LIE BRACKET EXTENSIONS AND AVERAGING: THE SINGLE-BRACKET CASE

Héctor J. Sussmann*and Wensheng Liu
Department of Mathematics
Rutgers University
New Brunswick, NJ 08903
Phone: (908) 932-5407 FAX: (908) 932-5530
e-mail: sussmann@hamilton.rutgers.edu
wliu@math.rutgers.edu

November 7, 1997

Abstract

We explain a general approximation technique for nonholonomic systems by discussing in detail a special example, chosen so as to illustrate some of the technical aspects of the general construction. The example considered is that of an extension of a two-input system obtained by adding a single bracket of degree five. This bracket is sufficiently complicated to exhibit some phenomena, such as multiplicity, that do not occur for brackets of lower degree.

*This author's work was supported in part by the National Science Foundation under NSF Grant DMS-8902994.

INTRODUCTION

In [19], we introduced a general approximation technique for nonholonomic systems which gives an explicit construction, for any given path γ in state space, of a sequence of admissible trajectories converging to γ, and in addition has a number of other desirable properties. The technique is based on an improved version of a theorem proved by Haynes and Hermes in [5], and uses a result on convergence of trajectories that extends earlier work ([7], [8], [9]) by Kurzweil and Jarnik. (Cf. Remark 2 for a discussion of this work.)

The purpose of this paper is to explain this technique and some of its properties, by discussing in detail a special example, selected so as to illustrate a number of nontrivial issues that arise in the general construction.

We will consider a two-input driftless control-linear system

$$\dot{x} = u_1(t)f_1(x) + u_2(t)f_2(x)\,, \tag{1}$$

together with a "Lie bracket extension" of (1) given by

$$\dot{x} = v_1(t)f_1(x) + v_2(t)f_2(x) + v_3(t)[[f_1, f_2], [f_1, [f_1, f_2]]](x)\,. \tag{2}$$

We will describe an explicit procedure for constructing, for each input $v(t) = (v_1(t), v_2(t), v_3(t))$ for (2), defined on an interval $[a, b]$, a sequence $\{u^j\}$ of inputs for (1) that generate, for each initial condition $x(a) = \bar{x}$, trajectories x^j that converge to the trajectory $\hat{x}$ of (2) generated by v. This is a particular case of the general algorithm presented in [19], which provides a similar approximation for arbitrary Lie bracket extensions of general m-input systems. We have selected this particular example because it is in a sense the simplest one whose structure is sufficiently rich to serve as a good illustration of the general situation. In particular, the bracket used in this extension has *multiplicity two* (cf. below), and is the simplest bracket whose multiplicity is nontrivial, i.e. > 1. As explained below, for brackets with nontrivial multiplicity some complications arise that would never be seen in the trivial case, and this is the main reason for our choice of the bracket $[[f_1, f_2], [f_1, [f_1, f_2]]]$.

The general definitions of [19] are as follows. A *Lie bracket extension* of an m-input system

$$\dot{x} = \sum_{k=1}^{m} u_k(t) f_k(x) , \tag{3}$$

where $f_1, \ldots, f_m$ are smooth vector fields on a smooth manifold M, is a system

$$\dot{x} = \sum_{k=1}^{r} v_k(t) f_k(x) , \tag{4}$$

where the first m vector fields f_k, $k = 1, \ldots, m$, are the same as those in (3), and the new directions $f_{m+1}, \ldots, f_r$ are Lie brackets of the f_k, $k \in \{1, \ldots, m\}$. A sequence of L^1 inputs $u^j : [a, b] \to \mathbb{R}^m$ for (3) *converges in the trajectory sense* (or, for short, *T-converges*) to an L^1 input $v : [a, b] \to \mathbb{R}^r$ for (4) if, for each initial condition $x(a) = p$, the trajectories x^j of (3) generated by the u^j converge uniformly to the trajectory $\hat{x}$ of (4) for v.[1] The algorithm of [19] produces, for each input v for (4), a sequence $\{u^j\}$ of inputs for (3) that T-converges to v. As formulated, the definition of T-convergence depends on the choice of M and of specific vector fields $f_1, \ldots, f_m$, but in fact our construction is "universal" in the sense that the u^j only depend on the formal structure of the Lie brackets used in the extension, but not on the specific choice of $M, f_1, \ldots, f_m$. (This will be given a precise meaning below, by defining the concept of "EI-convergence," but what it means in our special case is, simply, that for any given input v for (2) we will construct a sequence $\{u^j\}$ that T-converges to v for *all* M and *all* possible pairs (f_1, f_2) of sufficiently smooth vector fields on M.)

The situation studied in this paper is a special case of that analyzed in [19]. Here, rather than discuss arbitrary Lie bracket extensions, as was done in [19], we will consider the very simple "single-bracket" extension (2). We hope that a detailed description of the procedure for such a simple case will make the general

[1] In general, the trajectories of (3) and (4) might fail to be unique, or might have explosions. To take care of this, our precise definition of T-convergence is as follows: given a system (3) with state space M, and a Lie bracket extension (4), the sequence $\{u^j\}$ *T-converges to* v if for every point p, if $\hat{x} : [a, b] \to M$ is a trajectory of (4) for v that satisfies $\hat{x}(a) = p$, then for sufficiently large j there are trajectories $x^j : [a, b] \to M$ of (3) for u^j that satisfy $x^j(a) = p$, and $x^j \to \hat{x}$ uniformly as $j \to \infty$.

method clearer. We remark that the extension of the method from the single-bracket case to the general multibracket situation is nontrivial, since it gives rise to new technical issues. In the last section of this paper we briefly outline this generalization and some related technical problems, and we explain how our techniques also make it possible to handle certain systems with drift, and to approximate feedback control laws.

