SIAM J. CONTROL OPTIM.
Vol. 43, No. 1, pp. 32–57

PRACTICAL AND ASYMPTOTIC STABILIZATION OF CHAINED SYSTEMS BY THE TRANSVERSE FUNCTION CONTROL APPROACH*

PASCAL MORIN† AND CLAUDE SAMSON†

Abstract. A control approach for the practical and asymptotic stabilization of nonlinear driftless systems subjected to additive perturbations is proposed. Such perturbations arise naturally, for instance, in the modeling of trajectory stabilization problems for controllable driftless systems on Lie groups. The objective of the approach is to provide practical stability of an arbitrary given point in the state space, whatever the perturbations, and asymptotic stability (resp., convergence to the point) when the perturbations are absent (resp., tend to zero). A general framework is presented in this paper, and a control solution is proposed for the class of the chained systems.

Key words. nonlinear systems, stabilization, feedback control, Lie groups

AMS subject classifications. 93B05, 93B29, 93C10

DOI. 10.1137/S0363012903421868

1. Introduction. The development of the transverse function (t.f.) approach [18] finds its original motivation in the problem of *practical* stabilization of the origin of a control system in the form

$$\mathcal{S}: \qquad \dot{x} = \sum_{i=1}^{m} u_i X_i(x) + P(x,t), \tag{1}$$

with $x \in \mathbb{R}^n$, $n > m$, $\{X_1, \ldots, X_m\}$ a set of smooth vector fields (v.f.) that satisfy the Lie algebra rank condition (LARC) on an open ball centered at $x = 0$, and P an additive perturbation, continuous in x and t but otherwise *arbitrary.* Note that such a perturbation may well forbid the existence of any equilibrium point for the controlled system. The t.f. approach provides a general solution to this problem. Up to now, and to our knowledge, this solution is unique in its class, even though several other methods and many control laws have been devised during the last decade to address the stabilization problem when $P \equiv 0$. These studies were motivated in the first place by Brockett's theorem [5] according to which, if $m < n$ and the control v.f. evaluated at $x = 0$ are linearly independent, no smooth or even continuous pure state feedback can make the origin of the system asymptotically stable. Different types of feedback laws have been considered to circumvent this difficulty, although not all of them guarantee Lyapunov stability. Discontinuous feedback [1, 3, 6, 11] and hybrid feedback [2, 15, 23] are two possibilities. Another one, more related to the present approach, consists of using continuous time-varying feedback [21, 7, 20, 27, 22, 13, 19, 16, 14]. An early survey on the control of nonholonomic systems, whose kinematic models are nonlinear driftless systems, can also be found in [4]. The importance of considering the perturbed case in association with the objective of practical stabilization is well illustrated when $\mathcal{S}$ is a system on a Lie group and the control objective consists of tracking a trajectory. Indeed, it is shown in [18] that the *error system* associated with

*Received by the editors January 29, 2003; accepted for publication (in revised form) October 7, 2003; published electronically May 25, 2004.
http://www.siam.org/journals/sicon/43-1/42186.html
†INRIA, 2004 Route des Lucioles, B. P. 93, 06902 Sophia-Antipolis Cedex, France (Pascal.Morin@inria.fr, Claude.Samson@inria.fr)

this problem is in the same form as the original system, except for the presence of a perturbation P. Moreover, when the trajectory is not a solution of the control system, asymptotic stabilization is not possible. Other reasons for considering practical stabilization as a reasonable control objective, in the case of nonlinear driftless systems, are also pointed out in [18]: lack of robustness of exponential (continuous/time-varying or discontinuous) stabilizers, nonexistence of feedback controllers capable of stabilizing asymptotically every feasible trajectory [12], and incapacity of most existing asymptotic stabilizers to ensure ε-ultimate boundedness of the closed-loop trajectories when a destabilizing perturbation P is present. However, it is important to realize that practical stabilization is by no means opposed to asymptotic stabilization. It is merely a weaker requirement, whose interest resides precisely in the fact that it is weaker and thus applicable to more numerous situations. Once practical stabilization is granted, it may still be possible, and desirable in some cases, to achieve asymptotic stabilization, or at least convergence to zero—when, for instance, P vanishes after some time. For the same reasons, feedback controllers derived with the t.f. approach should not be considered as antagonistic to other controllers proposed for nonlinear driftless systems—asymptotic stabilizers, in particular. A more pertinent issue is the possibility of deriving a practical stabilizer which also ensures asymptotic stabilization when the perturbation P allows for it. For instance, can the t.f. approach be used for this purpose?

This question is addressed in the present paper, and a partial positive answer is obtained. More precisely, an extension of the approach in [18] is proposed in order to achieve asymptotic stabilization of the origin of $\mathcal{S}$ when $P \equiv 0$, and asymptotic convergence to the origin when P tends to zero as time tends to infinity. The main ingredient of this extension is the concept of a *generalized* t.f. introduced in section 2. The principles of the t.f. approach and design of stabilizers are also laid out in this section. A solution to the problem of practical and asymptotic stabilization for the popular class of the chained systems is proposed in section 3, and illustrated by simulation results in section 4. The practical relevance of this case comes from the widespread use of chained systems to model the kinematic equations of various mechanical systems subjected to nonholonomic constraints (unicycle and car-like mobile robots, for instance) and also the possibility of using them as homogeneous approximations of dynamics involved in several other physical systems (ships, induction motors, etc.). Finding a more general solution, which applies to a broader class of systems, remains an open subject of research. In order to facilitate the reading of the paper, we have distributed the proofs of our results into two sections: the cores are given in section 5, whereas intermediate technical results of lesser conceptual significance are regrouped in the appendix.

Since the t.f. approach finds its most natural exposition in the context of systems which are invariant on Lie groups, we have chosen to recast the systems and control problems evoked above in this framework. Let us recall the prominent role played by Lie groups in control theory [26, 10]. In particular, controllable driftless systems can always be approximated by controllable driftless homogeneous systems which are, after a possible dynamic extension, systems on Lie groups. The chained systems, which are more specifically addressed here, are systems on Lie groups.

The following notation is used throughout the paper. The tangent space of a manifold M at a point p is denoted as T_pM. The differential of a smooth mapping f between manifolds, at a point p, is denoted as $df(p)$. The torus of dimension $p \in \mathbb{N}$ is $\mathbb{T}^p$ with $\mathbb{T} \stackrel{\Delta}{=} \mathbb{R}/2\pi\mathbb{Z}$. An element $\theta \in \mathbb{T}$ is identified with the real number in $(-\pi, +\pi]$ which belongs to the class of equivalence of θ. Addition of angles makes $\mathbb{T}^p$

a Lie group. The ith component of $\sigma \in \mathbb{T}^p$ is denoted as σ_i, i.e., $\sigma = (\sigma_1, \ldots, \sigma_p)$. The canonical basis of $\mathbb{R}^p$ is the set of unitary vectors $\{e_i\}_{i=1,\ldots,p}$. Since this set is also the natural basis of the Lie algebra of $\mathbb{T}^p$, a vector field v on $\mathbb{T}^p$ is identified with its vector of coordinates in this basis, i.e., $v = (v_1, \ldots, v_p)'$ if $v = \sum_{i=1}^p v_i e_i$. If $\sigma(.)$ is a smooth curve on $\mathbb{T}^p$, this identification allows us to view $\dot{\sigma}(t)$ as a vector in $\mathbb{R}^p$. Consider a differentiable mapping f from $\mathbb{T}^p$ to a manifold M. By a slight abuse of notation, and for the sake of simplifying the writing of several forthcoming equations, we write the Lie derivative of f along e_i at σ as $\frac{\partial f}{\partial \sigma_i}(\sigma)$, or $\frac{\partial f}{\partial \sigma}(\sigma)e_i$, instead of $df(\sigma)(e_i)$ (or $L_{e_i}f(\sigma)$), even though the normal use of the partial derivative symbol refers to a system of coordinates on M. Accordingly, along an arbitrary v.f. v on $\mathbb{T}^p$, we write $\frac{\partial f}{\partial \sigma}(\sigma)v \triangleq df(\sigma)(v)$. We also use standard notation for Lie groups—see, e.g., [8] for more details on this topic. G denotes a Lie group of dimension n, with Lie algebra (of left-invariant v.f.) $\mathfrak{g}$. For simplicity, we assume that G is connected so that there exists a globally defined left-invariant distance d_G on G. The identity element of G is denoted by e. Left and right translations are denoted by l and r, respectively, i.e., $l_\sigma(\tau) = r_\tau(\sigma) = \sigma\tau$. As usual, if $X \in \mathfrak{g}$, then $\exp tX$ is the solution at time t of $\dot{g} = X(g)$ with initial condition $g(0) = e$. The adjoint representation of G is Ad; i.e., for $\sigma \in G$, $\mathrm{Ad}(\sigma) = dI_\sigma(e)$ with $I_\sigma : G \longrightarrow G$ defined by $I_\sigma(g) = \sigma g \sigma^{-1}$. By extension we define the v.f. $\mathrm{Ad}(\sigma)X$ on G by $\mathrm{Ad}(\sigma)X(g) = dl_g(e)(\mathrm{Ad}(\sigma)X(e))$. The differential of Ad is ad, and $(\mathrm{ad}X, Y) = [X, Y]$, the Lie bracket of X and Y.

2. Control of perturbed driftless systems by the t.f. approach. Consider a control system

$$\mathcal{S}(g): \qquad \dot{g} = \sum_{i=1}^m u_i X_i(g) + P(g,t) \tag{2}$$

on a Lie group G, with $X_1, \ldots, X_m$ independent left-invariant smooth v.f. that satisfy the LARC. We assume that the drift term $P(g,t)$ is a continuous function of g and t, and that[1]

$$\forall (g,t) \in G \times \mathbb{R}, \qquad P(g,t) \in \mathrm{span}\{X_1(g), \ldots, X_m(g)\}^\perp, \tag{3}$$

where orthogonality refers to an arbitrary Riemannian metric on G. The definition of a *transverse function*, as originally given in [17] for v.f. on an arbitrary manifold—i.e., not necessarily on a Lie group—is now recalled.

DEFINITION 1. *Let $X_1, \ldots, X_m$ denote smooth v.f. on a manifold M. A function $f \in \mathcal{C}^\infty(\mathbb{T}^p; M)$ is called a* transverse function *(for the v.f. $X_1, \ldots, X_m$) if*

$$\forall \sigma \in \mathbb{T}^p, \quad \mathrm{span}\{X_1(f(\sigma)), \ldots, X_m(f(\sigma))\} + df(\sigma)(T_\sigma\mathbb{T}^p) = T_{f(\sigma)}M. \tag{4}$$

Another way of writing the above relation (with the notation explained before) is

$$\forall \sigma \in \mathbb{T}^p, \quad \mathrm{span}\left\{X_1(f(\sigma)), \ldots, X_m(f(\sigma)), \frac{\partial f}{\partial \sigma_1}(\sigma), \ldots, \frac{\partial f}{\partial \sigma_p}(\sigma)\right\} = T_{f(\sigma)}M. \tag{5}$$

Note that, by this definition, the image set $\mathrm{Im}(f) = f(\mathbb{T}^p)$ is compact. The main contribution of [17] was to show that if a set of v.f. $X_1, \ldots, X_m$ satisfies the LARC at some point $q \in M$, then for any neighborhood $\mathcal{U}$ of q there exists a transverse function with values in $\mathcal{U}$.

[1] Note that (3) can always be obtained after the application of a suitable preliminary feedback.

In the context of stabilization, transverse functions allow to use $\dot{\sigma}$ as a new—virtual—control input vector. This leads us to introduce the following dynamic extension of $\mathcal{S}(g)$, which evolves on $G \times \mathbb{T}^p$:

$$\mathcal{S}(g,\sigma): \quad \begin{cases} \dot{g} = \sum_{i=1}^{m} u_i X_i(g) + P(g,t), \\ \dot{\sigma} = u_\sigma, \end{cases} \tag{6}$$

where (u, u_σ) is viewed as an extended control vector. In the following subsection, the practical stabilization of $g = e$ for $\mathcal{S}(g)$, based on the t.f. approach, is addressed. More details on the approach, as well as several examples with explicit derivations of t.f., can be found in [18].

2.1. Practical stabilization. The formulation of a general practical stabilization problem which can be associated with $\mathcal{S}(g)$ is as follows: given an arbitrary neighborhood $\mathcal{U}_G(e)$ of e, determine a (smooth, or at least continuous) feedback control (which depends on g and, eventually, on other variables) which asymptotically stabilizes some compact set $\mathcal{D}_G \subset \mathcal{U}_G(e)$. The t.f. control approach provides a solution to this problem. This solution is now recalled.

Consider the change of variables on $G \times \mathbb{T}^p$ defined by $\Psi_f(g,\sigma) = (f(\sigma)g^{-1}, \sigma)$, with f a t.f. such that $f(\mathbb{T}^p) \subset \mathcal{U}_G(e)$. From now on, in order to ease the notation, the element $f(\sigma)g^{-1} \in G$ associated with g and $f(\sigma)$ will be abbreviated as z, i.e., $z \triangleq f(\sigma)g^{-1}$. By differentiating both members of the equality $zg = f(\sigma)$, one easily verifies that, along any trajectory $(g,\sigma)(.)$ of $\mathcal{S}(g,\sigma)$,

$$\dot{z} = -dr_{g^{-1}}(f(\sigma))\left(\sum_{i=1}^{m} u_i X_i(f(\sigma)) - \frac{\partial f}{\partial \sigma}(\sigma)\dot{\sigma} + dl_z(g)P(g,t)\right). \tag{7}$$

Therefore, $\mathcal{S}(g,\sigma)$ is equivalent to the control system

$$\bar{\mathcal{S}}(z,\sigma): \begin{cases} \dot{z} = -dr_{g^{-1}}(f(\sigma))\left(\sum_{i=1}^{m} u_i X_i(f(\sigma)) - \frac{\partial f}{\partial \sigma}(\sigma)u_\sigma + dl_z(g)P(g,t)\right), \\ \dot{\sigma} = u_\sigma. \end{cases} \tag{8}$$

From the definition of Ψ_f, the asymptotic stability of $\{e\} \times \mathbb{T}^p$ for $\bar{\mathcal{S}}(z,\sigma)$ is equivalent to the asymptotic stability of $\{(f(\sigma),\sigma) : \sigma \in \mathbb{T}^p\}$ for $\mathcal{S}(g,\sigma)$. It is also equivalent to the asymptotic stability of $f(\mathbb{T}^p)$ for $\mathcal{S}(g)$, provided that, for some left-invariant distance on G, the initial value $\sigma(0)$ of σ is chosen so as to minimize the distance between $z(0) = f(\sigma(0))g(0)^{-1}$ and e.

