SERIES EXPANSIONS FOR THE EVOLUTION OF MECHANICAL CONTROL SYSTEMS*

FRANCESCO BULLO†

Abstract. This paper presents a series expansion that describes the evolution of a mechanical system starting at rest and subject to a time-varying external force. Mechanical systems are presented as second-order systems on a configuration manifold via the notion of affine connections. The series expansion is derived by exploiting the homogeneity property of mechanical systems and the variations of constant formula. A convergence analysis is obtained using some analytic functions and combinatorial analysys results. This expansion provides a rigorous mean of analyzing locomotion gaits in robotics, and lays the foundation for the design of motion control algorithms for a large class of underactuated mechanical systems.

Key words. series expansions, control of mechanical systems, nonlinear controllability

AMS subject classifications. 30B99, 70Q05, 93B05, 93B29

1. Introduction.

The general purpose of this work is to develop innovative and powerful control and analysis method for underactuated mechanical control systems. This paper introduces a series expansion that characterize the evolution of a mechanical system starting at rest and subject to an open loop time-varying force. This tool should prove useful in the study of robotic locomotion and in the design of motion control algorithms.

1.1. Series expansions and their control applications.

Original works on perturbation methods and series expansions in mechanics go back to Poincarè and Lagrange. Magnus [34] describes the evolution of systems on a Lie group. Chen [15], Fliess [18], and Sussmann [43] develop a general framework to describe the evolution of a nonlinear system via the so-called Chen-Fliess series and its factorization. Related work on the "chronological calculus" formalism was developed by Agračhev and Gamkrelidze [2].

Within the context of modern geometric control theory, series expansions play a key role in the study of nonlinear controllability. Small-time local controllability was studied for example by Sussmann [42, 44], Agračhev and Gamkrelidze [3], and Kawski [23, 25]. Controllability along trajectories is investigated by Bianchini and Stefani in [8]. Finally, the work by Lewis and Murray [33] on configuration controllability for mechanical control systems is very related to this work.

Motion planning problems provide a second important use of series expansions. A rich literature is available on the motion planning problem for kinematic systems, that is systems without drift. Numerous approaches include algorithms for chained systems by Murray and Sastry [37], for systems on Lie groups by Leonard and Krishnaprasad [31] and Kolmanovsky and McClamroch [27], and the very general solution proposed by Lafferriere and Sussmann [29]. These works rely on the following observation: an explicit expression for the "input history to final displacement" map

*Submitted, December 4, 1999. This research was supported by the Campus Research Board at University of Illinois at Urbana-Champaign. A preliminary short version of this work appeared in the Proceedings of the 1999 IFAC World Congress [11].

†Coordinated Science Laboratory and General Engineering Department, University of Illinois at Urbana-Champaign, 1308 W. Main St, Urbana, IL 61801, USA, Telephone: (217) 333-0656, Fax: (217) 244-1653, Email: bullo@uiuc.edu, Url: motion.csl.uiuc.edu

simplifies dramatically the two-point boundary value problem that defines the motion planning task. In other words, whenever an explicit expression (provided by a series expansion) for the evolution of the control system is available, the two-point boundary value problem is reduced to a low dimensional nonlinear program. Accordingly, motion control algorithms are designed by inverting this "input history to final displacement" map.

Finally, series expansions and the techniques developed in this paper have potential relevance in several areas including averaging and vibrational stabilization [6, 10], high order variations for use in optimal control [26], digital multirate sampling of nonlinear systems [19] and model reduction [20].

Series expansions that specifically exploit the structure of mechanical systems have so far not been computed. However, some preliminary progress in this direction has been obtained by Bullo, Leonard and Lewis [12, 14] via a perturbation analysis. Under the assumption of small amplitude forcing, the authors compute the initial terms of a Taylor series describing the forced evolution. The results are then found to be in agreement with the controllability analysis in [33]. A different but related research direction has focused on open loop vibrational control and the recent progress we described in [10] is related to this paper.

1.2. Summary of results. The main contribution of the paper is a series that describes the evolution of a forced mechanical system starting from rest. Mechanical systems are characterized as second-order systems on a configuration manifold using the theory of affine connections. By exploiting the problem's structure, the system's evolution is described as a flow on the configuration space (n-dimensional) instead of a series on the full phase space ($2n$-dimensional).

The treatment relies on some differential geometric tools to describe the homogeneity properties of nonlinear mechanical systems and the variations of constants formula; see [2] and [24]. The homogeneous structure of nonlinear mechanical systems leads to a recursive procedure to compute the forced solution to a mechanical system. The terms in the series are computed recursively via time integrals and certain Lie brackets called symmetric products [33].

The series is guaranteed to convergence in a strong sense for small amplitude inputs and bounded final time. The convergence analysis is sophisticated and relies on various concepts from complex and combinatorial analysis. Following the analysis by Agračhev and Gamkrelidze in [2, Proposition 2.1], a bound is computed for every term of the series so that a notion of order is established. However, as opposed to [2], only a recursive expression for the series terms is available and this much complicates the treatment. The key idea is to obtain a recursive bound not only on the terms of the expansion but also on their partial derivative.

The series expansion can be computed in simplified fashion in two settings. For simple Hamiltonian systems with integrable forces, the main theorem can be interpreted as a statement on gradient and Hamilton flows: the flow of a Hamiltonian system forced from rest can be written as a (time-varying) gradient flow. For invariant systems on groups, the series can be computed via algebraic manipulations (no differentiations). In other words, the computations are performed on the corresponding Lie algebra and the theorem reduces to a statement on the flow of polynomial control systems. These results agree and supersede the preliminary results in in [12, 14].

Finally, some numerical simulations of a three degree of freedom robotic manipulator are performed. Truncating the series expansion at increasingly higher order, various approximations are obtained and their accuracy is illustrated via some numerical data.

1.3. Organization. The paper is organized as follows. In Section 2 we present the model and the homogeneity properties of a large class of mechanical control systems. Most ideas are common in the literature, some are not. In Section 3 we present the main result of the paper, that is, a convergent series describing the evolution of a forced mechanical system. Section 4 contains some applications and extensions, including the simple Hamiltonian and the invariant system settings, as well as some simulations. We present our conclusions in Section 5.

2. Some Geometric and Analytical Properties of Mechanical Systems.

We present a geometric definition of mechanical control systems, study their homogeneous properties, and provide bounds using analytic function theory.

2.1. Natural objects on manifolds. We review some basic definitions to fix some notation, see [1]. All the objects we consider are smooth in the sense of analytic. Let Q be a finite dimensional, Hausdorff, second countable manifold, q be a point on it, v_q be a point on TQ, $I \subset \mathbb{R}$ be a real interval and $\gamma : I \to Q$ be a curve on Q. We let 0_q denote the zero velocity tangent vector at on the tangent space T_qQ. Let $\pi : TQ \to Q$ denote the usual projection on the tangent bundle, that is, $\pi(v_q) = q$. On the manifold Q, we will define scalar functions $q \mapsto f(q) \in \mathbb{R}$ and vector fields $q \mapsto X(q) \in T_qQ$. Lie derivatives of functions and Lie brackets of vector fields are denoted by

$$\mathscr{L}_X f, \qquad \text{and} \qquad \mathscr{L}_X Y = [X, Y].$$

2.2. Variation of constants formula in geometric terms. This section presents a quick review of the variation of constants formula within the chronological calculus formalism introduced in [2], see also [39]. Given a vector field Y and a diffeomorphism ϕ, the *pull-back of Y along ϕ*, denoted ϕ^*Y, is a vector field defined by

$$(\phi^*Y)(q) \triangleq T_q\phi^{-1} \circ Y \circ \phi,$$

where $T_q\phi^{-1}$ is the tangent map to ϕ^{-1}, see [1]. In a system of local coordinate $(q^1, \dots, q^n)$, a vector field is written as $Y(q) = Y^i(q)\ \partial/\partial q^i$, and the pull-back of Y along ϕ is

$$(\phi^*Y)^i(q) = \frac{\partial(\phi^{-1})^i}{\partial q^j} Y^j(\phi(q),$$

where the summation convention is enforced here and in what follows.

A time-varying vector field $(q, t) \mapsto X(q, t)$ gives rise to the initial value problem

$$\dot{q}(t) = X(q, t), \qquad q(0) = q_0,$$

and its solution at time T, which we refer to as the flow of X, is denoted by $q(T) = \Phi^X_{0,T}(q_0)$. We shall usually assume time-varying quantities to be integrable with respect to time. Given a time-varying vector field $X(q, t)$, we denote its definite time integral from time 0 to time T by:

$$\overline{X}(q, T) = \int_0^T X(q, \tau) d\tau. \tag{2.1}$$

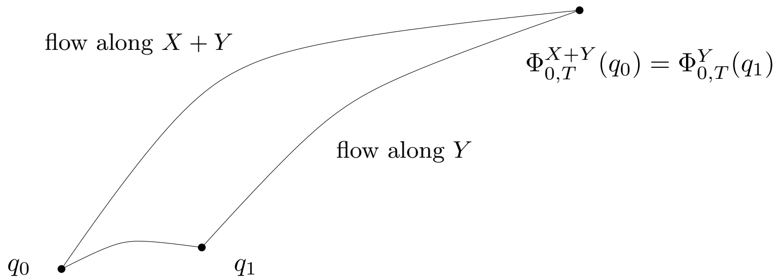

FIG. 2.1. *The flow along $X+Y$ is written as the flow along Y with initial condition q_q. The "variation" q_1 is computed via the variation of constants formula as the flow along $(\Phi^Y_{0,t})^*X$ for time $[0,T]$ with initial condition q_0.*

The integral takes place over the linear space T_qQ at fixed $q \in Q$. This operation can be defined in two ways. Given a coordinates chart about q, the integral is well-defined in the coordinate system (this definition suffices for the purpose of this paper, since the analysis is local). A global coordinate-free definition is obtained providing sufficient conditions in order for T_qQ to be a Banach space and introducing the Cauchy-Bochner integral, see [1, see the discussion at page 61].