Explicit approximation algorithms are of interest for the following reason. It is clear that an extended system such as (4) gives rise to many more trajectories than the original system (3) does. Therefore, when we look for paths of a special kind (e.g. paths that join two given points and avoid certain obstacles), it will be easier to find them for (4) than for (3). In particular, if the right-hand side of (4) is so rich that direct control in *every* direction is possible, then every curve of class C^1 will be a trajectory of (4). The existence of an extension which is "sufficiently rich" in this sense is guaranteed if (3) satisfies the *Lie algebra rank condition* (LARC)[2]. Precisely, if the LARC holds, and K is any prescribed compact subset of M, then we can choose r large enough such that the set of vectors $\{f_k(x),\, k = 1, \ldots, r\}$ spans T_xM for every $x \in K$. On the other hand, the LARC is a natural condition for this kind of problems, because under fairly general conditions —for instance, real analyticity— it is equivalent to *complete controllability*, i.e. the property that given any two points p, q there is a trajectory from p to q. For a system (3) that satisfies the LARC, every smooth curve $\gamma : [a, b] \to M$ is a trajectory of a suitably chosen Lie bracket extension (4) for some input $v : [a, b] \to \mathbb{R}^r$. If we know how to produce a sequence $\{u^j\}$ of inputs for (3) that T-converges to v as $j \to \infty$, then we get, in particular, a solution of the problem of approximating any given smooth curve in M by trajectories of (3). If there is an "obstacle," i.e. a closed set C that we wish to avoid, then any solution of the path finding problem with obstacle avoidance for totally unrestricted paths will yield an approximate solution of the similar restricted problem, in which the paths are in addition required to be trajectories of (3). This leads to an alter-

[2]System (3) satisfies the LARC if for every $x \in \mathbb{R}^n$ we have $\Lambda(x) = T_xM$, where Λ is the Lie algebra of vector fields generated by the $f_k, k = 1, \ldots, m$, $\Lambda(x) = \{X(x) : X \in \Lambda\}$, and T_xM is the tangent space of M at x.

native approach to the Motion Planning Problem for nonholonomic systems studied, e.g., by Brockett, Sastry, Hauser, Murray, Gurvits, Li, Lafferriere and Sussmann (cf. [2], [3], [4], [6], [10], [11], [12], [13], [14], [15]).

Our strategy for constructing the desired approximation relies on a *T-convergence theorem* that gives sufficient conditions for T-convergence of a sequence of inputs u^j for (3) to an input v for (4). Applying this theorem, we can explicitly compute the limits of certain special sequences. A particular case is that of *highly oscillatory sequences* (HOS's) of inputs, defined below. An HOS of inputs involves several arbitrary functions η_ω, and the limit turns out to be given by an explicit formula in terms of the η_ω. The final step is then to realize that this formula implies that any desired limit v can be achieved by a suitable choice of the η_ω.

Our T-convergence theorem generalizes a classical convergence result. This classical result says that, if $\{u^j\}$ is a sequence of inputs belonging to $L^1([a,b],\mathbb{R}^m)$, and if x^j is the solution of (3) corresponding to u^j, with an initial condition $x^j(a) = p$, then the x^j converge uniformly on $[a,b]$ to the solution x^∞ of

$$\begin{aligned} \dot{x}(t) &= \sum_{k=1}^{m} u_k^\infty(t) f_k(x(t)) \,, \\ x(a) &= p \,, \end{aligned} \tag{5}$$

provided that

(a) the vector fields f_k are locally Lipschitz,

(b) x^∞ exists on the whole interval $[a,b]$,

(c) the indefinite integrals $U^j(t) = \int_a^t u^j(s)\,ds$ converge uniformly on $[a,b]$ to $U^\infty(t) = \int_a^t u^\infty(s)\,ds$,

(d) the functions u^j are uniformly bounded in L^1, i.e. there exists a finite constant C such that $\int_a^b \|u^j(t)\|dt \le C$ for all j.

Remark 1. The most widely known form of the classical convergence theorem is a slightly weaker statement, in which conditions

(c) and **(d)** are replaced by the stronger hypothesis that the u^j converge weakly in L^1 to u^∞. It is not hard to see that **(c)** and **(d)** together are strictly weaker than weak convergence. Indeed, weak convergence is equivalent to uniform convergence of the integrals plus equiintegrability, so it implies **(c)** and **(d)**, but the converse need not be true, since a sequence satisfying **(c)** and **(d)** need not be equiintegrable. ∎

The above classical convergence theorem can be restated as a "first-order" T-convergence result: if **(a)**, **(c)** and **(d)** hold then the u^j T-converge to the *ordinary* input u^∞, so that the limiting equation contains no new Lie brackets. Our generalization is a high-order T-convergence theorem in which limiting equations involving brackets are obtained. It is clear that, for any such generalization, Condition **(c)** will have to be kept. Indeed, if the u^j T-converge to v for all possible choices of sufficiently smooth $f_1, \ldots, f_m$, then they have to T-converge to v in the special case when $f_1, \ldots, f_m$ are linearly independent constant vector fields on $\mathbb{R}^m$, and this suffices to imply that **(c)** holds if we let $u^\infty = (v_1, \ldots, v_m)$. So our extension will have to be based on dropping the boundedness condition **(d)**. It is not hard to see that, if **(d)** fails, then the x^j may fail to converge to x^∞, but it is possible for them to converge to a trajectory of a Lie bracket extension (4), which could fail to be a trajectory of (3). Moreover, this can happen in the "universal" way described above, so that in fact the u^j T-converge to some v for all possible choices of sufficiently smooth $f_1, \ldots, f_m$.

For a simple illustration of this phenomenon, consider first the system

$$\dot{x}_1 = u_1 \ , \ \dot{x}_2 = u_2 \ , \ \dot{x}_3 = u_2 x_1 \ , \tag{6}$$

on an interval $[0, T]$, with initial condition $(0, 0, 0)$. If we let

$$u^j(t) = \sqrt{j}(\cos jt, \sin jt) \ ,$$

then the u^j "converge to 0" in the sense that their indefinite integrals $U^j(t) = \int_0^t u^j(s)\, ds$ converge to 0 uniformly. So, if we let $u^\infty \equiv (0, 0)$, we might think that the u^j T-converge to u^∞, so that the $x^j(t)$ converge uniformly to $x^\infty(t) \equiv (0, 0, 0)$. However, a simple calculation shows that this is not so, and in fact $x^j(t)$ converges uniformly to $\hat{x}(t) = (0, 0, \frac{t}{2})$. On the other hand, the vector fields f_1, f_2 have components $(1,0,0)$ and $(0,1,x_1)$, respectively. So the Lie bracket $f_3 = [f_1, f_2]$ has components $(0,0,1)$, and the limiting function $\hat{x}$ turns out to satisfy $\dot{x} = \frac{1}{2}f_3(x)$. Moreover, one can verify by an integration by parts that for *every* two-input system (1) with C^2 vector fields f_1, f_2, and every initial condition $x(0)$, the trajectories x^j that correspond to the u^j converge to the solution of $\dot{x} = \frac{1}{2}[f_1, f_2](x)$, provided that the latter exists on $[0,T]$. In other words, the ordinary $\mathbb{R}^2$-valued inputs u^j for (1) T-converge to the $\mathbb{R}^3$-valued input $v = (0,0,\frac{1}{2})$ for $\dot{x} = v_1 f_1(x) + v_2 f_2(x) + v_3[f_1, f_2](x)$ for all choices of C^2 vector fields f_1, f_2. We can restate this in vector-field-free language by introducing *formal indeterminates* X_1, X_2, to be thought of as "slots" where vector fields can be plugged in. Our example then simply says that the ordinary inputs $u^j(t) = \sqrt{j}\cos jt\, X_1 + \sqrt{j}\sin jt\, X_2$ T-converge to the "extended input" $\frac{1}{2}[X_1, X_2]$ for all choices of C^2 vector fields f_1,f_2. We rise to an even higher level of abstraction and completely eliminate the vector fields, by just saying that the $u^j(t) = \sqrt{j}\cos jt\, X_1 + \sqrt{j}\sin jt\, X_2$ *EI-converge* (that is, "converge in the extended input sense") to $\frac{1}{2}[X_1, X_2]$.