Now, for any v.f. Z on G, the property of transversality of f ensures that the equation

$$\sum_{i=1}^{m} u_i X_i(f(\sigma)) - \frac{\partial f}{\partial \sigma}(\sigma)u_\sigma = -dl_z(g)P(g,t) - dr_g(z)Z(z) \tag{9}$$

admits a feedback solution $(u, u_\sigma)(g,\sigma,t)$. Applying any[2] such feedback law to $\bar{\mathcal{S}}(z,\sigma)$, and using the fact that $(dr_g(z))^{-1} = dr_{g^{-1}}(f)$, it follows from (7) that

$$\dot{z} = Z(z). \tag{10}$$

[2]The only (weak) requirement is that the solutions of $\mathcal{S}(g,\sigma)$ must be well defined for $t \in [0,\infty)$.

Therefore, provided that Z is chosen so as to asymptotically stabilize e for system (10), the feedback law (u, u_σ) defined by (9) makes the set $\{e\} \times \mathbb{T}^p$ asymptotically stable for $\bar{\mathcal{S}}(z, \sigma)$.

In general, the solution (u, u_σ) of (9) is not unique. It is shown in [18], however, that one can always find[3] t.f. $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m}; G)$, i.e., such that $p = n - m$ with the notation of Definition 1. It is clear from the transversality condition (4) that this value of p is minimal and that the solution (u, u_σ) of (9), given f, is unique in this case. Allowing the t.f. f to depend on a larger number of variables provides complementary control inputs which can be used to guarantee complementary control objectives. The asymptotic stabilization of e for $\mathcal{S}(g)$ when $P \equiv 0$ will, for instance, be addressed in this way.

2.2. A framework for asymptotic stabilization. Let us introduce, in the framework of Lie groups, the following specific class of transverse functions.

DEFINITION 2. *Consider a function $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m} \times \mathbb{T}^{n-m}; G)$ and the associated family of functions $\{f_\beta\}_{\beta \in \mathbb{T}^{n-m}}$ defined by $f_\beta(\theta) = f(\theta, \beta)$. The function f is called a* generalized *t.f. for the v.f. $X_1, \dots, X_m$ on the Lie group G if*

$$\forall \sigma = (\theta, \beta) \in \mathbb{T}^{n-m} \times \mathbb{T}^{n-m}, \quad \text{span}\{X_1(f(\sigma)), \dots, X_m(f(\sigma))\} + df_\beta(\theta)(T_\theta \mathbb{T}^{n-m}) = T_{f(\sigma)}G \tag{11}$$

and

$$\forall \beta \in \mathbb{T}^{n-m}, \quad f(0, \beta) = e. \tag{12}$$

From now on, variables in $\mathbb{T}^{n-m}$ will be indexed starting from $m+1$; i.e., if $\theta \in \mathbb{T}^{n-m}$, then $\theta = (\theta_{m+1}, \dots, \theta_n)$. With the notation specified in the introduction, another way of writing relation (11) is

$$\forall \sigma = (\theta, \beta) \in \mathbb{T}^{n-m} \times \mathbb{T}^{n-m}, \quad \text{span}\left\{X_1(f(\sigma)), \dots, X_m(f(\sigma)), \frac{\partial f}{\partial \theta_{m+1}}(\sigma), \dots, \frac{\partial f}{\partial \theta_n}(\sigma)\right\} = T_{f(\sigma)}G. \tag{13}$$

It is clear that any generalized t.f. is a t.f. It is also quite simple to build a generalized t.f. $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m} \times \mathbb{T}^{n-m}; G)$ from a t.f. $\bar{f} \in \mathcal{C}^\infty(\mathbb{T}^{n-m}; G)$. For example, define

$$\forall (\theta, \beta) \in \mathbb{T}^{n-m} \times \mathbb{T}^{n-m}, \quad f(\theta, \beta) = (\bar{f}(\beta))^{-1} \bar{f}(\theta + \beta).$$

Let us now consider any generalized t.f. We let

$$\dot{\theta} = v, \qquad \dot{\beta} = w, \tag{14}$$

so that $\dot{\sigma} = u_\sigma = (v, w)$. With this notation, (9)—whose satisfaction yields $\dot{z} = Z(z)$—is equivalent to

$$\sum_{i=1}^m u_i X_i(f(\sigma)) - \frac{\partial f}{\partial \theta}(\sigma) v = \frac{\partial f}{\partial \beta}(\sigma) w - dl_z(g) P(g, t) - dr_g(z) Z(z). \tag{15}$$

From (11), this equation has a unique feedback solution $(u, v)(g, \sigma, t)$ for any function w. The v.f. Z is again chosen so as to make $z = e$ asymptotically stable. Now the

[3] Expressions of such functions are given in that paper—see also the next subsection.

objective is to determine w in order to make θ tend to zero. Indeed, this latter property implies, in view of (12), that f tends to e so that, from the fact that $z = f(\sigma)g^{-1}$ tends to e, the asymptotic convergence of g to e follows. Note that such a convergence cannot be obtained without the drift term P satisfying some extra conditions. For instance, if $P(e,t)$ is periodically different from zero, then it follows from (3) that e cannot be an equilibrium for system (2), whatever the control u. Moreover, under mild complementary regularity conditions upon the function P, convergence of $P(g,t)$ to zero when g tends to e and t tends to infinity is necessary to the convergence of the system's solutions to e.

The feedback law (u,v) defined by (15) ensures the convergence of z to e independently of w. Hence, the asymptotic behavior of $\theta(t)$ and $\beta(t)$, for the controlled system, is described by the *zero-dynamics* obtained by setting $z = e$ in (15), i.e.,

$$\sum_{i=1}^{m} u_i(g,\sigma,t)X_i(f(\sigma)) - \frac{\partial f}{\partial \theta}(\sigma)v(g,\sigma,t) = \frac{\partial f}{\partial \beta}(\sigma)w - P(f(\sigma),t). \tag{16}$$

From the initial assumption that the v.f. $X_1,\dots,X_m$ are independent, there exist v.f. $X_{m+1},\dots,X_n$ such that $\operatorname{span}\{X_1,\dots,X_n\} = \mathfrak{g}$. For any such set of v.f., there exist smooth functions $a_{i,j}$ and $b_{i,j}$ such that

(17)
$$\forall j = m+1,\dots,n\,, \quad \frac{\partial f}{\partial \theta_j}(\sigma) = \sum_{i=1}^{n} a_{i,j}(\sigma)X_i(f(\sigma))\,, \quad \frac{\partial f}{\partial \beta_j}(\sigma) = \sum_{i=1}^{n} b_{i,j}(\sigma)X_i(f(\sigma)).$$

With d_i $(i = m+1,\dots,n)$ denoting the one-forms defined by $\langle d_i, X_k\rangle = \delta_{i,k}$ (the Kronecker delta), the application of d_i to each side of (16) yields, since $\dot\theta = v$,

$$A(\sigma)\dot\theta = -B(\sigma)w + \sum_{i=m+1}^{n} \langle d_i(f(\sigma)), P(f(\sigma),t)\rangle e_i \tag{18}$$

with

$$A(\sigma) \triangleq (a_{i,j}(\sigma))_{i,j=m+1,\dots,n}, \qquad B(\sigma) \triangleq (b_{i,j}(\sigma))_{i,j=m+1,\dots,n}, \tag{19}$$

and e_i the $(i-m)$th unit vector in $\mathbb{R}^{n-m}$. Note that the transversality condition (11) is equivalent to the matrix $A(\sigma)$ being invertible for any σ.

Equation (18) is important because it explicitly relates the control w (the time-derivative of β) to the variation of θ. In particular, the simplification obtained when $P \equiv 0$, i.e.,

$$\dot\theta = -A^{-1}(\sigma)B(\sigma)w, \tag{20}$$

suggests some ways of choosing w to make $|\theta(t)|$ nonincreasing on the zero-dynamics. However, a difficulty arising at this stage, to ensure the convergence of $\theta(t)$ to zero, comes from the fact that $B(\sigma)$ tends to the null matrix when θ tends to zero, since $f(0,\beta) = e\ \forall\beta \Rightarrow \frac{\partial f}{\partial \beta}(0,\beta) = 0,\ \forall\beta$. This difficulty is itself related to the well-known impossibility of ensuring *exponential* stabilization of e by means of a *smooth* feedback [13, Thm. 3]. The matter would still be easily settled if $B(\sigma)$ were invertible everywhere except at $\theta = 0$. Unfortunately, this is not true in general, and further inspection of this matrix, in relation to the way the structure of f combines with the structure of the Lie algebra $\mathfrak{g}$, is required. Although we do not know whether a solution always exists, we were able to use the specific structure of the Lie algebra associated with the chained systems and derive a solution in this case. Prior to reporting it in the next section, we propose below a formulation of the problem which, whereas it is restricted to the zero-dynamics (20), simplifies the search for a solution for the complete system.

PROBLEM 1. *Given a neighborhood $\mathcal{U}_G(e)$ of e, determine a triplet (f, w, V) consisting of*

(i) *a generalized t.f. $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m} \times \mathbb{T}^{n-m}; \mathcal{U}_G(e))$,*

(ii) *a function $w \in \mathcal{C}^1(\mathcal{U}_{\mathbb{T}^{n-m}}(0); \mathbb{R}^{n-m})$,*

(iii) *a function $V \in \mathcal{C}^1(\mathcal{U}_{\mathbb{T}^{n-m}}(0); \mathbb{R})$ with bounded first-order partial derivatives*

such that

1. *$\forall \theta \in V^{-1}([0, V_{\max}))$, $h_{V_m}(|\theta|) \leq V(\theta) \leq h_{V_M}(|\theta|)$ with h_{V_m} and h_{V_M} two $\mathcal{K}$-functions, and $V_{\max} > 0$ a real number such that $V^{-1}([0, V_{\max})) \subset \mathcal{U}_{\mathbb{T}^{n-m}}(0)$;*
2. *the following proposition is true:*

(21)
$$\forall \beta \in \mathbb{T}^{n-m},\ \forall \theta \in V^{-1}([0, V_{\max})), \quad L_F(\pi^* V)(\sigma) \leq -\gamma V(\theta)^l, \quad \gamma, l > 0,$$

with π and F defined by

$$\forall \sigma = (\theta, \beta), \quad \pi(\sigma) = \theta,\ F(\sigma) = -A^{-1}(\sigma) B(\sigma) w(\theta). \tag{22}$$

Note that (21) clearly implies that $\theta = 0$ is locally asymptotically stable for the system (20). Note also that the inclusion $\mathcal{U}_{\mathbb{T}^{n-m}}(0) \subset \mathbb{T}^{n-m}$ has to be strict since (21) would otherwise contradict the known nonexistence of global asymptotic stabilizers on $\mathbb{T}^{n-m}$. Once the above problem is solved, it is not difficult to infer a solution to the problem of asymptotic stabilization of e for system $\mathcal{S}(g)$ when $P \equiv 0$. Such a solution is pointed out in the following proposition.

PROPOSITION 1. *Let Z denote a smooth v.f. which asymptotically stabilizes e for the system $\dot{z} = Z(z)$. Assume that Problem 1 is solved by a triplet $(f, w^\star, V)$, and consider for $\mathcal{S}(g, \sigma)$ the feedback control (u, v, w) with (u, v) defined by (15) and w defined by*

$$w(\theta) = k\left(\frac{1}{V_{\max} - V(\theta)}\right) w^\star(\theta), \tag{23}$$

with k denoting any $\mathcal{K}_\infty$-function. Assume also that the initial condition $\theta(0)$ is chosen in $V^{-1}([0, V_{\max}))$. Then

1. *whatever P, the above-defined feedback control asymptotically stabilizes the set $\{(f(\sigma), \sigma) : \sigma \in V^{-1}([0, V_{\max})) \times \mathbb{T}^{n-m}\}$ for $\mathcal{S}(g, \sigma)$;*
2. *if $P \equiv 0$, then this control asymptotically stabilizes the set $\{e\} \times \{0\} \times \mathbb{T}^{n-m}$ for $\mathcal{S}(g, \sigma)$;*
3. *if $P(g, t)$ tends to zero as $t \longrightarrow +\infty$, uniformly w.r.t. g in compact sets, then $(g, \theta)(t) \longrightarrow (e, 0)$ as $t \longrightarrow +\infty$.*

Note that when P, Z, and k are differentiable, the stabilizing feedback control (u, v, w) so obtained is also differentiable. When $P \equiv 0$ and $\theta(0) = 0$, this control asymptotically stabilizes e for $\mathcal{S}(g)$. However, as in the case of a time-periodic Lipschitz-continuous asymptotic stabilizer of $\mathcal{S}(g)$, the control's differentiability rules out the possibility of a uniform convergence rate as fast as exponential. On the other hand, while the frequency of a time-periodic stabilizer is constant, the time-derivatives of θ and β, which may be interpreted as self-adapting frequencies in the case of a stabilizer derived with the t.f. approach, asymptotically tend to zero.