Next, consider the initial value problem

$$\dot{q}(t) = X(q,t) + Y(q,t), \qquad q(0) = q_0, \tag{2.2}$$

where X and Y are analytic time-varying vector fields. If we regard X as a perturbation to the vector field Y, we can describe the flow of $X+Y$ in terms of a nominal and perturbed flow. The following relationship is referred to as the *variation of constants* formula and describes the perturbed flow:

$$\Phi^{X+Y}_{0,t} = \Phi^{Y}_{0,t} \circ \Phi^{(\Phi^Y_{0,t})^*X}_{0,t}. \tag{2.3}$$

The result is illustrated in Figure 2.1 and proven in [2, equation (3.15)], see also [10, Appendix A.1]. The result can be alternatively stated as follows. For all $T \geq 0$, the final value $q(T)$ of the curve $q : [0,T] \to M$ solution to the initial value problem (2.2) is also the final value of the curve solution to

$$\dot{q}(s) = Y(q,s), \qquad q(0) = z(T), \tag{2.4}$$

where $z : [0,T] \to M$ is the solution to the initial value problem

$$\dot{z}(s) = \left((\Phi^Y_{0,s})^*X\right)(z), \qquad z(0) = q_0. \tag{2.5}$$

The differential equation (2.5) is referred to as the "pulled back" or the "adjoint" system in [21]. If both X and Y are time invariant, then the classic infinitesimal Campbell-Backer-Hausdorff formula, see [22], provides a mean of computing the pull-back:

$$(\Phi^Y_{0,t})^*X = \sum_{k=0}^{\infty} \mathrm{ad}^k_Y X \, \frac{t^k}{k!}.$$

If instead X and Y are time-varying, a generalized expression is, see [2]:

$$(\Phi^Y_{0,t})^* X(q,t) = X(q,t) \\ + \sum_{k=1}^{\infty} \int_0^t \cdots \int_0^{s_{k-1}} \left(\mathrm{ad}_{Y(q,s_k)} \dots \mathrm{ad}_{Y(q,s_1)} X(q,t)\right) \, ds_k \dots ds_1. \quad (2.6)$$

Just like in the classic Campbell-Backer-Hausdorff formula, see [45], the convergence of the series expansion in the previous equation is a delicate manner. Sufficient conditions for local convergence are given in [2, Proposition 2.1 and 3.1]. For our analysis, the following straightforward statement suffices. If all the Lie brackets $\mathrm{ad}_{Y(s_k)} \dots \mathrm{ad}_{Y(s_1)} X$ vanish for all k greater than a given N, then the series in equation (2.6) becomes a finite sum and it readily converges.

2.3. Affine connections. We refer to [17, 30] for a comprehensive treatment on affine connections and Riemannian geometry. An *affine connection* on Q is a smooth map that assigns to a pair of vector fields X, Y a vector field $\nabla_X Y$ such that for any function f and for any third vector field Z:

(i) $\nabla_{fX+Y} Z = f\nabla_X Z + \nabla_Y Z$,
(ii) $\nabla_X (fY + Z) = (\mathscr{L}_X f)Y + f\nabla_X Y + \nabla_X Z$.

We also say that $\nabla_X Y$ is the covariant derivative of Y with respect to X. Vector fields can also be covariantly differentiated along curves, and this concept will be instrumental in writing the Euler-Lagrange equations. Consider a smooth curve $\gamma : [0,1] \to Q$ and a vector field along γ, that is, a map $v : [0,1] \to TQ$ such that $\pi(v(t)) = \gamma(t)$ for all $t \in [0,1]$. Let V be a smooth vector field satisfying $V(\gamma(t)) = v(t)$. The covariant derivative of the vector field v along γ is defined by

$$\frac{Dv(t)}{dt} \triangleq \nabla_{\dot\gamma(t)} v(t) = \nabla_{\dot\gamma(t)} V(q)\big|_{q=\gamma(t)}.$$

It can be shown that this definition is independent of the choice of V. In a system of local coordinate $(q^1, \dots, q^n)$, an affine connection is uniquely determined by its Christoffel symbols[1] Γ^i_{ij}

$$\nabla_{\frac{\partial}{\partial q^i}} \left(\frac{\partial}{\partial q^j}\right) = \Gamma^k_{ij} \frac{\partial}{\partial q^k},$$

and accordingly, the covariant derivative of a vector field is written as

$$\nabla_X Y = \left(\frac{\partial Y^i}{\partial q^j} X^j + \Gamma^i_{jk} X^j Y^k\right) \frac{\partial}{\partial q^i}.$$

2.4. Control systems described by affine connections. We introduce a class of control systems that is a generalization of Lagrangian control systems. This approach to modeling of vehicles and robotic manipulators is common to a number of recent works; see [9, 33, 32, 10]. A *control system described by an affine connection* is defined by the following objects:

(i) an n-dimensional configuration manifold Q, with $q \in Q$ being the configuration of the system and $v_q \in T_q Q$ being the system's velocity,

[1] We here refer to the Γ^k_{ij} functions as Christoffel symbols even without requiring ∇ to be a Levi-Civita connection.

(ii) an affine connection ∇ on Q, whose Christoffel symbols are $\{\Gamma^i_{jk} : i, j, k \in \{1, \dots, n\}\}$,
(iii) a time-varying vector field Y on Q defining the input force.

The corresponding equations of motion are written as

$$\frac{Dv_q}{dt} = Y(q,t) \tag{2.7}$$

or equivalently in coordinates as

$$\dot{q}^i = v^i, \qquad \dot{v}^i + \Gamma^i_{jk}(q)v^j v^k = Y^i(q,t), \tag{2.8}$$

where the indices i, j, k run from 1 to n, and where $v_q = v^i \frac{\partial}{\partial q^i}$. These equations are a generalized form of the Euler-Lagrange equations.

REMARK 2.1. *This definition of control systems described by an affine connection provides a convenient mean of treating various classes of Lagrangian mechanical systems. For example, systems with nonholonomic constraints are described within this framework in [32]. We will treat in more details "simple Hamiltonian systems" in Section 4.2 and "invariant systems on Lie groups" in Section 4.3. A more detailed exposition in presented in [10].*

The second-order system in equation (2.7) can be written as a first-order differential equation on the tangent bundle TQ. Using $\{\frac{\partial}{\partial q^i}, \frac{\partial}{\partial v^i}\}$ as basis for the tangent bundle to TQ, we define

$$Z(v_q) = v^i \frac{\partial}{\partial q^i} - \Gamma(q)^i_{jk} v^j v^k \frac{\partial}{\partial v^i}, \quad \text{and} \quad Y^{\text{lift}}(v_q, t) = Y^i(q,t) \frac{\partial}{\partial v^i},$$

so that the control system is rewritten as

$$\dot{v}_q = Z(v_q) + Y^{\text{lift}}(v_q, t). \tag{2.9}$$

We refer to [33, 30] for coordinate independent definitions of the lifting operation $Y \to Y^{\text{lift}}$ and of the drift vector field Z.

2.5. Homogeneity and Lie algebraic structure. One fundamental structure of the control system in equation (2.7) is the polynomial dependence of the vector fields Z and Y^{lift} on the velocity variables v^i_q. This structure is reflected in the Lie brackets computations involving Z and Y^{lift}; see related ideas in [41, 10].

We here rely on the notion of *geometric homogeneity*[2] as described in [24]. Given two vector fields X and X_E, we say that the vector field X is homogeneous with degree m with respect to X_E if

$$[X_E, X] = mX.$$

For control systems described by an affine connection, we introduce the Liouville vector field on TQ, see [7, pages 19 and 29], as

$$L(v_q) = v^i \frac{\partial}{\partial v^i},$$

[2]Geometric homogeneity corresponds to the existence of an (infinitesimal) symmetry in the equations of motion. For control systems described by an affine connection the symmetry is invariance under affine time-scaling transformations.

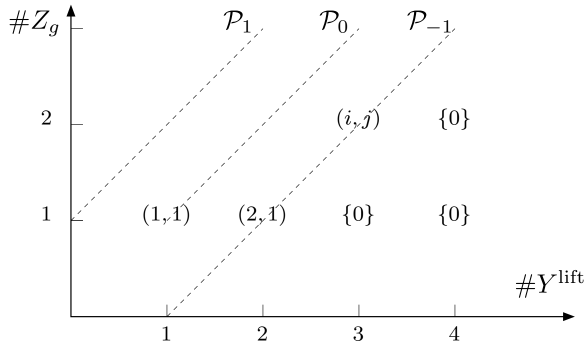

FIG. 2.2. *Table of Lie brackets between the drift vector field Z and the input vector field Y^{lift}. The (i,j)th position contains Lie brackets with i copies of Y^{lift} and j copies of Z. The corresponding homogeneous degree is $j-i$. All Lie brackets to the right of $\mathcal{P}_{-1}$ exactly vanish. All Lie brackets to the left of $\mathcal{P}_{-1}$ vanish when evaluated at $v_q = 0_q$.*

where we recall $v_q = v^i \frac{\partial}{\partial q^i}$. The key mathematical relationships between vector fields on TQ are

$$[L, Z] = (+1)Z, \quad \text{and} \quad [L, Y^{\text{lift}}] = (-1)Y^{\text{lift}}.$$

Hence, the vector field Z is homogeneous of degree $+1$, and the vector field Y^{lift} is homogeneous of degree -1 with respect to the Liouville vector field. Let $\mathcal{P}_j$ be the set of vector fields on TQ of homogeneous degree j, so that

$$Z \in \mathcal{P}_1, \quad \text{and} \quad Y^{\text{lift}} \in \mathcal{P}_{-1}.$$

The sets $\mathcal{P}_j$ enjoy various interesting properties. Table 2.2 illustrates them, their proof is via direct computation, and they are listed next:

(i) $[\mathcal{P}_i, \mathcal{P}_j] \subset \mathcal{P}_{i+j}$, that is, the Lie bracket between a vector field in $\mathcal{P}_i$ and a vector field in $\mathcal{P}_j$ belongs to $\mathcal{P}_{i+j}$.
(ii) $\mathcal{P}_k = \{0\}$ for all $k \leq -2$,
(iii) for all $X \in \mathcal{P}_k$ with $k \geq 1$, $X(0_q) = 0_q$,
(iv) every $X \in \mathcal{P}_{-1}$ is the lift of a vector field on Q.

It is helpful to provide an interpretation of $\mathcal{P}_i$ in coordinates. In a system of local coordinates, let $\mathcal{H}_i(q, v_q)$ be the set of scalar functions on $TQ = \mathbb{R}^{2n}$, which are arbitrary functions of q and which are homogeneous polynomials in $\{v^1, \ldots, v^n\}$ of degree i. $\mathcal{P}_i$ is the set of vector fields on $\mathbb{R}^{2n}$ with the first n components in $\mathcal{H}_i$ and the second n components in $\mathcal{H}_{i+1}$.

Finally, it is of interest to focus on the Lie bracket $[Y_b^{\text{lift}}, [Z, Y_a^{\text{lift}}]]$, where Y_a, Y_b are two vector fields on Q. This operation will play an important role in later computations. Since this Lie bracket belongs to $\mathcal{P}_{-1}$, there must exist a vector field on Q, which we denote $\langle Y_a : Y_b \rangle$, such that

$$\langle Y_a : Y_b \rangle^{\text{lift}} = [Y_b^{\text{lift}}, [Z, Y_a^{\text{lift}}]].$$

Such a vector field is called *symmetric product* between Y_b and Y_a and a direct computation shows that it satisfies

$$\langle Y_b : Y_a \rangle = \nabla_{Y_a} Y_b + \nabla_{Y_b} Y_a,$$

or equivalently in coordinates

$$\langle Y_b : Y_a \rangle^i = \frac{\partial Y_a^i}{\partial q^j} Y_b^j + \frac{\partial Y_b^i}{\partial q^j} Y_a^j + \Gamma^i_{jk} \left(Y_a^j Y_b^k + Y_a^k Y_b^j \right).$$

The adjective "symmetric" comes from the equality $\langle Y_a : Y_b \rangle = \langle Y_b : Y_a \rangle$.