Our general convergence theorem gives sufficient conditions for EI-convergence of a sequence of ordinary inputs u^j to an extended input v. An application of this theorem will give us a sufficient condition for inputs $u^j = u_1^j X_1 + u_2^j X_2$ to EI-converge on an interval $[a,b]$ to an extended input

$$\mathbf{v} = v_1 X_1 + v_2 X_2 + v_3[[X_1, X_2], [X_1, [X_1, X_2]]],$$

where v_1, v_2, v_3 are C^1 functions on $[a,b]$. Using this, we will explicitly compute the limit of certain HOS's involving twelve unknown functions, and then it will be obvious how to choose these functions so as to match any given $\mathbf{v}$ of the above form, for any C^1 functions v_1, v_2, v_3.

Remark 2. Haynes and Hermes proved, in [5], that, if a driftless system (3) with state space M satisfies the Lie algebra rank condition, then every continuous curve in M can be uniformly approximated by trajectories of (3). Kurzweil and Jarnik found ([7], [8], [9]) more restrictive versions of our general convergence result (Theorem 2). The idea of making systematic use of Lie bracket extensions was introduced in Lafferriere-Sussmann [11]. ∎

SOME ALGEBRAIC PRELIMINARIES

To clarify the presentation, it will be convenient to introduce some algebraic machinery. We follow the notations and definitions of [17]. For simplicity, we restrict ourselves to the two-input case.

We let $\mathbf{X} = \{X_1, X_2\}$ be a sequence of two objects, that will be called *indeterminates*. We use $A(\mathbf{X})$, $L(\mathbf{X})$ to denote, respectively, the free associative algebra in X_1, X_2 and the Lie subalgebra of $A(\mathbf{X})$ generated by X_1, X_2. (Then $L(\mathbf{X})$ can be identified with the free Lie algebra in X_1, X_2.) In simpler terms, the elements of $A(\mathbf{X})$ are *formal noncommutative polynomials* in X_1, X_2, i.e. finite linear combinations such as $3X_1 - 7X_2 + 4X_1^2 + 8X_1X_2 + 5X_2X_1$ of *formal monomials* (cf. below) in X_1, X_2. The elements of $L(\mathbf{X})$ are the *Lie polynomials* in X_1, X_2, i.e. the $P \in A(\mathbf{X})$ that can be written as linear combinations of Lie brackets of X_1 and X_2. (For example, X_1X_2 is not in $L(\mathbf{X})$, but $X_1X_2 - X_2X_1$, which by definition is equal to $[X_1, X_2]$, is in $L(\mathbf{X})$. Other examples of members of $L(\mathbf{X})$: $3X_1 - 2X_2 - [X_1, X_2]$, $[[X_1, X_2], [X_1, [X_1, X_2]]]$.)

We also let $\hat{A}(\mathbf{X})$, $\hat{L}(\mathbf{X})$ be, respectively, the algebra of *formal noncommutative power series* and the Lie algebra of *Lie series* in X_1, X_2. The elements of $\hat{A}(\mathbf{X})$ are the *noncommutative power series* in X_1, X_2, i.e. arbitrary, not necessarily finite, linear combinations of monomials, and the elements of $\hat{L}(\mathbf{X})$ are those power series that are actually *Lie series*, i.e. possibly infinite linear combinations of Lie brackets.

A *monomial* in X_1, X_2 is a formal product such as $X_1X_2X_1$, or $X_2X_1X_1X_1X_1X_2X_1$. (As usual, we use power notation for repeated products, so the latter monomial is written more simply as $X_2X_1^4X_2X_1$.) Monomials are multiplied by concatenation so that, for instance, if $M_1 = X_1X_2X_1$ and $M_2 = X_2X_1^4X_2X_1$, then $M_1M_2 = X_1X_2X_1X_2X_1^4X_2X_1$. (Notice that $M_1M_2 \neq M_2M_1$.) Multiplication of monomials is associative, i.e. $(MN)P = M(NP)$ for any monomials M, N, P. More precisely, monomials can be labeled by multiindices $I = (\ell_1, \ldots, \ell_k)$ whose entries ℓ_s are integers equal to 1 or 2. We write

$$X_I \stackrel{\text{def}}{=} X_{\ell_1} X_{\ell_2} \cdots X_{\ell_k}. \tag{7}$$

Use $|I|$ to denote the length of I, i.e. $|I| = k$ if $I = (\ell_1, \ldots, \ell_k)$. Then it is clear that, as I ranges over all possible multiindices I of lengths $|I| \geq 0$, the X_I form a basis of $A(\mathbf{X})$. (The only index of length 0 is the empty index $\emptyset$, and then $X_\emptyset \stackrel{\text{def}}{=} 1$.) So every element P of $A(\mathbf{X})$ has a unique expression as a finite linear combination $\sum_I P_I X_I$ of the X_I. In both $A(\mathbf{X})$ and $\hat{A}(\mathbf{X})$, addition is done componentwise, and multiplication is carried out using the formula $X_I X_J = X_{IJ}$, where IJ is the concatenation of I and J, i.e. the multiindex obtained by writing, in order, first the components of I and then those of J.

The Lie Algebra $L(\mathbf{X})$ is spanned by the *brackets* of X_1 and X_2. Precisely we define $\mathcal{B}r(\mathbf{X})$ to be the smallest subset of $L(\mathbf{X})$ that contains X_1, X_2 and is closed under bracketing. The elements of $\mathcal{B}r(\mathbf{X})$ will be referred to as *brackets* of $\mathbf{X}$.