2.3. A class of generalized t.f. In this section, we introduce a class of generalized t.f. which is instrumental in solving Problem 1 for the class of the chained systems. First, we need to recall the definition of a graded basis of $\mathfrak{g}$ (see [18]). This definition is similar to the one of a *basis adapted to the control filtration* [9, 25]; a complementary requirement is that some elements of the basis be expressed as Lie brackets of other elements of the basis.

DEFINITION 3. *Let* $X_1, \dots, X_m \in \mathfrak{g}$ *denote independent v.f. such that* $\mathrm{Lie}(X_1, \dots, X_m) = \mathfrak{g}$. *Let* $\mathfrak{u} = \mathrm{span}\{X_1, \dots, X_m\}$, *and define inductively, for* $k = 2, \dots, K$, $\mathfrak{u}^k = \mathfrak{u}^{k-1} + [\mathfrak{u}, \mathfrak{u}^{k-1}]$ *with* $K = \min\{k : \mathfrak{u}^k = \mathfrak{g}\}$. *A* graded basis *of* $\mathfrak{g}$ *associated with* $X_1, \dots, X_m$ *is an ordered basis* $\{X_1, \dots, X_n\}$ *of* $\mathfrak{g}$ *associated with two mappings* $\lambda, \rho : \{m+1, \dots, n\} \longrightarrow \{1, \dots, n\}$ *such that*

1. *for any* $k \in \{1, \dots, K\}$, $\mathfrak{u}^k = \mathrm{span}\{X_1, X_2, \dots, X_{\dim \mathfrak{u}^k}\}$;
2. *for* $k \geq 2$ *and* $\dim \mathfrak{u}^{k-1} < i \leq \dim \mathfrak{u}^k$, $X_i = [X_{\lambda(i)}, X_{\rho(i)}]$ *with* $X_{\lambda(i)} \in \mathfrak{u}^a$, $X_{\rho(i)} \in \mathfrak{u}^b$, *and* $a + b = k$.

With any graded basis of $\mathfrak{g}$, one can associate a *weight-vector* $(r_1, \dots, r_n)$ defined by

$$r_i = k \Longleftrightarrow X_i \in \mathfrak{u}^k \setminus \mathfrak{u}^{k-1} \Longleftrightarrow \dim \mathfrak{u}^{k-1} < i \leq \dim \mathfrak{u}^k.$$

Note that $1 = r_1 \leq r_2 \leq \cdots \leq r_n = K$, and, from Definition 3, $\forall i > m$, $r_i = r_{\lambda(i)} + r_{\rho(i)}$.

With $\{X_1, \dots, X_n\}$ any graded basis of $\mathfrak{g}$, let us define $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m} \times \mathbb{T}^{n-m}; G)$ by

$$\forall \sigma = (\theta, \beta) \in \mathbb{T}^{n-m} \times \mathbb{T}^{n-m}, \quad f(\sigma) = f_n(\sigma_n) \cdots f_{m+1}(\sigma_{m+1}), \tag{24}$$

with $f_j : \mathbb{T} \times \mathbb{T} \longrightarrow G$ defined by

$$\forall \sigma_j = (\theta_j, \beta_j), \quad f_j(\sigma_j) = \exp\left(\alpha_j(\sigma_j) X_j\right) \exp\left(\alpha_{j,\lambda}(\sigma_j) X_{\lambda(j)} + \alpha_{j,\rho}(\sigma_j) X_{\rho(j)}\right), \tag{25}$$

where

$$\alpha_{j,\lambda}(\sigma_j) = \varepsilon_j^{r_{\lambda(j)}} \left(\sin(\theta_j + \beta_j) - \sin \beta_j\right), \qquad \alpha_{j,\rho}(\sigma_j) = \varepsilon_j^{r_{\rho(j)}} \left(\cos(\theta_j + \beta_j) - \cos \beta_j\right),$$

$$\alpha_j(\sigma_j) = \frac{\varepsilon_j^{r_j}}{2} \sin \theta_j, \tag{26}$$

and the ε_j's are positive real numbers. This function obviously satisfies (12). As for the transversality condition (11), we have the following result.

PROPOSITION 2. *Let* $X_1, \dots, X_m$ *denote independent v.f. on a Lie group* G *of dimension* n. *Assume that* $\mathrm{Lie}(X_1, \dots, X_m) = \mathfrak{g}$. *Let* $f \in \mathcal{C}^\infty(\mathbb{T}^{n-m} \times \mathbb{T}^{n-m}; G)$ *be defined by* (24), (25), (26), *with* $\{X_1, \dots, X_n\}$ *a graded basis of* $\mathfrak{g}$. *Then there exist real positive numbers* $\eta_{m+1}, \dots, \eta_n$ *and* ε_0 *such that, for* $(\varepsilon_{m+1}, \dots, \varepsilon_n) = \varepsilon(\eta_{m+1}, \dots, \eta_n)$ *with* $\varepsilon \in (0, \varepsilon_0)$, f *satisfies* (11). *More precisely, the* η_k*'s can be defined recursively by choosing any* $\eta_{m+1} > 0$ *and for* $k = m+2, \dots, n$, *choosing* η_k *large enough w.r.t.* $\eta_{m+1}, \dots, \eta_{k-1}$.

3. Asymptotic stabilization of chained systems.

A solution to Problem 1 is provided in the case where $G = \mathbb{R}^n$, $m = 2$, and the control v.f. X_1, X_2 are defined by

$$X_1(x) = (1, 0, x_2, \dots, x_{n-1})', \quad X_2 = (0, 1, 0, \dots, 0)' \tag{27}$$

with $g = x = (x_1, \ldots, x_n)'$ and $e = 0$.

The v.f. X_1 and X_2 defined by (27) are left-invariant w.r.t. the group operation

$$(xy)_i = \begin{cases} x_i + y_i & \text{if } i = 1, 2, \\ x_i + y_i + \displaystyle\sum_{j=2}^{i-1} \frac{y_1^{i-j}}{(i-j)!} x_j & \text{otherwise,} \end{cases}$$

with $x, y \in \mathbb{R}^n$ (see [24], for instance). Furthermore, $\mathrm{Lie}(X_1, X_2) = \mathfrak{g}$, so that chained systems (with $P \equiv 0$) are controllable, and the v.f.

$$X_1, X_2, X_k \triangleq [X_1, X_{k-1}] \quad (k = 3, \ldots, n) \tag{28}$$

define a graded basis. The associated weight-vector r is given by

$$r_1 = r_2 = 1, \qquad r_k = k - 1 \quad (k = 3, \ldots, n). \tag{29}$$

Since the underlying Lie group G is $\mathbb{R}^n$, a simple example of v.f. which globally exponentially stabilizes the origin of $\dot{z} = Z(z)$ on $\mathbb{R}^n$ is defined by $Z(z) = Kz$, with K denoting any $n \times n$ Hurwitz-stable matrix. The main result is stated next.

THEOREM 1. *When $m = 2$ and the v.f. X_1, X_2 are given by* (27)*, there exist real positive numbers $\eta_{m+1}, \ldots, \eta_n$ such that a solution to Problem* 1 *is the triplet* (f, w, V) *consisting of*

1. *the candidate generalized t.f. defined by* (24)–(26) *with* $(\varepsilon_{m+1}, \ldots, \varepsilon_n) = \varepsilon(\eta_{m+1}, \ldots, \eta_n)$ *and $\varepsilon > 0$ chosen small enough so that f ranges in* $\mathcal{U}_{\mathbb{R}^n}(0)$,
2. *the function $w \in \mathcal{C}^1((-\pi, \pi)^{n-2}; \mathbb{R}^{n-2})$ defined by*

$$w_i(\theta_i) = \frac{1}{\eta_i^{i-2}} |\theta_i|^{(i-3)} \theta_i \qquad (i = 3, \ldots, n), \tag{30}$$

3. *the function $V \in \mathcal{C}^1((-\pi, \pi)^{n-2}; \mathbb{R})$ defined by*

$$V(\theta) \triangleq \sum_{i=3}^{n} \eta_i^{i-3/2} |\theta_i|^{n+2-i} \quad \text{with} \quad V_{\max} = \min_{i=3,\ldots,n} \{\eta_i^{i-3/2} \pi^{n+2-i}\}.$$

Remark 1. The proof of this theorem in section 5.3 involves a recursive procedure for the determination of the numbers $\eta_{m+1}, \ldots, \eta_n$, which is similar to the one indicated in Proposition 2.

Remark 2. The solution to Problem 1 given in Theorem 1 applies also to a unicycle-like mobile robot without having to transform its kinematic equations into the chain form—the only restriction is that ε must be smaller than some finite upper bound $\varepsilon_0 > 0$, whatever $\mathcal{U}_G(e)$, whereas, in the case of a chained system, $\varepsilon_0 = +\infty$. One only has to check that the proof of Theorem 1 works as well in this case with $n = 3$, $G = \mathbb{R}^2 \times S^1$, $g = (x, y, \alpha)'$, and the system's control v.f. defined by

$$X_1(g) = (\cos\alpha, \sin\alpha, 0)', \quad X_2(g) = (0, 0, 1)'. \tag{31}$$

These v.f. are left-invariant w.r.t. the group operation

$$g_1 g_2 = \begin{pmatrix} \begin{pmatrix} x_1 \\ y_1 \end{pmatrix} + R(\alpha_1) \begin{pmatrix} x_2 \\ y_2 \end{pmatrix} \\ \alpha_1 + \alpha_2 \end{pmatrix}$$

with $g_i = (x_i, y_i, \alpha_i)'$ and $R(\alpha_1)$ the rotation matrix of angle α_1. Also, $\text{Lie}(X_1, X_2) = \mathfrak{g}$ and $\{X_1, X_2, X_3 = [X_1, X_2]\}$ constitutes a graded basis of $\mathfrak{g}$ with weight vector $(r_1 = r_2 = 1, r_3 = 2)$.

Let us comment on the rate of convergence provided by a feedback control derived according to Proposition 1 and Theorem 1 when $P \equiv 0$. This will be the starting point of a more general discussion about what the t.f. approach can offer in comparison with other control design methods, its limitations and assets. Assuming that the v.f. Z used in the expression of (u, v) is chosen so as to stabilize the origin of $\dot{z} = Z(z)$ exponentially, the rate of convergence of $g(t)$ to e coincides with the slower rate of convergence of $\theta(t)$ to zero on the zero-dynamics. This latter rate is itself given by the rate of convergence of $V(\theta(t))$ to zero, and is thus related to the integer l in relation (21). From (47) in the proof of Theorem 1, we have $l = \frac{n+1}{2}$, and one deduces that $V(\theta(t))$ tends to zero as quickly as $t^{-\frac{2}{n-1}}$. In fact, a complementary analysis would show that $V(\theta(t))$ cannot tend to zero faster. Now, since $k_1|\theta|^{n-1} \leq V(\theta) \leq k_2|\theta|^2$ in the neighborhood of $\theta = 0$, this in turn implies that $|\theta(t)|$ may (and will usually) not tend to zero faster than $t^{-\frac{2}{(n-1)^2}}$. The same rate holds for the convergence of $|g(t)|$ towards e. This polynomial rate of convergence is similar to the one which can be obtained by applying a smooth time-periodic stabilizer to $\mathcal{S}(g)$. Therefore, one can conclude that, as far as asymptotic stabilization is concerned, no clear advantage results from designing a stabilizer with the t.f. approach. In the authors' opinion this conclusion is correct, but it conveys only a partial picture of the properties granted by the approach. Indeed, the primary feature of such a controller, which motivated the development of the t.f. approach in the first place, is the capacity of ensuring practical stabilization, with easily tunable arbitrary small ultimate bound of the state error, independently of the "perturbation" P acting on the system. As shown in [18], this allows, for example, the tracking of *any* trajectory in the state space (it does not have to be a solution to the system's equations) with arbitrarily good precision, in the sense that tracking errors are ultimately bounded by a prespecified (nonzero, but otherwise as small as desired) threshold. To our knowledge, no other controller proposed so far in the literature has this capacity. Our motivation for the present paper was to show that such a controller can also be endowed with the extra property of ensuring asymptotic point-stabilization when such a feature is desirable. This is achieved via the concept of a generalized t.f. depending upon two sets of variables whose time-derivatives are used as extra control inputs. Transversality is maintained with respect to the first set θ, while the second set β is used to enforce some type of "phase-tuning," which allows us to reduce the size of the t.f. when the perturbation P vanishes.

4. Simulation results. The control law proposed in the previous section has been tested by simulation on the four-dimensional (4d) chained system. The following parameters for the definition of the transverse function have been used: $\varepsilon = 0.2$, $\eta_3 = 1, \eta_4 = 8$. The v.f. $Z(z)$ in (15) has been chosen as $Z(z) = -0.3z$. Finally, the $\mathcal{K}_\infty$-function k in (23) has been defined by $k(s) = 10V_{\max}s$, with $V_{\max}$ as specified in Theorem 1. The initial condition for the simulation was $x(0) = (0, 0, 0, 10)'$, and $\sigma(0) = 0$. Figure 1 displays the state variables versus time. As discussed in the previous section, the convergence rate to zero is slow. For comparison, Figure 2 displays the same variables when no attempt is made to achieve convergence to zero, i.e., with $w = 0$ and $\beta = 0$ in the control law defined by (15). In this case $\theta(t)$ exponentially converges to some $\theta_{\text{lim}} \in \mathbb{T}^{n-m}$, and $x(t)$ exponentially converges to $f(\theta_{\text{lim}}, 0)$. Note that the solution to Problem 1 given by Theorem 1 is only one of its kind, and that much room is left for improving the proposed stabilization method.