2.6. Integrable flows. Here we compute solutions to a few differential equations defined by certain homogeneous vector fields. In particular, significant simplifications take place in the following two cases. First, let $(q,t) \mapsto X(q,t)$ be a time-varying vector field on Q, and consider the differential equation on TQ

$$\dot{v}_q = X^{\text{lift}}(v_q, t), \tag{2.10}$$

with initial condition $v_q(0) = v_0 \in T_{q_0}Q$. It can be seen that

$$\Phi_{0,t}^{X^{\text{lift}}}(v_0) = v_0 + \int_0^t X(q_0, s) ds, \tag{2.11}$$

that is, in coordinates

$$\Phi_{0,t}^{X^{\text{lift}}} \left(\begin{bmatrix} q_0 \\ v_0 \end{bmatrix} \right) = \begin{bmatrix} q_0 \\ v_0 + \int_0^t X(q_0, s) ds \end{bmatrix}.$$

Next, let $X_0 \in \mathcal{P}_0$ and $X_1 \in \mathcal{P}_1$ and consider the differential equation

$$\dot{v}_q = X_0(v_q, t) + X_1(v_q, t), \tag{2.12}$$

with initial condition $v_q(0) = 0_{q_0} \in T_{q_0}Q$. Define the vector field $X_{0,1}$ on Q and its flow $\zeta : [0,T] \mapsto Q$ via

$$\begin{aligned} X_{0,1} &= T\pi \circ X_0, \\ \zeta(t) &= \Phi_{0,t}^{X_{0,1}}(q_0), \end{aligned}$$

where $T\pi : TTQ \to TQ$ is the tangent map to the projection map $\pi : TQ \to Q$. In coordinates, this vector field consists of the first n components of the vector field $X_0 = \big[X_{0,1}(q,t)', \quad X_{0,2}(q,v,t)' \big]'$ on TQ. It can be seen that

$$\Phi_{0,t}^{X_0 + X_1}(0_{q_0}) = 0_{\zeta(t)},$$

that is, in coordinates

$$\Phi_{0,t}^{X_0 + X_1} \left(\begin{bmatrix} q_0 \\ 0 \end{bmatrix} \right) = \begin{bmatrix} \zeta(t) \\ 0 \end{bmatrix}.$$

The key observation in proving this statement is that the components of $X_{0,2}$ and X_1 are polynomials in $\{v^1, \dots, v^n\}$ of degree at least 1. Since the initial velocity is assumed zero, v_q remains zero for all time.

2.7. Analyticity and bounds over complex neighborhoods. In this section we introduce a norm on the set of analytic vector fields over a compact subset of Q. We also provide bounds to partial derivatives of analytic functions. The bounds are *not* coordinate-free, i.e., they depend on the specific selection of coordinate system. Accordingly, the treatment here assumes $Q = \mathbb{R}^n$.

Let q_0 be a point on $\mathbb{R}^n$, and let σ be a positive scalar, and define the complex σ-neighborhood of q_0 in $\mathbb{C}^n$ as

$$B_\sigma(q_0) = \{z \in \mathbb{C}^n : \|z - q_0\| < \sigma\}.$$

Let f be a real analytic function on $\mathbb{R}^n$ that admits a bounded analytic continuation over $B_\sigma(q_0)$. The norm of f is defined as

$$\|f\|_\sigma \triangleq \max_{z \in B_\sigma(q_0)} |f(z)|,$$

where f denotes both the function over $\mathbb{R}^n$ and its analytic continuation. Given a time-varying vector field $(q,t) \mapsto Y(q,t) = Y_t(q)$, let Y_t^i be its ith component with respect to the usual basis on $\mathbb{R}^n$. Assuming $t \in [0,T]$, and assuming that every component function Y_t^i is analytic over $B_\sigma(q_0)$, we define the norm of Y as

$$\|Y\|_{\sigma,T} \triangleq \max_{t\in[0,T]} \max_{i\in\{1,\dots,n\}} \|Y_t^i\|_\sigma.$$

In what follows, we will often simplify notation by neglecting the subscript T in the norm of a time-varying vector field. Finally, given an affine connection ∇ with Christoffel symbols $\{\Gamma^i_{jk} : i,j,k \in \{1,\dots,n\}\}$, introduce the notation:

$$\|\Gamma\|_\sigma \triangleq \max_{ijk} \left\|\Gamma^i_{jk}\right\|_\sigma .$$

Next, we examine the norm of partial derivatives of these objects. Recall that the Cauchy integral representation of analytic functions leads to bounds on high order derivatives of analytic functions in terms of the norm of the functions themselves, see the so-called Cauchy estimates in [28, Section 2.3] and [38]. Let $(i_1,\dots,i_m)$ be a collection of integers belonging to $\{1,\dots,n\}$, and let σ' be a positive real strictly less than σ. It is known that

$$\|\partial^m f\|_{\sigma'} \triangleq \max_{i_1,\dots,i_m} \left\| \frac{\partial^m f}{\partial q_{i_1}\cdots\partial q_{i_m}} \right\|_{\sigma'} \le m!\,\delta^m \|f\|_\sigma,$$

where $\delta = n/(\sigma-\sigma')$. The quantity $\partial^m f/\partial q_{i_1}\cdots\partial q_{i_m}$ is a real function; it is bounded by bounding its analytic continuation over $B_\sigma(q_0)$. Similarly, for vector fields

$$\|\partial^m Y\|_{\sigma'} \triangleq \max_{t\in[0,T]} \max_{i,i_1,\dots,i_m} \left\| \frac{\partial^m Y_t^i}{\partial q_{i_1}\cdots\partial q_{i_m}} \right\|_{\sigma'} \le m!\,\delta^m \|Y\|_\sigma,$$

and for the Christoffel symbols

$$\|\partial^m \Gamma\|_{\sigma'} \triangleq \max_{i,j,k,i_1,\dots,i_m} \left\| \frac{\partial^m \Gamma^i_{jk}}{\partial q_{i_1}\cdots\partial q_{i_m}} \right\|_{\sigma'} \le m!\,\delta^m \|\Gamma\|_\sigma.$$

3. A Series Expansion for Mechanical Control Systems. This section describes first a preliminary bound, then the main result of the paper, that is, a series expansion describing the evolution of a forced control system starting at rest.

Problem 3.1. Assume the functions $q \mapsto \Gamma^i_{jk}(q)$ and the vector field $(q,t) \mapsto Y(q,t)$ analytic in $q \in Q$, and integrable in $t \in [0,T]$, for some positive time T. Let $\gamma : [0,T] \mapsto Q$ be the solution to the differential equation (2.7) with initial condition $\dot\gamma(0) = 0_{q_0}$. Characterize γ as a series expansion containing iterated symmetric products and time integrals of Y.

We start with a conservative bound.

LEMMA 3.2 (Bound on evolution). *Consider the system as described in Problem 3.1. Select a coordinate system about the point $q_0 \in Q$ and let σ be a positive constant. A sufficient condition for $\gamma([0,T])$ to be a subset of $B_\sigma(q_0)$ is that*

$$\|Y\|_\sigma T^2 < \frac{\eta^2(\sigma n^2 \|\Gamma\|_\sigma)}{n^2\|\Gamma\|_\sigma}, \tag{3.1}$$

where the function $\eta : x \in \mathbb{R}_+ \to [0, \pi/2]$ is the unique solution to $\eta \tan(\eta) = x$.

Proof. Let $T_0 < T$ be the smallest time at which the solution γ reaches the distance $\|\gamma(T_0) - q_0\| = \sigma$. If the solution never reaches this distance, then $\gamma([0,T])$ is obviously a subset of $B_\sigma(q_0)$. Since $\gamma([0,T_0]) \subset B_\sigma(q_0)$, for all $t \in [0,T_0]$ we have the bound $\|\dot\gamma(t)\| \leq y(t)$, where

$$\dot y = n^2\|\Gamma\|_\sigma y^2 + \|Y\|_\sigma, \qquad y(0) = 0.$$

The solution to this initial value problem is

$$y(t) = \sqrt{\frac{\|Y\|_\sigma}{n^2\|\Gamma\|_\sigma}} \tan\left(\sqrt{\|Y\|_\sigma n^2 \|\Gamma\|_\sigma}\, t\right).$$

Straightforward manipulations show that the condition in equation (3.1) is equivalent to $Ty(T) < \sigma$. But since y is a monotone function, also $T_0 y(T_0) < \sigma$. Note that $\|\gamma(0) - q_0\| = 0$ and

$$\frac{d}{dt}\|\gamma(t) - q_0\| \leq \|\dot\gamma\| \leq y(t) < \sigma/T_0$$

for all $t \in [0,T_0]$. Therefore $\|\gamma(T_0) - q_0\| < T_0\sigma/T_0$ and the contradiction is now immediate. ◻

We are now ready to present the main theorem.

THEOREM 3.3 (Evolution of a forced mechanical system starting at rest). *Consider the system as described in Problem 3.1. Define recursively the time-varying vector fields V_k:*

$$V_1(q,t) = \int_0^t Y(q,s)ds \tag{3.2}$$

$$V_k(q,t) = -\frac{1}{2}\sum_{j=1}^{k-1} \int_0^t \Big\langle V_j(q,s) :\ V_{k-j}(q,s)\Big\rangle ds, \qquad k \geq 2. \tag{3.3}$$

Select a coordinate system about the point $q_0 \in Q$, let $\sigma > \sigma'$ be two positive constants, and assume that

$$\|Y\|_\sigma T^2 < L \triangleq \min\left\{ \frac{\sigma - \sigma'}{2^4 n^2(n+1)}, \frac{1}{2^4 n(n+1)\|\Gamma\|_\sigma}, \frac{\eta^2(\sigma' n^2\|\Gamma\|_{\sigma'})}{n^2\|\Gamma\|_{\sigma'}} \right\}. \tag{3.4}$$

Then the solution $\gamma : [0,T] \to Q$ satisfies

$$\dot\gamma(t) = \sum_{k=1}^{+\infty} V_k(\gamma(t),t), \tag{3.5}$$

where V_k satisfies the bound

$$\|V_k\|_{\sigma'} \leq L^{1-k} \|Y\|_\sigma^k \, t^{2k-1}, \tag{3.6}$$

and the series $(q,t) \mapsto \sum_{k=1}^\infty V_k(q,t)$ converges absolutely and uniformly for $q \in B_{\sigma'}(q_0)$ and for $t \in [0,T]$.

A few comments on the various steps of the proof are appropriate. First, we investigate how to write the flow of a mechanical control system as the composition of more elementary flows. Two observations play a key role: the homogeneity of system (2.7) renders the computations tractable, the simplifying procedure can be easily repeated giving rise to an iterative procedure. Second, we prove absolute and uniform convergence of the series expansion resulting from the first formal part of the proof. The proof of the bounds is inspired by the treatment in [2, Proposition 2.1], but it is considerably more complicated here by the recursive nature of the series expansion. Once the series is formally derived and it is proven to be convergent, a limiting argument leads to the final statement in equation (3.5).

Proof.