Remark 3. Later on we will also have to consider the set $\mathcal{FB}r(\mathbf{X})$ of *formal brackets*, which are certain formal expressions in the symbols X_1, X_2, the comma and the left and right square brackets. (Precisely, $\mathcal{FB}r(\mathbf{X})$ is the smallest set S of such expressions that contains X_1 and X_2, and is such that, whenever P and Q are in S, it follows that '$[P, Q]$' is in S.) The elements of $\mathcal{FB}r(\mathbf{X})$ are *purely formal expressions*, whereas those of $\mathcal{B}r(\mathbf{X})$ are elements of $L(\mathbf{X})$. To see the difference, notice that, for instance, $[X_1, [X_1, X_2]]$ and $[[X_2, X_1], X_1]$ are the same element of $\mathcal{B}r(\mathbf{X})$ —as can be seen by two applications of the skew-symmetry law $[P, Q] = -[Q, P]$— but are different as elements of $\mathcal{FB}r(\mathbf{X})$. Formal brackets have a *unique factorization property*: every $B \in \mathcal{FB}r(\mathbf{X})$ except X_1 and X_2 can be written in a unique way as a bracket $[B_1, B_2]$, with $B_1, B_2 \in \mathcal{FB}r(\mathbf{X})$. (The brackets B_1, B_2 are, respectively, the *left and right factors* of B.) This fails for "real brackets." For example, if $B = [X_1, [X_1, X_2]]$ then B has the unique factorization $B = [B_1, B_2]$ in $\mathcal{FB}r(\mathbf{X})$, with $B_1 = X_1$ and $B_2 = [X_1, X_2]$, but the factorization in $\mathcal{B}r(\mathbf{X})$ is not unique, because we can also write $B = [\tilde{B}_1, \tilde{B}_2]$ with $\tilde{B}_1 = [X_2, X_1]$, $\tilde{B}_2 = X_1$. ∎

For each bracket $B \in \mathcal{B}r(\mathbf{X})$ we use $\delta(B)$ to denote the total degree of B, and let $\delta_k(B), k = 1, 2$ be the degree of B with respect to X_k, so $\delta_1(B) + \delta_2(B) = \delta(B)$. For example if $B = [X_1, [X_1, X_2]]$, then $\delta_1(B) = 2$, $\delta_2(B) = 1$, and $\delta(B) = 3$.

An *extended input* on an interval $[a, b]$ is an $\hat{L}(\mathbf{X})$-valued, integrable function on $[a, b]$. (If $\mathbf{v}$ is any $\hat{A}(\mathbf{X})$-valued function on $[a, b]$, we say that $\mathbf{v}$ is *integrable* if all the coefficient functions v_I in the expression $\mathbf{v}(t) = \sum_I v_I(t) X_I$ are integrable on $[a, b]$. This makes sense in particular if $\mathbf{v}$ is $\hat{L}(\mathbf{X})$-valued.) An ordinary input $u = (u_1, u_2)$ can be regarded as an extended input by identifying it with the function $\mathbf{u}(t) = u_1(t)X_1 + u_2(t)X_2$. As explained above, an extended input such as $v_1(t)X_1 + v_2(t)X_2 + v_3(t)[X_1, X_2]$ should be thought of as an object that contains "slots" (the indeterminates X_1, X_2) where vector fields can be plugged in, giving rise to an ordinary differential equation. For example, the extended input $v_1(t)X_1 + v_2(t)X_2 + v_3(t)[X_1, X_2]$ gives rise, for each choice of sufficiently smooth vector fields f_1, f_2, to the differential equation

$$\dot{x} = v_1(t) f_1(x) + v_2(t) f_2(x) + v_3(t)[f_1, f_2](x) .$$

We say that an extended input $\mathbf{v}$ is *of order* $\leq k$ if it only involves brackets of degree $\leq k$. (If $\mathbf{v}$ is of order $\leq k$ for some integer k then we say that v is *of finite order*. If k is the smallest integer for which this happens, then we say that v is of order k.) An extended input v of order $\leq k$ can be "plugged" into any system (1), if the vector fields f_1, f_2 are of class C^κ with $\kappa \geq k - 1$. The result is a time-varying differential equation whose right-hand side is of class $C^{\kappa+1-k}$ with respect to the state variable x. (In particular, if $\kappa \geq k$ then the right-hand side is locally Lipschitz in x and satisfies the conditions of the existence and uniqueness theorem of Carathéodory.) Since, for each choice of sufficiently smooth vector fields, extended inputs of finite order give rise to trajectories for each initial condition, the concept of *T-convergence* of such inputs is well defined.

Let $\mathbf{v}$ be an extended input of finite order k, and let r be an integer such that $r \geq k - 1$. A sequence $\{u^j\}$ of ordinary inputs will be said to *EI-converge to* $\mathbf{v}$ *for systems of class* C^r if it T-converges to $\mathbf{v}$ for all choices of vector fields f_1, f_2 of class C^r. (The requirement that $r \geq k - 1$ is needed to make sure that the vector fields can be plugged into $\mathbf{v}$.)

With this terminology, we can reformulate the approximation problem as follows:

(AP) Given an extended input $\mathbf{v}$ of order $\leq r$, find a sequence of ordinary inputs that EI-converges to $\mathbf{v}$ for systems of class C^r.

As indicated in the introduction, the basic idea for solving **(AP)** is to use *highly oscillatory sequences* (HOS's) of inputs. An HOS is a sequence $\{u^j\}$ whose components u_1^j, u_2^j are finite linear combinations of expressions such as $j^\alpha \eta_\omega(t)\cos(j\omega t)$ and $j^\alpha \eta_\omega(t)\sin(j\omega t)$. The HOS's used in our general approximation algorithm involve a finite set of frequencies ω chosen so as to satisfy some special resonance conditions. In the general treatment given in [19], where $\mathbf{v}$ is an arbitrary extended input of finite order, we proceed by trying to handle each bracket in $\mathbf{v}$ separately. We split $\mathbf{v}$ into a sum $\sum_{\nu_1,\nu_2} \mathbf{v}_{\nu_1,\nu_2}$ of "totally homogeneous components" (cf. below) and then try to produce, for each $\mathbf{v}_{\nu_1,\nu_2}$, an HOS $\{u^j_{\nu_1,\nu_2}\}$ that EI-converges to $\mathbf{v}_{\nu_1,\nu_2}$. We then let $u^j = \sum_{\nu_1,\nu_2} u^j_{\nu_1,\nu_2}$, and hope that this will work. It turns out that, even though the input-to-trajectory map is highly nonlinear, a kind of "high frequency superposition principle" holds, and the u^j EI-converge to $\mathbf{v}$, *provided that the frequencies associated to the various components* $\mathbf{v}_{\nu_1,\nu_2}$ *are independent* in a sense that will be made precise below.