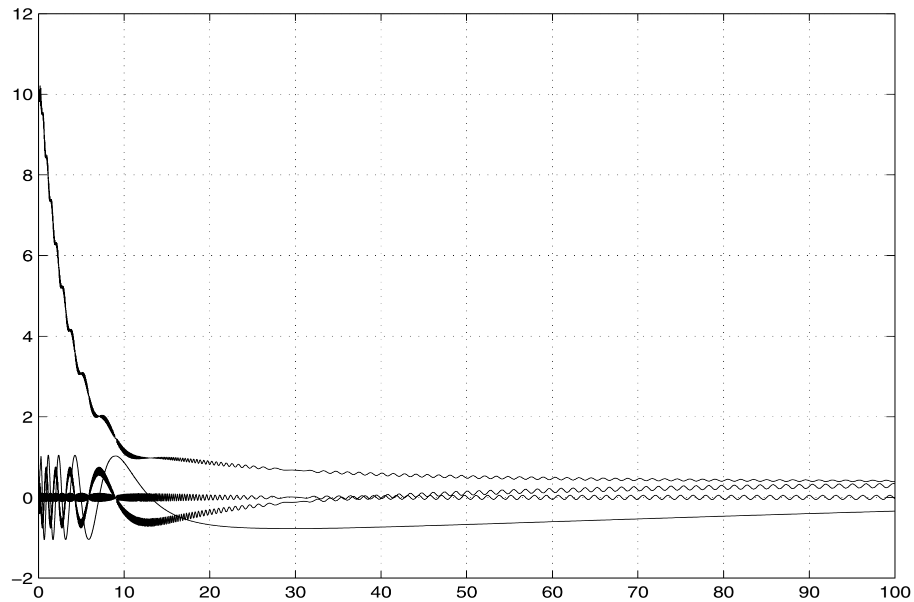

FIG. 1. *State variables for the* 4*d chained system, asymptotic stabilization.*

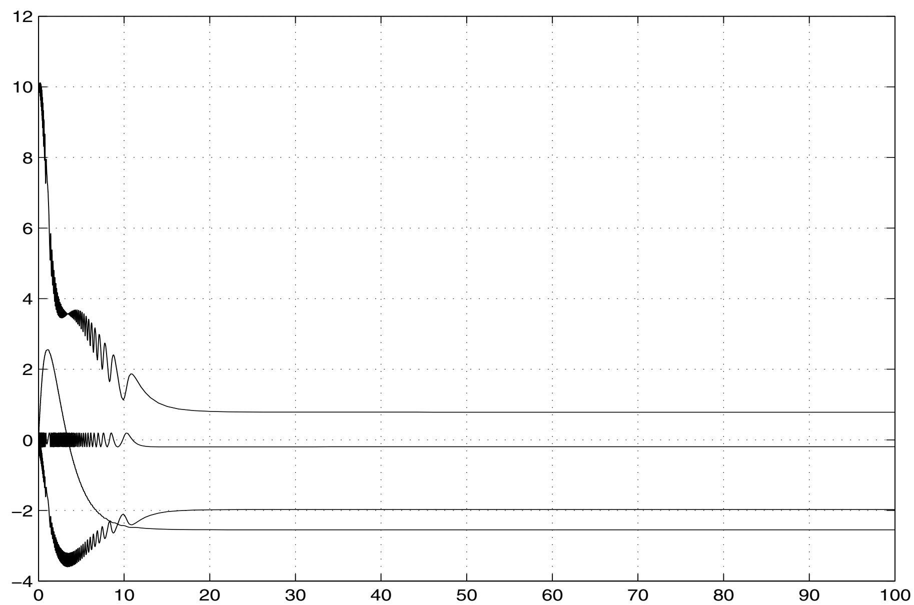

FIG. 2. *State variables for the* 4*d chained system, practical stabilization.*

5. Proofs.

5.1. Proof of Proposition 1. Let us first recall, as shown in section 2.2, that the control $(u, v)(g, \sigma, t)$ defined by (15) yields

$$\dot{z} = Z(z), \tag{32}$$

with $z = fg^{-1}$ and Z chosen so as to ensure the asymptotic stability of e for the above system. Now, applying the one-forms d_i ($i = m+1, \dots, n$) to each side of the equality (15) yields (compare with (18))

$$\dot{\theta} = -A^{-1}(\sigma)B(\sigma)w(\theta) + A^{-1}(\sigma)\sum_{i=m+1}^{n}\langle d_i(f), dl_z(g)P(g,t) + dr_g(z)Z(z)\rangle e_i, \tag{33}$$

where we have used the same notation as in section 2.2. Using the fact that $z = fg^{-1}$, we rewrite this equation as

$$\begin{aligned}\dot{\theta} = -A^{-1}(\sigma)B(\sigma)w(\theta) + A^{-1}(\sigma)\sum_{i=m+1}^{n}\langle d_i(f), dl_z(z^{-1}f)P(z^{-1}f,t)\\ + dr_{z^{-1}f}(z)Z(z)\rangle e_i.\end{aligned} \tag{34}$$

By using (23), this in turn implies that, along a solution of the controlled system,

$$\frac{d}{dt}V(\theta) = k\left(\frac{1}{V_{\max} - V(\theta)}\right) L_F(\pi^* V)(\sigma) + Q(g,\sigma,t),$$

with

$$F(\sigma) \triangleq -A^{-1}(\sigma)B(\sigma)w^\star(\theta),$$
$$Q(g,\sigma,t) \triangleq \frac{\partial V}{\partial \theta}(\theta)A^{-1}(\sigma)\sum_{i=m+1}^{n}\langle d_i(f), dl_z(z^{-1}f)P(z^{-1}f,t) + dr_{z^{-1}f}(z)Z(z)\rangle e_i.$$

Therefore, in view of (21),

$$\frac{d}{dt}V(\theta) \le -\xi(V(\theta)) + Q(g,\sigma,t), \qquad \xi(V) \triangleq k\left(\frac{1}{V_{\max} - V}\right)\gamma V^l. \tag{35}$$

Let us show that $\theta(t)$ cannot leave the set $V^{-1}([0, V_{\max}))$. We first remark that, on any time-interval $[0, T)$ such that $\theta(t)$ stays in this set, there exists a constant M_T, independent of the trajectory $\theta(.)$, such that $|Q(g(t), \sigma(t), t)| \le M_T$ because (i) by assumption, $\frac{\partial V}{\partial \theta}$ is bounded on $\mathcal{U}_{\mathbb{T}^{n-m}}(0) \supset V^{-1}([0, V_{\max}))$; (ii) z, and subsequently z^{-1}, are bounded due to the asymptotic stability of e for the system (32); (iii) P is continuous. Since, by (35), ξ is a bijective increasing function from $[0, V_{\max})$ to $[0, +\infty)$, we deduce from (35) that on any such interval $[0, T)$

$$V(\theta(t)) \le \max\{\xi^{-1}(M_T), V(\theta(0))\} < V_{\max}.$$

This implies that $V(\theta(t))$ cannot tend to $V_{\max}$ in finite time, so that $\theta(t)$ remains in the set $V^{-1}([0, V_{\max}))$. This in turn implies that the control law is well defined along any trajectory of the closed-loop system with initial conditions $(g(0), \theta(0), \beta(0))$ such that $z(0) = f(\sigma(0))g(0)^{-1}$ is in the stability domain of e for the system $\dot{z} = Z(z)$ and $\theta(0) \in V^{-1}([0, V_{\max}))$, and that such a trajectory is complete. Point 1 of Proposition 1 then follows directly from the asymptotic stability of $z = e$, as ensured by (32), and the invariance of the set $V^{-1}([0, V_{\max}))$ for the variable θ.

As for point 2, which assumes that $P \equiv 0$, it is sufficient to consider trajectories with initial conditions $(z(0), \theta(0))$ in a small neighborhood of the point $(e, 0)$. From the definition of Q, the asymptotic stability of $z = e$, combined with the invariance of the set $V^{-1}([0, V_{\max}))$ for the variable θ and the fact that V has bounded partial derivatives on this set, yields the existence of a $\mathcal{K}$-function h_z such that

$$\forall t \geq 0, \qquad |Q(g(t), \sigma(t), t)| \leq h_z(d_G(z(0), e)).$$

Therefore, in view of (35),

$$\forall t \geq 0, \qquad \frac{d}{dt} V(\theta(t)) \leq -k_V V(\theta(t))^l + h_z(d_G(z(0), e)),$$

with $k_V = k(\frac{1}{V_{\max}})\gamma(> 0)$. This in turn implies

$$\forall t \geq 0, \qquad V(\theta(t)) \leq \left(\frac{h_z(d_G(z(0), e))}{k_V}\right)^{1/l} + V(\theta(0)). \tag{36}$$

In view of (36) and property 1 in Problem 1,

$$\forall t \geq 0, \qquad |\theta(t)| \leq h_{V_m}^{-1}\left(\left(\frac{h_z(d_G(z(0), e))}{k_V}\right)^{1/l} + h_{V_M}(|\theta(0)|)\right).$$

This relation, combined with the asymptotic stability of $z = e$, implies the stability of the set $\{e\} \times \{0\} \times \mathbb{T}^{n-m}$ for $\mathcal{S}(g, \sigma)$. The convergence of the closed-loop trajectories to this set simply results from the convergence of $Q(g(t), \sigma(t), t)$ to zero when $z(t)$ tends to e, since $Z(z(t))$ then converges to zero. In view of (35), this yields the convergence of $V(\theta(t))$ to zero.

When it is assumed only that $P(g, t)$ tends to zero when t tends to infinity—uniformly w.r.t. g in compact sets—the term $Q(g(t), \sigma(t), t)$ in (35) still converges to zero, because the asymptotic stability of $z = e$ implies that $Z(z(t))$ converges to $Z(e) = 0$. Hence, the convergence of $V(\theta(t))$ to zero is still ensured, so that $\theta(t)$ tends to zero and $f(\sigma(t))$ tends to e (using the property (12) of a generalized t.f.). Therefore $(g, \theta)(t)$ tends to the point $(e, 0)$, as announced in point 3 of the proposition.

5.2. Proof of Proposition 2. The following notation is used in the forthcoming proofs. With v denoting a smooth function of the real variables x and y—possibly vector-valued—we write $v = \text{o}(x^k)$ (resp., $v = \text{O}(x^k)$) if $(|v(x, y)|/|x|^k) \longrightarrow 0$ as $|x| \longrightarrow 0$ (resp., if $(|v(x, y)|/|x|^k) \leq K < \infty$ in some neighborhood of $x = 0$) uniformly w.r.t. the y variable which takes values in a compact set. Finally, for indexed variables x_i with $i = k, \ldots, n$, we define the set of indexed vectors $\{\bar{x}_p\}_{p \in \{k, \ldots, n\}}$ by setting $\bar{x}_p = (x_k, \ldots, x_p)$.

Remark 3. Various results in the paper, starting with Proposition 2, refer to t.f. which depend on a vector of parameters $\varepsilon \in \mathbb{R}^{n-m}$, used as a means to monitor the "size" of the functions. Relations (24)–(26) define such a family of transverse functions. A member of this family could have been denoted as f_ε or $f(\varepsilon, .)$ in order to point out the functional dependence upon ε explicitly. However, for the sake of simplifying the (already cumbersome) notation used in the paper, we have chosen to systematically omit the argument ε when referring to t.f. It is nonetheless important to keep this dependence in mind when reading the forthcoming proofs. In particular, several functions associated with an arbitrary member of the family of t.f. defined by (24)–(26) will be introduced in Lemmas 1 and 2. Each of them is thus also a function of ε. For the sake of keeping the notation coherent throughout the paper, the index is again omitted when referring to such a function.

The proof of Proposition 2 consists of three steps summarized in the form of three lemmas, which are proved in the appendix.

LEMMA 1. *Assume that the assumptions of Proposition* 2 *are satisfied. Then, for each* $j \in \{m+1,\ldots,n\}$ *and* $i \in \{1,\ldots,n\}$*, there exist analytic functions* $v_{i,j}$ *and* $w_{i,j}$ *of* $\varepsilon_j \in \mathbb{R}$ *and* $\sigma_j \in \mathbb{T}\times\mathbb{T}$ *such that*

$$\frac{\partial f_j}{\partial \theta_j}(\sigma_j) = \sum_{i=1}^{n} v_{i,j}(\sigma_j) X_i(f_j(\sigma_j)), \quad \frac{\partial f_j}{\partial \beta_j}(\sigma_j) = \sum_{i=1}^{n} w_{i,j}(\sigma_j) X_i(f_j(\sigma_j)), \tag{37}$$

with

$$v_{i,j} = \begin{cases} \mathrm{O}(\varepsilon_j^{r_i}) & \forall i, \\ \mathrm{o}(\varepsilon_j^{r_i}) & \text{if } i<j \text{ and } r_i = r_j, \\ \dfrac{\varepsilon_j^{r_j}}{2} + \mathrm{o}(\varepsilon_j^{r_j}) & \text{if } i=j, \end{cases} \tag{38}$$

and

$$w_{i,j} = \begin{cases} \mathrm{O}(\varepsilon_j^{r_i})\mathrm{O}(\theta_j) & \forall i, \\ \varepsilon_j^{r_j}(1-\cos\theta_j) + \mathrm{o}(\varepsilon_j^{r_j})\mathrm{o}(\theta_j^2) & \text{if } i=j. \end{cases} \tag{39}$$

In the following lemma, $\mathrm{O}(\bar{\varepsilon}_m)$ formally appears when setting $j = m+1$ in $\mathrm{O}(\bar{\varepsilon}_{j-1})$, although ε_m has not been defined previously. The lemma's statement is nonetheless valid, provided that $\mathrm{O}(\bar{\varepsilon}_m)$ is identified with the null function.

