1}^d is of the form $\text{Subgoal}(\bar{x}, \eta, j)$. Let $x^d = \text{Subgoal}(\bar{x}, \eta, j+1)$. One can write:

$$d\left(x_{i-1}, x^d\right) \leqslant d\left(x_{i-1}, x_{i-1}^d\right) + d\left(x_{i-1}^d, x^d\right).$$

By construction, it is $\|z(x_{i-1}^d) - z(x^d)\|_r \leqslant \eta < \bar{\eta}$, which implies $d(x_{i-1}^d, x^d) < \varepsilon_{\mathcal{V}^c}/2$. The induction hypothesis (4.3) implies that $d(x_{i-1}, x_{i-1}^d) \leqslant \frac{1}{2} d(x_{i-2}, x_{i-1}^d)$.

Finally, one gets

$$d\left(x_{i-1}, x^d\right) \leqslant \frac{1}{2} d\left(x_{i-2}, x_{i-1}^d\right) + d\left(x_{i-1}^d, x^d\right) \leqslant \left(1/2 + \cdots + (1/2)^i\right)\varepsilon_{\mathcal{V}^c}.$$

In view of (2.8), the conditional statement of Line 7 is not true, and so $x_i^d = x^d$. This ends the induction.

Notice that, at some step i, $\eta \geqslant \bar{\eta}$, the conditional statement of Line 7 could be false. In this case, η is decreased as in Line 8. The updating law of η guarantees that after a finite number of iterations of Line 8, there holds $\eta < \bar{\eta}$. This ends the proof. $\quad \square$

When the working space Ω is equal to the whole $\mathbb{R}^n$, the assumption that the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant 0}$ constructed by Algorithm 3 both stay within a compact set $\mathcal{V}^c$ can be removed. This requires a simple modification of Lines 11 and 12 of Algorithm 3.

We choose a real number R close to one, precisely $(\frac{1}{2})^{1/(r+1)^2} < R < 1$, where r is the maximum value of the degree of nonholonomy of system (2.1) on $\mathcal{V}^c$. For every non-negative integer k, we set $R_k = 1 + R + \cdots + R^k$. The algorithm is modified as follows. Introduce first a new variable k, and add the initialization $k := 0$. Replace then Lines 11 and 12 of Algorithm 3 by the procedure below.

This procedure guarantees that the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant 0}$ of the algorithm both belong to the compact set

$$K = \left\{ x \in \mathbb{R}^n \colon \left\| z(x) \right\|_r \leqslant \frac{1}{1 - R} \left\| z\left(x^{\text{initial}}\right) \right\|_r \right\}.$$

1: **if** $\|z(x)\|_r \geqslant R_{k+1}\|z(x^{\text{initial}})\|_r$ **then**
2:	$\eta := \frac{\eta}{2}$;
3: **else if** $R_k\|z(x^{\text{initial}})\|_r \leqslant \|z(x)\|_r < R_{k+1}\|z(x^{\text{initial}})\|_r$ **then**
4:	$i := i + 1$; $j := j + 1$;
5:	$x_i := x$; $x_i^d := x^d$;
6:	$\eta := \frac{\eta}{2}$;
7:	$k := k + 1$;
8: **else if** $\|z(x)\|_r \leqslant R_k\|z(x^{\text{initial}})\|_r$ **then**
9:	$i := i + 1$; $j := j + 1$;
10:	$x_i := x$; $x_i^d := x^d$;
11: **end if**

Moreover, at each iteration of the algorithm, the new variable k is such that

$$\left\|z(x_i)\right\|_r \geqslant R_k\left\|z(x^{\text{initial}})\right\|_r \quad \Longrightarrow \quad \eta \leqslant \frac{\|z(x^{\text{initial}})\|_r}{2^k}.$$

For the sake of clarity, we state here the complete modified algorithm named as Algorithm 4.

Algorithm 4 GlobalFree_Modified(x^{initial}, x^{final}, e, $\mathcal{V}^c$, AppSteer).

1: $i := 0$; $j := 1$;
2: $x_i := x^{\text{initial}}$; $\bar{x} := x^{\text{initial}}$;
3: $\eta := \|z(x^{\text{initial}})\|_r$;	{initial choice of the maximum step size;}
4: **while** $\|z(x_i)\|_r > e$ **do**
5:	$x^d :=$ Subgoal $(\bar{x}, \eta, j)$;
6:	$x :=$ AppSteer (x_i, x^d);
7:	**if** $\|\Phi^X(x^d, x)\|_r > \frac{1}{2}\|\Phi^X(x^d, x_i)\|_r$ **then** {if the system is not approaching the subgoal,}
8:		$\eta := \frac{\eta}{2}$;	{reduce the maximum step size,}
9:		$\bar{x} := x_i$; $j := 1$;	{change the path $\delta_{0,t}(\bar{x})$.}
10:	**else if** $\|z(x)\|_r \geqslant R_{k+1}\|z(x^{\text{initial}})\|_r$ **then**
11:		$\eta := \frac{\eta}{2}$;
12:	**else if** $R_k\|z(x^{\text{initial}})\|_r \leqslant \|z(x)\|_r < R_{k+1}\|z(x^{\text{initial}})\|_r$ **then**
13:		$i := i + 1$; $j := j + 1$;
14:		$x_i := x$; $x_i^d := x^d$;
15:		$\eta := \frac{\eta}{2}$;
16:		$k := k + 1$;
17:	**else if** $\|z(x)\|_r \leqslant R_k\|z(x^{\text{initial}})\|_r$ **then**
18:		$i := i + 1$; $j := j + 1$;
19:		$x_i := x$; $x_i^d := x^d$;
20:	**end if**
21: **end while**
22: **return** x_i.

Proposition 4.4. *Let $\mathcal{V}^c \subset \Omega$ be a connected compact set equal to the closure of its interior. Under the assumptions* (i) *and* (ii) *of Theorem 4.3, $\forall(x^{\text{initial}}, x^{\text{final}}) \in \mathcal{V}^c \times \mathcal{V}^c$, Algorithm 4 terminates in a finite number of iterations for any choice of the tolerance $e > 0$.*

Proof. Notice that Lines 17, 18, and 19 in Algorithm 4 are identical to Lines 10, 11, and 12 in Algorithm 3. Therefore, it is enough to show that, after a finite number of iterations, the condition of Line 17 in Algorithm 4 holds true. Another preliminary remark is that the distance $\|z(x) - z(y)\|_r$ gives a rough estimate of the sub-Riemannian distance. Indeed it follows from Corollary 2.3 that, for every pair of close enough points $(x, y) \in \mathcal{V}^c \times \mathcal{V}^c$, one has

$$\frac{1}{C_0}\left\|z(x) - z(y)\right\|_r^{r+1} \leqslant d(x, y) \leqslant C_0\left\|z(x) - z(y)\right\|_r^{1/(r+1)}, \tag{4.4}$$

where C_0 is a positive constant. As a consequence, Eq. (4.2) holds true if $\bar{\eta} \leqslant (\varepsilon_{\mathcal{V}^c}/(2C_0))^{r+1}$.

Let us choose a positive $\bar{\eta}$ smaller than $(\varepsilon_{\mathcal{V}^c}/(2C_0))^{r+1}$. We next show that if, at some step i_0, $\eta < \bar{\eta}$, then the case of Line 10 and the one of Line 12 occur only in a finite number of iterations. Recall first that, from the proof of Theorem 4.3, one gets, for every $i > i_0$,

$$\left\| z\bigl(x_i^d\bigr) \right\|_r \leqslant \left\| z(x_{i_0}) \right\|_r \quad \text{and} \quad d\bigl(x_i, x_i^d\bigr) \leqslant \varepsilon_{\mathcal{V}^c}.$$

In view of Eq. (4.4), an obvious adaptation of the latter proof yields, for every $i > i_0$, $d(x_i, x_i^d) \leqslant 2C_0 \eta^{1/(r+1)}$, and thus

$$\left\| z(x_i) - z\bigl(x_i^d\bigr) \right\|_r \leqslant \bigl(2C_0^2\bigr)^{1/(r+1)} \eta^{1/(r+1)^2}.$$

Finally one gets

$$
\begin{aligned}
\left\| z(x_i) \right\|_r &\leqslant \left\| z\bigl(x_i^d\bigr) \right\|_r + \left\| z(x_i) - z\bigl(x_i^d\bigr) \right\|_r \\
&\leqslant \left\| z(x_{i_0}) \right\|_r + \bigl(2C_0^2\bigr)^{1/(r+1)} \eta^{1/(r+1)^2}.
\end{aligned}
\tag{4.5}
$$

On the other hand, there exists an integer k_0 such that $\eta \geqslant \frac{\| z(x^{\text{initial}}) \|_r}{2^{k_0}}$. This implies that $\| z(x_{i_0}) \|_r \leqslant R_{k_0} \| z(x^{\text{initial}}) \|_r$. Up to reducing $\bar{\eta}$, and so increasing k_0, assume

$$\bigl(2C_0^2\bigr)^{1/(r+1)} \left(\frac{\| z(x^{\text{initial}}) \|_r}{2^{k_0}} \right)^{1/(r+1)^2} \leqslant R^{k_0+1} \left\| z\bigl(x^{\text{initial}}\bigr) \right\|_r,$$

since one has chosen $R > (\frac{1}{2})^{1/(r+1)^2}$. Using Eq. (4.5), it holds, for every $i \geqslant i_0$, $\| z(x_i) \|_r \leqslant R_{k_0} \| z(x^{\text{initial}}) \|_r + R^{k_0+1} \| z(x^{\text{initial}}) \|_r = R_{k_0+1} \| z(x^{\text{initial}}) \|_r$. Therefore, the case of Line 10 and the one of Line 12 occur in at most $k_0 + 1$ iterations. Applying again the arguments of the proof of Theorem 4.3, the conclusion follows. $\quad \square$

Remark 4.6. It is worth pointing out that the additional steps involved in Algorithm 4 are designed to prevent the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant 0}$ from accumulating toward the boundary of the compact $\mathcal{V}^c$. There exist other numerical artifacts of probabilistic nature which solve this problem. One also deduces from the proof of Proposition 4.4 that if the points x^{initial} and x^{final} are *far* enough from the boundary of $\mathcal{V}^c$, the sequences $(x_i)_{i \geqslant 0}$ and $(x_i^d)_{i \geqslant}$ will remain in $\mathcal{V}^c$.

5. Exact steering method for nilpotent systems

In this section, we devise an exact steering method for general nilpotent systems. Without loss of generality, we assume that the system $X = \{X_1, \ldots, X_m\}$ is nilpotent of step r, free up to step r, and given in the canonical form in coordinates x. Recall that, under this assumption, the dynamics is written as follows

$$
\begin{aligned}
\dot{x}_i &= u_i, \quad \text{if } i = 1, \ldots, m; \\
\dot{x}_I &= \frac{1}{k!} x_{I_L} \dot{x}_{I_R}, \quad \text{if } X_I = \text{ad}_{X_{I_L}}^k X_{I_R}, \ I_L, I_R \in \mathcal{H}^r,
\end{aligned}
\tag{5.1}
$$

where the components of x are numbered by the elements of $\mathcal{H}^r$, i.e., for $I \in \mathcal{H}^r$, the component x_I corresponds to the element X_I. We also assume that we want to steer the system (5.1) from any point $x \in \mathbb{R}^{\tilde{n}_r}$ to the origin 0 of $\mathbb{R}^{\tilde{n}_r}$.

Remark 5.1. Note that these two assumptions are not restrictive since, for general nilpotent systems, in order to steer from x^{initial} to x^{final}, it suffices to apply the Desingularization Algorithm at the final point x^{final} (see also Remark 3.6).

This method can also be applied for the construction of a sub-optimal steering law for the approximate system $\mathcal{A}^X$ defined in Section 4.1. For practical uses, we require that the inputs give rise to regular trajectories (i.e., at least C^1), which are not too "complex" in the sense that, during the control process, we do not want the system to stop too many times or to make a large number of maneuvers.