The precise definition of the $\mathbf{v}_{\nu_1,\nu_2}$ is as follows. Let $L(\mathbf{X})_{\nu_1,\nu_2}$ denote the linear span of all the brackets $B \in \mathcal{B}r(\mathbf{X})$ such that $\delta_1(B) = \nu_1$ and $\delta_2(B) = \nu_2$. Then $L(\mathbf{X})$ is the direct sum of the spaces $L(\mathbf{X})_{\nu_1,\nu_2}$, so every element $w \in L(\mathbf{X})$ has a unique decomposition as a sum $\sum_{\nu_1,\nu_2} w_{\nu_1,\nu_2}$, $w_{\nu_1,\nu_2} \in L(\mathbf{X})_{\nu_1,\nu_2}$. Similarly, an extended input $\mathbf{v} : [a,b] \to L(\mathbf{X})$ has a unique decomposition as a sum $\sum_{\nu_1,\nu_2} v_{\nu_1,\nu_2}$, where each v_{ν_1,ν_2} is $L(\mathbf{X})_{\nu_1,\nu_2}$-valued. These v_{ν_1,ν_2} are the *totally homogeneous components* of $\mathbf{v}$. If $B \in \mathcal{B}r(\mathbf{X})$, $\delta_1(B) = \nu_1$ and $\delta_2(B) = \nu_2$, then the *multiplicity* of B is the dimension of $L(\mathbf{X})_{\nu_1,\nu_2}$.

For the special case of the Lie bracket extension (2) to be discussed in this paper, the right-hand side of the extension only involves one bracket. Let us write $B_0 = [[X_1, X_2], [X_1, [X_1, X_2]]]$. Our goal is to exhibit, for every triple of functions v_1, v_2, v_3 of class C^1 on some interval $[a, b]$, a sequence of ordinary inputs that EI-converges to the extended input $\bar{\mathbf{v}} \stackrel{\text{def}}{=} v_1(t)X_1 + v_2(t)X_2 + v_3(t)B_0$. The bracket B_0 has degrees $\delta_1(B_0) = 3$, $\delta_2(B_0) = 2$. It is easy to verify that the component $L(\mathbf{X})_{3,2}$ is *two-dimensional*, with a basis given by the two brackets B_0 and B_1, where $B_1 = [X_2, [X_1, [X_1, [X_1, X_2]]]]$. So B_0 has multiplicity two.

Unfortunately, *the "high frequency superposition principle" mentioned above does not make it possible to separate linearly independent brackets in the same totally homogeneous component of* $L(\mathbf{X})$. This creates an additional technical complication, that is resolved in [19] as follows. We declare two brackets *equivalent* if they belong to the same totally homogeneous component. Then, for each component $E = L(\mathbf{X})_{\nu_1,\nu_2}$ we use an HOS $\{u^j\}$ with a number of frequencies equal to $\mu(\nu_1 + \nu_2)$, where $\mu = \dim E$, and the frequencies are grouped into μ groups Ω^k of $\nu_1 + \nu_2$ frequencies each, of which ν_1 are allocated to u_1 and ν_2 to u_2. These frequencies are required to satisfy certain resonance and nonresonance conditions. If $P_1, \ldots, P_\mu$ is a suitably chosen basis of E (a P. Hall basis, to be precise, cf. below), then it turns out that we can produce a sequence $\{u^{E,j}\}$ that EI-converges to a limit $\sum_\ell \psi_\ell P_\ell$, where (i) each ψ_ℓ is in turn a sum $\sum_k \xi_k g_{P_\ell}(\Omega^k)$, (ii) the ξ_k are certain functions that can be chosen arbitrarily, and (iii) each $\Omega \to g_{P_\ell}(\Omega)$ is a rational function of $(\nu_1 + \nu_2)$-tuples Ω of frequencies. One then needs the algebraic fact that the matrix $\Big(g_{P_\ell}\big(\Omega^k\big)\Big)_{1\le k,\ell\le\mu}$ can be chosen to be invertible. Using this, one can clearly find ξ_k such that the limit of the $u^{E,j}$ is any desired E-valued extended input $\mathbf{v}$. (The general proof of this algebraic fact seems to be quite complicated, but in our special case we will be able to verify it easily.)

THE APPROXIMATING INPUTS

We now turn to the details, and give a precise description of the approximating sequence $\{u^j\}$.

Without loss of generality, we will fix an interval $[0, T]$ and assume that all the inputs, ordinary or extended, are defined on $[0, T]$.

We follow the prescription of [19]. Since $\delta_1(B_0) = 3$, $\delta_2(B_0) = 2$, and the multiplicity of B_0 is two, we take two 5-tuples

$$\bar{\Omega} = (w_1,\, w_2\,, w_3\,, w_4\,, w_5)\ ,\ \ \bar{\Omega}' = (w_1',\, w_2'\,, w_3'\,, w_4'\,, w_5')\ , \tag{8}$$

each one consisting of five different frequencies, three of which (the w_k and w_k' for $k = 1, 2, 3$) are allocated to u_1, while the other two are assigned to u_2. We require:

(ch1) $\bar{\Omega}$ and $\bar{\Omega}'$ are minimally canceling.

(ch2) the 5-tuples $\bar{\Omega}$ and $\bar{\Omega}'$ are independent.

(An m-tuple $\Omega = (\omega_1, \ldots, \omega_m)$ of real numbers is called *minimally canceling* (MC) if, whenever $\{a_k\}_{k=1}^m$ is a collection of integers such that $\sum_{k=1}^m |a_k| \leq m$, then it follows that $\sum_{k=1}^m a_k\omega_k = 0$ if and only if all the a_k are all equal. Two m-tuples $\Omega = (\omega_1, \ldots, \omega_m)$, $\Omega' = (\omega_1', \ldots, \omega_m')$ are *independent* if, whenever $\{a_k\}_{k=1}^m$ and $\{a_k'\}_{k=1}^m$ are two collections of integers such that $\sum_{k=1}^m (|a_k| + |a_k'|) \leq m$ and $\sum_{k=1}^m (a_k\omega_k + a_k'\omega_k') = 0$, it follows that $\sum_{k=1}^m a_k\omega_k = \sum_{k=1}^m a_k'\omega_k' = 0$.)