LEMMA 2. *Assume that the assumptions of Proposition* 2 *are satisfied. Then, for each* $j \in \{m+1,\ldots,n\}$ *and* $i \in \{1,\ldots,n\}$*, there exist analytic functions* $a_{i,j}$ *and* $b_{i,j}$ *of* $\bar{\varepsilon}_j \in \mathbb{R}^{j-m}$ *and* $\sigma \in \mathbb{T}^{n-m}\times\mathbb{T}^{n-m}$ *such that*

$$\frac{\partial f}{\partial \theta_j}(\sigma) = \sum_{i=1}^{n} a_{i,j}(\sigma) X_i(f(\sigma)), \quad \frac{\partial f}{\partial \beta_j}(\sigma) = \sum_{i=1}^{n} b_{i,j}(\sigma) X_i(f(\sigma)), \tag{40}$$

with

$$a_{i,j} = \begin{cases} \mathrm{O}(\bar{\varepsilon}_j^{r_i}) & \forall i, \\ \mathrm{O}(\bar{\varepsilon}_{j-1})\mathrm{O}(\bar{\varepsilon}_j^{r_i-1}) + \mathrm{o}(\bar{\varepsilon}_j^{r_i}) & \text{if } i<j \text{ and } r_i = r_j, \\ \dfrac{\varepsilon_j^{r_j}}{2} + \mathrm{O}(\bar{\varepsilon}_{j-1})\mathrm{O}(\bar{\varepsilon}_j^{r_j-1}) + \mathrm{o}(\bar{\varepsilon}_j^{r_j}) & \text{if } i=j, \end{cases} \tag{41}$$

and

$$b_{i,j} = \begin{cases} \mathrm{O}(\bar{\varepsilon}_j^{r_i})\mathrm{O}(\bar{\theta}_j) & \forall i, \\ \varepsilon_j^{r_j}(1-\cos\theta_j) + \mathrm{O}(\bar{\varepsilon}_{j-1})\mathrm{O}(\bar{\varepsilon}_j^{r_j-1})\mathrm{O}(\theta_j)\mathrm{O}(\bar{\theta}_{j-1}) + \mathrm{o}(\bar{\varepsilon}_j^{r_j})\mathrm{o}(\bar{\theta}_j^2) & \text{if } i=j. \end{cases} \tag{42}$$

Note that, if all O and o terms in the above expressions were equal to zero, then the transversality property would simply follow from (40)–(41) and the fact that $\{X_1,\ldots,X_n\}$ is a basis of $\mathfrak{g}$. Although this is not the case, one can show that these terms can be neglected, provided that the ε_j's are adequately chosen.

LEMMA 3. *Assume that the assumptions of Proposition* 2 *are satisfied. Then there exist* $n-m$ *numbers* $\eta_{m+1},\ldots,\eta_n$ *and* $\varepsilon_0 > 0$ *such that choosing*

$$(\varepsilon_{m+1},\ldots,\varepsilon_n) = \varepsilon(\eta_{m+1},\ldots,\eta_n)$$

with $\varepsilon \in (0, \varepsilon_0)$ *yields*

$$\forall \sigma \in \mathbb{T}^{n-m} \times \mathbb{T}^{n-m}, \quad \operatorname{Det} A(\sigma) \neq 0 \quad \textit{with } A(\sigma) = (a_{i,j}(\sigma))_{i,j=m+1,\dots,n}. \tag{43}$$

5.3. Proof of Theorem 1. One easily verifies that for any positive real numbers $\eta_3, \dots, \eta_n$ the functions w and V satisfy (ii) and (iii) of Problem 1 with $\mathcal{U}_{\mathbb{T}^{n-m}}(0) = (-\pi, \pi)^{n-m}$. It is also clear that property 1 of Problem 1 is verified. We show below that, for an adequate choice of positive $\eta_3, \dots, \eta_n$, properties (i) and 2 are also satisfied. The proof relies on the following lemma, proved in the appendix, which points out complementary properties of the functions $a_{i,j}$ and $b_{i,j}$ in Lemma 2 in the case of the chained systems.

LEMMA 4. *In the case of chained systems, the functions* $a_{i,j}$ *and* $b_{i,j}$ $(i = 1, \dots, n,$ $j = 3, \dots, n)$ *are homogeneous polynomials of degree* r_i *in* $\varepsilon_3, \dots, \varepsilon_j$. *Furthermore,*

$$a_{i,j} = \mathrm{O}(\bar{\theta}_j^{r_i - r_j}), \quad b_{i,j} = \mathrm{O}(\bar{\theta}_j^{\max(1, r_i - r_j + 2)}). \tag{44}$$

Let

$$A_p(\sigma) \triangleq (a_{i,j}(\sigma))_{i,j=3,\dots,p}, \quad B_p(\sigma) \triangleq (b_{i,j}(\sigma))_{i,j=3,\dots,p}, \tag{45}$$

and note that $A_n = A$ and $B_n = B$, with A and B defined by (19).

PROPOSITION 3. *For any* $p = 3, \dots, n$, *there exists a set of positive numbers* $\{\eta_3, \dots, \eta_p\}$ *such that setting* $(\varepsilon_3, \dots, \varepsilon_p) = \varepsilon(\eta_3, \dots, \eta_p)$ *with* $\varepsilon > 0$ *implies that*

(i) *the matrix* $A_p(\sigma)$ *is invertible for any* σ, *and*

$$\forall i, j = 3, \dots, p, \quad (A_p^{-1}(\sigma))_{i,j} = \mathrm{O}(\bar{\theta}_p^{r_i - r_j}); \tag{46}$$

(ii) *the following is true:*

$$V_p(\bar{\theta}_p) < V_{p,max} \Longrightarrow L_{F_p} V_p(\sigma) \leq -\alpha_p |\bar{\theta}_p|^{n+1} \quad (\alpha_p > 0) \tag{47}$$

with

$$V_p(\bar{\theta}_p) \triangleq \sum_{i=3}^{p} \eta_i^{i-3/2} |\theta_i|^{n+2-i}, \qquad V_{p,max} = \min_{i=3,\dots,p} \{\eta_i^{i-3/2} \pi^{n+2-i}\} \tag{48}$$

and

$$F_p(\sigma) = -A_p^{-1}(\sigma) B_p(\sigma) \bar{w}_p(\bar{\theta}_p). \tag{49}$$

With $p = n$, property (i) of this proposition implies that the function f satisfies the transversality condition (11). Since (12) is trivially verified from (24), (25), (26), property (i) of Problem 1 follows. As for property 2 in Problem 1, it is true by (ii) in the above proposition. Note that, to be fully precise, in (47) one should write $L_{F_p}(\pi_p^* V_p)(\sigma)$ instead of $L_{F_p} V_p(\sigma)$, with $\pi_p(\sigma) = \bar{\theta}_p$ (compare with (21)). For the sake of simplifying the notation in the forthcoming proof, we have chosen to keep this small abuse of notation.

Proof of Proposition 3. We proceed by induction. For $p = 3$, it follows from (29) and Lemma 2 that

$$a_{3,3}(\sigma) = \frac{\varepsilon_3^2}{2} + \mathrm{o}(\varepsilon_3^2), \quad b_{3,3}(\sigma) = \varepsilon_3^2 (1 - \cos\theta_3) + \mathrm{o}(\varepsilon_3^2)\mathrm{o}({\theta_3}^2). \tag{50}$$

Lemma 4 implies that the o terms in the above equation are identically equal to zero, since $a_{3,3}$ and $b_{3,3}$ are homogeneous polynomials of degree $r_3 = 2$ in ε_3. Therefore, $a_{3,3}(\sigma) > 0$ for any $\varepsilon_3 > 0$, and the point (i) of the proposition is verified.

Take $\eta_3 = 1$. From (30), (49), (50), and the fact that the o terms in (50) are equal to zero,

$$F_3(\sigma) = -a_{3,3}^{-1}(\sigma)\left(\varepsilon_3^2(1-\cos\theta_3)\right)\theta_3 = -2(1-\cos\theta_3)\theta_3. \tag{51}$$

From (48), one easily checks that

$$L_{F_3}V_3(\sigma) = -2(n-1)(1-\cos\theta_3)|\theta_3|^{n-1}. \tag{52}$$

Since $V_3(\theta_3) = |\theta_3|^{n-1}$ and $V_{3,max} = \pi^{n-1}$, we deduce from (52) that

$$V_3(\theta_3) < \pi^{n-1} \Longrightarrow \qquad L_{F_3}V_3(\sigma) \leq -\alpha_3|\theta_3|^{n+1}$$

for some $\alpha_3 > 0$. Point (ii) of the proposition is thus verified with $\eta_3 = 1$ and $\varepsilon = \varepsilon_3 > 0$, and this concludes the proof of Proposition 3 for $p = 3$.

Let us now assume that points (i) and (ii) of the proposition hold true up to some $p < n$, with $\bar{\varepsilon}_p = \bar{\eta}_p$, and show that they are also true for $p+1$, with $\bar{\varepsilon}_{p+1} = \bar{\eta}_{p+1}$. This will in turn imply that they are true when $\bar{\varepsilon}_{p+1} = \varepsilon\bar{\eta}_{p+1}$ with $\varepsilon > 0$, thanks to the homogeneity properties of the $a_{i,j}$'s and $b_{i,j}$'s—see Lemma 4. Indeed, when $\bar{\eta}_{p+1}$ is multiplied by ε, then A_{p+1} and B_{p+1} are just premultiplied by the diagonal matrix $\mathrm{Diag}(\varepsilon^{r_3}, \ldots, \varepsilon^{r_{p+1}})$, thus leaving $F_{p+1}(\sigma)$ and the subsequent analysis unchanged.

From (45), A_{p+1} and B_{p+1} can be written as

$$A_{p+1} = \begin{pmatrix} A_p & a_{*,p+1} \\ a_{p+1,*} & a_{p+1,p+1} \end{pmatrix}, \quad B_{p+1} = \begin{pmatrix} B_p & b_{*,p+1} \\ b_{p+1,*} & b_{p+1,p+1} \end{pmatrix}, \tag{53}$$

with the star denoting the indexes from 1 to p, i.e., $a_{p+1,*} = (a_{p+1,1}, \ldots, a_{p+1,p})$ and $a_{*,p+1} = (a_{1,p+1}, \ldots, a_{p,p+1})'$. Let us recall (see, e.g., [29, Chap. 2]) that if A_{11} and A_{22} are square matrices with A_{11} nonsingular, the matrix

$$A = \begin{pmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{pmatrix}$$

is invertible if and only if the Schur complement $S \triangleq A_{22} - A_{21}A_{11}^{-1}A_{12}$ of A_{11} in A is invertible. Then

$$A^{-1} = \begin{pmatrix} A_{11}^{-1} + A_{11}^{-1}A_{12}S^{-1}A_{21}A_{11}^{-1} & -A_{11}^{-1}A_{12}S^{-1} \\ -S^{-1}A_{21}A_{11}^{-1} & S^{-1} \end{pmatrix}. \tag{54}$$

From (53), the Schur complement of A_p in A_{p+1} is $S = a_{p+1,p+1} - a_{p+1,*}A_p^{-1}a_{*,p+1}$, and, in view of (29) and Lemmas 2 and 4,

$$S = \frac{\varepsilon_{p+1}^p}{2} + q^{p-1}(\varepsilon_{p+1}), \tag{55}$$

with $q^{p-1}(\varepsilon_{p+1})$ a polynomial of degree $p-1$ in ε_{p+1}. (Note, from the domain definition of the functions $a_{i,j}$ in Lemma 2, that the term $a_{p+1,*}A_p^{-1}a_{*,p+1}$ depends on ε_{p+1} only through $a_{*,p+1}$ so that, by Lemma 4, it is a polynomial of degree $r_p = p-1$ in ε_{p+1}.) This implies that S, and thus A_{p+1}, are invertible for ε_{p+1} large enough. In order to prove (i), it remains to show that (46) holds true for $p+1$. Since (46) is true for p, for any $p = 3, \ldots, n-1$ and $\varepsilon_3, \ldots, \varepsilon_{p+1}$ such that A_p and A_{p+1} are invertible, let us use (54) to decompose A_{p+1}^{-1} as follows:

$$A_{p+1}^{-1} = \begin{pmatrix} A_p^{-1} & 0 \\ 0 & S^{-1} \end{pmatrix} + \begin{pmatrix} \Xi_{11} & \Xi_{12} \\ \Xi_{21} & 0 \end{pmatrix} \triangleq \begin{pmatrix} A_p^{-1} & 0 \\ 0 & S^{-1} \end{pmatrix} + \Xi, \tag{56}$$

with

$$\begin{aligned} \Xi_{11} &= A_p^{-1} a_{*,p+1} S^{-1} a_{p+1,*} A_p^{-1}, \\ \Xi_{12} &= -A_p^{-1} a_{*,p+1} S^{-1}, \\ \Xi_{21} &= -S^{-1} a_{p+1,*} A_p^{-1}. \end{aligned}$$

By Lemma 4, $a_{i,j}$ is a homogeneous polynomial in $\varepsilon_3, \ldots, \varepsilon_j$ of degree r_i and satisfies (44). Therefore, there exists a constant C such that

$$\forall \sigma, \quad |a_{i,j}(\sigma)| \le C |\bar{\varepsilon}_j|^{r_i} |\bar{\theta}_j|^{r_i - r_j}. \tag{57}$$

Then, from Lemma 4, relations (46), (55), and (57), and using the fact that neither A_p nor $a_{p+1,*}$ depend on ε_{p+1}, one infers that

$$\varepsilon_{p+1} \ge 1 \Longrightarrow \begin{cases} |\xi_{i,j}| \le C \varepsilon_{p+1}^{-1} |\bar{\theta}_{p+1}|^{r_i - r_j} & \text{for } i \le p, \\ |\xi_{i,j}| \le C \varepsilon_{p+1}^{-p} |\bar{\theta}_{p+1}|^{r_i - r_j} & \text{for } i = p+1, \end{cases} \tag{58}$$

with $\Xi = \{\xi_{i,j}\}_{i,j=3,\ldots,p+1}$. The fact that (46) holds true for $p+1$, provided that it is true up to p, directly follows from (56) and (58). Note that a relation similar to (57) holds for $b_{i,j}$, i.e.,

$$\forall \sigma, \quad |b_{i,j}(\sigma)| \le C |\bar{\varepsilon}_j|^{r_i} |\bar{\theta}_j|^{\max(1, r_i - r_j + 2)}. \tag{59}$$

This relation will be used later on.