Several methods were proposed in the literature for steering nilpotent systems. In [24], the authors make use of piecewise constant controls and obtain smooth controls by imposing some special parameterization (namely by requiring the control system to stop during the control process). In that case, the regularity of the inputs is recovered by using a reparameterization of the time, which cannot prevent in general the occurrence of cusps or corners for the corresponding trajectories. However, regularity of the *trajectories* is generally mandatory for robotic applications. Therefore, the method proposed in [25] is not adapted to such applications. In [24], the proposed inputs are polynomial functions in time, but an algebraic system must be inverted in order to access to these inputs. Moreover, the size and the degree of this algebraic system increase exponentially with respect to the dimension of state space, and there does not exist a general efficient exact method to solve it. Even the existence of solutions is a nontrivial issue. Furthermore, the methods [24] and [25] both make use of exponential coordinates which are not explicit and thus require in general numerical integrations of nonlinear differential equations. That prevents the use of these methods in an iterative scheme such as Algorithm 1. Let us also mention the path approximation method by Liu and Sussmann [27], which uses unbounded sequences of sinusoids. Even though this method bears similar theoretical aspects with our method, it is not adapted from a numerical point of view to the motion planning issue since it relies on a limit process of highly oscillating inputs.

5.1. Steering by sinusoids

We consider input functions in the form of linear combinations of sinusoids with integer frequencies. In [29], authors used this family of inputs to control the chained-form systems.

We first note that if every component of the input $u = (u_1, \ldots, u_m)$ in Eq. (5.1) is a linear combination of sinusoids with integer frequencies, then the dynamics of every component in Eq. (5.1) is also a linear combination of sinusoids with integer frequencies, which are themselves linear combinations of frequencies involved in the input u. One may therefore expect to move some components during a 2π time-period without modifying others if the frequencies in u are properly chosen. Due to the triangular form of Eq. (5.1), it is reasonable to expect to move the components of x one after another according to the order "$\prec$" induced by the P. Hall basis. In that case, one must ensure that all the components already moved to their preassigned values return to the same values after each 2π-period of control process, while the component under consideration arrives to its preassigned position. However, all the components cannot be moved independently by using sinusoids. For that purpose, we introduce the following notion of *equivalence*.

Definition 5.1 (*Equivalence*). Two elements X_I and X_J in the P. Hall family are said to be *equivalent* if $\Delta_i(X_I) = \Delta_i(X_J)$ for $i = 1, \ldots, m$, where we use $\Delta_i(X_I)$ to denote the number of times X_i occurs in X_I. We write $X_I \sim X_J$ if X_I and X_J are equivalent and *equivalence classes* will be denoted by

$$\mathcal{E}_X(\ell_1, \ldots, \ell_m) := \big\{ X_I \mid \Delta_i(X_I) = l_i, \text{ for } i = 1, \ldots, m \big\}.$$

We say that the components x_I and x_J are *equivalent* if the corresponding brackets X_I and X_J are equivalent and *equivalent classes for components* are defined as follows,

$$\mathcal{E}_x(\ell_1, \ldots, \ell_m) := \big\{ x_I \mid X_I \in \mathcal{E}_X(\ell_1, \ldots, \ell_m) \big\}.$$

Remark 5.2. We will see in the following subsections that the frequencies occurring in the dynamics of x_I only depend on the equivalence class of x_I, and not on the structure of the bracket X_I. Therefore, the equivalent components (in the sense of Definition 5.1) cannot be moved separately by using sinusoids.

Definition 5.2 (*Ordering of equivalence classes*). Let $\mathcal{E}_x(\ell_1, \ldots, \ell_m)$ and $\mathcal{E}_x(\tilde{\ell}_1, \ldots, \tilde{\ell}_m)$ be two equivalence classes. $\mathcal{E}_x(\ell_1, \ldots, \ell_m)$ is said to be smaller than $\mathcal{E}_x(\tilde{\ell}_1, \ldots, \tilde{\ell}_m)$ if the smallest element (in the sense of "$\prec$") in $\mathcal{E}_x(\ell_1, \ldots, \ell_m)$ is smaller than the one in $\mathcal{E}_x(\tilde{\ell}_1, \ldots, \tilde{\ell}_m)$, and we write (by abuse of notation) $\mathcal{E}_x(\ell_1, \ldots, \ell_m) \prec \mathcal{E}_x(\tilde{\ell}_1, \ldots, \tilde{\ell}_m)$.

Let $\{\mathcal{E}_x^1, \mathcal{E}_x^2, \ldots, \mathcal{E}_x^{\tilde{N}}\}$ be the partition of the set of the components of x induced by Definition 5.1. Assume that, for every pair $(i, j) \in \{1, \ldots, \tilde{N}\}^2$ with $i < j$, one has $\mathcal{E}_x^i \prec \mathcal{E}_x^j$. Our control strategy consists in displacing these equivalence classes one after another according to the ordering "$\prec$" by using sinusoidal inputs. For every $j = 1, \ldots, \tilde{N}$, the key point is to determine how to construct an input u^j defined on $[0, 2\pi]$ such that the two following conditions are verified:

(C1) under the action of u^j, every element of $\mathcal{E}_x^j$ reaches its preassigned value at $t = 2\pi$;
(C2) under the action of u^j, for every $i < j$, every element of $\mathcal{E}_x^i$ returns at $t = 2\pi$ to its value taken at $t = 0$.

Once one knows how to construct an input u^j verifying (C1) and (C2) for every $j = 1, \ldots, \tilde{N}$, it suffices to *concatenate* them to control the complete system.

Definition 5.3 (*Concatenation*). The concatenation of $u^1, \ldots, u^{\tilde{N}}$ is defined on the interval $[0, 2\tilde{N}\pi]$ by

$$u^1 * \cdots * u^{\tilde{N}}(t) := u^j\big(t - 2(j-1)\pi\big), \tag{5.2}$$

for $t \in [2(j-1)\pi, 2j\pi]$ and $j \in \{1, \ldots, \tilde{N}\}$.

Remark 5.3. As we will show later (see Remark 5.8), for every positive integer k, it is possible to make C^k concatenations such that the inputs are globally of class C^k and the corresponding trajectories are not only piecewise smooth, but also globally of class C^{k+1}.

5.2. Choice of frequencies

In this section, we fix an equivalence class $\mathcal{E}_x^j$. We choose frequencies in u^j such that Conditions (C1) and (C2) are verified. For the sake of clarity, we first treat the case $m = 2$ in Sections 5.2.1 and 5.2.2, and we show, in Section 5.2.3, how to adapt the method to greater values of m.

5.2.1. A simple case: $m = 2$ and $\mathrm{Card}(\mathcal{E}_x^j) = 1$

Let x_I be the only element of $\mathcal{E}_x^j$ and X_I the corresponding bracket. Let $m_1 := \Delta_1(X_I)$ and $m_2 := \Delta_2(X_I)$.

Proposition 5.1. *Consider three positive integers ω_1, ω_2, ω_3, and $\varepsilon \in \{0, 1\}$ such that*

$$\begin{cases} \omega_3 = m_1\omega_1 + (m_2 - 1)\omega_2, \\ \varepsilon = m_1 + m_2 - 1 \pmod 2, \end{cases} \tag{5.3}$$

and

$$\omega_2 > (m_1 + m_2)m_1. \tag{5.4}$$

By choosing properly ζ, the control

$$u_1 = \cos \omega_1 t, \qquad u_2 = \cos \omega_2 t + \zeta \cos\left(\omega_3 t - \varepsilon \frac{\pi}{2}\right) \tag{5.5}$$

steers, during $[0, 2\pi]$, the component x_I from any initial value to any preassigned final value without modifying any component x_J, with $J \prec I$. Moreover, $x_I(2\pi) - x_I(0)$ gives rise to a non-zero linear function of ζ, where ζ is the coefficient in front of $\cos(\omega_3 t - \varepsilon \frac{\pi}{2})$ in Eq. (5.5).

The key point is to understand the frequencies occurring in the dynamics $\dot{x}_I$.

Lemma 5.2. *For $J \leqslant I$, the dynamics $\dot{x}_J$ is a linear combination of cosine functions of the form*

$$\cos\left\{(\ell_1 \omega_1 + \ell_2 \omega_2 + \ell_3 \omega_3)t - (\ell_3 \varepsilon + \ell_1 + \ell_2 + \ell_3 - 1)\frac{\pi}{2}\right\}, \tag{5.6}$$

where $\ell_1, \ell_2, \ell_3 \in \mathbb{Z}$ satisfy $|\ell_1| \leqslant m_1$, $|\ell_2| + |\ell_3| \leqslant m_2$.
   In particular, the term

$$\cos\left[(m_1 \omega_1 + (m_2 - 1)\omega_2 - \omega_3)t - (-\varepsilon + m_1 + m_2 - 1)\frac{\pi}{2}\right]$$

occurs in $\dot{x}_I$ with a zero coefficient depending linearly on ζ.

Proof. The proof goes by induction on $|J|$.

- $|J| = 1$, the result is true since $\dot{x}_{I_1} = u_1$ and $\dot{x}_{I_2} = u_2$.
- *Inductive step*: Assume that the result holds true for all $\tilde{J}$ such that $|\tilde{J}| < s$. We show that it remains true for J such that $|J| = s$.
  By construction, we have $X_J = \mathrm{ad}_{X_{J_1}}^k X_{J_2}$ with $|J_1| < s$ and $|J_2| < s$. Then,

$$\dot{x}_J = \frac{1}{k!} x_{J_1}^k \dot{x}_{J_2}, \tag{5.7}$$

  $\dot{x}_{J_2}$ is given by the inductive hypothesis and x_{I_1} is obtained by integration of Eq. (5.6). By using product formulas for cosine function, the result still holds true for J of length s. This ends the proof of Lemma 5.2. $\square$

Proof of Proposition 5.1. First note that integrating between 0 and 2π a function of the form $\cos(\gamma t + \overline{\gamma} \frac{\pi}{2})$ with $(\gamma, \overline{\gamma}) \in \mathbb{N}^2$ almost always gives 0 except for $\gamma = 0$ and $\overline{\gamma} = 0 \pmod 2$ at the same time. Therefore, in order to obtain a nontrivial contribution for x_I, $\dot{x}_I$ must contain some cosine functions verifying the following condition

$$\begin{cases} \ell_1 \omega_1 + \ell_2 \omega_2 + \ell_3 \omega_3 = 0, \\ \ell_3 \varepsilon + \ell_1 + m_2 + \ell_3 - 1 \equiv 0 \pmod 2, \end{cases} \tag{5.8}$$

and this condition must not be satisfied by $J \prec I$ in order to avoid a change in the component x_J.
   Under conditions (5.3) and (5.4), we claim that

(1) $(m_1, m_2 - 1, -1, \varepsilon)$ is the only 4-tuple verifying (5.8) for x_I, and $x_I(2\pi) - x_I(0)$ is a non-zero *linear* function of ζ;
(2) Eq. (5.8) is never satisfied for x_J with $J < I$.

Indeed, consider $(\ell_1, \ell_2, \ell_3) \in \mathbb{Z}^3$ verifying $|\ell_1| \leqslant m_1$, $|\ell_2| + |\ell_3| \leqslant m_2$. One has

$$
\begin{aligned}
\ell_1 \omega_1 + \ell_2 \omega_2 + \ell_3 \omega_3 &= \ell_1 \omega_1 + \ell_2 \omega_2 + \ell_3 \big((m_2 - 1)\omega_2 + m_1 \omega_1\big) \\
&= \big(\ell_3(m_2 - 1) + \ell_2\big)\omega_2 + (\ell_1 + \ell_3 m_1)\omega_1.
\end{aligned}
\tag{5.9}
$$

Assume that $\omega_2 > (m_1 + m_2)m_1 \omega_1$. Then, except for the 4-tuple $(m_1, m_2, m_3, \varepsilon)$ verifying Eq. (5.3), the only possibility to have the right-hand side of Eq. (5.9) equal to 0 is $\ell_1 = \ell_2 = \ell_3 = 0$. In that case,

$$
\ell_1 + \ell_2 + \ell_3 \neq 1 \pmod 2.
$$

Then, Eq. (5.8) is not satisfied, and (2) is proved.