Part I: Here we write the solution to equation (2.7) *as composition of the flow of two separate vector fields, one of which is defined recursively.*

Let k be a strictly positive integer, let X_k, Y_k, W_k be time-varying vector fields on Q, and let $v_{q,k}$ be a smooth curve on TQ that satisfies the differential equation

$$\begin{aligned} \dot{v}_{q,k} &= \left(Z + [X_k^{\text{lift}}, Z] + Y_k^{\text{lift}} + W_k^{\text{lift}}\right)(v_{q,k}, t) \\ v_{q,k}(0) &= 0_{q_0}. \end{aligned} \tag{3.7}$$

The mechanical system in equation (2.7) corresponds to setting $k = 1$, $X_1 = W_1 = 0$, $Y_1 = Y(q,t)$, and accordingly $\dot{\gamma}(t) = v_{q,1}(t)$. Using the formula in equations (2.4) and (2.5) discussed in Section 2.2, we set

$$v_{q,k}(t) = \Phi_{0,t}^{Y_k^{\text{lift}}}(v_{q,k+1}(t)) \tag{3.8}$$

and

$$\begin{aligned} \dot{v}_{q,k+1} &= \left(\left(\Phi_{0,t}^{Y_k^{\text{lift}}}\right)^* (Z + [X_k^{\text{lift}}, Z] + W_k^{\text{lift}})\right)(v_{q,k+1}) \\ v_{q,k+1}(0) &= 0_{q_0}, \end{aligned} \tag{3.9}$$

where we compute the pull-back along the flow by means of the infinite series in equation (2.6). Remarkably, this series reduces to a finite sum. From the discussion in Section 2.5 on the Lie algebraic structure of the various vector fields, we have

$$\begin{gathered} \mathrm{ad}_{Y_k^{\text{lift}}}^{m+2} Z = 0 \\ \mathrm{ad}_{Y_k^{\text{lift}}}^{m+1} [X_k^{\text{lift}}, Z] = 0, \quad \mathrm{ad}_{Y_k^{\text{lift}}}^{m} W_k^{\text{lift}} = 0, \end{gathered}$$

for all $m \geq 1$. With a little book-keeping we can exploit these equalities and compute

$$\begin{aligned}
\left(\Phi_{0,t}^{Y_k^{\text{lift}}}\right)^* & (Z + [X_k^{\text{lift}}, Z] + W_k^{\text{lift}}) \\
&= Z + [X_k^{\text{lift}}, Z] + W_k^{\text{lift}} + \int_0^t [Y_k^{\text{lift}}(s), \big(Z + [X_k^{\text{lift}}, Z]\big)] ds \\
&\qquad + \int_0^t \int_0^{s_1} [Y_k^{\text{lift}}(s_2), [Y_k^{\text{lift}}(s_1), Z]] ds_2 ds_1 \\
&= Z + [X_k^{\text{lift}} + \overline{Y}_k^{\text{lift}}, Z] + [\overline{Y}_k^{\text{lift}}(s), [X_k^{\text{lift}}, Z]] + W_k^{\text{lift}} \\
&\qquad + \int_0^t \int_0^{s_1} [Y_k^{\text{lift}}(s_2), [Y_k^{\text{lift}}(s_1), Z]] ds_2 ds_1 \\
&= Z + [X_k^{\text{lift}} + \overline{Y}_k^{\text{lift}}, Z] - \left\langle \overline{Y}_k : X_k \right\rangle^{\text{lift}} + W_k^{\text{lift}} - \frac{1}{2} \left\langle \overline{Y}_k : \overline{Y}_k \right\rangle^{\text{lift}},
\end{aligned}$$

where we have used the $\bar{\cdot}$ notation introduced in equation (2.1). The last equality also relies on

$$\int_0^t \int_0^{s_1} [Y_k^{\text{lift}}(s_2), [Y_k^{\text{lift}}(s_1), Z]] ds_2 ds_1 = -\frac{1}{2} \left\langle \overline{Y}_k : \overline{Y}_k \right\rangle^{\text{lift}},$$

which follows from an integration by part and the symmetry of the symmetric product. Remarkably, the differential equation describing the evolution of $v_{k+1}(t)$ is of the same form as equation (3.7) describing the evolution of $v_{q,k}(t)$, where

$$\begin{aligned}
X_{k+1} &= X_k + \overline{Y}_k \\
Y_{k+1} + W_{k+1} &= -\left\langle \overline{Y}_k \;\;:\;\; X_k + \tfrac{1}{2} \overline{Y}_k \right\rangle + W_k.
\end{aligned}$$

The vector field X_k can be computed and substituted in as:

$$\begin{aligned}
X_k &= \sum_{j=1}^{k-1} \overline{Y}_j \\
Y_{k+1} + W_{k+1} &= -\left\langle \overline{Y}_k \;\;:\;\; \sum_{j=1}^{k-1} \overline{Y}_j + \frac{1}{2} \overline{Y}_k \right\rangle + W_k.
\end{aligned} \tag{3.10}$$

Notice that the quantities Y_k and W_k are not yet uniquely determined. Equation (3.10) is verified for all k, if and only if for all m:

$$(Y_2 + Y_3 + \ldots + Y_{m+1}) + W_{m+1} = -\sum_{k=1}^{m} \left\langle \overline{Y}_k : \sum_{j=1}^{k-1} \overline{Y}_j + \frac{1}{2} \overline{Y}_k \right\rangle, \tag{3.11}$$

where we used $W_1 = 0$. Some further manipulation leads to:

$$\begin{aligned}
\sum_{k=1}^{m} \left\langle \overline{Y}_k : \sum_{j=1}^{k-1} \overline{Y}_j + \frac{1}{2} \overline{Y}_k \right\rangle &= \sum_{k=1}^{m} \sum_{j=1}^{k-1} \left\langle \overline{Y}_k : \overline{Y}_j \right\rangle + \frac{1}{2} \sum_{k=1}^{m} \left\langle \overline{Y}_k : \overline{Y}_k \right\rangle \\
&= \frac{1}{2} \sum_{j,k=1, j \neq k}^{m} \left\langle \overline{Y}_k : \overline{Y}_j \right\rangle + \frac{1}{2} \sum_{k=1}^{m} \left\langle \overline{Y}_k : \overline{Y}_k \right\rangle \\
&= \frac{1}{2} \sum_{j,k=1}^{m} \left\langle \overline{Y}_k : \overline{Y}_j \right\rangle.
\end{aligned}$$

A selection of $\{Y_i : i \in \{1, \dots, m\}\}$, and W_{m+1} that satisfies equation (3.11) is

$$\begin{aligned} Y_i &= -\frac{1}{2} \sum_{j,k=1,j+k=i}^{m} \left\langle \overline{Y}_k : \overline{Y}_j \right\rangle = -\frac{1}{2} \sum_{j=1}^{i-1} \left\langle \overline{Y}_j : \overline{Y}_{i-j} \right\rangle \\ W_{m+1} &= -\frac{1}{2} \sum_{j,k=1,j+k>m}^{m} \left\langle \overline{Y}_k : \overline{Y}_j \right\rangle . \end{aligned} \tag{3.12}$$

Note that equation (3.12) is a well defined recursive relationship, and note that the recursive definition of V_k in equation (3.3) and (3.2) inside the theorem statement corresponds to setting $V_k(q,t) = \overline{Y}_k(q,t)$. The iteration procedure proves that, for any $k \geq 2$, the solution to the original mechanical system $\dot{\gamma} = v_{q,1} : [0,T] \mapsto TQ$ satisfies

$$\dot{\gamma}(t) = \left(\Phi_{0,t}^{Y_1^{\text{lift}}} \circ \Phi_{0,t}^{Y_2^{\text{lift}}} \circ \ldots \circ \Phi_{0,t}^{Y_{k-1}^{\text{lift}}} \right) (v_{q,k}(t)),$$

where $v_{q,k} : [0,T] \mapsto TQ$ is the solution to equation (3.7). The flow $\dot{\gamma}$ is now written as the composition of k flows and a first simplification is immediate. For all integers i, j and for all times s_1, s_2 the vector fields Y_i^{lift} and Y_j^{lift} commute, that is

$$[Y_i^{\text{lift}}(v_q, s_1), Y_j^{\text{lift}}(v_q, s_2)] = 0,$$

so that γ is the solution to

$$\dot{\gamma}(t) = \Phi_{0,t}^{\sum_{j=1}^{k-1} Y_j^{\text{lift}}} (v_{q,k}(t)). \tag{3.13}$$

A second simplification is also straightforward. The vector field in equation (3.13) is homogeneous of degree 0, i.e., it is in the form of equation (2.10) in Section 2.6. According to the result in equation (2.11) we have for all $t \in [0,T]$

$$\dot{\gamma}(t) = v_{q,k}(t) + \sum_{j=1}^{k-1} \overline{Y}_j(\pi(v_{q,k}(t)), t), \tag{3.14}$$

where the sequence of vector fields Y_j is defined via equation (3.12) and where the curve $v_{q,k} : [0,T] \mapsto TQ$ is the solution to

$$\begin{aligned} \frac{dv_{q,k}}{dt} &= \left(Z + \left[\sum_{j=1}^{k-1} \overline{Y}_j^{\text{lift}}, Z \right] + Y_k^{\text{lift}} + W_k^{\text{lift}} \right) (v_{q,k}, t) \\ v_{q,k}(0) &= 0_{q_0}. \end{aligned} \tag{3.15}$$

Part II: Here we show absolute and uniform convergence of the series $\sum_{k=1}^{\infty} Y_k(q,t)$ *over all* q *in a compact neighborhood of* q_0 *and for all* $t \leq T$.