Let us now examine the case of (ii). Throughout the rest of the proof, we assume that $\varepsilon_{p+1} \ge 1$. From (53) and (56),

$$A_{p+1}^{-1} B_{p+1} \bar{w}_{p+1} = \begin{pmatrix} A_p^{-1} B_p \bar{w}_p \\ S^{-1} b_{p+1,p+1} w_{p+1} \end{pmatrix} + D_2, \tag{60}$$

with

$$D_2 = \Xi B_{p+1} \bar{w}_{p+1} + \begin{pmatrix} A_p^{-1} b_{*,p+1} w_{p+1} \\ S^{-1} b_{p+1,*} \bar{w}_p \end{pmatrix}. \tag{61}$$

From (42) and Lemma 4, it is not difficult to deduce that

$$b_{p+1,p+1} = \varepsilon_{p+1}^p (1 - \cos\theta_{p+1}) + R_b, \tag{62}$$

with

$$|R_b| \le C \varepsilon_{p+1}^{p-1} |\bar{\theta}_{p+1}|^2 \tag{63}$$

for some constant C—recall that $\varepsilon_{p+1} \ge 1$. From the definition of $F_{p+1}(\sigma)$ in Proposition 3 and from relations (60) and (62),

(64)

$$F_{p+1}(\sigma) = \underbrace{\begin{pmatrix} F_p(\sigma) \\ -S^{-1} \varepsilon_{p+1}^p (1 - \cos\theta_{p+1}) w_{p+1}(\theta) \end{pmatrix}}_{D_0} - \underbrace{\begin{pmatrix} 0 \\ S^{-1} R_b w_{p+1}(\theta) \end{pmatrix}}_{D_1} - D_2.$$

We claim that the Lie derivative $L_{D_0}V_{p+1}$ of V_{p+1} along D_0 defined by (64) satisfies

$$L_{D_0}V_{p+1}(\sigma) \leq -\alpha_p|\bar{\theta}_p|^{n+1} - \alpha_1\varepsilon_{p+1}^{1/2}|\theta_{p+1}|^{n+1} \quad (\alpha_p, \alpha_1 > 0). \tag{65}$$

Indeed, by (48), $V_{p+1} = V_p + \varepsilon_{p+1}^{p-1/2}|\theta_{p+1}|^{n-p+1}$ (recall that $\bar{\varepsilon}_{p+1} = \bar{\eta}_{p+1}$), and it follows from (64) that

$$L_{D_0}V_{p+1}(\sigma) = L_{F_p}V_p(\sigma) - (n-p+1)S^{-1}\varepsilon_{p+1}^{2p-1/2}(1-\cos\theta_{p+1})w_{p+1}(\theta)\theta_{p+1}^{\{n-p\}}, \tag{66}$$

with the notation $x^{\{n\}} = |x|^{n-1}x$, also used in subsequent relations. From (47),

$$L_{F_p}V_p(\sigma) \leq -\alpha_p|\bar{\theta}_p|^{n+1}, \tag{67}$$

and, proceeding as for $a_{3,3}$, it is simple to verify, by using (30), (55), and the fact that $\varepsilon_{p+1} = \eta_{p+1} \geq 1$, that

$$-(n-p+1)S^{-1}\varepsilon_{p+1}^{2p-1/2}(1-\cos\theta_{p+1})w_{p+1}(\theta)\theta_{p+1}^{\{n-p\}} \leq -\alpha_1\varepsilon_{p+1}^{1/2}|\theta_{p+1}|^{n+1}. \tag{68}$$

Then, (65) follows from (66), (67), and (68).

From (30), (55), (63), and (64), it is straightforward to verify—by again using the condition $\varepsilon_{p+1} \geq 1$—that

$$|L_{D_1}V_{p+1}(\sigma)| \leq \alpha_2\varepsilon_{p+1}^{-1/2}|\bar{\theta}_p|^{n+1} + \alpha_2|\theta_{p+1}|^{n+1}. \tag{69}$$

Finally, we claim that

$$|L_{D_2}V_{p+1}(\sigma)| \leq \left(\frac{\alpha_p}{2} + \alpha_3\varepsilon_{p+1}^{-1/2}\right)|\bar{\theta}_p|^{n+1} + \alpha_4|\theta_{p+1}|^{n+1}. \tag{70}$$

Indeed, from (53), (56), and (61),

$$D_2 = \begin{pmatrix} (\Xi_{11}B_p + \Xi_{12}b_{p+1,*})\bar{w}_p + (\Xi_{11}b_{*,p+1} + \Xi_{12}b_{p+1,p+1})w_{p+1} + A_p^{-1}b_{*,p+1}w_{p+1} \\ \Xi_{21}B_p\bar{w}_p + \Xi_{21}b_{*,p+1}w_{p+1} + S^{-1}b_{p+1,*}\bar{w}_p \end{pmatrix}.$$

By using (29), (30), (46), (57), (58), and (59), it is tedious but not difficult to show that

$$\left\{ \begin{array}{rcl} |(D_2)_i| & \leq & C\varepsilon_{p+1}^{-1}|\bar{\theta}_{p+1}|^i + C|\bar{\theta}_{p+1}|^{i-p+1}|\theta_{p+1}|^{p-1} \quad \text{for } i = 3, \dots, p, \\ |(D_2)_{p+1}| & \leq & C\varepsilon_{p+1}^{-p}|\bar{\theta}_{p+1}|^{p+1}. \end{array} \right. \tag{71}$$

We infer from (48) and (71) that

$$|L_{D_2}V_p(\sigma)| \leq \alpha_5\varepsilon_{p+1}^{-1}|\bar{\theta}_{p+1}|^{n+1} + \alpha_6|\bar{\theta}_{p+1}|^{n-p+2}|\theta_{p+1}|^{p-1}. \tag{72}$$

By using Young's inequality, one shows that

$$\begin{aligned} \alpha_6|\bar{\theta}_{p+1}|^{n-p+2}|\theta_{p+1}|^{p-1} &\leq \frac{\alpha_p}{2}|\bar{\theta}_{p+1}|^{n+1} + \alpha_7|\theta_{p+1}|^{n+1} \\ &\leq \frac{\alpha_p}{2}|\bar{\theta}_p|^{n+1} + \alpha_8|\theta_{p+1}|^{n+1} \end{aligned} \tag{73}$$

for other constants α_7, α_8. We deduce from (72) and (73) that

$$|L_{D_2}V_p(\sigma)| \leq \left(\frac{\alpha_p}{2} + \alpha_9\varepsilon_{p+1}^{-1}\right)|\bar{\theta}_p|^{n+1} + \alpha_{10}|\theta_{p+1}|^{n+1}. \tag{74}$$

We also deduce from (71) that

$$|L_{D_2}(V_{p+1} - V_p)(\sigma)| \leq \alpha_{11}\varepsilon_{p+1}^{-1/2}|\bar{\theta}_{p+1}|^{n+1}, \tag{75}$$

and (70) then follows from (74), (75), and the condition $\varepsilon_{p+1} \geq 1$.

Let us now use (65), (69), and (70) to get an upper bound for $L_{F_{p+1}}V_{p+1}$. We obtain

$$\begin{aligned} L_{F_{p+1}}V_{p+1}(\sigma) &= (L_{D_0}V_{p+1} - L_{D_1}V_{p+1} - L_{D_2}V_{p+1})(\sigma) \\ &\leq -\left(\frac{\alpha_p}{2} - \alpha_{12}\varepsilon_{p+1}^{-1/2}\right)|\bar{\theta}_p|^{n+1} - \left(\alpha_1\varepsilon_{p+1}^{1/2} - \alpha_{13}\right)|\theta_{p+1}|^{n+1}. \end{aligned}$$

Since, by (65), α_p and α_1 are strictly positive, for ε_{p+1} large enough there exists $\alpha_{p+1} > 0$ such that

$$L_{F_{p+1}}V_{p+1}(\sigma) \leq -\alpha_{p+1}|\bar{\theta}_{p+1}|^{n+1}.$$

This concludes the proofs of Proposition 3 and Theorem 1.

Appendix: Proofs of Lemmas 1–4. The proofs of these lemmas rely on the following two properties.

Claim 1. *Let Y and Z denote two time-dependent left-invariant v.f. on G, and g, h solutions of $\dot{g} = Y(g,t)$ and $\dot{h} = Z(h,t)$, respectively. Then $\nu \overset{\Delta}{=} gh$ is a solution of $\dot{\nu} = \mathrm{Ad}(h^{-1})Y(\nu,t) + Z(\nu,t)$.*

This is simple to verify. Indeed one has

$$\begin{aligned} \tfrac{d}{dt}(gh) &= dl_g(h)\dot{h} + dr_h(g)\dot{g} \\ &= dl_g(h)Z(h,t) + dr_h(g)Y(g,t) \\ &= Z(gh,t) + dr_h(g)dl_g(e)Y(e,t), \end{aligned}$$

so that one has to show only that $dr_h(g)dl_g(e) = dl_{gh}(e)\mathrm{Ad}(h^{-1})$. For this purpose, it suffices to use the definition of the Ad operator, i.e.,

$$\begin{aligned} \mathrm{Ad}(h) &= d(l_h \circ r_{h^{-1}})(e) \\ &= dl_h(r_{h^{-1}}(e))dr_{h^{-1}}(e) \\ &= dl_h(h^{-1})dr_{h^{-1}}(e), \end{aligned}$$

and well-known relations obtained by differentiating both members of the identities $l_g \circ r_h = r_h \circ l_g$, $l_g \circ l_{g^{-1}} = i_d$, and $l_{gh} = l_g \circ l_h$. The desired result is then obtained as follows:

$$\begin{aligned} dr_h(g)dl_g(e) &= dl_g(h)dr_h(e) \\ &= dl_g(h)dl_h(e)dl_{h^{-1}}(h)dr_h(e) \\ &= dl_{gh}(e)dl_{h^{-1}}(h)dr_h(e) \\ &= dl_{gh}(e)\mathrm{Ad}(h^{-1}). \end{aligned}$$

Claim 2. *Let $\{X_1, \ldots, X_n\}$ denote a graded basis of the Lie algebra $\mathfrak{g}$ of a Lie group G. Let $\lambda, \rho, q \in \{1, \ldots, n\}$, $\alpha_p \in \mathbb{R}$, and $s \in \mathbb{N}$. Then, there exist analytic functions $g_1, \ldots, g_n$ such that, for any $\alpha_\lambda, \alpha_\rho \in \mathbb{R}$,*

$$\sum_{j=s}^{\infty} \frac{1}{j!}(\mathrm{ad}^j(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho), X_q) = \sum_{k=1}^{n} g_k(\alpha_\lambda, \alpha_\rho)X_k.$$

Furthermore, if $\alpha_\lambda, \alpha_\rho$ are analytic functions of x and y such that $\alpha_\lambda = \mathrm{O}(x^{r_\lambda})$ and $\alpha_\rho = \mathrm{O}(x^{r_\rho})$, then $g_k(\alpha_\lambda, \alpha_\rho)$ is an analytic function of x and y and $g_k(\alpha_\lambda, \alpha_\rho) = \mathrm{O}(x^{\max\{s \min\{r_\lambda, r_\rho\}, r_k - r_q\}})$.

The proof of this claim, which can be viewed as a direct adaptation of [28, sect. 2], is given in [18, App. A, Claim 2].

Proof of Lemma 1. In order to simplify the notation, let

$$X_\lambda = X_{\lambda(j)}\,,\ X_\rho = X_{\rho(j)}\,,\ \ \alpha_\lambda = \alpha_{j,\lambda}\,,\ \alpha_\rho = \alpha_{j,\rho}. \tag{76}$$

With this notation, it follows from (25) that $f_j(\sigma_j) = g_j(\sigma_j)h_j(\sigma_j)$, with $g_j(\sigma_j) = \exp(\alpha_j(\sigma_j)X_j)$ and $h_j(\sigma_j) = \exp(\alpha_\lambda(\sigma_j)X_\lambda + \alpha_\rho(\sigma_j)X_\rho)$.