Due to Eq. (5.5), the power of ζ is equal to the number of times ω_3 occurs in the resonance condition (5.3). The latter is clearly equal to 1. Thus, $x_I(2\pi) - x_I(0)$ gives rise to a linear function of ζ. It remains to show that the coefficient in front of ζ is not equal to zero. By Lemma 5.2, one knows that

$$
\begin{aligned}
\dot{x}_I = {}& g_I \cos\left\{(m_1 \omega_1 + m_2 \omega_2)t - (m_1 + m_2 - 1)\frac{\pi}{2}\right\} \\
&+ f_I a \cos\left\{\big(m_1 \omega_1 + (m_2 - 1)\omega_2 - \omega_3\big)t - (m_1 + m_2 - 1 - \varepsilon)\frac{\pi}{2}\right\} + \mathcal{R},
\end{aligned}
\tag{5.10}
$$

where we gathered all other terms into $\mathcal{R}$. Note that the numerical coefficients f_I and g_I depend on the frequencies ω_1, ω_2, and ω_3. The goal is to show that f_I is not equal to zero if we want to move the component x_I, i.e., when $\omega_3 = (m_2 - 1)\omega_2 + m_1 \omega_1$. If we consider f_I as a function of ω_1, ω_2, and ω_3, it suffices to show that this function is not identically equal to zero over the hyperplane of $\mathbb{R}^3$ defined by the resonance condition $\omega_3 = (m_2 - 1)\omega_2 + m_1 \omega_1$. We assume that the next lemma holds true, and we will provide an argument immediately after finishing the proof of Proposition 5.1.

Lemma 5.3. *For all $J \leqslant I$, let $m_1^J := \Delta_1(X_J)$ and $m_2^J := \Delta_2(X_J)$. If f_J is the coefficient in front of the term $\cos\{(m_1^J \omega_1 + (m_2^J - 1)\omega_2 - \omega_3)t - (m_1^J + m_2^J - 1 - \varepsilon)\frac{\pi}{2}\}$, and g_J the one in front of the term $\cos\{(m_1^J \omega_1 + m_2^J \omega_2)t - (m_1^J + m_2^J - 1)\frac{\pi}{2}\}$, then, the quotient $\alpha_J := f_J / g_J$ verifies the following inductive formula.*

- *If $X_J = X_1, \alpha_J = 0$; if $X_J = X_2, \alpha_J = 1$;*
- *if $X_J = [X_{J_1}, X_{J_2}]$, α_J is defined by*

$$
\alpha_J = \frac{m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2}{m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3}\alpha_{J_1} + \alpha_{J_2},
$$

where $m_i^{J_1} = \Delta_i(X_{J_1})$ for $i = 1, 2$.

Let us take $\omega_3 = -\omega_2$. It results from Lemma 5.3 that, for every $J \leqslant I$, one has

$$
\alpha_J = \alpha_{J_1} + \alpha_{J_2}, \quad \text{if } X_J = [X_{J_1}, X_{J_2}].
$$

Since $\alpha_1 = 0$ and $\alpha_2 = 1$, then, over the hyperplane of $\mathbb{R}^3$ defined by $\omega_3 = -\omega_2$, the function α_J is a strictly positive number independent of ω_1 and ω_2.

Let us show now that $\alpha_J(\omega_1, \omega_2, \omega_3)$ is not identically equal to zero over the hyperplane of $\mathbb{R}^3$ defined by $\omega_3 = m_1\omega_1 + (m_2 - 1)\omega_2$. Let $\hat{\omega}_2 := -m_1\omega_1/m_2$. One has $m_1\omega_1 + (m_2 - 1)\hat{\omega}_2 = -\hat{\omega}_2$. It implies that

$$\alpha_I(\omega_1, \hat{\omega}_2, m_1\omega_1 + (m_2 - 1)\hat{\omega}_2) = \alpha_I(\omega_1, \hat{\omega}_2, -\hat{\omega}_2).$$

Since the function $\alpha_I(\omega_1, \omega_2, -\omega_2)$ is never equal to zero, and it coincides with the function $\alpha_I(\omega_1, \omega_2, m_1\omega_1 + (m_2 - 1)\omega_2)$ at the point $(\omega_1, \hat{\omega}_2)$, which is not identically equal to zero. Therefore, $f_I(\omega_1, \omega_2, \omega_3)$ is not identically equal to zero over the hyperplane $\omega_3 = (m_2 - 1)\omega_2 + m_1\omega_1$. Moreover, as it is a nontrivial rational function, it eventually vanishes at a finite number of integer points. Then, we obtain a non-zero linear function of ζ, and (1) is now proved. Proposition 5.1 results from (1) and (2). □

Proof of Lemma 5.3. The proof goes by induction on $|I|$. Since $\dot{x}_1 = u_1$ and $\dot{x}_2 = u_2$, by Eq. (5.5), one has $\alpha_1 = 0$ and $\alpha_2 = 1$.

Assume that $|J| \geqslant 2$. By construction, one has $X_J = [X_{J_1}, X_{J_2}]$ with $|J_1| \leqslant |J_2| < |J|$. According to the inductive hypothesis, one has

$$\dot{x}_{J_1} = g_{J_1} \cos\left\{ (m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2)t - (m_1^{J_1} + m_2^{J_1} - 1)\frac{\pi}{2} \right\}$$
$$+ f_{J_1} \cos\left\{ (m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3)t - (m_1^{J_1} + m_2^{J_1} - 1 - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_{J_1},$$

$$\dot{x}_{J_2} = g_{J_2} \cos\left\{ (m_1^{J_2}\omega_1 + m_2^{J_2}\omega_2)t - (m_1^{J_2} + m_2^{J_2} - 1)\frac{\pi}{2} \right\}$$
$$+ f_{J_2} \cos\left\{ (m_1^{J_2}\omega_1 + (m_2^{J_2} - 1)\omega_2 - \omega_3)t - (m_1^{J_2} + m_2^{J_2} - 1 - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_{J_2}.$$

This implies that

$$\dot{x}_J = \left(\frac{1}{m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2} g_{J_1} \cos\left\{ (m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2)t - (m_1^{J_1} + m_2^{J_1})\frac{\pi}{2} \right\} + \frac{1}{m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3} \right.$$
$$\times f_{J_1} a \cos\left\{ (m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3)t - (m_1^{J_1} + m_2^{J_2} - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_{J_1} \right)$$
$$\times \left(g_{J_2} \cos\left\{ (m_1^{J_2}\omega_1 + m_2^{J_2}\omega_2)t - (m_1^{J_2} + m_2^{J_2} - 1)\frac{\pi}{2} \right\} \right.$$
$$\left. + f_{J_2} a \cos\left\{ (m_1^{J_2}\omega_1 + (m_2^{J_2} - 1)\omega_2 - \omega_3)t - (m_1^{J_2} + m_2^{J_2} - 1 - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_{J_2} \right)$$
$$= \frac{1}{2} \frac{g_{J_1} g_{J_2}}{m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2} \cos\left\{ (m_1^J\omega_1 + m_2^J\omega_2)t - (m_1^J + m_2^J - 1)\frac{\pi}{2} \right\}$$
$$+ \frac{1}{2} \left(\frac{g_{J_1} f_{J_2}}{m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2} + \frac{g_{J_2} f_{J_1}}{m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3} \right)$$
$$\times \cos\left\{ (m_1^J\omega_1 + m_2^J\omega_2 - \omega_3)t - (m_1^J + m_2^J - 1 - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_J$$

$$= g_J \cos\left\{ (m_1^J \omega_1 + m_2^J \omega_2)t - (m_1^J + m_2^J - 1)\frac{\pi}{2} \right\}$$

$$+ f_J \cos\left\{ (m_1^J \omega_1 + m_2^J \omega_2 - \omega_3)t - (m_1^J + m_2^J - 1 - \varepsilon)\frac{\pi}{2} \right\} + \mathcal{R}_J.$$

Therefore, one obtains $\alpha_J = \dfrac{m_1^{J_1}\omega_1 + m_2^{J_1}\omega_2}{m_1^{J_1}\omega_1 + (m_2^{J_1} - 1)\omega_2 - \omega_3} \alpha_{J_1} + \alpha_{J_2}$. $\square$

5.2.2. A more general case: $m = 2$ and $\mathrm{Card}(\mathcal{E}_x^j) > 1$

In general, given a pair (m_1, m_2), the equivalence class $\mathcal{E}_x(m_1, m_2)$ contains more than one element. This situation first occurs for Lie brackets of length 5. For instance, given the pair $(3, 2)$, one has both $X_I = [X_2, [X_1, [X_1, [X_1, X_2]]]]$ and $X_J = [[X_1, X_2], [X_1, [X_1, X_2]]]$. By Lemma 5.2, if one chooses a 4-tuple verifying the resonance condition (5.3) for $\dot{x}_I$, the same resonance occurs in $\dot{x}_J$. Such two components cannot be independently steered by using resonance. The idea is to move simultaneously these components. For instance, one can choose (u_1, u_2) as follows:

$$u_1(t) = \cos\omega_1 t,$$

$$u_2(t) = \cos\omega_2 t + a_I \cos\omega_3 t + \cos\omega_4 t + a_J \cos\omega_5 t,$$

where $\omega_1 = 1$, ω_2 is chosen according to Eq. (5.4), $\omega_3 = (m_2 - 1)\omega_2 + m_1\omega_1$, and $\omega_5 = (m_2 - 1)\omega_4 + m_1\omega_1$, with ω_4 large enough to guarantee Condition (C2). After explicit integration of Eq. (5.1), one obtains

$$\begin{pmatrix} f_I(\omega_1, \omega_2) & f_I(\omega_1, \omega_4) \\ f_J(\omega_1, \omega_2) & f_J(\omega_1, \omega_4) \end{pmatrix} \begin{pmatrix} a_I \\ a_J \end{pmatrix} = A \begin{pmatrix} a_I \\ a_J \end{pmatrix} = \begin{pmatrix} x_I(2\pi) - x_I(0) \\ x_J(2\pi) - x_J(0) \end{pmatrix},$$

where f_I and f_J are two rational functions of frequencies. Thus, the pair (u_1, u_2) controls exactly and simultaneously x_I and x_J, provided that the matrix A is invertible. We generalize this strategy in the following paragraphs. Assume that $\mathcal{E}_x^j(m_1, m_2) = \{x_{I_1}, \ldots, x_{I_N}\}$. The main result is given next.

Proposition 5.4. *Consider*

$$\{\omega_{11}^1, \ldots, \omega_{11}^{m_1}\}, \ldots, \{\omega_{1N}^1, \ldots, \omega_{1N}^{m_1}\},$$

$$\{\omega_{21}^1, \ldots, \omega_{21}^{m_2-1}, \omega_{21}^*\}, \ldots, \{\omega_{2N}^1, \ldots, \omega_{2N}^{m_2-1}, \omega_{2N}^*\}$$

belonging to $\mathbb{N}^{m_1 N} \times \mathbb{N}^{m_2 N}$ such that

$$\begin{cases} \forall j = 1, \ldots, N, \quad \omega_{2j}^* = \displaystyle\sum_{i=1}^{m_1} \omega_{1j}^i + \sum_{i=1}^{m_2-1} \omega_{2j}^i, \\ \varepsilon = m_1 + m_2 - 1 \pmod 2, \end{cases} \tag{5.11}$$

and

$$\forall j = 1, \ldots, N-1, \quad \begin{cases} \omega_{11}^1 \in \mathbb{N}; \\ \omega_{1j}^{i+1} > m_1 \omega_{1j}^i; & i = 1, \ldots, m_1, \\ \omega_{2j}^1 > m_1 \omega_{1j}^{m_1}; \\ \omega_{2j}^i > m_2 \omega_{2j}^{i-1} + m_1 \omega_{1j}^{m_1}; & i = 2, \ldots, m_2 - 1, \\ \omega_{1j+1}^1 > m_2 \omega_{2j}^{m_2-1} + m_1 \omega_{1j}^{m_1}. \end{cases} \tag{5.12}$$

Then, the input $u := (u_1, u_2)$ *defined by*

$$
\begin{cases}
u_1 := \displaystyle\sum_{j=1}^{N} \sum_{i=1}^{m_1} \cos \omega_{1j}^i t, \\[4mm]
u_2 := \displaystyle\sum_{j=1}^{N} \sum_{i=1}^{m_2-1} \cos \omega_{2j}^i t + a_j \cos\left(\omega_{2j}^* t - \varepsilon \frac{\pi}{2}\right),
\end{cases}
\tag{5.13}
$$

steers the components $(x_{I_1}, \ldots, x_{I_N})$ *from an arbitrary initial condition* $(x_{I_1}(0), \ldots, x_{I_N}(0))$ *to an arbitrary final one* $(x_{I_1}(2\pi), \ldots, x_{I_N}(2\pi))$, *without modifying any other component having been previously moved to its final value.*

This result generalizes Proposition 5.1. The proof is decomposed in two parts as follows:

Part I: we show that, if (5.12) holds and the control functions are of the form (5.13), then (5.11) is the *only* resonance occurring in $(\dot{x}_{I_1}, \ldots, \dot{x}_{I_N})$.