Given the vector field Y, let $\Omega_1 = \{Y\}$, and define recursively the set Ω_k to be the collection of vector fields $-\frac{1}{2}\left\langle \overline{B}_i : \overline{B}_{k-i} \right\rangle$, for all $B_i \in \Omega_i$ and $B_{k-i} \in \Omega_{k-i}$. The first few sets are:

$$\begin{aligned} &\Omega_1 = \{Y\}, \quad \Omega_2 = \left\{ -\frac{1}{2} \left\langle \overline{Y} : \overline{Y} \right\rangle \right\}, \quad \Omega_3 = \left\{ \frac{1}{4} \left\langle \overline{Y} : \overline{\left\langle \overline{Y} : \overline{Y} \right\rangle} \right\rangle \right\}, \\ &\Omega_4 = \left\{ -\frac{1}{8} \left\langle \overline{Y} : \overline{\left\langle \overline{Y} : \overline{\left\langle \overline{Y} : \overline{Y} \right\rangle} \right\rangle} \right\rangle, -\frac{1}{8} \left\langle \overline{\left\langle \overline{Y} : \overline{Y} \right\rangle} : \overline{\left\langle \overline{Y} : \overline{Y} \right\rangle} \right\rangle \right\}. \end{aligned} \tag{3.16}$$

Next, we prove by induction that, for all k, the vector field Y_k is the sum of N_k vector fields belonging to Ω_k. The statement is true at $k=1$ with $N_1=1$. We assume it true for all $j<k$ and prove it for k. Because of the induction assumption, we write $Y_j=\sum_{a=1}^{N_j} B_{j,a}$, where the $B_{j,a}$ are elements in Ω_j. We compute

$$\begin{aligned} Y_k &= -\frac{1}{2}\sum_{j=1}^{k-1}\left\langle \overline{Y}_j : \overline{Y}_{k-j}\right\rangle \\ &= -\frac{1}{2}\sum_{j=1}^{k-1}\left\langle \sum_{a=1}^{N_j}\overline{B_{j,a}} : \sum_{b=1}^{N_{k-j}}\overline{B_{k-j,b}}\right\rangle \\ &= \sum_{j=1}^{k-1}\sum_{a=1}^{N_j}\sum_{b=1}^{N_{k-j}} \underbrace{-\frac{1}{2}\left\langle \overline{B_{j,a}} : \overline{B_{k-j,b}}\right\rangle}_{\in\ \Omega_k}, \end{aligned}$$

This concludes the proof by induction and the recursive relation on N_k is

$$N_1=1, \qquad N_k=\sum_{j=1}^{k-1} N_j N_{k-j}, \quad k\geq 2. \tag{3.17}$$

As we discuss in Appendix A equation (A.1), the sequence N_k can be explicitly computed and bounded as

$$N_k=\frac{1}{k}\binom{2k-2}{k-1}\leq \frac{2^{2(k-1)}}{k-\frac{1}{2}}. \tag{3.18}$$

We now focus our attention on bounding the generic time-varying vector field $(q,s)\mapsto B_k(q,s)$ in Ω_k. Recall that the symbols $\delta, \|\cdot\|_\sigma, \|\partial^m\cdot\|_\sigma$ introduced in Section 2.7. We claim that there exist sequences of real and integer coefficients $\{c_k : k\in\mathbb{R}\}$ and $\{d_k : k\in\mathbb{N}\}$, such that

$$\|\partial^m B_k\|_{\sigma'} \leq c_k \frac{(m+d_k)!}{d_k!}\,\delta^{m+k-1}\|Y\|_\sigma^k t^{2(k-1)}. \tag{3.19}$$

For convenience, we redefine δ to $\delta=\max\left\{\frac{n}{\sigma-\sigma'}, \|\Gamma\|_\sigma\right\}$, so that $\|\partial^m\Gamma\|_{\sigma'}\leq m!\,\delta^{m+1}$. As discussed in that section, the bound in equation (3.19) is satisfied at $k=1$ for all $m\in\mathbb{N}$, with $c_1=1$, $d_1=0$. In what follows we provide a proof by induction on $k\geq 2$.

Any time-varying vector field B_k at $k\geq 2$ can be written as $B_k=-\frac{1}{2}\left\langle\overline{B_a}:\overline{B_b}\right\rangle$ for some $1\leq a,b\leq k-1, a+b=k$ and $B_a\in\Omega_a$, $B_b\in\Omega_b$. Accordingly, we compute:

$$\begin{aligned} &\|\partial^m\left\langle\overline{B_a}:\overline{B_b}\right\rangle\|_{\sigma'} \\ &\quad= \max_{i,i_1,\dots,i_m}\left\|\frac{\partial^m}{\partial q_{i_1}\cdots\partial q_{i_m}}\left(\frac{\partial\overline{B_a^i}}{\partial q^j}\overline{B_b^j}+\frac{\partial\overline{B_b^i}}{\partial q^j}\overline{B_a^j}+\Gamma^i_{jl}(\overline{B_a^j}\,\overline{B_b^l}+\overline{B_b^j}\,\overline{B_a^l})\right)\right\|_{\sigma'} \\ &\quad\leq \max_i\left(\left\|\partial^m\left(\frac{\partial\overline{B_a^i}}{\partial q^j}\overline{B_b^j}\right)\right\|_{\sigma'}+\left\|\partial^m\left(\frac{\partial\overline{B_b^i}}{\partial q^j}\overline{B_a^j}\right)\right\|_{\sigma'}+2\left\|\partial^m\left(\Gamma^i_{jl}\overline{B_a^j}\,\overline{B_b^l}\right)\right\|_{\sigma'}\right). \end{aligned}$$

Relying on the equality

$$\frac{d^m}{dx^m} f(x)g(x) = \sum_{\alpha=0}^{m} \binom{m}{\alpha} \frac{d^\alpha f(x)}{dx^\alpha} \frac{d^{m-\alpha} g(x)}{dx^{m-\alpha}},$$

the first term is bounded according to:

$$\begin{aligned}
\left\| \partial^m \left(\frac{\partial \overline{B_a^i}}{\partial q^j} \overline{B_b^j} \right) \right\|_{\sigma'} &\leq n \sum_{\alpha=0}^{m} \frac{m!}{\alpha!(m-\alpha)!} \left\|\partial^{\alpha+1} \overline{B_a}\right\|_{\sigma'} \left\|\partial^{m-\alpha} \overline{B_b}\right\|_{\sigma'} \\
&\leq n \sum_{\alpha=0}^{m} \frac{m!}{\alpha!(m-\alpha)!} \left(c_a \frac{(\alpha+1+d_a)!}{d_a!} \delta^{\alpha+a} \|Y\|_\sigma^a \frac{t^{2a-1}}{2a-1} \right) \\
&\qquad \cdot \left(c_b \frac{(m-\alpha+d_b)!}{d_b!} \delta^{m-\alpha+b-1} \|Y\|_\sigma^b \frac{t^{2b-1}}{2b-1} \right) \\
&= \frac{n c_a c_b}{(2a-1)(2b-1)} \left(\frac{m!}{d_a! d_b!} \sum_{\alpha=0}^{m} \frac{(\alpha+1+d_a)!(m-\alpha+d_b)!}{\alpha!(m-\alpha)!} \right) \\
&\qquad \cdot \delta^{m+a+b-1} \|Y\|_\sigma^{a+b} t^{2(a+b-1)}.
\end{aligned}$$

The third term is bounded according to:

$$\begin{aligned}
\left\| \partial^m \left(\Gamma^i_{jl} \overline{B_a^j}\, \overline{B_b^l} \right) \right\|_{\sigma'} &= n^2 \sum_{\alpha=0}^{m} \sum_{\beta=0}^{\alpha} \frac{m! \, \|\partial^{m-\alpha}\Gamma\|_{\sigma'}}{(m-\alpha)!\beta!(\alpha-\beta)!} \|\partial^\beta \overline{B_a}\|_{\sigma'} \|\partial^{\alpha-\beta} \overline{B_b}\|_{\sigma'} \\
&= n^2 \sum_{\alpha=0}^{m} \sum_{\beta=0}^{\alpha} \frac{m!}{(m-\alpha)!\beta!(\alpha-\beta)!} \left((m-\alpha)!\, \delta^{m-\alpha+1}\right) \\
&\qquad \cdot \left(c_a \frac{(\beta+d_a)!}{d_a!} \delta^{\beta+a-1} \|Y\|_\sigma^a \frac{t^{2a-1}}{2a-1} \right) \\
&\qquad \cdot \left(c_b \frac{(\alpha-\beta+d_b)!}{d_b!} \delta^{\alpha-\beta+b-1} \|Y\|_\sigma^b \frac{t^{2b-1}}{2b-1} \right) \\
&= \frac{n^2 c_a c_b}{(2a-1)(2b-1)} \left(\frac{m!}{d_a! d_b!} \sum_{\alpha=0}^{m} \sum_{\beta=0}^{\alpha} \frac{(\beta+d_a)!(\alpha-\beta+d_b)!}{\beta!(\alpha-\beta)!} \right) \\
&\qquad \cdot \delta^{m+a+b-1} \|Y\|_\sigma^{a+b} t^{2(a+b-1)}.
\end{aligned}$$

To simplify notation, let us define

$$S(l, d_1, d_2) \triangleq \sum_{a=0}^{l} \frac{(a+d_1)!(l-a+d_2)!}{a!(l-a)!}.$$

Putting it all together:

$$\begin{aligned}
\|\partial^m \langle \overline{B_a} : \overline{B_b} \rangle \|_{\sigma'} \leq{}& \frac{n c_a c_b}{(2a-1)(2b-1)} \frac{m!}{d_a! d_b!} \delta^{m+a+b-1} \|Y\|_\sigma^{a+b} t^{2(a+b-1)} \\
&\cdot \left(S(m, d_a+1, d_b) + S(m, d_a, d_b+1) + 2n \sum_{\alpha=0}^{m} S(\alpha, d_a, d_b) \right).
\end{aligned}$$

Equation (A.2) in Appendix A implies that

$$S(l, d_1, d_2) = \frac{d_1! d_2! (l + 1 + d_1 + d_2)!}{l! (1 + d_1 + d_2)!}, \tag{3.20}$$

so that we compute

$$\begin{aligned}
&\frac{m!}{d_a! d_b!} \left(S(m, d_a + 1, d_b) + S(m, d_a, d_b + 1) + 2n \sum_{\alpha=0}^{m} S(\alpha, d_a, d_b) \right) \\
&\quad = \frac{m!}{d_a! d_b!} \Bigg(\Big((d_a + 1)! d_b! + d_a! (d_b + 1)! \Big) \frac{(m + 2 + d_a + d_b)!}{m! (2 + d_a + d_b)!} \\
&\qquad\qquad + 2n \sum_{\alpha=0}^{m} \frac{d_a! d_b! (\alpha + 1 + d_a + d_b)!}{\alpha! (1 + d_a + d_b)!} \Bigg) \\
&\quad = \frac{(m + 2 + d_a + d_b)!}{(1 + d_a + d_b)!} + \frac{2nm!}{(1 + d_a + d_b)!} \underbrace{\sum_{\alpha=0}^{m} \frac{(\alpha + 1 + d_a + d_b)!}{\alpha!}}_{S(m, 1 + d_a + d_b, 0)},
\end{aligned}$$

and again applying equation (3.20) with $(l, d_1, d_2) = (m, 1 + d_a + d_b, 0)$

$$\begin{aligned}
&\quad = \frac{(m + 2 + d_a + d_b)!}{(1 + d_a + d_b)!} + \frac{2nm!}{(1 + d_a + d_b)!} \frac{(1 + d_a + d_b)! (m + 2 + d_a + d_b)!}{m! (2 + d_a + d_b)!} \\
&\quad = \frac{(m + 2 + d_a + d_b)!}{(2 + d_a + d_b)!} (2 + 2n + d_a + d_b).
\end{aligned}$$

Substituting in

$$\begin{aligned}
&\|\partial^m \langle \overline{B_a} : \overline{B_b} \rangle \|_{\sigma'} \\
&\qquad \leq \frac{n c_a c_b (2 + 2n + d_a + d_b)(m + 2 + d_a + d_b)!}{(2a - 1)(2b - 1)(2 + d_a + d_b)!} \delta^{m + a + b - 1} \|Y\|_{\sigma}^{a+b} t^{2(a+b-1)}.
\end{aligned}$$