Let $\sigma_j(.)$ denote an arbitrary smooth curve on $\mathbb{T}^2$. By using the fact that $\frac{d}{dt}\exp X(t) = \frac{d}{ds}\exp(X(t) + s\frac{d}{dt}X(t))_{|s=0}$ and that (see, e.g., [8, p. 105])

$$\frac{d}{ds}\exp(X + sY)_{|s=0} = (\phi(\mathrm{ad}X), Y)(\exp X)\,, \quad \phi(z) \triangleq \sum_{k=0}^{\infty} \frac{(-1)^k}{(k+1)!} z^k,$$

one infers that

$$\begin{aligned} \dot{h}_j &\triangleq \frac{d}{dt} h_j(\sigma_j(t)) \\ &= \frac{d}{ds}\exp\left(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho + s\frac{d}{dt}(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho)\right)_{|s=0} \\ &= (\phi(\mathrm{ad}(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho)), \dot{\alpha}_\lambda X_\lambda + \dot{\alpha}_\rho X_\rho)(h_j). \end{aligned}$$

One has also $\dot{g}_j = \dot{\alpha}_j X_j(g_j)$. The application of Claim 1 then yields

$$\dot{f}_j = (\phi(\mathrm{ad}(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho)), \dot{\alpha}_\lambda X_\lambda + \dot{\alpha}_\rho X_\rho)(f_j) + \dot{\alpha}_j \mathrm{Ad}(\exp(-\alpha_\lambda X_\lambda - \alpha_\rho X_\rho))X_j(f_j). \tag{77}$$

Let us now use the fact (see, e.g., [8, p. 128]) that $\mathrm{Ad}(\exp Y)Z = (\exp \mathrm{ad}Y, Z)$. From (77),

$$\begin{aligned} \dot{f}_j &= (\phi(\mathrm{ad}(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho)), \dot{\alpha}_\lambda X_\lambda + \dot{\alpha}_\rho X_\rho)(f_j) \\ &\quad + \dot{\alpha}_j(\exp \mathrm{ad}(-\alpha_\lambda X_\lambda - \alpha_\rho X_\rho), X_j)(f_j) \\ &= \dot{\alpha}_\lambda X_\lambda(f_j) + \dot{\alpha}_\rho X_\rho(f_j) - \frac{1}{2}[\alpha_\lambda X_\lambda + \alpha_\rho X_\rho, \dot{\alpha}_\lambda X_\lambda + \dot{\alpha}_\rho X_\rho](f_j) \\ &\quad + \sum_{k=2}^{\infty} \frac{(-1)^k}{(k+1)!}(\mathrm{ad}^k(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho), \dot{\alpha}_\lambda X_\lambda + \dot{\alpha}_\rho X_\rho)(f_j) \\ &\quad + \dot{\alpha}_j X_j(f_j) + \dot{\alpha}_j \sum_{k=1}^{\infty} \frac{1}{k!}(\mathrm{ad}^k(-\alpha_\lambda X_\lambda - \alpha_\rho X_\rho), X_j)(f_j). \end{aligned} \tag{78}$$

Since $X_j = [X_\lambda, X_\rho]$ by Definition 3, it comes from (78) that

$$\begin{aligned} \dot{f}_j &= \dot{\alpha}_\lambda X_\lambda(f_j) + \dot{\alpha}_\rho X_\rho(f_j) + \left(\dot{\alpha}_j - \frac{1}{2}(\alpha_\lambda \dot{\alpha}_\rho - \alpha_\rho \dot{\alpha}_\lambda)\right) X_j(f_j) \\ &\quad + (\alpha_\lambda \dot{\alpha}_\rho - \alpha_\rho \dot{\alpha}_\lambda) \sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{(k+2)!}(\mathrm{ad}^k(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho), X_j)(f_j) \\ &\quad + \dot{\alpha}_j \sum_{k=1}^{\infty} \frac{1}{k!}(\mathrm{ad}^k(-\alpha_\lambda X_\lambda - \alpha_\rho X_\rho), X_j)(f_j). \end{aligned} \tag{79}$$

It follows from (26) that

$$\alpha_\lambda, \frac{\partial\alpha_\lambda}{\partial\theta_j}, \frac{\partial\alpha_\lambda}{\partial\beta_j} = \mathrm{O}(\varepsilon_j^{r_\lambda}); \quad \alpha_\rho, \frac{\partial\alpha_\rho}{\partial\theta_j}, \frac{\partial\alpha_\rho}{\partial\beta_j} = \mathrm{O}(\varepsilon_j^{r_\rho}); \quad \alpha_j, \frac{\partial\alpha_j}{\partial\theta_j}, \frac{\partial\alpha_j}{\partial\beta_j} = \mathrm{O}(\varepsilon_j^{r_j}). \tag{80}$$

Therefore, by application of Claim 2 (with $x = \varepsilon_j$ and $y = \theta_j$),

$$\sum_{k=1}^{\infty} \frac{(-1)^{k+1}}{(k+2)!}(\mathrm{ad}^k(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho), X_j)(f_j) = \sum_{k=1}^{n} g_k(\alpha_\lambda, \alpha_\rho) X_k(f_j) \tag{81}$$

for some analytic functions $g_1, \dots, g_n$ which verify

$$g_k(\alpha_\lambda, \alpha_\rho) = \mathrm{O}(\varepsilon_j^{\max\{1, r_k - r_j\}}). \tag{82}$$

Similarly, by applying Claim 2 again,

$$\sum_{k=1}^{\infty} \frac{1}{k!}(\mathrm{ad}^k(-\alpha_\lambda X_\lambda - \alpha_\rho X_\rho), X_j)(f_j) = \sum_{k=1}^{n} h_k(\alpha_\lambda, \alpha_\rho) X_k, \tag{83}$$

with

$$h_k(\alpha_\lambda, \alpha_\rho) = \mathrm{O}(\varepsilon_j^{\max\{1, r_k - r_j\}}). \tag{84}$$

From (79), (81), and (83), we get

(85)

$$\begin{aligned} \dot f_j = & (\dot\alpha_\lambda + (\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda)\, g_\lambda(\alpha_\lambda, \alpha_\rho) + \dot\alpha_j\, h_\lambda(\alpha_\lambda, \alpha_\rho))\, X_\lambda(f_j) \\ & + (\dot\alpha_\rho + (\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda)\, g_\rho(\alpha_\lambda, \alpha_\rho) + \dot\alpha_j\, h_\rho(\alpha_\lambda, \alpha_\rho))\, X_\rho(f_j) \\ & + \left(\dot\alpha_j - \frac{1}{2}(\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda) + (\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda)\, g_j(\alpha_\lambda, \alpha_\rho) + \dot\alpha_j\, h_j(\alpha_\lambda, \alpha_\rho)\right) X_j(f_j) \\ & + \sum_{k \notin \{\lambda, \rho, j\}} ((\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda)\, g_k(\alpha_\lambda, \alpha_\rho) + \dot\alpha_j\, h_k(\alpha_\lambda, \alpha_\rho))\, X_k(f_j). \end{aligned}$$

Since this equality holds along any smooth curve $\sigma_j(.)$ on $\mathbb{T}^2$, it is also true when the time-derivatives are replaced by partial derivatives w.r.t. either θ_j or β_j.

Now, it follows from (26) that

$$d\alpha_j - \frac{1}{2}(\alpha_\lambda d\alpha_\rho - \alpha_\rho d\alpha_\lambda) = \frac{\varepsilon_j^{r_j}}{2} d\theta_j + \varepsilon_j^{r_j}(1 - \cos\theta_j) d\beta_j \tag{86}$$

and

$$\alpha_\lambda,\ \frac{\partial\alpha_\lambda}{\partial\beta_j},\ \alpha_\rho,\ \frac{\partial\alpha_\rho}{\partial\beta_j},\ \alpha_j = \mathrm{O}(\theta_j); \quad \frac{\partial\alpha_j}{\partial\beta_j} = 0. \tag{87}$$

Furthermore, if f is an analytic function of ε and θ such that $f = \mathrm{O}(|\varepsilon|^p)$ and $f = \mathrm{O}(|\theta|^q)$, then $f = \mathrm{O}(|\varepsilon|^p)\mathrm{O}(|\theta|^q)$. Therefore, by using (80), (82), (84), (86), and (87) in (85), it is tedious but simple to recover all relations in Lemma 1. (For the last relation of (39), note that g_j in (85) is an $\mathrm{O}(\theta_j)$ because it is a function of α_λ and α_ρ, which vanishes when $\alpha_\lambda = \alpha_\rho = 0$).

Proof of Lemma 2. From Claim 1 and relations (24) and (37),

(88)
$$\frac{\partial f}{\partial \theta_j} = \sum_{k=1}^{n} v_{k,j} \mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1}) X_k(f)\,, \quad \frac{\partial f}{\partial \beta_j} = \sum_{k=1}^{n} w_{k,j} \mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1}) X_k(f).$$

From the fact that $\mathrm{Ad}(g_1 g_2) = \mathrm{Ad}(g_1)\mathrm{Ad}(g_2)$ and (25),

(89)
$$\begin{aligned}\mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1}) &= \prod_{p=m+1}^{j-1} \mathrm{Ad}(f_p^{-1}) \\ &= \prod_{p=m+1}^{j-1} \mathrm{Ad}(\exp(-\alpha_{p,\lambda} X_{\lambda(p)} - \alpha_{p,\rho} X_{\rho(p)}))\mathrm{Ad}(\exp - \alpha_p X_p).\end{aligned}$$

By application of Claim 2, for any $p, q, k = 1, \ldots, n$ and $(\alpha_p, \alpha_q) \in \mathbb{R}^2$,

$$\mathrm{Ad}(\exp - \alpha_p X_p - \alpha_q X_q) X_k = X_k + \sum_{i=1}^{n} h_{p,q}^i(\alpha_p, \alpha_q) X_i$$

for some analytic functions $h_{p,q}^i$. Moreover, if $\alpha_p = \mathrm{O}(\varepsilon^{r_p})$ and $\alpha_q = \mathrm{O}(\varepsilon^{r_q})$ are analytic functions, then $h_{p,q}^i(\alpha_p, \alpha_q) = \mathrm{O}(\varepsilon^{\max(1, r_i - r_k)})$. This is used to infer from (80) and (89) that

(90)
$$\mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1}) X_k = X_k + \sum_i g_{j,k}^i X_i \quad \text{with } g_{j,k}^i = \mathrm{O}(\bar{\varepsilon}_{j-1}^{\max(1, r_i - r_k)}).$$

From (90),

$$\sum_{k=1}^{n} v_{k,j} \mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1}) X_k(f) = \sum_{i=1}^{n} \left(v_{i,j} + \sum_{k=1}^{n} v_{k,j} g_{j,k}^i \right) X_i(f),$$

and a similar expression holds when replacing v by w. Therefore, in view of (88), equation (40) holds with

(91)
$$a_{i,j} \triangleq v_{i,j} + \sum_{k=1}^{n} v_{k,j}\, g_{j,k}^i = A + B + C\,,$$
$$A = \sum_{r_k \le r_i} v_{k,j}\, g_{j,k}^i\,,\ B = v_{i,j}\,,\ C = \sum_{r_k > r_i} v_{k,j}\, g_{j,k}^i,$$

and

(92)
$$b_{i,j} \triangleq w_{i,j} + \sum_{k=1}^{n} w_{k,j}\, g_{j,k}^i = D + E + F\,,$$
$$D = \sum_{r_k \le r_i} w_{k,j}\, g_{j,k}^i\,,\ E = w_{i,j}\,,\ F = \sum_{r_k > r_i} w_{k,j}\, g_{j,k}^i.$$

Lemma 2 follows from this decomposition. Let us first show how (41) is obtained. From (38) and (90), A, B, and C in (91) are $\mathrm{O}(\bar{\varepsilon}_j^{r_i})$. This gives the first relation of (41).

For $i < j$ and $r_i = r_j$, A vanishes at $\bar{\varepsilon}_{j-1} = 0$ because of (90), and in view of (38), $B = \mathrm{o}(\bar{\varepsilon}_j^{r_i})$ and $C = \mathrm{o}(\bar{\varepsilon}_j^{r_i})$. This gives the second relation of (41).

For $i = j$, the only difference with the previous case comes from the B term, which, in view of (38), is equal to $\varepsilon_j^{r_j}/2 + \mathrm{o}(\varepsilon_j^{r_j})$. This gives the third relation of (41).

Let us now show how (42) is obtained. From (90),

$$g^i_{j,k} = \mathrm{O}(\bar{\theta}_{j-1}) \tag{93}$$

because, by (25) and (26),

$$\bar{\theta}_{j-1} = 0 \Longrightarrow f_{m+1} = \cdots = f_{j-1} = e \Longrightarrow \mathrm{Ad}(f_{m+1}^{-1} \cdots f_{j-1}^{-1})X_k = X_k.$$

The first relation of (42) is then simply obtained from (39), (90), (92), and (93).

For $i = j$, E in (92) accounts for the term $\varepsilon_j^{r_j}(1 - \cos\theta_j)$—up to higher order terms—in the second relation of (42), whereas D and F account for the remaining term by inspection of (39), (90), and (93).

Proof of Lemma 3. The lemma is a direct consequence of the following property, which can be proved by induction exactly as in the proof of [18, Lem. 3]:

$$\begin{array}{l} \forall k = m+1, \ldots, n, \ \exists \bar{\eta}_k \in \mathbb{R}^{k-m}, \ \exists \alpha_k > 0 \ : \\ \bar{\varepsilon}_k = \varepsilon_k \bar{\eta}_k \Rightarrow D_k \geq \alpha_k \varepsilon_k^{\bar{r}_k} + o(|\varepsilon_k|^{\bar{r}_k}), \end{array} \tag{94}$$

with $\bar{r}_k = r_{m+1} + \cdots + r_k$ and D_k the function defined by

$$D_k(\sigma) \overset{\Delta}{=} \mathrm{Det}(a_{i,j}(\sigma))_{i,j=m+1,\ldots,k}.$$

The first step consists of showing that (94) holds for $k = m+1$. From relation (41) in Lemma 2, $a_{m+1,m+1} = \frac{1}{2}\varepsilon_{m+1}^{r_{m+1}} + o(\varepsilon_{m+1}^{r_{m+1}})$. Since $D_{m+1} = a_{m+1,m+1}$ and $\bar{r}_{m+1} = r_{m+1}$, (94) is verified with $\eta_{m+1} = 1$ and $\alpha_{m+1} = \frac{1}{2}$. For the subsequent steps of the proof, the reader is referred to [18].

Proof of Lemma 4. Let us first show how Lemma 4—relation (44), in particular—is obtained from the following two claims.