In other words, Condition **(ch1)** says that the five frequencies in $\bar{\Omega}$ and those in $\bar{\Omega}'$ "resonate" (in the sense that their sum is zero) but no resonance occurs with other combinations of five or fewer frequencies, even with repetitions. (So that, for instance, $w_1 + w_2 + w_3 + 2w_4 \neq 0$, $2w_1 + w_2 - w_4 \neq 0$, $2w_2 - 3w_5 \neq 0$, $w_2 + w_3 \neq 0$, etc.) Condition **(ch2)** says that no resonance can be produced by picking five frequencies in a "mixed" way, i.e. some from $\bar{\Omega}$ and some from $\bar{\Omega}'$.

From now on, the ten numbers w_k, w_k' will be fixed. The letter ω, and symbols such as ω_1, ω_2, will be used to denote *variables* ranging over frequencies.

We now let

$$\begin{aligned} \Omega(1) &= \{\pm w_1, \pm w_2, \pm w_3, \pm w_1', \pm w_2', \pm w_3'\}\,, & (9)\\ \Omega(2) &= \{\pm w_4, \pm w_5, \pm w_4', \pm w_5'\}\,, & (10)\\ \Omega &= \Omega(1) \cup \Omega(2)\,, & (11) \end{aligned}$$

and think of the 12 frequencies in $\Omega(1)$ as being assigned to u_1, and the 8 frequencies in $\Omega(2)$ as allocated to u_2.

We then take the approximating control sequence to be

$$u_1^j(t) = \eta_{1,0}(t) + j^{\frac{4}{5}} \sum_{\omega\in\Omega(1)} \eta_\omega(t) e^{ij\omega t}, \tag{12}$$

$$u_2^j(t) = \eta_{2,0}(t) + j^{\frac{4}{5}} \sum_{\omega\in\Omega(2)} \eta_\omega(t) e^{ij\omega t}, \tag{13}$$

where

(ch3) the functions $\eta_{k,0}, \eta_\omega$ are of class C^1 on $[0,T]$, $\eta_{k,0}$ are real valued, η_ω are complex valued, and $\eta_{-\omega} = \bar{\eta}_\omega$.

(As usual, $\bar{\cdot}$ stands for "complex conjugate," and $i = \sqrt{-1}$. Notice that the various terms of (12) and (13) are in principle complex, but Condition **(ch3)** implies that the u_k^j are real. The exponent $\frac{4}{5}$ is chosen because our bracket B_0 has degree five. In general, for brackets of degree ν the exponent would be $\frac{\nu-1}{\nu}$.)

We then have:

Theorem 1 *Let $u^j(t) = (u_1^j(t), u_2^j(t))$ be the control sequence defined in (12) and (13) with $\bar{\Omega}$, $\bar{\Omega}'$, $\Omega(1)$, $\Omega(2)$, the w's and the η's satisfying (8), (9), (10),* **(ch1)**, **(ch2)**, **(ch3)**. *Then $\{u^j\}$ EI-converges to*

$$\begin{aligned} \mathbf{u}^\infty(t) &= \eta_{1,0}(t)X_1 + \eta_{2,0}(t)X_2 \\ &\quad + \Big(g_{B_0}(\bar{\Omega})\xi(t) + g_{B_0}(\bar{\Omega}')\tilde{\xi}(t)\Big)B_0 \\ &\quad + \Big(g_{B_1}(\bar{\Omega})\xi(t) + g_{B_1}(\bar{\Omega}')\tilde{\xi}(t)\Big)B_1 \,, \end{aligned} \tag{14}$$

where

$$\xi(t) = \eta_{w_1}(t)\ldots\eta_{w_5}(t) + \overline{\eta_{w_1}(t)\ldots\eta_{w_5}(t)}, \tag{15}$$

$$\tilde{\xi}(t) = \eta_{w_1'}(t)\ldots\eta_{w_5'}(t) + \overline{\eta_{w_1'}(t)\ldots\eta_{w_5'}(t)}, \tag{16}$$

and the functions g_{B_0}, g_{B_1} are defined by

$$g_{B_0}(\omega_1,\ldots,\omega_5) = \frac{1}{\omega_1\omega_2\omega_3}\Big[\frac{1}{\omega_1+\omega_4} + \frac{1}{\omega_1+\omega_5} + \frac{1}{\omega_2+\omega_4}$$

$$+\frac{1}{\omega_2+\omega_5}+\frac{1}{\omega_3+\omega_4}+\frac{1}{\omega_3+\omega_5}\Big],$$

$$g_{B_1}(\omega_1,\ldots,\omega_5) = \frac{1}{\omega_1\omega_2\omega_3}\left[\frac{1}{\omega_4}+\frac{1}{\omega_5}\right]. \blacksquare$$

It is now easy to see how to choose the $\eta_{k,0},\eta_\omega$ so that the limiting input $\mathbf{u}^\infty$ has a desired value $\bar{\mathbf{v}} = v_1X_1 + v_2X_2 + v_3B_0$. It suffices to take $\eta_{1,0} = v_1$, $\eta_{2,0} = v_2$, $\eta_{w_\ell}(t) = \eta_{w'_\ell}(t) = 1$ for $\ell = 1,2,3,4$, and $\eta_{w_5}(t) = \frac{\zeta(t)}{2}$, $\eta_{w'_5}(t) = \frac{\tilde{\zeta}(t)}{2}$, where the functions ζ, $\tilde{\zeta}$ satisfy the system of linear equations

$$g_{B_0}(\bar{\Omega})\zeta(t) + g_{B_0}(\bar{\Omega}')\tilde{\zeta}(t) = v_3(t)\,, \tag{17}$$

$$g_{B_1}(\bar{\Omega})\zeta(t) + g_{B_1}(\bar{\Omega}')\tilde{\zeta}(t) = 0\,. \tag{18}$$

To be sure that the solution of (17) and (18) exists for arbitrary v_3, we need the extra condition

$$\det\begin{pmatrix} g_{B_0}(\bar{\Omega}) & g_{B_0}(\bar{\Omega}') \\ g_{B_1}(\bar{\Omega}) & g_{B_1}(\bar{\Omega}') \end{pmatrix} \neq 0\,. \tag{19}$$