Part II: as the resonance gives rise to a system of linear equations on $(a_1, \ldots, a_N)$, we recover the invertibility of this system by choosing suitable frequencies in the control function (5.13).

Part I: Frequencies and resonance

Consider inputs of the form (5.13). Generalizing Lemma 5.2, we give a general form of frequencies involved in $\dot{x}_J$.

Lemma 5.5. *The dynamics* $\dot{x}_J$ *is a linear combination of cosine functions of the form*

$$
\big(\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^*\big) t - \big(\ell_1 + \ell_2 + m_2^* - 1 + \ell_2^* \varepsilon\big) \frac{\pi}{2},
\tag{5.14}
$$

where

$$
\ell_1 \cdot \omega_1 = \sum_{j=1}^{N} \sum_{i=1}^{m_1} \ell_{1j}^i \omega_{1j}^i, \qquad \ell_2 \cdot \omega_2 = \sum_{j=1}^{N} \sum_{i=1}^{m_2-1} \ell_{2j}^i \omega_{2j}^i, \qquad \ell_2^* \cdot \omega_2^* = \sum_{j=1}^{N} \ell_{2j}^* \omega_{2j}^*,
$$

$$
\ell_1 = \sum_{j=1}^{N} \sum_{i=1}^{m_1} \ell_{1j}^i, \qquad \ell_2 = \sum_{j=1}^{N} \sum_{i=1}^{m_2-1} \ell_{2j}^i, \qquad \ell_2^* = \sum_{j=1}^{N} m_{2j}^*,
$$

with $(\ell_{1j}^i, \ell_{2j}^i, \ell_{2j}^*) \in \mathbb{Z}^3$.

Let

$$
|\ell_1| = \sum_{j=1}^{N} \sum_{i=1}^{m_1} |\ell_{1j}^i|, \qquad |\ell_2| = \sum_{j=1}^{N} \sum_{i=1}^{m_2-1} |\ell_{2j}^i|, \quad \text{and} \quad |\ell_2^*| = \sum_{j=1}^{N} |\ell_{2j}^*|,
$$

then, one has $|\ell_1| \leqslant \Delta_1(X_J)$, $|\ell_2| + |\ell_2^*| \leqslant \Delta_2(X_J)$.

Proof. The proof goes by induction on $|J|$.

- $|J| = 1$: the result is true since $\dot{x}_1 = u_1$ and $\dot{x}_2 = u_2$.

- *Inductive step*: Assume that the result holds true for all $x_{\tilde{J}}$ such that $1 \leqslant |\tilde{J}| < s$. We show that it remains true for x_J with $|J| = s$. By construction, we have $X_J = \mathrm{ad}_{X_{J_1}}^k X_{J_2}$, and

$$\dot{x}_J = \frac{1}{k!} x_{J_1}^k \dot{x}_{J_2}, \tag{5.15}$$

with $|J_1| < |J|$, $|J_2| < |J|$, and $k|J_1| + |J_2| = |J|$.
Then, by the inductive hypothesis, we have

$$\dot{x}_{J_1} = \mathrm{LinCom}\left\{ \cos\left\{ (\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^*)t - (\ell_1 + \ell_2 + \ell_2^* - 1 + \ell_2^* \varepsilon)\frac{\pi}{2} \right\} \right\}, \tag{5.16}$$

$$\dot{x}_{J_2} = \mathrm{LinCom}\left\{ \cos\left\{ (\tilde{\ell}_1 \cdot \omega_1 + \tilde{\ell}_2 \cdot \omega_2 + \tilde{\ell}_2^* \cdot \omega_2^*)t - (\tilde{\ell}_1 + \tilde{\ell}_2 + \tilde{\ell}_2^* - 1 + \tilde{\ell}_2^* \varepsilon)\frac{\pi}{2} \right\} \right\}, \tag{5.17}$$

where $\mathrm{LinCom}\{\cdot\}$ stands for "linear combination".
Eq. (5.16) implies that

$$x_{J_1} = \mathrm{LinCom}\left\{ \cos\left\{ (\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^*)t - (\ell_1 + \ell_2 + \ell_2^* - 1 + \ell_2^* \varepsilon)\frac{\pi}{2} - \frac{\pi}{2} \right\} \right\}$$

$$= \mathrm{LinCom}\left\{ \cos\left\{ (\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^*)t - (\ell_1 + \ell_2 + \ell_2^* + \ell_2^* \varepsilon)\frac{\pi}{2} \right\} \right\}. \tag{5.18}$$

For notational ease, we will only write down the case $\dot{x}_J = x_{J_1} \dot{x}_{J_2}$. Using product formulas for cosine function, one has

$$\dot{x}_J = \mathrm{LinCom}\left\{ \cos\left\{ \left[(\ell_1 \pm \tilde{\ell}_1) \cdot \omega_1 + (\ell_2 \pm \tilde{\ell}_2) \cdot \omega_2 + (\ell_2^* \pm \tilde{\ell}_2^*) \cdot \omega_2^* \right]t \right.\right.$$

$$\left.\left. - \left[(\ell_1 \pm \tilde{\ell}_1) + (\ell_2 \pm \tilde{\ell}_2) + (\ell_2^* \pm \tilde{\ell}_2^*) - 1 + (\ell_2^* \pm \tilde{\ell}_2^*)\varepsilon \right]\frac{\pi}{2} \right\} \right\}. \tag{5.19}$$

Moreover, according to the inductive hypothesis, one has

$$|\ell_1| \leqslant \Delta_1(X_{J_1}), \qquad |\ell_2| + |\ell_2^*| \leqslant \Delta_2(X_{J_1}),$$

and

$$|\tilde{\ell}_1| \leqslant \Delta_1(X_{J_2}), \qquad |\tilde{\ell}_2| + |\tilde{\ell}_2^*| \leqslant \Delta_2(X_{J_2}).$$

Then, one gets

$$|\tilde{\ell}_1 \pm \tilde{\ell}_1| \leqslant \Delta_1(X_J) \quad \text{and} \quad |\tilde{\ell}_2 \pm \tilde{\ell}_2| + |\ell_2^* \pm \tilde{\ell}_2^*| \leqslant \Delta_2(X_J).$$

This concludes the proof of Lemma 5.5. □

By Lemma 5.5, one gets a nontrivial contribution for x_J if the resonance condition

$$\begin{cases} \ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^* = 0, \\ \ell_2^* \varepsilon + \ell_1 + \ell_2 + \ell_2^* - 1 \equiv 0 \pmod 2 \end{cases} \tag{5.20}$$

is verified by the frequencies of some cosine functions involved in $\dot{x}_J$.

Lemma 5.6. *Under conditions* (5.11) *and* (5.12) *in Proposition* 5.4, *one gets a nontrivial contribution on* x_{I_j} *depending linearly on* a_j *for all* $j = 1, \ldots, N$.

Proof. It is clear that the resonance condition (5.20) holds for

$$\{\omega_{11}^1, \ldots, \omega_{11}^{m_1}\}, \ldots, \{\omega_{1N}^1, \ldots, \omega_{1N}^{m_1}\},$$
$$\{\omega_{21}^1, \ldots, \omega_{21}^{m_2-1}, \omega_{21}^*\}, \ldots, \{\omega_{2N}^1, \ldots, \omega_{2N}^{m_2-1}, \omega_{2N}^*\},$$

and $\varepsilon \in \{0, 1\}$ verifying (5.11). We show that it is the only resonance occurring in $\dot{x}_{I_j}$. Indeed, by Lemma 5.5, the integer part of frequencies in $\dot{x}_{I_j}$ is in the following form

$$\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^*$$
$$= \sum_{j=1}^N \sum_{i=1}^{m_1} \ell_{1j}^i \omega_{1j}^i + \sum_{j=1}^N \sum_{i=1}^{m_2-1} \ell_{2j}^i \omega_{2j}^i + \sum_{j=1}^N \ell_{2j}^* \omega_{2j}^*$$
$$= \sum_{j=1}^N \sum_{i=1}^{m_1} \ell_{1j}^i \omega_{1j}^i + \sum_{j=1}^N \sum_{i=1}^{m_2-1} \ell_{2j}^i \omega_{2j}^i + \sum_{j=1}^N \ell_{2j}^* \left(\sum_{i=1}^{m_1} \omega_{1j}^i + \sum_{i=1}^{m_2-1} \omega_{2j}^i \right)$$
$$= \sum_{j=1}^N \sum_{i=1}^{m_1} (\ell_{1j}^i + \ell_{2j}^*) \omega_{1j}^i + \sum_{j=1}^N \sum_{i=1}^{m_2-1} (\ell_{2j}^i + \ell_{2j}^*) \omega_{2j}^i. \tag{5.21}$$

By condition (5.12), Eq. (5.21) is equal to zero if and only if

$$\ell_{1j}^i + \ell_{2j}^* = 0, \quad \text{for } i = 1, \ldots, m_1,$$
$$\ell_{2j}^i + \ell_{2j}^* = 0, \quad \text{for } i = 1, \ldots, m_2 - 1.$$

Then, one has

$$|\ell_1| = \sum_{j=1}^N \sum_{i=1}^{m_1} |\ell_{1j}^i| = \sum_{j=1}^N \sum_{i=1}^{m_1} |\ell_{2j}^*| = m_1 \sum_{j=1}^N |\ell_{2j}^*|,$$
$$|\ell_2| = \sum_{j=1}^N \sum_{i=1}^{m_2-1} |\ell_{2j}^i| = \sum_{j=1}^N \sum_{i=1}^{m_2-1} |\ell_{2j}^*| = (m_2 - 1) \sum_{j=1}^N |\ell_{2j}^*|.$$

However, by Lemma 5.5, one knows that $|\ell_1| \leqslant m_1$ and $|\ell_2| + |\ell_2^*| \leqslant m_2$. Then, one necessarily has $m_{2j}^* = 0$ for all $j = 1, \ldots, N$. In that case, one obtains

$$\ell_2^* \varepsilon + \ell_1 + \ell_2 + \ell_2^* - 1 = -1 \neq 0 \pmod{2}.$$

In conclusion, the resonance condition (5.20) does not hold for any 4-tuple $(\ell_1, \ell_2, \ell_2^*, \varepsilon)$ different from $(m_1, m_2 - 1, -1, m_1 + m_2 - 1 \pmod 2)$.

By Eq. (5.13), the power of a_j is equal to the number of times ω_{2j}^* occurs in the resonance condition (5.3). Since the latter is equal to 1, we obtain a linear function of a_j. This ends the proof of Lemma 5.6. □

Lemma 5.7. *If* $x_J \in \mathcal{E}_x^i$ *and* $i < j$, *then* $x_J(2\pi) - x_J(2\pi) = 0$.