Next, we express everything back in terms of $k = a + b$ and $B_k = -\frac{1}{2} \langle \overline{B_a} : \overline{B_b} \rangle$. We have that:

$$\|\partial^m B_k\|_{\sigma'} \leq \max_{a+b=k} \left(\frac{n c_a c_b (2 + 2n + d_a + d_b)}{2(2a - 1)(2b - 1)} \frac{(m + 2 + d_a + d_b)!}{(2 + d_a + d_b)!} \right) \delta^{m+k-1} \|Y\|_{\sigma}^{k} t^{2(k-1)}$$

Equation (3.19) is proven by defining sequences c_k and d_k such that $c_1 = 1$, $d_1 = 0$ together with

$$\begin{aligned}
d_k &\leq \max_{a+b=k} 2 + d_a + d_b, \\
c_k &\leq \max_{a+b=k} \frac{n c_a c_b (2 + 2n + d_a + d_b)}{2(2a - 1)(2b - 1)}.
\end{aligned}$$

It is immediate to see that $d_k = 2(k - 1)$ satisfies the recursive requirement, so that we require c_k to satisfy $c_1 = 1$ together with the requirement

$$c_k \leq \max_{a+b=k} \frac{n(k + n - 1) c_a c_b}{(2a - 1)(2b - 1)} = \max_{a \in \{1, \ldots k-1\}} \frac{n(k + n - 1) c_a c_{k-a}}{(2a - 1)(2k - 2a - 1)}.$$

Consider the polynomial $p(a) = (2a-1)(2k-2a-1)$ in $a \in [1, k-1]$, it assumes its minimum value $(2k-3)$ at $a = 1$, or equivalently $a = k-1$. Accordingly, a stricter requirement on c_k is

$$c_k \leq \max_{a \in \{1, \ldots k-1\}} \frac{n(k+n-1)}{2k-3} c_a c_{k-a}.$$

Since $(k-1)/(2k-3) \leq 1$ and $n/(2k-3) \leq n$ for all $k \geq 2$, a conservative selection of c_k that satisfies this requirement is provided by the sequence

$$c_1 = 1, \qquad c_k = n(1+n) \sum_{a=1}^{k-1} c_a c_{k-a}, \quad k \geq 2.$$

Recalling the definition in equation (3.17), one can show that $c_k = (n(1+n))^{k-1} N_k$.

Finally, we summarize all the analysis in Part II and prove convergence. Evaluating at $m = 0$ the bound in equation (3.19), we have

$$\|B_k\|_{\sigma'} \leq \, (n(1+n))^{k-1} N_k \, \delta^{k-1} \|Y\|_\sigma^k \, t^{2(k-1)},$$

and recalling the bound in equation (3.18), we compute

$$\begin{aligned} \|Y_k\|_{\sigma'} &\leq N_k \, \|B_k\|_{\sigma'} \leq (n(1+n))^{k-1} N_k^2 \, \delta^{k-1} \, \|Y\|_\sigma^k \, t^{2(k-1)} \\ &\leq \frac{(2^4 n(1+n)\,\delta)^{k-1}}{(k-1/2)^2} \, \|Y\|_\sigma^k \, t^{2(k-1)}. \end{aligned}$$

An immediate consequence is that for $(2^4 n(n+1)\delta) \, \|Y\|_\sigma T^2 < 1$, the series

$$Y_\infty(q,t) \triangleq \lim_{K \to \infty} \sum_{k=1}^{K} Y_k(q,t)$$

converges absolutely and uniformly in $t \in [0, T]$ and $q \in B_{\sigma'}(q_0)$.

Part III: Here we provide the final limiting argument by collecting various results in Part I, Part II and in Lemma 3.2

We start by studying the behavior as $k \to \infty$ of the equation (3.14) and of the initial value problem (3.15) from Part I. We shall exploit a variation of a standard result on the continuous dependence of solutions of differential equations with respect to parameter changes, see [16, Chapter I, Section 3]. Uniform convergence of the vector field describing a differential equation, say for example $\sum_{k=1}^{K} Y_k$, implies the uniform convergence of the solution to the Kth differential equation to the solution of the limiting differential equation. In order to apply this result to the differential equation (3.15) we need to ensure that the vector field on right hand side converges uniformly and absolutely.

Assume that the time length T and input vector field Y satisfy the bound in equation (3.4) inside the theorem statement. Then Lemma 3.2 guarantees that $\gamma([0,T]) \subset B_{\sigma'}(q_0)$, and the analysis in Part II guarantees that series $\sum_{k=1}^{\infty} Y_k$ converges absolutely and uniformly over $q \in B_{\sigma'}(q_0)$. Therefore the series converges uniformly and absolutely along the curve γ. From equation (3.14) one can deduce that $\gamma(t) = \pi\,(v_{q,k}(t))$, so that the series $\sum_{k=1}^{\infty} Y_k$ converges also along $\pi \circ v_{q,k} : [0, T] \mapsto Q$. Accordingly, we can take the limit as $k \to \infty$ in equation (3.15).

Notice that uniformly in $t \in [0,T]$ and $q \in B_{\sigma'}(q_0)$

$$\lim_{k\to\infty} Y_k(q,t) = 0, \qquad \text{and} \qquad \lim_{k\to\infty} W_k(q,t) = 0,$$

and define the time-varying vector field

$$V_\infty = \sum_{k=1}^{\infty} V_k = \sum_{k=1}^{\infty} \overline{Y}_k.$$

Taking the limit as $k \to \infty$ in equations (3.14) and (3.15) one obtains

$$\dot\gamma(t) = v_{q,\infty}(t) + V_\infty(\pi(v_{q,\infty}(t)), t),$$

where the curve $v_{q,\infty} : [0,T] \mapsto TQ$ is the solution to

$$\begin{aligned} \frac{dv_{q,\infty}}{dt} &= \left(Z + \left[V_\infty^{\text{lift}},\, Z\right]\right)(v_{q,\infty}, t) \\ v_{q,\infty}(0) &= 0_{q_0}. \end{aligned} \tag{3.21}$$

According to the discussion in Section 2.6, the initial value problem in equation (3.21) can be explicitly integrated. Because $Z \in \mathcal{P}_1$, $[V_\infty^{\text{lift}}, Z] \in \mathcal{P}_0$, and because of the equality

$$T\pi \circ \left[V_\infty^{\text{lift}},\, Z\right] = V_\infty,$$

the curve $v_{q,\infty}$ satisfies

$$v_{q,\infty}(t) = 0_{\zeta(t)}, \qquad \text{where} \quad \zeta(t) = \Phi_{0,t}^{V_\infty}(q_0),$$

and plugging in

$$\dot\gamma(t) = 0_{\zeta(t)} + V_\infty(\pi(0_{\zeta(t)}), t) = V_\infty(\zeta(t), t).$$

The last two statements imply $\gamma = \zeta$ and are equivalent to statement in equation (3.5) inside the theorem. ☐

Two brief comments are appropriate. First, it is interesting to emphasize an intermediate result proved in *Part II*: the V_k term in the series is the sum of known number of vector fields belonging to the set Ω_k, see the definition preceding equation (3.16). This additional structure might be useful in controllability or motion planning studies. Second, it is unpleasant to remark that while the series expansion is stated in a coordinate-free context, its convergence properties rely on the introduction of a coordinate system.

4. Applications and Extensions. We present a few diverse comments in order to relate the theorem to various earlier works, as well as obtain stronger results under specific additional assumptions on the system.

4.1. The first few order terms and small amplitude forcing. Equation (3.5) is well-defined in the sense that, at fixed q, the integration is performed with respect to the time variable. Using the abbreviated notation introduced in equation (2.1), the first few terms of the sequence $\{V_k : k \in \mathbb{N}\}$ are computed as

$$\begin{aligned}
V_1 &= \overline{Y} \\
V_2 &= -\frac{1}{2}\overline{\langle \overline{Y} : \overline{Y} \rangle} \\
V_3 &= \frac{1}{2}\overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{Y} \right\rangle} \\
V_4 &= -\frac{1}{2}\overline{\left\langle \overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{Y} \right\rangle} : \overline{Y} \right\rangle} - \frac{1}{8}\overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{\langle \overline{Y} : \overline{Y} \rangle} \right\rangle},
\end{aligned}$$

so that we can write

$$\begin{aligned}
\dot{\gamma}(t) = \overline{Y}(\gamma, t) &- \frac{1}{2}\overline{\langle \overline{Y} : \overline{Y} \rangle}(\gamma, t) + \frac{1}{2}\overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{Y} \right\rangle}(\gamma, t) \\
&- \frac{1}{2}\overline{\left\langle \overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{Y} \right\rangle} : \overline{Y} \right\rangle}(\gamma, t) - \frac{1}{8}\overline{\left\langle \overline{\langle \overline{Y} : \overline{Y} \rangle} : \overline{\langle \overline{Y} : \overline{Y} \rangle} \right\rangle}(\gamma, t) + O(\|Y\|^5 t^9).
\end{aligned}$$

This series converges under the assumption that the product of final time T and input magnitude $\|Y\|_\sigma$ be small. Typically, in controllability studies [42] it is the final time that is assumed to be small (the famous acronym STLC stands for small-time local controllability). Within the context of motion planning problems [31, 14], it is instead convenient to study the small input magnitude case. Motivated by the treatment in [14], let ϵ be a small positive constant, and consider a total acceleration of the form

$$Y(q, t, \epsilon) = \epsilon X_1(q, t) + \epsilon^2 X_2(q, t) + \epsilon^3 X_3(q, t), \qquad t \in [0, 1].$$

Accordingly, equation (3.5) is equivalent to

$$\begin{aligned}
\dot{\gamma}(t) = \epsilon \overline{X}_1(\gamma, t) &+ \epsilon^2 \left(\overline{X}_2 - \frac{1}{2}\overline{\langle \overline{X}_1 : \overline{X}_1 \rangle} \right)(\gamma, t) \\
&+ \epsilon^3 \left(\overline{X}_3 - \frac{1}{2}\overline{\langle \overline{X}_1 : \overline{X}_2 \rangle} + \frac{1}{2}\overline{\left\langle \overline{\langle \overline{X}_1 : \overline{X}_1 \rangle} : \overline{X}_1 \right\rangle} \right)(\gamma, t) + O(\epsilon^4).
\end{aligned}$$

This expression generalizes the results presented in Proposition 4.1 in [14]. Note that those results were proven via a perturbation theory argument that is not as general and powerful as the treatment in Theorem 3.3.

4.2. Simple Hamiltonian systems with integrable forces. In this and the following section we analyze systems with more structure both in the affine connections ∇ as well as in the input forces Y. Here we consider systems with Lagrangian equal to "kinetic minus potential" and with integrable forces. In the interest of brevity, we refer to the textbooks [17, 35] for a detailed presentation and review here only the necessary notation. The affine connection of a simple system is the Levi-Civita connection associated with the kinetic energy matrix M, that is, the Christoffel symbols are defined according to the usual relationship

$$\Gamma_{ij}^k = \frac{1}{2} M^{mk} \left(\frac{\partial M_{mj}}{\partial q^i} + \frac{\partial M_{mi}}{\partial q^j} - \frac{\partial M_{ij}}{\partial q^m} \right), \tag{4.1}$$

where M_{ij} and M^{mk} are the components of the matrix representation of M and of its inverse. An integrable time-varying force is written as

$$Y(q, t) = \operatorname{grad} \varphi(q, t), \qquad \text{where} \qquad (\operatorname{grad} \varphi)^i = M^{ij} \frac{\partial \varphi}{\partial q_j}, \tag{4.2}$$

and where φ is a scalar function on $\mathbb{R}^n \times \mathbb{R}$.