CLAIM 3. *For any* i, j,

$$\left\{ \begin{array}{ll} v_{i,j} = \varepsilon_j^{r_i} \tilde{v}_{i,j} & \textit{with} \ \ \tilde{v}_{i,j} = \mathrm{O}(\theta_j^{r_i - r_j}), \\ w_{i,j} = \varepsilon_j^{r_i} \tilde{w}_{i,j} & \textit{with} \ \ \tilde{w}_{i,j} = \mathrm{O}(\theta_j^{\max(1, r_i - r_j + 2)}), \end{array} \right. \tag{95}$$

where the functions $\tilde{v}_{i,j}$ *and* $\tilde{w}_{i,j}$ *do not depend on the* ε_k*'s.*

CLAIM 4. *Each function* $g^i_{j,k}$ *in* (90) *is a polynomial in* $\varepsilon_3, \ldots, \varepsilon_{j-1}$ *homogeneous of degree* $r_i - r_k$. *Furthermore,*

$$g^i_{j,k} = \left\{ \begin{array}{ll} \mathrm{O}(\bar{\theta}_{j-1}^{r_i - r_k}) & \textit{if } r_j \leq r_k < r_i, \\ \mathrm{O}(\bar{\theta}_{j-1}^{r_i - r_j + 1}) & \textit{if } r_k < r_j < r_i. \end{array} \right. \tag{96}$$

From these claims, and from (91) and (92), it is straightforward to show that $a_{i,j}$ and $b_{i,j}$ are polynomials homogeneous of degree r_i in $\varepsilon_3, \ldots, \varepsilon_j$. Then, by (95), E and F in (92) are $\mathrm{O}(\bar{\theta}_j^{\max(1, r_i - r_j + 2)})$. As for the term D, it can be decomposed as

$$D = \sum_{r_k < r_i} w_{k,j} \, g^i_{j,k} + \sum_{r_k = r_i} w_{k,j} \, g^i_{j,k}. \tag{97}$$

From (95) and (96), the first sum in (97) is an $O(\theta_j^{\max(1,r_k-r_j+2)})O(\bar\theta_{j-1}^{r_i-r_k})$ if $r_j \le r_k < r_i$, and an $O(\theta_j)O(\bar\theta_{j-1}^{r_i-r_j+1})$ if $r_k < r_j < r_i$. Therefore, in both cases, it is an $O(\bar\theta_j^{\max(1,r_i-r_j+2)})$. As for the second sum in (97), it follows from (95) that it is an $O(\bar\theta_j^{\max(1,r_i-r_j+2)})$. This proves (44) for the term $b_{i,j}$. The proof for $a_{i,j}$ is similar.

It remains to prove Claims 3 and 4. In the case of a chained system, each element X_j of the graded basis, for $j = 3, \dots, n$, is equal to $[X_{\lambda(j)}, X_{\rho(j)}]$ with $\lambda(j) = 1$ and $\rho(j) = j - 1$. It is also a constant v.f. With the notation used in the proof of Lemma 1, these two facts imply that

$$(\mathrm{ad}(\alpha_\lambda X_\lambda + \alpha_\rho X_\rho), X_j) = \begin{cases} \alpha_\lambda X_{j+1} & \text{if } j < n, \\ 0 & \text{if } j = n. \end{cases}$$

Relation (79) in Lemma 1 then becomes

$$\begin{aligned} \dot f_j = {} & \dot\alpha_\lambda X_\lambda(f_j) + \dot\alpha_\rho X_\rho(f_j) + \left(\dot\alpha_j - \frac{1}{2}(\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda)\right) X_j(f_j) \\ & + (\alpha_\lambda\dot\alpha_\rho - \alpha_\rho\dot\alpha_\lambda) \sum_{k=1}^{n-j} \frac{(-1)^{k+1}}{(k+2)!} \alpha_\lambda^k X_{j+k}(f_j) + \dot\alpha_j \sum_{k=1}^{n-j} \frac{(-\alpha_\lambda)^k}{k!} X_{j+k}(f_j). \end{aligned}$$

Claim 3 is easily obtained by identifying this equality with (37), and by using (26) and (29).

Let us now prove Claim 4 by showing how relation (96) is obtained. The first step involves the evaluation of $\mathrm{Ad}(f_p^{-1})X_k$, for $p \in \{3, \dots, n-1\}$ and $k \in \{1, \dots, n\}$. We distinguish two cases.

Case 1. $k \ne 1$. From the definition (28) of $X_1, \dots, X_n$ and from (25),

$$\begin{aligned} \mathrm{Ad}(f_p^{-1})X_k &= \mathrm{Ad}(\exp(-\alpha_{p,\lambda}X_1 - \alpha_{p,\rho}X_{p-1}))\mathrm{Ad}(\exp -\alpha_p X_p)X_k \\ &= \mathrm{Ad}(\exp(-\alpha_{p,\lambda}X_1 - \alpha_{p,\rho}X_{p-1}))X_k \\ &= X_k + \sum_{j=1}^{n-k} \frac{(-\alpha_{p,\lambda})^j}{j!} X_{k+j} \\ &= X_k + \sum_{j=1}^{n-k} \varepsilon_p^j h_{p,k}^{k+j} X_{k+j} \quad \text{with } h_{p,k}^{k+j} = O(\theta_p^j), \end{aligned} \tag{98}$$

where the last equality comes from (26) and (29), and $h_{p,k}^{k+j}$ is a function which does not depend on ε_p. From (29), $r_{k+j} = r_k + j$ for $k > 1$ and $0 \le j \le n-k$. Therefore, from (26) and (98),

$$\mathrm{Ad}(f_p^{-1})X_k = X_k + \sum_{i>k} \varepsilon_p^{r_i-r_k} h_{p,k}^i X_i \quad \text{with } h_{p,k}^i = O(\theta_p^{r_i-r_k}). \tag{99}$$

By applying (99) recursively, it follows that, for any $k \ne 1$,

$$\mathrm{Ad}(f_3^{-1} \cdots f_{j-1}^{-1})X_k = X_k + \sum_{i>k} g_{j,k}^i X_i \quad \text{with } g_{j,k}^i = O(\bar\theta_{j-1}^{r_i-r_k}), \tag{100}$$

where each $g_{j,k}^i$ is a polynomial homogeneous of degree $r_i - r_k$ in $\varepsilon_3, \dots, \varepsilon_{j-1}$. This yields (96) for $r_j \le r_k \le r_i$, and also for $r_k < r_j < r_i$ (and $k \ne 1$) after noticing that, in this case, $r_i - r_k \ge r_i - r_j + 1$.

Case 2. $k=1$. We have

(101)

$$\begin{aligned}\mathrm{Ad}(f_p^{-1})X_1 &= \mathrm{Ad}(\exp(-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1}))\mathrm{Ad}(\exp -\alpha_pX_p)X_1\\ &= \mathrm{Ad}(\exp(-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1}))(X_1+\alpha_pX_{p+1})\\ &= X_1+\alpha_pX_{p+1}+[-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1},X_1+\alpha_pX_{p+1}]\\ &\quad+\sum_{k=2}^{\infty}\frac{1}{k!}\Big(\mathrm{ad}^{k-1}(-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1}),\\ &\qquad\qquad [-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1},X_1+\alpha_pX_{p+1}]\Big)\\ &= X_1+\alpha_pX_{p+1}-\alpha_{p,\lambda}\alpha_pX_{p+2}+\alpha_{p,\rho}X_p\\ &\quad+\sum_{k=2}^{\infty}\frac{1}{k!}(\mathrm{ad}^{k-1}(-\alpha_{p,\lambda}X_1-\alpha_{p,\rho}X_{p-1}),-\alpha_{p,\lambda}\alpha_pX_{p+2}+\alpha_{p,\rho}X_p)\\ &= X_1+\alpha_pX_{p+1}-\alpha_{p,\lambda}\alpha_pX_{p+2}+\alpha_{p,\rho}X_p\\ &\quad-\alpha_{p,\lambda}\alpha_p\sum_{k=2}^{\infty}\frac{(-\alpha_{p,\lambda})^{k-1}}{k!}X_{p+2+k-1}+\alpha_{p,\rho}\sum_{k=2}^{\infty}\frac{(-\alpha_{p,\lambda})^{k-1}}{k!}X_{p+k-1}.\end{aligned}$$

It follows from (26) and (101) that

$$\mathrm{Ad}(f_p^{-1})X_1 = X_1+\sum_{i>1}\varepsilon_p^{r_i-r_1}h_p^iX_i \quad \text{with } h_p^i=\mathrm{O}(\theta_p^{r_i-r_p}), \tag{102}$$

and h_p^i does not depend on ε_p. By applying (102) recursively and by using (100), it follows that

$$\mathrm{Ad}(f_3^{-1}\cdots f_{j-1}^{-1})X_1 = X_1+\sum_{i>1}g_{j,1}^iX_i \quad \text{with } g_{j,1}^i=\mathrm{O}(\bar\theta_{j-1}^{r_i-r_{j-1}})=\mathrm{O}(\bar\theta_{j-1}^{r_i-r_j+1}),$$

where each $g_{j,1}^i$ is polynomial homogeneous of degree r_i-r_1 in $\varepsilon_3,\dots,\varepsilon_{j-1}$. This concludes the proof of Claim 4.

Acknowledgments. The authors wish to thank the anonymous reviewers who read the first version of the paper in depth and provided numerous relevant remarks and suggestions to improve the presentation.

REFERENCES

[1] A. Astolfi, *Discontinuous control of nonholonomic systems*, Systems Control Lett., 27 (1996), pp. 37–45.

[2] M. K. Bennani and P. Rouchon, *Robust stabilization of flat and chained systems*, in Proceedings of the European Control Conference (ECC), Rome, 1995, pp. 2642–2646.

[3] A. Bloch and S. Drakunov, *Stabilization and tracking in the nonholonomic integrator via sliding modes*, Systems Control Lett., 29 (1996), pp. 91–99.

[4] A. Bloch, M. Reyhanoglu, and N. McClamroch, *Control and stabilization of nonholonomic dynamic systems*, IEEE Trans. Automat. Control, 37 (1992), pp. 1746–1757.

[5] R. Brockett, *Asymptotic stability and feedback stabilization*, in Differential Geometric Control Theory, R. W. Brockett, R. S. Millman, and H. Sussmann, eds., Birkhäuser Boston, Cambridge, MA, 1983, pp. 181–191.

[6] C. Canudas de Wit and O. J. Sørdalen, *Exponential stabilization of mobile robots with nonholonomic constraints*, IEEE Trans. Automat. Control, 37 (1992), pp. 1791–1797.

[7] J.-M. Coron, *Global asymptotic stabilization for controllable systems without drift*, Math. Control Signals Systems, 5 (1992), pp. 295–312.
[8] S. Helgason, *Differential Geometry, Lie Groups, and Symmetric Spaces*, Academic Press, New York, 1978.
[9] H. Hermes, *Nilpotent and high-order approximations of vector field systems*, SIAM Rev., 33 (1991), pp. 238–264.
[10] M. Kawski and H. Sussmann, *Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory*, in Operators, Systems, and Linear Algebra, D. Prätzel-Wolters, U. Helmke, and E. Zerz, eds., Teubner, Leipzig, 1997, pp. 111–128.
[11] W. Lin and P. Radom, *Recursive design of discontinuous controllers for uncertain driftless systems in power chained form*, IEEE Trans. Automat. Control, 45 (2000), pp. 1886–1892.
[12] D. Lizárraga, *Obstructions to the existence of universal stabilizers for smooth control systems*, Math. Control Signals Systems, 16 (2004), pp. 255–277.
[13] R. M'Closkey and R. Murray, *Exponential stabilization of driftless nonlinear control systems using homogeneous feedback*, IEEE Trans. Automat. Control, 42 (1997), pp. 614–628.
[14] P. Morin, J.-B. Pomet, and C. Samson, *Design of homogeneous time-varying stabilizing control laws for driftless controllable systems via oscillatory approximation of Lie brackets in closed loop*, SIAM J. Control. Optim., 38 (1999), pp. 22–49.
[15] P. Morin and C. Samson, *Exponential stabilization of nonlinear driftless systems with robustness to unmodeled dynamics*, ESAIM Control. Optim. Calc. Var. 4 (1999), pp. 1–36.
[16] P. Morin and C. Samson, *Control of non-linear chained systems. from the Routh-Hurwitz stability criterion to time-varying exponential stabilizers*, IEEE Trans. Automat. Control, 45 (2000), pp. 141–146.
[17] P. Morin and C. Samson, *A characterization of the Lie algebra rank condition by transverse periodic functions*, SIAM J. Control. Optim., 40 (2001), pp. 1227–1249.
[18] P. Morin and C. Samson, *Practical stabilization of driftless systems on Lie groups: The transverse function approach*, IEEE Trans. Automat. Control, 48 (2003), pp. 1496–1508.
[19] J.-B. Pomet and C. Samson, *Exponential stabilization of nonholonomic systems in power form*, in Proceedings of the IFAC Symposium on Robust Control Design, Rio de Janeiro, Brazil, 1994, pp. 447–452.
[20] J.-B. Pomet, *Explicit design of time-varying stabilizing control laws for a class of controllable systems without drift*, Systems Control Lett., 18 (1992), pp. 467–473.
[21] C. Samson, *Velocity and torque feedback control of a nonholonomic cart*, Int. Workshop in Adaptive and Nonlinear Control: Issues in Robotics (1990); also in Lecture Notes in Control and Inform. Sci. 162, Springer-Verlag, Berlin, 1991, pp. 125–151.
[22] C. Samson, *Control of chained systems. Application to path following and time-varying point-stabilization*, IEEE Trans. Automat. Control, 40 (1995), pp. 64–77.
[23] O. J. Sørdalen and O. Egeland, *Exponential stabilization of nonholonomic chained systems*, IEEE Trans. Automat. Control, 40 (1995), pp. 35–49.
[24] H. Struemper and P. Krishnaprasad, *Tracking and Stabilization for Control Systems on Matrix Lie Groups*, Tech. report 97-34, Institute for Systems Research (ISR), University of Maryland, College Park, MD, 1997.
[25] H. Struemper, *Nilpotent approximation and nilpotentization for under-actuated systems on matrix Lie groups*, in Proceedings of the IEEE Conference on Decision and Control (CDC), Phoenix, AZ, 1998, pp. 4188–4193.
[26] H. J. Sussmann, *Lie brackets and local controllability: A sufficient condition for scalar-input systems*, SIAM J. Control Optim., 21 (1983), pp. 686–713.
[27] A. Teel, R. Murray, and G. Walsh, *Nonholonomic control systems: From steering to stabilization with sinusoids*, Internat. J. Control, 62 (1995), pp. 849–870.
[28] J. Wei and E. Norman, *On global representations of the solutions of linear differential equations as a product of exponentials*, Proc. Amer. Math. Soc., 15 (1964), pp. 327–334.
[29] K. Zhou and J. Doyle, *Essentials of Robust Control*, Prentice–Hall, Englewood Cliffs, NJ, 1998.