Proof. We first note that Eq. (5.21) still holds true. Recall its expression here:

$$\ell_1 \cdot \omega_1 + \ell_2 \cdot \omega_2 + \ell_2^* \cdot \omega_2^* = \sum_{j=1}^{N} \sum_{i=1}^{m_1} (\ell_{1j}^i + \ell_{2j}^*) \cdot \omega_{1j}^i + \sum_{j=1}^{N} \sum_{i=1}^{m_2-1} (\ell_{2j}^i + \ell_{2j}^*) \cdot \omega_{2j}^i. \quad (5.22)$$

By condition (5.12) in Proposition 5.4, Eq. (5.22) is equal to zero if and only if $\ell_{1j}^i + \ell_{2j}^* = 0$ for $i = 1, \ldots, m_1$, $j = 1, \ldots, N$, and $\ell_{2j}^i + \ell_{2j}^* = 0$ for $i = 1, \ldots, m_2 - 1$, $j = 1, \ldots, N$. In that case, one has

$$|\ell_1| = m_1 \sum_{j=1}^{N} |\ell_{2j}^*|, \qquad |\ell_2| + |\ell_2^*| = m_2 \sum_{j=1}^{N} |\ell_{2j}^*|.$$

One also knows that $|\ell_1| \leqslant \Delta_1(X_J)$, $|\ell_2| + |\ell_2^*| \leqslant \Delta_2(X_J)$, with $\Delta_1(X_J) < m_1$ or $\Delta_2(X_J) < m_2$. Therefore, one has $\ell_{2j}^* = 0$ for all $j = 1, \ldots, N$. This implies that

$$\ell_2^* \varepsilon + \ell_1 + \ell_2 + \ell_2^* - 1 = -1 \neq 0 \quad (\mathrm{mod}\ 2).$$

In conclusion, the resonance condition (5.20) does not hold true. This ends the proof of Lemma 5.7. □

Part II: Invertibility

As a consequence of Lemma 5.6, one has

$$\begin{pmatrix} x_{I_1}(2\pi) - x_{I_1}(0) \\ \vdots \\ x_{I_N}(2\pi) - x_{I_N}(0) \end{pmatrix} = A(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-1}, \omega_{2N}^*) \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}$$

$$= \begin{pmatrix} f_{I_1}^X(\omega_{11}^1, \ldots, \omega_{21}^*), & \cdots & , f_{I_1}^X(\omega_{1N}^1, \ldots, \omega_{2N}^*) \\ \vdots & \ddots & \vdots \\ f_{I_N}^X(\omega_{11}^1, \ldots, \omega_{21}^*), & \cdots & , f_{I_N}^X(\omega_{1N}^1, \ldots, \omega_{2N}^*) \end{pmatrix} \begin{pmatrix} a_1 \\ \vdots \\ a_N \end{pmatrix}, \quad (5.23)$$

where $f_{I_j}^X : \mathbb{R}^m \to \mathbb{R}$ are rational functions of frequencies, and every ω_{2j}^* verifies Eq. (5.11) for $j = 1, \ldots, N$.

Definition 5.4 (*Control matrix and control vector*). The matrix A and the vector $(a_1, \ldots, a_N)$ occurring in Eq. (5.23) are called respectively *control matrix* and *control vector* associated with the equivalence class $\mathcal{E}_x^j$.

We show in the sequel that it is possible to choose integer frequencies

$$\{\omega_{11}^1, \ldots, \omega_{11}^{m_1}\}, \ldots, \{\omega_{1N}^1, \ldots, \omega_{1N}^{m_1}\},$$
$$\{\omega_{21}^1, \ldots, \omega_{21}^{m_2-1}, \omega_{21}^*\}, \ldots, \{\omega_{2N}^1, \ldots, \omega_{2N}^{m_2-1}, \omega_{2N}^*\},$$

so that the invertibility of the control matrix A is guaranteed, as well as the non-resonance of every component x_J belonging to a class smaller than $\mathcal{E}_x^j$.

For $j = 1, \ldots, N$, we use P_j to denote the hyperplane in $\mathbb{R}^M$ with $M := m_1 + m_2$ defined by Eq. (5.11), for which we recall the expression next,

$$\omega_{2j}^* = \sum_{i=1}^{m_1} \omega_{1j}^i + \sum_{i=1}^{m_2-1} \omega_{2j}^i.$$

We start by showing that the function $\det A(\omega_{11}^1, \ldots, \omega_{2N}^*)$ is not identically equal to zero on $\bigcap_{j=1}^N P_j$. This is a consequence of the following lemma.

Lemma 5.8. *The family of functions*

$$\left\{ f_{I_1}^X(\omega_1^1, \ldots, \omega_2^{m_2-1}, \omega_2^*), \ldots, f_{I_N}^X(\omega_1^1, \ldots, \omega_2^{m_2-1}, \omega_2^*) \right\}$$

is linearly independent on the hyperplane P in $\mathbb{R}^M$ defined by the equation

$$\omega_2^* = \sum_{i=1}^{m_1} \omega_1^i + \sum_{i=1}^{m_2-1} \omega_2^i.$$

Proof. The first part of the argument consists in considering a family of M indeterminates $Y = \{Y_1, \ldots, Y_M\}$ and the associated control system

$$\dot{y} = \sum_{i=1}^{M} v_i Y_i. \tag{5.24}$$

Let H_Y be a *P. Hall family* over Y. Consider the elements $\{Y_{J_1}, \ldots, Y_{J_{\tilde{N}}}\}$ in H_Y of length M such that $\Delta_i(Y_{J_j}) = 1$, for $i = 1, \ldots, M$, and $j = 1, \ldots, \tilde{N}$, and the corresponding components $\{y_{J_1}, \ldots, y_{J_{\tilde{N}}}\}$ in exponential coordinates.

If we apply one control of the form $\{v_i = \cos v_i t\}_{i=1,\ldots,M}$, with $v_m = \sum_{i=1}^{m-1} v_i$, to system (5.24), then, by explicit integration, there exists, for each component y_{J_j}, a fractional function $f_{J_j}^Y : \mathbb{R}^m \to \mathbb{R}$ such that

$$y_{J_j}(2\pi) - y_{J_j}(0) = f_{J_j}^Y(v_1, \ldots, v_M), \quad \text{for } v_M = \sum_{i=1}^{M-1} v_i. \tag{5.25}$$

Claim 11. *The family of functions $\{f_{J_1}^Y, \ldots, f_{I_{\tilde{N}}}^Y\}$ is linearly independent on the hyperplane in $\mathbb{R}^M$ defined by $v_M = \sum_{i=1}^{M-1} v_i$.*

Proof. We first define $\tilde{f}_{J_j}^Y$ by

$$\tilde{f}_{J_j}^Y(v_1, \ldots, v_M) = f_{J_j}^Y(v_1, \ldots, -v_M). \tag{5.26}$$

Then, it is easy to see that $\tilde{f}_{J_j}^Y$ verifies the following inductive formula:

1. for $J = i = 1, \ldots, M$, $\tilde{f}_J^Y(v_i) = \frac{1}{v_i}$;
2. for $|J| > 1$, $Y_J = [Y_{J_1}, Y_{J_2}]$, there exists an injective function

$$\sigma_J : \{1, \ldots, m^J\} \to \{1, \ldots, M\}$$

such that

$$\tilde{f}_J^Y(v_{\sigma_J(1)}, \ldots, v_{\sigma_J(m^J)}) = \frac{\tilde{f}_{J_1}^Y(v_{\sigma_J(1)}, \ldots, v_{\sigma_J(m^{J_1})})}{\sum_{i=1}^{m^{J_1}} v_{\sigma_J(i)}} \tilde{f}_{J_2}^Y(v_{\sigma_J(m^{J_1}+1)}, \ldots, v_{\sigma_J(m^J)}), \qquad (5.27)$$

where $m^J := \Delta(Y_J)$, $m^{J_1} := \Delta(Y_{J_1})$, and $m^{J_2} := \Delta(Y_{J_2})$.

We note that the family of rational functions $\tilde{f}_J^Y$ is well defined for all the Lie brackets Y_J such that $\Delta_i(Y_J) \leqslant 1$, $i = 1, \ldots, M$. The algebraic construction could be extended to all the Lie brackets, but it is not necessary for our purpose. We also note that Claim 11 is equivalent to the fact that the family of rational functions

$$\{\tilde{f}_{J_j}^Y(v_1, \ldots, v_M)\}_{j=1,\ldots,\tilde{N}}$$

is linearly independent over the hyperplane $\sum_{i=1}^{M} v_i = 0$.

Recall that every element Y_{J_j} in the family $\{Y_{J_1}, \ldots, Y_{J_{\tilde{N}}}\}$ writes uniquely as

$$Y_{J_j} = [Y_{J_{j_1}}, Y_{J_{j_2}}]. \qquad (5.28)$$

Definition 5.5 *(Left and right factors).* For $J \in \{J_1, \ldots, J_{\tilde{N}}\}$, the *left factor* $L(J)$ and the *right factor* $R(J)$ of J are defined in such a way that $Y_J = [Y_{L(J)}, Y_{R(J)}]$.

Let L^* be defined by

$$L^* := \max_{j=1,\ldots,\tilde{N}} \{L(J_j)\}. \qquad (5.29)$$

The integer L^* is well defined since a P. Hall family is totally ordered. Thus, there exists $J^* \in \{J_1, \ldots, J_{\tilde{N}}\}$ such that $L^* = L(J^*)$. Then, define $R^* := R(J^*)$ and set $m^* = |L^*|$. Let

$$\Lambda = \Lambda_L \cup \Lambda_R \quad \text{and} \quad \overline{\Lambda} = \{1, \ldots, \tilde{N}\} \setminus \Lambda,$$

with Λ_L and Λ_R defined by

$$\Lambda_L := \{j \in \{1, \ldots, \tilde{N}\}, \text{ such that } Y_{L(J_j)} \sim Y_L\}, \qquad (5.30)$$

$$\Lambda_R := \{j \in \{1, \ldots, \tilde{N}\}, \text{ such that } Y_{L(J_j)} \sim Y_R\}. \qquad (5.31)$$

Then, for all $j \in \Lambda$, there exists an injection function

$$\sigma_j : \{1, \ldots, M\} \to \{1, \ldots, M\}$$

such that one has

$$\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \frac{\tilde{f}^Y_{L(J_j)}(\nu_{\sigma_j(1)}, \ldots, \nu_{\sigma_j(m^*)})}{\sum_{i=1}^{m^*} \nu_{\sigma_j(i)}} \tilde{f}^Y_{R(J_j)}(\nu_{\sigma_j(m^*+1)}, \ldots, \nu_{\sigma_j(M)}), \quad \text{if } j \in \Lambda_L, \quad (5.32)$$

$$\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \frac{\tilde{f}^Y_{L(J_j)}(\nu_{\sigma_j(m^*+1)}, \ldots, \nu_{\sigma_j(M)})}{\sum_{i=m^*+1}^{M} \nu_{\sigma_j(i)}} \tilde{f}^Y_{R(J_j)}(\nu_{\sigma_j(1)}, \ldots, \nu_{\sigma(m^*)}), \quad \text{if } j \in \Lambda_R. \quad (5.33)$$

Note that, for all j_1 and j_2 in Λ_L, one has $\{\nu_{\sigma_{j_1}(1)}, \ldots, \nu_{\sigma_{j_1}(m^*)}\} = \{\nu_{\sigma_{j_2}(1)}, \ldots, \nu_{\sigma_{j_2}(m^*)}\}$. Denote by Ξ_L the set of variables involved in $\tilde{f}^Y_{L(J_j)}$ with $j \in \Lambda_L$. A similar property holds for Λ_R. For all j_1 and j_2 in Λ_R, one has $\{\nu_{\sigma_{j_1}(m^*+1)}, \ldots, \nu_{\sigma_{j_1}(M)}\} = \{\nu_{\sigma_{j_2}(m^*+1)}, \ldots, \nu_{\sigma_{j_2}(M)}\}$. Denote by Ξ_R the set of all variables occurring in $\tilde{f}^Y_{L(J_j)}$ with $j \in \Lambda_R$. Then one has $\Xi_L \cup \Xi_R = \{\nu_1, \ldots, \nu_M\}$. By abuse of notation, we re-write Eqs. (5.32) and (5.33) in the following form:

$$\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \frac{\tilde{f}^Y_{L(J_j)}(\Xi_L)}{\sum_{\tilde{\nu}_k \in \Xi_L} \tilde{\nu}_k} \tilde{f}^Y_{R(J_j)}(\Xi_R), \quad \text{if } j \in \Lambda_L; \quad (5.34)$$

$$\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \frac{\tilde{f}^Y_{L(J_j)}(\Xi_R)}{\sum_{\tilde{\nu}_k \in \Xi_R} \tilde{\nu}_k} \tilde{f}^Y_{R(J_j)}(\Xi_L), \quad \text{if } j \in \Lambda_R. \quad (5.35)$$

Moreover, by the resonance condition $\sum_{i=1}^{M} \nu_i = 0$, Eq. (5.35) becomes

$$\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \frac{\tilde{f}^Y_{L(J_j)}(\Xi_R)}{-\sum_{\tilde{\nu}_k \in \Xi_L} \tilde{\nu}_k} \tilde{f}^Y_{R(J_j)}(\Xi_L), \quad \text{if } j \in \Lambda_R. \quad (5.36)$$

We now prove that the family of rational functions $\{\tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M)\}_{j=1,\ldots,\tilde{N}}$ is linearly independent over the hyperplane $\sum_{i=1}^{M} \nu_i = 0$. The proof goes by induction over the length of the Lie brackets under consideration. For the brackets of length 1, the result is obviously true. Assume that the result holds for all brackets of length smaller than $M-1$, $M \geqslant 2$.