One remarkable simplification takes place for a simple system described by a Levi-Civita connection: the set of gradient vector fields is closed under the operation of symmetric product. Let φ_1, φ_2 be scalar functions on $\mathbb{R}^n$ and define a symmetric product between functions according to

$$\langle \varphi_1 : \varphi_2 \rangle \triangleq \frac{\partial \varphi_1}{\partial q} M^{-1} \frac{\partial \varphi_2}{\partial q}. \tag{4.3}$$

Then the symmetric product of the corresponding gradient vector fields equals the gradient of the symmetric product of the functions. In equations:

$$\langle \operatorname{grad} \varphi_1 : \operatorname{grad} \varphi_2 \rangle = \operatorname{grad} \langle \varphi_1 : \varphi_2 \rangle .$$

We refer to [10] for the proof. Accordingly, the main theorem can be restated as follows.

THEOREM 4.1. *Consider the system as described in Problem 3.1. Additionally, let the Christoffel symbols and the input vector field be defined as in equation* (4.1) *and* (4.2). *Define recursively the time-varying functions:*

$$\begin{aligned} \varphi_1(q,t) &= \int_0^t \varphi(q,s) ds \\ \varphi_k(q,t) &= -\frac{1}{2} \sum_{j=1}^{k-1} \int_0^t \Big\langle \varphi_j(q,s) :\ \varphi_{k-j}(q,s) \Big\rangle ds, \qquad k \geq 2. \end{aligned}$$

Then the solution $\gamma : [0,T] \to Q$ satisfies

$$\dot\gamma(t) = \operatorname{grad} \sum_{k=1}^{+\infty} \varphi_k(\gamma(t), t). \tag{4.4}$$

In other words, the flow of a simple Hamiltonian system forced from rest is written as a (time-varying) gradient flow. For completeness, we include a convergence treatment derived from the one in the main theorem.

REMARK 4.2. *Given $0 < \sigma'' < \sigma' < \sigma$, we assume M and φ to be analytic in a neighborhood $B_\sigma(q_0)$ of q_0 and uniformly integrable in $t \in [0,T]$. Two immediate bounds are*

$$\begin{aligned} \|\operatorname{grad} \varphi\|_{\sigma'} &\leq n \|M^{-1}\|_{\sigma'} \left\| \frac{\partial \varphi}{\partial q} \right\|_{\sigma'} \\ \|\Gamma\|_{\sigma'} \leq A &\triangleq \frac{3n^2}{2(\sigma - \sigma')} \|M^{-1}\|_{\sigma'} \|M\|_{\sigma}. \end{aligned}$$

Accordingly, the bounds in the main theorem can be restated (in a more conservative manner) as follows. If

$$\left\| \frac{\partial \varphi}{\partial q} \right\|_{\sigma'} T^2 < \frac{1}{n\|M^{-1}\|_{\sigma'}} \min \left\{ \frac{\sigma' - \sigma''}{2^4 n^2 (n+1)}, \frac{1}{2^4 n (n+1) A}, \frac{\eta^2 (n^2 \sigma'' A)}{n^2 A} \right\},$$

the series $\sum_{k=1}^{\infty} \varphi_k(q,t)$ converges absolutely and uniformly in t and q for all $t \in [0,T]$ and for all q in a neighborhood $B_{\sigma''}(q_0)$ of q_0.

4.3. Invariant systems on Lie groups. In this section we briefly investigate systems with kinetic energy and input forces invariant under a certain group action. These systems have a configuration space G with the structure of an n dimensional matrix Lie group. Systems in this class include satellites, hovercraft, and underwater vehicles.

The equation of motion (2.7) decouples into a kinematic and dynamic equation in the configuration variable $g \in G$ and the body velocity[3] $v \in \mathbb{R}^n$. The kinematic equation can be written [4] as a matrix differential equation using matrix group notation $\dot{g} = g\widehat{v}$; we refer to [36] for the details. The dynamic equation, sometimes referred to as Euler-Poincarè, is

$$\dot{v}^i + \gamma^i_{jk} v^j v^k = y^i(t), \tag{4.5}$$

where the coefficients γ^i_{jk} are determined by the group and metric structure. The curve $y : [0, T] \mapsto \mathbb{R}^n$ denotes the time-varying forcing.

Within this setting, the result in Theorem 3.3 is summarized as follows. The solution to the equation (4.5) with initial condition $v(0) = 0$ is $v(t) = \sum_{k=1}^{\infty} v_k(t)$, where

$$\begin{aligned} v_1(t) &= \int_0^t y(s) ds \\ v_k(t) &= -\frac{1}{2} \sum_{j=1}^{k-1} \int_0^t \Big\langle v_j(s) :\ v_{k-j}(s) \Big\rangle ds, \qquad k \geq 2, \end{aligned}$$

and where the symmetric product between velocity vectors is $\langle x : y \rangle^i = -2\gamma^i_{jk} x^j y^k$. Local convergence for the series expansion can be easily established in this setting.

This result agrees and indeed supersedes the ones presented in [14] obtained via the perturbation method. The relationship of this case to the more general setting studied in Theorem 3.3 is clarified via the notion of invariant connection; see [5, Appendix B] and [40, Section 27, "Variations on a theme by Euler"] for more details.

4.4. Simulations for a three degree of freedom manipulator. In this section we illustrate the approximations derived in Theorem 3.3 by applying them to an example system. We consider a three link planar manipulator. The configuration is described by three angles $(\theta_1, \theta_2, \theta_3)$. A constant (integrable) force is applied to the first variable. Specifically, we set $\varphi(q, t) = \epsilon\theta_1$ and we let the parameter ϵ vary in the range 10^{-2} to 1. The integration time is $T = 1$ seconds. Setting all lengths, masses and moments of inertia to unity, the kinetic energy matrix is:

$$M = \frac{1}{16} \begin{bmatrix} 25 & 6\cos(\theta_1 - \theta_2) & 2\cos(\theta_1 - \theta_3) \\ 6\cos(\theta_1 - \theta_2) & 21 & 2\cos(\theta_2 - \theta_3) \\ 2\cos(\theta_1 - \theta_3) & 2\cos(\theta_2 - \theta_3) & 17 \end{bmatrix}.$$

The initial condition is assumed to be $q(0) = (0, \pi/4, 0)$. We investigate the error value $e_{\epsilon,N} = \|\gamma(T) - \gamma_N(T)\|$, where γ_N is the solution to the Nth order truncation: $\dot{\gamma}_N(t) = \operatorname{grad} \sum_{k=1}^{N} \varphi_k(\gamma_N, t)$. An empirical forecast of the $e_{\epsilon,N}$ is computed as

[3] More precisely, the body velocity v lives in the Lie algebra of the group G.

[4] Alternatively, the kinematic equation can be written in a system of local coordinates q (e.g., Euler angles in the case of rotation matrices) as $\dot{q} = J(q)v$, where $J(q)$ is an appropriate Jacobian matrix.

TABLE 4.1
Numerical comparison of various degrees of approximations. The entries in the table are the error values $e_{\epsilon,N}$ that provide a measure of the accuracy of the Nth order truncated approximation.

ϵ	1	.1	.01
$N=1$	$5.3 \cdot 10^{-3}$	$4.6 \cdot 10^{-5}$	$4.5 \cdot 10^{-7}$
$N=2$	$2.4 \cdot 10^{-4}$	$4.2 \cdot 10^{-7}$	$4.2 \cdot 10^{-10}$
$N=3$	$1.4 \cdot 10^{-4}$	$3.0 \cdot 10^{-9}$	$2.3 \cdot 10^{-13}$
$N=4$	$5.2 \cdot 10^{-5}$	$2.4 \cdot 10^{-10}$	$3.5 \cdot 10^{-15}$

follows. Since $T = 1$ and $\|Y\| = O(\epsilon)$, there exist two constants c, d such that the kth term in series is bounded by $c(d\epsilon)^k$. Summing the neglected contributions from $k = N+1$ to infinity and assuming that $d\epsilon \ll 1$, one can compute $e_{\epsilon,N} \approx c(d\epsilon)^{N+1}$. We summarize the results of the numerical investigation[5] in Table 4.1. The results are in qualitative agreement with the theoretical forecasts.

5. Conclusions. We have presented a series expansion that describes the evolution of a forced mechanical system. Our result provides a first order description to the solutions of a second order initial value problem. Both the series and the proof method provide insight into the geometry of mechanical control systems. The treatment expands on our previous work [10] on high amplitude high frequency averaging and vibrational stabilization.

Series expansions are the underlying technique for controllability and motion planning. For mechanical systems moving in the low velocity regime, these two problems have been tackled with various degrees of success in [33, 13]. Future research will rely on the contributions in this work to develop more general motion planning algorithms than the ones in [14], and sharper sufficient controllability tests than the ones in [33].

Acknowledgments. The author thanks Jim Radford and Andrew D. Lewis for helpful and stimulating discussions.

REFERENCES

[1] R. Abraham, J. E. Marsden, and T. S. Ratiu, *Manifolds, Tensor Analysis, and Applications*, vol. 75 of AMS, Springer Verlag, New York, NY, second ed., 1988.
[2] A. A. Agračhev and R. V. Gamkrelidze, *The exponential representation of flows and the chronological calculus*, Math. USSR Sbornik, 35 (1978), pp. 727–785.
[3] ———, *Local controllability and semigroups of diffeomorphisms*, Acta Applicandae Mathematicae, 32 (1993), pp. 1–57.
[4] G. E. Andrews, *Number Theory*, Dover Publications, New York, NY, 1994.
[5] V. I. Arnol'd, *Mathematical Methods of Classical Mechanics*, vol. 60 of GTM, Springer Verlag, New York, NY, second ed., 1989.
[6] J. Baillieul, *Stable average motions of mechanical systems subject to periodic forcing*, in Dynamics and Control of Mechanical Systems: The Falling Cat and Related Problems, M. J. Enos, ed., vol. 1, Field Institute Communications, 1993, pp. 1–23.
[7] A. L. Besse, *Manifolds all of whose Geodesics are Closed*, Springer Verlag, New York, NY, 1978.
[8] R. M. Bianchini and G. Stefani, *Graded approximations and controllability along a trajectory*, SIAM Journal on Control and Optimization, 28 (1990), pp. 903–924.
[9] A. M. Bloch and P. E. Crouch, *Nonholonomic control systems on Riemannian manifolds*, SIAM Journal on Control and Optimization, 33 (1995), pp. 126–148.

[5] The numerical integration is performed inside the Mathematica environment, specifying 16 digits of accuracy and 32 digits of working precision.