So we need to choose $\bar{\Omega}$ and $\bar{\Omega}'$ such that (19) holds in addition to (**ch1**) and (**ch2**). It is easy to see that this is always possible. Indeed, if we let H be the hyperplane in $\mathbb{R}^5$ consisting of those 5-tuples $\omega = (\omega_1,\ldots,\omega_5)$ such that $\omega_1+\ldots+\omega_5 = 0$, then g_{B_0} and g_{B_1} are rational functions on H, and are well defined and analytic on a relatively open and dense subset of H. If we take $\omega = (1,1,1,1,1)$, then $g_{B_0}(\omega) = 3$ and $g_{B_1}(\omega) = 2$. If we choose instead $\omega = (1,1,1,2,2)$, then $g_{B_0}(\omega) = 2$ and $g_{B_1}(\omega) = 1$. So (19) holds for at least one pair of 5-tuples of frequencies in H, and hence for all pairs in a relatively open dense subset of $H\times H$. Conditions **(ch1)** and **(ch2)** also hold on a relatively open dense subset of $H\times H$, so it is possible to satisfy them simultaneously with (19).

So we impose the extra requirement that

(ch4) $\bar{\Omega}$ and $\bar{\Omega}'$ satisfy (19).

Then for this choice of the frequencies $\bar{\Omega}$ and $\bar{\Omega}'$ we can always choose $\eta_{k,0},\eta_\omega$ as above to make the limiting input $\mathbf{u}^\infty$ in (14) be identical to $\bar{\mathbf{v}}$. This completes the description of the approximating HOS $\{u^j\}$. In the following sections we will show that the u^j EI-converge to $\mathbf{u}^\infty$.

THE CONVERGENCE THEOREM

Our first step is to state and prove a general convergence theorem, giving conditions under which a sequence $\{u^j\}$ of ordinary inputs EI-converges to an extended input $\mathbf{v}$ of finite order.

Given any sequence $\{u^j\}$ in $L^1([0,T],\mathbb{R}^2)$, let x^j be the solution of the initial value problem

$$\dot{x}(t) = u_1^j(t)f_1(x) + u_2^j(t)f_2(x) \ , \ \ x(0) = x_0 \,, \tag{20}$$

where f_1, f_2 are smooth vector fields. Rewrite (20) as an integral equation

$$x^j(t) = x_0 + \sum_{\ell=1}^{2} \int_0^t u_\ell^j(s) f_\ell\Big(x^j(s)\Big) ds \,.$$

Assume that we can "remove the convergent part" from $\int_0^t u_\ell^j(s)ds$ by finding a function $v_\ell \in L^1[0,T]$ with the property that the integral $\int_0^t (v_\ell(s) - u_\ell^j(s))ds$ converges to 0 uniformly as $j \to \infty$. Let $\tilde{u}_\ell^j = v_\ell - u_\ell^j$, $\tilde{U}_\ell^j(t) = \int_0^t \tilde{u}_\ell^j(s)ds$. Integration by parts gives

$$\begin{aligned}
x^j(t) = x_0 \ &+ \ \sum_{\ell=1}^{2} \int_0^t u_\ell^j(s) f_\ell\Big(x^j(s)\Big) ds \\
= x_0 \ &+ \ \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\Big(x^j(s)\Big) ds - \sum_{\ell=1}^{2} \int_0^t \tilde{u}_\ell^j(s) f_\ell\Big(x^j(s)\Big) ds \\
= x_0 \ &+ \ \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\Big(x^j(s)\Big) ds - \sum_{\ell=1}^{2} \tilde{U}_\ell^j(t)\, f_\ell\Big(x^j(t)\Big) \\
&+ \ \sum_{\ell_1,\ell_2=1}^{2} \int_0^t u_{\ell_1}^j(s) \tilde{U}_{\ell_2}^j(s)\, (Df_{\ell_2}) f_{\ell_1}\Big(x^j(s)\Big) ds \,,
\end{aligned}$$

where $Df(x)$ denotes the Jacobian matrix of the vector field f, evaluated at x. If the "residual functions" $\tilde{u}_\ell^j$ are uniformly bounded in L^1 then, since the $\tilde{U}^j_\ell$ converge to zero uniformly, it follows that the x^j converge to a solution of

$$x(t) = x_0 + \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\Big(x(s)\Big)\, ds,$$

i.e. of

$$\dot{x}(t) = \sum_{\ell=1}^{2} v_\ell(t) f_\ell\Big(x(t)\Big)\,,$$

with initial condition $x(0) = x_0$. Equivalently, $\{u^j\}$ EI-converges to

$$\mathbf{v} = \sum_{\ell=1}^{2} v_\ell(t) X_\ell.$$

Now let us no longer assume that the $\tilde{u}^j_\ell$ are uniformly bounded in L^1, but suppose that we can "remove the convergent part" from each function $u^j_{\ell_1}\tilde{U}^j_{\ell_2}$, by choosing $v_{\ell_1,\ell_2} \in L^1[0,T]$ such that

$$\tilde{U}^j_{\ell_1,\ell_2}(t) = \int_0^t \tilde{u}^j_{\ell_1,\ell_2}(s) \stackrel{\text{def}}{=} \int_0^t \Big(v_{\ell_1,\ell_2}(s) - u^j_{\ell_1}(s)\tilde{U}^j_{\ell_2}(s)\Big)\, ds \to 0$$

uniformly. (If the $\tilde{u}^j_\ell$ were uniformly bounded in L^1, then it is easy to see that v_{ℓ_1,ℓ_2} have to be $\equiv 0$.)