Assume that there exist $\ell_j \in \mathbb{R}^{\tilde{N}}$ such that

$$\sum_{j=1}^{\tilde{N}} \ell_j \tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = 0, \quad \text{with } \sum_{i=1}^{M} \nu_i = 0. \quad (5.37)$$

One has

$$\sum_{j=1}^{\tilde{N}} \ell_j \tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = \sum_{j \in \Lambda} \ell_j \tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) + \sum_{j \in \overline{\Lambda}} \ell_j \tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M)$$

$$= \sum_{j \in \Lambda_L} \ell_j \frac{\tilde{f}^Y_{L(J_j)}(\Xi_L)}{\sum_{\tilde{\nu}_k \in \Xi_L} \tilde{\nu}_k} \tilde{f}^Y_{R(J_j)}(\Xi_R) - \sum_{j \in \Lambda_R} \ell_j \frac{\tilde{f}^Y_{L(J_j)}(\Xi_R)}{\sum_{\tilde{\nu}_k \in \Xi_L} \tilde{\nu}_k} \tilde{f}^Y_{R(J_j)}(\Xi_L)$$

$$+ \sum_{j \in \overline{\Lambda}} \ell_j \tilde{f}^Y_{J_j}(\nu_1, \ldots, \nu_M) = 0. \quad (5.38)$$

Multiplying Eq. (5.38) by the factor $\sum_{\tilde{\nu}_k \in \Xi_L} \tilde{\nu}_k$, one gets

$$\sum_{j\in\Lambda_L}\ell_j\tilde{f}^Y_{L(J_j)}(\varXi_L)\tilde{f}^Y_{R(J_j)}(\varXi_R) - \sum_{j\in\Lambda_R}\ell_j\tilde{f}^Y_{L(J_j)}(\varXi_R)\tilde{f}^Y_{R(J_j)}(\varXi_L) + \left(\sum_{\tilde{v}_k\in\varXi_L}\tilde{v}_k\right)\sum_{j\in\overline{\Lambda}}\ell_j\tilde{f}^Y_{J_j}(v_1,\dots,v_M)$$

$$= 0. \tag{5.39}$$

Since L^* is the maximal element among the left factors of Lie brackets of length M, the fraction $\tilde{f}^Y_{J_j}$ does not contain the factor $\sum_{\tilde{v}_k\in\varXi_L}\tilde{v}_k$ for all $j\in\overline{\Lambda}$. Therefore, on the hyperplane of $\mathbb{R}^{m^*}$ defined by $\sum_{\tilde{v}_k\in\varXi_L}\tilde{v}_k = 0$, one has

$$\sum_{j\in\Lambda_L}\ell_j\tilde{f}^Y_{L(J_j)}(\varXi_L)\tilde{f}^Y_{R(J_j)}(\varXi_R) - \sum_{j\in\Lambda_R}\ell_j\tilde{f}^Y_{L(J_j)}(\varXi_R)\tilde{f}^Y_{R(J_j)}(\varXi_L) = 0. \tag{5.40}$$

Fixing variables belonging to $\varXi_R$, Eq. (5.40) is a linear combination of elements of the family

$$\left\{\tilde{f}^Y_{L(J_j)}(\varXi_L)\right\}_{j\in\Lambda_L} \cup \left\{\tilde{f}^Y_{R(J_j)}(\varXi_L)\right\}_{j\in\Lambda_R}$$

associated with elements of length m^* in the P. Hall family. By the inductive hypothesis, this family is linearly independent over the hyperplane of $\mathbb{R}^{m^*}$ defined by $\sum_{\tilde{v}_k\in\varXi_L}\tilde{v}_k = 0$. We therefore obtain that

$$\ell_j\tilde{f}^Y_{R(J_j)}(\varXi_R) = 0, \quad \text{for all } j\in\Lambda_L, \tag{5.41}$$

$$\ell_j\tilde{f}^Y_{L(J_j)}(\varXi_R) = 0, \quad \text{for all } j\in\Lambda_R. \tag{5.42}$$

Since Eqs. (5.41) and (5.42) hold true over the whole hyperplane of $\mathbb{R}^{M-m^*}$ defined by $\sum_{\tilde{v}_k\in\varXi_R}\tilde{v}_k = 0$, one has $\ell_j = 0$ for every $j\in\Lambda$. Therefore, Eq. (5.38) becomes

$$\sum_{j\in\overline{\Lambda}}\ell_j\tilde{f}^Y_{J_j}(v_1,\dots,v_M) = 0. \tag{5.43}$$

Consider now the maximum left factor for $j\in\overline{\Lambda}$ and iterate the same reasoning used for Eq. (5.37). We deduce that $\ell_j = 0$ for every $j\in\overline{\Lambda}$. Therefore, the family $\{\tilde{f}^Y_{J_j}(v_1,\dots,v_M)\}_{j=1,\dots,\tilde{N}}$ is linearly independent over the hyperplane $\sum_{i=1}^M v_i = 0$ and this concludes the proof of Claim 11. $\square$

We are now in a position to proceed with the argument of Lemma 5.8. Let X_I be an element of $\mathcal{E}_X(m_1,m_2)$, $M := m_1 + m_2$ and $N := \operatorname{Card}\mathcal{E}_X(m_1,m_2)$. Consider also another family of M indeterminates $Y = \{Y_1,\dots,Y_M\}$ and let H_Y be the P. Hall family over Y. Finally, consider all the elements of the class $\mathcal{E}_Y(1,\dots,1) = \{Y_{J_1},\dots,Y_{J_{\tilde{N}}}\}$ in H_Y.

Let Π be the algebra homomorphism from $L(Y)$ to $L(X)$ defined by

$$\Pi(Y_i) = X_1, \quad \text{for } i = 1,\dots,m_1, \tag{5.44}$$

$$\Pi(Y_i) = X_2, \quad \text{for } i = m_1+1,\dots,M. \tag{5.45}$$

Note that the map Π is surjective from $\mathcal{E}_Y$ onto $\mathcal{E}_X$. Consider the following vector fields

$$V_Y = \{v_1Y_1 + \dots + v_MY_M\},$$

where

$$v_i = \cos \omega_i t, \quad \text{for } i = 1, \ldots, M-1, \quad \text{and} \quad v_M = \cos\left(\omega_M t + \varepsilon \frac{\pi}{2}\right), \tag{5.46}$$

with $\omega_M = \sum_{i=1}^{M-1} \omega_i$, and ω_i verifying the non-resonance conditions.

Then, the non-autonomous flow of V_Y between 0 and 2π is given by

$$\overrightarrow{\exp}(V_Y)(0, 2\pi) = e^{f_{J_1}^Y Y_{J_1}} \circ \cdots \circ e^{f_{J_{\tilde{N}}}^Y Y_{J_{\tilde{N}}}} \circ \prod_{J > J_1} e^{f_J^Y Y_J}. \tag{5.47}$$

Let us now apply Π to V_Y; we get

$$\Pi(V_Y) := V^X = \left\{ v_1 \Pi(Y_1) + \cdots + v_m \Pi(Y_m) \right\} = \{u_1 X_1 + u_2 X_2\}, \tag{5.48}$$

where

$$u_1 = \sum_{i=1}^{m_1} v_i = \sum_{i=1}^{m_1} \cos \omega_i t, \tag{5.49}$$

$$u_2 = \sum_{i=m_1+1}^{m} v_i = \sum_{i=m_1+1}^{m-1} \cos \omega_i t + \cos\left(\omega_m t + \varepsilon \frac{\pi}{2}\right). \tag{5.50}$$

Then, the non-autonomous flow of V_X between 0 and 2π is given by

$$\overrightarrow{\exp}(V_X)(0, 2\pi) = e^{f_{J_1}^Y \Pi(Y_{J_1})} \circ \cdots \circ e^{f_{J_{\tilde{N}}}^Y \Pi(Y_{J_{\tilde{N}}})} \circ \prod_{J > J_1} e^{f_J^Y \Pi(Y_J)}$$

$$= e^{\sum_{j=1}^{\tilde{N}} f_{J_j}^Y \Pi(Y_{J_j})} \circ \prod_{J > J_1} e^{\overline{f}_J^Y \Pi(Y_J)}. \tag{5.51}$$

We also know that

$$\overrightarrow{\exp}(V_X)(0, 2\pi) = e^{f_{I_1}^X X_{I_1}} \circ \cdots \circ e^{f_{I_N}^X X_{I_N}} \circ \prod_{I > I_1} e^{f_I^X X_I} = e^{\sum_{j=1}^{N} f_{I_j}^X X_{I_j}} \circ \prod_{I > I_1} e^{\overline{f}_I^X X_I}. \tag{5.52}$$

Recall that Π is surjective from $\mathcal{E}_Y(1, \ldots, 1)$ onto $\mathcal{E}_X(m_1, m_2)$. Therefore, by identifying Eqs. (5.51) and (5.52), we obtain that, for all $j = 1, \ldots, N$, $f_{I_j}^X$ is a linear combination of $f_{J_i}^Y$ with $i = 1, \ldots, \tilde{N}$, i.e.,

$$f_{I_j}^X = \sum_{i=1}^{\tilde{N}} \alpha_i^j f_{J_i}^Y. \tag{5.53}$$

Since the family $(f_{J_i}^Y)_{i=1,\ldots,\tilde{N}}$ is linearly independent and the matrix $A := (\alpha_i^j)_{i=1,\ldots,N;\, j=1,\ldots,\tilde{N}}$ is surjective, we conclude that the family $(f_{I_j}^X)_{j=1,\ldots,N}$ is also linearly independent. This ends the proof of Lemma 5.8. $\square$

A consequence of Lemma 5.8 is the following.

Corollary 5.9. *With the above notations, the function* $\det A$ *is not identically equal to zero on* $\bigcap_{j=1}^{N} P_j$.

Proof. For $j = 1, \ldots, N$, we define the vector L_j by

$$L_j = \left(f_{I_j}^X(\omega_{11}^1, \ldots, \omega_{21}^{m_2-1}, \omega_{21}^*), \ldots, f_{I_j}^X(\omega_{1N}^1, \ldots, \omega_{2N}^{m_2-1}, \omega_{2N}^*) \right)^T.$$

Assume that $\sum_{j=1}^N \ell_j L_j = 0$ with $l_j \in \mathbb{R}$. Then, for all $i = 1, \ldots, N$, we have

$$\sum_{j=1}^N \ell_j f_{I_j}^X(\omega_{1i}^1, \ldots, \omega_{1i}^{m_1}, \omega_{2i}^1, \ldots, \omega_{2i}^{m_2-1}, \omega_{2i}^*) = 0. \tag{5.54}$$

By Lemma 5.8, we have $l_j = 0$ for $j = 1, \ldots, N$. Then, the family $(L_j)_{j=1,\ldots,N}$ is linearly independent. We conclude that $\det A$ is not equal to zero. This ends the proof of Corollary 5.9. □

We still need another technical lemma which guarantees that there exist integer frequencies such that Eq. (5.12) is satisfied and the matrix A in Eq. (5.23) is invertible.

Lemma 5.10. *There exist integer frequencies such that* (5.12) *is satisfied and* $\det A$ *is not equal to zero.*

Proof. For $j = 1, \ldots, N$, we set

$$f_j(\omega_1, \ldots, \omega_{m-1}) = f_{I_j}^X\left(\omega_1, \ldots, \omega_{m-1}, \sum_{i=1}^{m-1} \omega_i \right), \tag{5.55}$$

then, we have

$$\det A = \begin{vmatrix} f_1(\omega_{11}^1, \ldots, \omega_{21}^{m_2-1}), & \ldots & , f_1(\omega_{1N}^1, \ldots, \omega_{2N}^{m_2-1}) \\ \vdots & \ddots & \vdots \\ f_N(\omega_{11}^1, \ldots, \omega_{21}^{m_2-1}), & \ldots & , f_N(\omega_{1N}^1, \ldots, \omega_{2N}^{m_2-1}) \end{vmatrix} = \frac{P(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-1})}{Q(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-1})}, \tag{5.56}$$

where P and Q are two polynomials of $(m-1)N$ variables.