[10] F. Bullo, *Averaging and vibrational control of mechanical systems*, SIAM Journal on Control and Optimization, (1999). Submitted.
[11] ———, *A series describing the evolution of mechanical control systems*, in IFAC World Conference, vol. E, Beijing, China, July 1999, pp. 479–485.
[12] F. Bullo and N. E. Leonard, *Motion control for underactuated mechanical systems on Lie groups*, in European Control Conference, Brussels, Belgium, July 1997, p. 480.
[13] F. Bullo, N. E. Leonard, and A. D. Lewis, *Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups.* To appear, *IEEE Transactions on Automatic Control*. Also Technical Report CDS 97–013, California Institute of Technology, Nov. 1997.
[14] ———, *Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups*, IEEE Transactions on Automatic Control, 45 (2000), pp. 1437–1454.
[15] K. T. Chen, *Integration of paths, geometric invariants, and a generalized Baker-Hausdorff formula*, Annals of Mathematics, 67 (1957), pp. 164–178.
[16] W. A. Coppel, *Stability and Asymptotic Behavior of Differential Equations*, D.C. Heath, Boston, MA, 1965.
[17] M. P. Do Carmo, *Riemannian Geometry*, Birkhäuser, Boston, MA, 1992.
[18] M. Fliess, *Fonctionelles causales non linéares et indeterminées non commutatives*, Bull. Soc. Math., 109 (1981), pp. 3–40.
[19] P. D. Giamberardino, M. Djemai, S. Monaco, and D. Normand-Cyrot, *Exact steering and stabilization of a PVTOL aircraft*, in IEEE Conf. on Decision and Control, San Diego, CA, Dec. 1997, pp. 2049–2054.
[20] W. S. Gray and J. M. A. Scherpen, *Hankel operators and Gramians for nonlinear systems*, in IEEE Conf. on Decision and Control, Tampa, FL, Dec. 1998, pp. 1416–1421.
[21] H. Hermes, *Nilpotent and high-order approximations of vector field systems*, SIAM Review, 33 (1991), pp. 238–264.
[22] A. Isidori, *Nonlinear Control Systems*, Springer Verlag, New York, NY, third ed., 1995.
[23] M. Kawski, *High-order small-time local controllability*, in Nonlinear Controllability and Optimal Control, H. J. Sussmann, ed., Dekker, 1990, pp. 441–477.
[24] ———, *Geometric homogeneity and applications to stabilization*, in Nonlinear Control Systems Design Symposium (NOLCOS), Tahoe City, CA, July 1995, pp. 251–256.
[25] ———, *Nonlinear control and combinatorics of words*, in Geometry of Feedback and Optimal Control, B. Jakubczyk and W. Respondek, eds., Dekker, New York, NY, 1998, pp. 305–346.
[26] M. Kawski and H. J. Sussmann, *Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory*, in Operators, Systems, and Linear Algebra, U. Helmke, D. Pratzel-Wolters, and E. Zerz, eds., Teubner, Stuttgart, Germany, 1997, pp. 111–128.
[27] I. Kolmanovsky and N. H. McClamroch, *Stabilizing feedback laws for internally actuated multibody systems in-space*, Nonlinear Analysis, Theory, Methods & Applications, 26 (1996), pp. 1461–1479.
[28] S. G. Krantz, *Function Theory of Several Complex Variables*, Pure and Applied Mathematics, John Wiley and Sons, New York, NY, 1982.
[29] G. Lafferriere and H. J. Sussmann, *A differential geometric approach to motion planning*, in Nonholonomic Motion Planning, Z. Li and J. F. Canny, eds., Kluwer Academic Publishers, 1993, pp. 235–270.
[30] S. Lang, *Differentiable and Riemannian Manifolds*, Springer Verlag, third ed., 1995.
[31] N. E. Leonard and P. S. Krishnaprasad, *Motion control of drift-free, left-invariant systems on Lie groups*, IEEE Transactions on Automatic Control, 40 (1995), pp. 1539–1554.
[32] A. D. Lewis, *Simple mechanical control systems with constraints*, IEEE Transactions on Automatic Control, 45 (2000), pp. 1420–1436.
[33] A. D. Lewis and R. M. Murray, *Configuration controllability of simple mechanical control systems*, SIAM Journal on Control and Optimization, 35 (1997), pp. 766–790.
[34] W. Magnus, *On the exponential solution of differential equations for a linear operator*, Communications on Pure and Applied Mathematics, VII (1954), pp. 649–673.
[35] J. E. Marsden and T. S. Ratiu, *Introduction to Mechanics and Symmetry*, Springer Verlag, New York, NY, second ed., 1999.
[36] R. M. Murray, Z. X. Li, and S. S. Sastry, *A Mathematical Introduction to Robotic Manipulation*, CRC Press, Boca Raton, FL, 1994.
[37] R. M. Murray and S. S. Sastry, *Nonholonomic motion planning: Steering using sinusoids*, IEEE Transactions on Automatic Control, 5 (1993), pp. 700–726.
[38] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, third ed., 1987.
[39] J. A. Sanders and F. Verhulst, *Averaging Methods in Nonlinear Dynamical Systems*, Springer Verlag, New York, NY, 1985.
[40] D. H. Sattinger and O. L. Weaver, *Lie Groups and Algebras, with Applications to Physics,*

Geometry and Mechanics, vol. 61 of AMS, Springer Verlag, New York, NY, 1986.
[41] E. D. SONTAG AND H. J. SUSSMANN, *Time-optimal control of manipulators*, in IEEE Int. Conf. on Robotics and Automation, San Francisco, CA, Apr. 1986, pp. 1692–1697.
[42] H. J. SUSSMANN, *A sufficient condition for local controllability*, SIAM Journal on Control and Optimization, 16 (1978), pp. 790–802.
[43] ———, *A product expansion of the Chen series*, in Theory and Applications of Nonlinear Control Systems, C. I. Byrnes and A. Lindquist, eds., Elsevier, Oxford, UK, 1986, pp. 323–335.
[44] ———, *A general theorem on local controllability*, SIAM Journal on Control and Optimization, 25 (1987), pp. 158–194.
[45] V. S. VARADARAJAN, *Lie Groups, Lie Algebras, and Their Representations*, vol. 102 of GTM, Springer Verlag, New York, NY, 1984.

Appendix A. Some Basic Identities in Combinatorial Analysis. We here present a basic result and derive a useful expression that is needed in the proof of the main theorem. The main reference is the method of generating functions as described in Section 3.4 in [4]. The first identity is explicitly proven in the reference. If $N_1 = 1$ and

$$N_k = \sum_{j=1}^{k-1} N_j N_{k-j}, \qquad k \geq 2,$$

then

$$N_k = \frac{1}{k}\binom{2k-2}{k-1} = \frac{2^k}{(4k-2)} \frac{1 \cdot 3 \cdot 5 \cdots (2k-1)}{1 \cdot 2 \cdot 3 \cdots k} \leq \frac{4^k}{4k-2}. \tag{A.1}$$

The second equality needed in the proof of Theorem 3.3 is

$$\sum_{a=0}^{k} \binom{a+d_1}{d_1}\binom{k-a+d_2}{d_2} = \binom{k+1+d_1+d_2}{k}. \tag{A.2}$$

To prove it we use the method of generating functions, see [4]. We claim that, for all real x with $|x| < 1$

$$\sum_{k=0}^{\infty}\left(\sum_{a=0}^{k} \binom{a+d_1}{d_1}\binom{k-a+d_2}{d_2}\right) x^k = \sum_{k=0}^{\infty} \binom{k+1+d_1+d_2}{k} x^k. \tag{A.3}$$

The first step is to notice that

$$\begin{aligned}
\sum_{k=0}^{\infty}\left(\sum_{a=0}^{k} \binom{a+d_1}{d_1}\binom{k-a+d_2}{d_2}\right) x^k &= \sum_{k=0}^{\infty} \sum_{m+n=k} \binom{n+d_1}{d_1}\binom{m+d_2}{d_2} x^{n+m} \\
&= \left(\sum_{n=0}^{\infty}\binom{n+d_1}{d_1} x^n\right)\left(\sum_{m=0}^{\infty}\binom{m+d_2}{d_2} x^m\right).
\end{aligned}$$

Accordingly, we define

$$f_a(x) = \sum_{m=0}^{\infty} \binom{m+a}{a} x^m, \tag{A.4}$$

and the thesis in equation (A.3) is equivalent to proving that

$$f_{d_1}(x) f_{d_2}(x) = f_{d_1+d_2+1}(x). \tag{A.5}$$

In passing, we also note that the convergence radius of f is $|x| < 1$.

The second step is to study the properties of f. First of all,

$$\begin{aligned} f_a(x) &= \sum_{m=0}^{\infty} \frac{(m+a)!}{m!a!} x^m = \frac{1}{a!} \sum_{m=0}^{\infty} (m+a)\cdots(m+1) x^m \\ &= \frac{1}{a!} \frac{d^a}{dx^a} \sum_{m=0}^{\infty} x^{m+a} = \frac{1}{a!} \frac{d^a}{dx^a} \left(x^a \sum_{m=0}^{\infty} x^m \right) = \frac{1}{a!} \frac{d^a}{dx^a} \frac{x^a}{1-x}. \end{aligned}$$

Additionally, it is immediate to see that

$$f_0(x) = \frac{1}{1-x}, \qquad x f_0(x) = f_0(x) - 1,$$

and consequently

$$f_a(x) = \frac{1}{a!} \frac{d^a}{dx^a} \frac{1}{1-x} = \frac{1}{a!} \frac{d^a}{dx^a} f_0(x).$$

Finally, we prove by induction that

$$f_a(x) = f_0(x)^{a+1}. \tag{A.6}$$

At $a = 0$ the statement is obvious. We assume it true up to a and compute:

$$\begin{aligned} f_{a+1}(x) &= \frac{1}{(a+1)!} \frac{d^{a+1}}{dx^{a+1}} \frac{1}{1-x} = \frac{1}{(a+1)!} \frac{d^a}{dx^a} \left(\frac{1}{1-x} \right)^2 \\ &= \frac{1}{(a+1)!} \sum_{b=0}^{a} \binom{a}{b} \left(\frac{d^b}{dx^b} \frac{1}{1-x} \right) \left(\frac{d^{a-b}}{dx^{a-b}} \frac{1}{1-x} \right) \\ &= \frac{a!}{(a+1)!} \sum_{b=0}^{a} \left(\frac{1}{b!} \frac{d^b}{dx^b} \frac{1}{1-x} \right) \left(\frac{1}{(a-b)!} \frac{d^{a-b}}{dx^{a-b}} \frac{1}{1-x} \right) \\ &= \frac{1}{a+1} \sum_{b=0}^{a} \left(\frac{1}{1-x} \right)^{b+1} \left(\frac{1}{1-x} \right)^{a-b+1} = \left(\frac{1}{1-x} \right)^{a+2}. \end{aligned}$$

This concludes the proof of equation (A.6), which immediately implies equation (A.5) and the main thesis in equation (A.3).