We first note that Q never vanishes over integer frequencies. Assume, by contradiction, that P is always equal to zero for integer frequencies verifying Eq. (5.12). Consider P as a polynomial in one variable $\omega_{2N}^{m_2-1}$, i.e.,

$$P(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-1}) = \sum_{j=0}^M P_j(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-2})(\omega_{2N}^{m_2-1})^j. \tag{5.57}$$

Given integer frequencies $(\omega_{11}^1, \ldots, \omega_{2N}^{m_2-2})$, if Eq. (5.57) is not identically equal to zero, then this polynomial in the variable $\omega_{2N}^{m_2-1}$ most has a finite number of roots. However, for a given choice of $(m-1)N - 1$ first frequencies, there exist an infinite number of $\omega_{2N}^{m_2-1}$ verifying (5.12). Then, $P_j = 0$ over all integer frequencies, and P_M is not identically equal to zero. We note that all P_j are polynomials of $(m-1)N - 1$ variables. Proceeding by induction on the number of variables, it is easy to see that, at the end, we obtain a polynomial in the variable ω_{11}^1 which is equal to zero over all integer ω_{11}^1, and which is not identically equal to zero according to Corollary 5.9. That contradiction ends the proof of Lemma 5.10. □

5.2.3. General case: m > 2

Notice that the proof of Proposition 5.4 does not really depend on the number of vector fields involved in the control system (2.1). Indeed, for $m > 2$, if the control functions are linear combination of sinusoids with integer frequencies, then the state variables in the canonical form are also linear combinations of sinusoids so that the frequencies are $\mathbb{Z}$-linear combinations of the frequencies occurring in the control functions. The proof is the same as that of Lemma 5.5, up to extra notation. Since Claim 11 depends only on the length of the Lie brackets, but not on the number of vector fields, the proof of Lemma 5.8 does not depend on m, either. In order to prove a similar result for $m > 2$, we just need to re-project Eqs. (5.44) and (5.45) to m vector fields instead of 2.

5.3. Exact and sub-optimal steering law

In this section, we explain how we can devise, from Proposition 5.4, an *exact* and *sub-optimal* steering law (cf. Definition 2.17) $\text{Exact}_{m,r}$ for the approximate system, which is already in the canonical form and how $\text{Exact}_{m,r}$ can be incorporated into the *global approximate steering algorithm* (cf. Section 4.2). Note that $\text{Exact}_{m,r}$ only depends on the number of controlled vector fields m and on the maximum degree of nonholonomy r.

Recall that the components of $x \in \mathbb{R}^{\tilde{n}^r}$ are partitioned into equivalence classes $\{\mathcal{E}_x^1, \mathcal{E}_x^2, \ldots, \mathcal{E}_x^{\tilde{N}}\}$ according to Definition 5.1 in such a way that $\mathcal{E}_x^i \prec \mathcal{E}_x^j$ for $(i, j) \in \{1, \ldots, \tilde{N}\}^2$ and $i < j$. For every equivalence class $\mathcal{E}_x^j$, Proposition 5.4 and Section 5.2.3 guarantee that we can choose frequencies such that the corresponding control matrix A_j (cf. Definition 5.4) is invertible and the corresponding control function u^j obtained by Eq. (5.13) steers all the elements of $\mathcal{E}_x^j$ from an arbitrary initial value to the origin 0 (see Remark 4.5) without modifying any elements belonging to smaller classes.

Let $x^{\text{initial}} \in \mathbb{R}^{\tilde{n}^r}$. Let $B_j := A_j^{-1}$ and $N_j := \text{Card}(\mathcal{E}_x^j)$, $j = 1, \ldots, \tilde{N}$. For $x \in \mathbb{R}^{\tilde{n}^r}$, we will use $[x]_{i,\ldots,k}$ with $1 \leqslant i < k \leqslant \tilde{n}_r$ to denote the vector $(x_i, \ldots, x_k)$, and $\|x\|$ to denote the *pseudo-norm* of x defined by the *free* weights (cf. Definitions 2.13 and 3.7). We also define an intermediate function $\text{Position}(u)$ as follows: if system (5.1) starts from $x = 0$ and evolves under the action of u, then $\text{Position}(u)$ returns its position at $t = 2\pi$.

Algorithm 5 Exact Steering Law: $\text{Exact}_{m,r}(x^{\text{initial}})$.

Require: $B_1, \ldots, B_{\tilde{N}}$, and $N_1, \ldots, N_{\tilde{N}}$;

1: $\lambda := \|x^{\text{initial}}\|_0$;
2: $x^{\text{new}} := \delta_{0, \frac{1}{\lambda}}(x^{\text{initial}})$;
3: $\hat{u}_{\text{norm}} := 0$;
4: $i := 0$;
5: **for** $j = 1, \ldots, \tilde{N}$ **do**
6:     $x := [x^{\text{new}}]_{i+1,\ldots,i+N_j}$;
7:     $a^j := B_j \, x$;
8:     construct u^j from a^j by Eq. (5.13);
9:     $x^{\text{new}} := x^{\text{new}} + \text{Position}(u^j)$;
10:     $\hat{u}_{\text{norm}} := \hat{u}_{\text{norm}} * u^j$ (cf. Definition 5.3);
11:     $i = i + N_j$;
12: **end for**
13: **return** $\hat{u} := \lambda \hat{u}_{\text{norm}}$.

Proposition 5.11. *For every $x^{\text{initial}} \in \mathbb{R}^{\tilde{n}^r}$, the input given by $\text{Exact}_{m,r}(x^{\text{initial}})$ steers system (5.1) from x^{initial} to 0 exactly. Moreover, there exists a constant $C > 0$ such that*

$$\ell\big(\text{Exact}_{m,r}\big(x^{\text{initial}}\big)\big) \leqslant C d\big(x^{\text{initial}}, 0\big), \quad \forall x^{\text{initial}} \in \mathbb{R}^{\tilde{n}^r}, \tag{5.58}$$

where we use d to denote the sub-Riemannian distance defined by the family X.

Proof. The fact that the procedure described by the Lines 5–12 in Algorithm 5 produces an input $\hat{u}_{\text{norm}}$ steering system (5.1) from $\delta_{0,\frac{1}{\lambda}}(x^{\text{initial}})$ to 0 is a consequence of Proposition 5.4 and Section 5.2.3. We also note that, due to the homogeneity of system (5.1), if an input u steers (5.1) from x to 0, then, for every $\lambda > 0$, the input λu steers (5.1) from $\delta_{0,\lambda}(x)$ to 0. Therefore, the input computed by $\text{Exact}_{m,r}(x^{\text{initial}})$ steers system (5.1) from x^{initial} to 0. Let us now show (5.58). In the sequel, the application $\text{Exact}_{m,r} : x \to \text{Exact}_{m,r}(x)$ will be simply denoted by $\hat{u} : x \to \hat{u}(x)$. Let $S(0,1) := \{y, \|y\|_0 = 1\}$ and $x \in \mathbb{R}^{\tilde{n}^r}$. Then, there exists $x_{\text{norm}} \in S(0,1)$ such that $x = \delta_{0,\lambda}(x_{\text{norm}})$ with $\lambda := \|x\|_0$. We have:

$$\ell\big(\hat{u}(x)\big) = \ell\big(\lambda \hat{u}(x_{\text{norm}})\big) = \lambda \ell\big(\hat{u}(x_{\text{norm}})\big) \leqslant \lambda \sup_{y \in S(0,1)} \ell\big(\hat{u}(y)\big).$$

We also know that, since the sub-Riemannian distance $d(0, \cdot)$ from 0 and the pseudo-norm $\| \cdot \|_0$ at 0 are both homogeneous of degree 1 with respect to the dilation $\delta_{0,t}(\cdot)$, there exists a constant $\tilde{C} > 0$ such that $\tilde{C}\lambda \leqslant d(0,x)$. Since the application $y \to \hat{u}(y)$ is continuous from $S(0,1)$ to $\mathbb{R}^m$ and $S(0,1)$ is compact, then, $\sup_{y \in S(0,1)} \ell(\hat{u}(y))$ is bounded, thus the inequality (5.58) holds true. □

The following theorem is a consequence of Proposition 5.11 and Remark 4.4.

Theorem 5.12. *The function* $\text{Exact}_{m,r}(\cdot)$ *constructed by Algorithm 5 provides the approximate system* $\mathcal{A}^X$ *defined in Section 4.1 with a sub-optimal steering law.*

Proof. It suffices to note that, for every $a \in \Omega$, $\mathcal{A}^X(a)$ has the same form (cf. Remark 4.4), thus defines the same sub-Riemannian distance d. Therefore, the inequality (5.58) holds uniformly with respect to the approximate point a, and this terminates the proof of Theorem 5.12. □

Remark 5.4. Frequencies choices and the construction of the corresponding control matrix A_j, as well as its inverse B_j, translate to off-line computations. We note that Proposition 5.4 only gives sufficient conditions to prevent resonance (by choosing widely spaced frequencies, cf. Eq. (5.12)) and guarantee the invertibility of the corresponding matrix (by using a sufficiently large number of independent frequencies). These conditions tend to produce high frequencies while it is desirable to find smaller ones for practical use. We can prove that two independent frequencies suffice to steer one component (cf. Section 5.2.1), and we *conjecture* that $2N$ independent frequencies suffice to control one equivalence class of cardinal N by producing an invertible matrix. One can implement a searching algorithm for finding the optimal frequencies for each equivalence class such that they prevent all resonances in smaller classes and produce an invertible matrix for the class under consideration. Proposition 5.4 guarantees the finiteness of such an algorithm. Moreover, one can construct, once for all, a table containing the choice of frequencies and the corresponding matrices for each equivalence class in the free canonical system.

Remark 5.5. Recall that the key point in our control strategy consists in choosing suitable frequencies such that, during each 2π-period, the corresponding input function displaces components of one equivalence class to their preassigned positions while all the components of smaller classes (according to the ordering in Definition 5.2) return at the end of this control period to the values taken at the beginning of the period. In order to achieve the previous task, special resonance conditions must be verified by the appropriate components, and these conditions must not hold for all the other smaller components (according to the ordering in Definition 5.2). Note that two categories of frequencies have been picked in Proposition 5.4: the basic frequencies $\{\omega_{ij}^k\}$, and the resonance frequencies $\{\omega_{ij}^*\}$. Since frequencies occurring in the dynamics of the state variables are just $\mathbb{Z}$-linear combinations of $\{\omega_{ij}^k\} \cup \{\omega_{ij}^*\}$, and the resonance frequencies $\{\omega_{ij}^*\}$ are chosen to be special $\mathbb{Z}$-linear combinations of basic frequencies (resonance condition), the frequencies in the dynamics of the state variables are special $\mathbb{Z}$-linear combinations of $\{\omega_{ij}^k\}$.

Remark 5.6. Once the frequencies and matrices are obtained, the on-line computation Exact$_{m,r}$ is only a series of matrix multiplications which can be performed on-line without any numerical difficulty.

Remark 5.7. The Desingularization Algorithm presented in Section 3.3 (see also Remarks 3.6 and 5.1) together with Algorithm 5 provides general nilpotent control systems with an exact steering method, which is also sub-optimal.

Remark 5.8. We note that the inputs constructed in this section are piecewise C^∞ during each time interval $[2i\pi, 2(i+1)\pi]$, for $i = 1, \ldots, \tilde{N} - 1$, but they are not globally continuous during the entire control period $[0, \ 2\tilde{N}\pi]$, due to discontinuity at $t = 2\pi, 4\pi, \ldots, 2(\tilde{N}-1)\pi$. However, it is not difficult to devise (globally) continuous inputs using interpolation techniques. We illustrate the idea with a simple example. Assume that we use $u^i = (u_1^i, u_2^i)$ and $u^j = (u_1^j, u_2^j)$ defined by

$$u_1^i(t) = \cos \omega_{1i} t,$$

$$u_2^i(t) = \cos \omega_{2i} t + a^i \cos\left(\omega_{2I}^* t + \varepsilon^i \frac{\pi}{2} \right), \quad t \in \left[2(i-1)\pi, 2i\pi \right],$$

$$u_1^j(t) = \cos \omega_{1j} t,$$

$$u_2^j(t) = \cos \omega_{2j} t + a^j \cos\left(\omega_{2j}^* t + \varepsilon^j \frac{\pi}{2} \right), \quad t \in \left[2(j-1)\pi, 2j\pi \right],$$

to steer two consecutive classes $\mathcal{E}_x^i$ and $\mathcal{E}_x^j$ (i.e. $j = i + 1$) which are both of cardinal equal to 1.

If we require their concatenation $u^i * u^j$ to be continuous, i.e.

$$u_1^i(2\pi) = u_1^j(2\pi), \tag{5.59}$$

$$u_2^i(2\pi) = u_2^j(2\pi), \tag{5.60}$$

we can proceed as follows.

For Eq. (5.59), it suffices to modify slightly u_1^j. We take

$$\tilde{u}_1^j(t) = u_1^i(2\pi) \cos \omega_{1J} t. \tag{5.61}$$

For Eq. (5.60), we distinguish two cases:

- if $\varepsilon^j = 1$, we can take

$$\tilde{u}_2^j(t) = u_2^i(2\pi) \cos \omega_{2j} t + a^j \cos\left(\omega_{2j}^* t - \frac{\pi}{2} \right)$$

$$= u_2^i(2\pi) \cos \omega_{2j} t + a^j \sin \omega_{2j}^* t; \tag{5.62}$$

- if $\varepsilon^j = 0$, we add a frequency ω_c to u_2^j, large enough to avoid any additional resonances,

$$\tilde{u}_2^j(t) = \cos \omega_{2j} t + a^j \cos \omega_{2j}^* t + \left(u_2^i(2\pi) - a^j - 1 \right) \cos \omega_c t. \tag{5.63}$$

Let $\tilde{u}^j := (\tilde{u}_1^j, \ \tilde{u}_2^j)$. Then, by construction, the new input $u^i * \tilde{u}^j$ is continuous over the time interval $[2i\pi, 2j\pi]$.

It is clear that this idea of interpolation by adding suitable frequencies can be used to construct continuous inputs over the entire control period $[0, 2\tilde{N}\pi]$. In fact, by using more refined interpolations, one can get inputs of class C^k for arbitrary finite integer k.

Remark 5.9. Using the idea presented in Remark 5.8 together with Remark 5.7, it is easy to conclude that, for general nilpotent systems, the resulting trajectories are globally C^1 curves and the regularity does not depend on the time-parameterization of the trajectories.

For the sake of completeness, we give in this short appendix the proof of Theorem 1.1 together with some comments on Algorithm 2.

Appendix A. Proof of Theorem 1.1

We first note that Steps 1 through 5 in Algorithm 2 are straightforward.

Theorem 3.5 guarantees that the Desingularization Algorithm (Section 3.3) provides us with a new family of vector fields $\xi = \{\xi_1, \dots, \xi_m\}$, which is regular and free up to step r with r denoting the maximum value of the degree of nonholonomy of the original system $X = \{X_1, \dots, X_m\}$, on the corresponding compact set $\mathcal{V}^c_{\mathcal{J}_i}$. Then, we construct the approximate system $\mathcal{A}^\xi$ using the procedure presented in Section 4.1 and provide it with the sub-optimal steering law $\text{Exact}_{m,r}$ defined in Algorithm 5. The sub-optimality of $\text{Exact}_{m,r}$ is guaranteed by Theorem 5.12. Therefore, by Theorem 4.2, the LAS method AppSteer associated with $\mathcal{A}^\xi$ and its steering law is uniformly contractive on the compact set $\mathcal{V}^c_{\mathcal{J}_i} \times \bar{B}_R(0)$. Then, by Theorem 4.3, $\text{GlobalFree}(\tilde{x}_0, \tilde{x}_1, e, \mathcal{V}^c, \text{AppSteer})$ provided by Algorithm 3 terminates in a finite number of steps and stops at a point $\tilde{x}$ such that $d(\tilde{x}, \tilde{x}_1) < e$. Since there is a finite number of compacts to be explored, we conclude that Algorithm 2 terminates in a finite number of steps and steers the system (Σ) from x^{initial} to some point x such that $d(x, x^{\text{final}}) < e$. This ends the proof of Theorem 1.1.

Appendix B. About the control set

Let $U \subset \mathbb{R}^m$ be any neighborhood of the origin. Then every trajectory of (2.1) corresponding to the inputs produced by Algorithm 2 can be time-reparameterized so that the resulting trajectory of (2.1) is associated with an input taking values in U.

Appendix C. Getting trajectories of class C^1 for the original control system

We can slightly modify Algorithm 2 to get trajectories of class C^1 for the original control system (Σ). This is equivalent to ask for continuous inputs produced by the algorithm. According to Remark 5.8, inputs can be made continuous within each iteration step in Algorithm 3 since they are computed based the nilpotent approximate system. By using the same interpolation technique as presented in Remark 5.8, we can still produce inputs which remain continuous from one step to another in Algorithm 3. Therefore, trajectories of class C^1 for the control system (Σ) are obtained.

References

[1] A.A. Agrachev, Yu.L. Sachkov, Control Theory from the Geometric Viewpoint, Springer, 2004.
[2] F. Alouges, Y. Chitour, R. Long, A motion planning algorithm for rolling-body problem, IEEE Trans. Robot. 26 (2010) 827–836.
[3] A. Bellaïche, The tangent space in sub-Riemannian geometry, in: A. Bellaïche, J.-J. Risler (Eds.), Sub-Riemannian Geometry, Birkhäuser, 1996, pp. 1–78.
[4] A. Bellaïche, F. Jean, J.J. Risler, Geometry of nonholonomic systems, in: J.P. Laumond (Ed.), Robot Motion Planning and Control, Springer, 1998, Chapter 2.
[5] F. Bullo, A. Lewis, N. Leonard, Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups, IEEE Trans. Automat. Control 45 (8) (2000) 1437–1454.
[6] J.F. Bonnans, J.Ch. Gilbert, C. Lemaréchal, C. Sagastizábal, Numerical Optimization – Theoretical and Practical Aspects, second ed., Universitext, Springer, Berlin, 2006.

[7] N. Bourbaki, Groupe et Algèbre de Lie, Hermann, Paris, 1972.
[8] A. Chelouah, Y. Chitour, On the motion planning of rolling surfaces, Forum Math. 15 (5) (2003) 727–758.
[9] Y. Chitour, Path planning on compact Lie groups using a homotopy method, Systems Control Lett. 47 (5) (2002) 383–391.
[10] Y. Chitour, A continuation method for motion planning problems, ESAIM Control. Optim. Calc. Var. 12 (2006).
[11] Y. Chitour, F. Jean, E. Trélat, Genericity results for singular curves, J. Differential Geom. 73 (2006).
[12] Y. Chitour, F. Jean, E. Trélat, Singular trajectories of control-affine systems, SIAM J. Control Optim. 47 (2) (2008).
[13] Y. Chitour, H.J. Sussmann, Line-integral estimates and motion planning using the continuation method, in: Essays on Mathematical Robotics, Minneapolis, MN, 1993, in: IMA Vol. Math. Appl., vol. 104, Springer, New York, 1998, pp. 91–125.
[14] W.L. Chow, Über Systeme von linearen partiellen Differentialgleichungen erster Ordnung, Math. Ann. 117 (1940) 98–115.
[15] M. Fliess, J. Lévine, P. Martin, P. Rouchon, Flatness and defect of nonlinear systems: Introductory theory and examples, Internat. J. Control 61 (1995) 1327–1361.
[16] N. Goodman, Nilpotent Lie groups, Springer Lecture Notes in Math., vol. 562, Springer, 1976.
[17] M. Grayson, R. Grossman, Nilpotent Lie algebras and vector fields, in: R. Grossman (Ed.), Symbolic Computation: Applications to Scientific Computing, SIAM, Philadelphia, 1989, pp. 77–96.
[18] M. Grayson, R. Grossman, Models for free nilpotent Lie algebra, J. Algebra 135 (1991) 177–191.
[19] H. Hermes, Nilpotent and high-order approximation of vector field systems, SIAM Rev. 33 (2) (1991) 238–264.
[20] F. Jean, Uniform estimation of sub-Riemannian balls, J. Dyn. Control Syst. 7 (4) (2001) 473–500.
[21] F. Jean, G. Oriolo, M. Vendittelli, A global convergent steering algorithm for regular nonholonomic systems, in: Proceedings of 44th IEEE CDC-ECC'05, Séville, Spain, 2005.
[22] F. Jean, R. Long, G. Oriolo, M. Vendittelli, An approximate algorithm for nonholonomic motion planning, preprint, R.I. 651, CMAP, Ecole Polytechnique, France, 2008.
[23] V. Jurdjevic, Geometric Control Theory, Cambridge Stud. Adv. Math., vol. 51, Cambridge University Press, 1997.
[24] G. Lafferriere, A general strategy for computing steering controls of systems without drift, in: Proceedings of 30th IEEE CDC, Brighton, England, 1991.
[25] G. Lafferriere, H.J. Sussmann, A differential geometry approach to motion planning, in: Z. Li, J.F. Canny (Eds.), Nonholonomic Motion Planning, Kluwer, 1992.
[26] N. Leonard, P.S. Krishnaprasad, Motion control of drift-free, left-invariant systems on Lie groups, IEEE Trans. Automat. Control 40 (9) (1995) 1539–1554.
[27] W.S. Liu, An approximation algorithm for nonholonomic systems, SIAM J. Control Optim. 35 (4) (1997) 1328–1365.
[28] P. Martin, R.M. Murray, P. Rouchon, Flat systems: Open problems, infinite dimensional extension, symmetries and catalog, in: Advances in the Control of Nonlinear Systems, in: Lecture Notes in Control and Inform. Sci., Springer, 2001, pp. 33–57.
[29] R.M. Murray, S. Sastry, Nonholonomic motion planning: Steering using sinusoids, IEEE Trans. Automat. Control 38 (5) (1993) 700–716.
[30] R.M. Murray, Nilpotent bases for a class of non-integrable distributions with applications to trajectory generation for nonholonomic systems, Math. Control Signals Systems 7 (1) (1994) 58–75.
[31] G. Oriolo, M. Vendittelli, A framework for the stabilization of general nonholonomic systems with an application to the plate–ball mechanism, IEEE Trans. Robot. 21 (2) (2005) 162–175.
[32] P.K. Rashevsky, Any two points of a totally nonholonomic space may be connected by an admissible line, Uch. Zap. Ped. Inst. im. Liebknechta, Ser. Phys. Math. 2 (1938) (in Russian).
[33] L.P. Rothschild, E.M. Stein, Hypoelliptic differential operators and nilpotent groups, Acta Math. 137 (1976) 247–320.
[34] E.D. Sontag, Control of systems without drift via generic loops, IEEE Trans. Automat. Control 40 (7) (1995) 1210–1219.
[35] P. Stefan, Accessible sets, orbits, and foliations with singularities, Proc. Lond. Math. Soc. 29 (1974) 699–713.
[36] H.J. Sussmann, Orbits of families of vector fields and integrability of distributions, Trans. Amer. Math. Soc. 180 (1973) 171–188.
[37] H.J. Sussmann, A product expansion for the Chen series, in: C.I. Byrnes, A. Lindquist (Eds.), Theory and Applications of Nonlinear Control Systems, Elsevier Science Publisher B.V., North-Holland, 1986.
[38] H.J. Sussmann, A general theorem on local controllability, SIAM J. Control Optim. 25 (1987) 158–194.
[39] H.J. Sussmann, A continuation method for nonholonomic path finding, in: 32th IEEE CDC, 1993, pp. 2718–2723.
[40] M. Vendittelli, G. Oriolo, F. Jean, J.-P. Laumond, Non-homogeneous nilpotent approximation for systems with singularities, IEEE Trans. Automat. Control 49 (2) (2004) 261–266.
[41] A.M. Vershik, V.Ya. Gershkovich, Nonholonomic dynamical systems, geometry of distributions and variational problems, in: V.I. Arnold, S.P. Novikov (Eds.), Dynamical Systems VII, in: Encyclopaedia Math. Sci., vol. 16, Springer, 1994.