We then get

$$\begin{aligned}
x^j(t) = x_0 \ &+ \ \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\Big(x^j(s)\Big)\, ds - \sum_{\ell=1}^{2} \tilde{U}^j_\ell(t)\, f_\ell\Big(x^j(t)\Big) \\
&+ \sum_{\ell_1,\ell_2=1}^{2} \int_0^t u^j_{\ell_1}(s)\tilde{U}^j_{\ell_2}(s)\,(Df_{\ell_2}) f_{\ell_1}\Big(x^j(s)\Big)\, ds \\
= x_0 \ &+ \ \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\Big(x^j(s)\Big)\, ds - \sum_{\ell=1}^{2} \tilde{U}^j_\ell(t)\, f_\ell\Big(x^j(t)\Big) \\
&+ \sum_{\ell_1,\ell_2=1}^{2} \int_0^t v_{\ell_1,\ell_2}(s)(Df_{\ell_2}) f_{\ell_1}\Big(x^j(s)\Big)\, ds - B
\end{aligned}$$

where the "bad part" B is given by

$$B = \sum_{\ell_1,\ell_2=1}^{2} \int_0^t \tilde{u}^j_{\ell_1,\ell_2}(s)(Df_{\ell_2}) f_{\ell_1}\Big(x^j(s)\Big)\, ds. \tag{21}$$

If the functions $\tilde{u}^j_{\ell_1,\ell_2}$ are uniformly bounded in L^1, then it can be shown that the x^j will converge to a solution of

$$\begin{aligned} x(t) &= x_0 + \sum_{\ell=1}^{2} \int_0^t v_\ell(s) f_\ell\big(x(s)\big)\, ds \\ &+ \sum_{\ell_1,\ell_2=1}^{2} \int_0^t v_{\ell_1,\ell_2}(s)(Df_{\ell_2})f_{\ell_1}\big(x(s)\big)\, ds\,, \end{aligned}$$

i.e. of

$$\dot{x}(t) = \sum_{\ell=1}^{2} v_\ell(t) f_\ell\big(x(t)\big) + \sum_{\ell_1,\ell_2=1}^{2} v_{\ell_1,\ell_2}(t)(Df_{\ell_2})f_{\ell_1}\big(x(t)\big). \qquad (22)$$

(This is not completely obvious at first sight, but is a particular case of the convergence theorem to be proved later, so we will not give a separate proof at this point.)

Remark 4. The right-hand side of (22) is not manifestly a vector field, but one can verify by integration by parts that $v_{\ell_1,\ell_2} = -v_{\ell_2,\ell_1}$. Hence, if we use $[f,g] = (Dg)f - (Df)g$, we can rewrite (22) as

$$\dot{x}(t) = \sum_{\ell=1}^{2} v_\ell(t) f_\ell\big(x(t)\big) + v_{1,2}(t)[f_1, f_2]\big(x(t)\big). \qquad (23)$$

Equivalently, $\{u^j\}$ EI-converges to

$$\mathbf{v}(t) = v_1(t)X_1 + v_2(t)X_2 + v_{1,2}(t)[X_1, X_2]\,. \qquad (24)$$

Our convergence theorem will generalize this formula to the higher-order case. ∎

If the $\tilde{u}^j_{\ell_1,\ell_2}$ are not uniformly bounded in L^1, then we may rewrite B using integration by parts, as

$$\begin{aligned} B = &\textstyle\sum_{\ell_1,\ell_2=1}^{2} \tilde{U}^j_{\ell_1,\ell_2}(t)(Df_{\ell_2})f_{\ell_1}\big(x^j(t)\big) \\ &- \textstyle\sum_{\ell_1,\ell_2,\ell_3=1}^{2} \int_0^t u_{\ell_1}(s)\tilde{U}^j_{\ell_2,\ell_3}(s) D\Big((Df_{\ell_3})f_{\ell_2}\Big) f_{\ell_1}\big(x^j(s)\big)\, ds\,. \end{aligned}$$

We might then once again be able to "remove the convergent part" from the $u^j_{\ell_1}\tilde{U}^j_{\ell_2,\ell_3}$ and repeat the above process one more time. To make a systematic study of the general process just described, we introduce the new concept of "extended difference."

Let $\mathbf{u}(t) = u_1(t)X_1 + u_2(t)X_2$, $\mathbf{v}(t) = \sum_{|I|>0} v_I(t)X_I$ be, respectively, an ordinary input and an integrable $\hat{A}(\mathbf{X})$-valued function. (Notice that $\mathbf{v}$ need not be an extended input, because it is only assumed to be $\hat{A}(\mathbf{X})$-valued, and may fail to take values in $\hat{L}(\mathbf{X})$.) We define an *integrated extended difference (IED) of* $\mathbf{u}$ *and* $\mathbf{v}$ to be an $\hat{A}(\mathbf{X})$-valued function $\widetilde{\mathbf{UV}}(t) = \sum_{|I|>0} \widetilde{UV}_I(t)X_I$, with absolutely continuous coefficients, that satisfies the equations

$$\begin{aligned} \widetilde{uv}_\ell(t) &= v_\ell(t) - u_\ell(t)\,, && (25)\\ \widetilde{uv}_{\ell_1,\ldots,\ell_k}(t) &= v_{\ell_1,\ldots,\ell_k}(t) - u_{\ell_1}(t)\widetilde{UV}_{\ell_2,\ldots,\ell_k}(t)\,, && (26)\end{aligned}$$

where the $\widetilde{\mathbf{uv}}^j_I$ are the derivatives of the $\widetilde{\mathbf{UV}}^j_I$.

Remark 5. The words "extended difference" are natural because of the following algebraic characterization: an $\hat{A}(\mathbf{X})$-valued function W with absolutely continuous coefficients and no zeroth-order term is an IED of $\mathbf{u}$ and $\mathbf{v}$ if and only if W satisfies the differential equation

$$\dot{W}(t) = -\mathbf{u}(t)W(t) + \mathbf{v}(t) - \mathbf{u}(t)\ . \tag{27}$$

If $W(0) = 0$, then it is clear that W has some of the properties of a "difference of $\mathbf{u}$ and $\mathbf{v}$." (For instance, $W \equiv 0$ iff $\mathbf{u} \equiv \mathbf{v}$.) In general, if $W(0)$ is "small" in some sense, then it is reasonable to expect that $W(t)$ is "small" for all t if and only if $\mathbf{u}$ and $\mathbf{v}$ are "close." The convergence theorem will make this precise. ∎

Now let $\{\mathbf{u}^j\}$ be a sequence of ordinary inputs and $\{\mathbf{v}^j\}$ a sequence of integrable $\hat{A}(\mathbf{X})$-valued functions, given by $\mathbf{v}^j(t) = \sum_{0<|I|\le r} v^j_I(t)X_I$. Write $\widetilde{\mathbf{UV}}^j$ rather than $\widetilde{\mathbf{U}^j\mathbf{V}^j}$ to denote an IED $\sum_I \widetilde{UV}^j_I X_I$ of $\mathbf{u}^j$ and $\mathbf{v}^j$, and $\widetilde{\mathbf{uv}}^j$ for the time derivative of $\widetilde{\mathbf{UV}}^j$. (The IED $\widetilde{\mathbf{UV}}$ is not unique, so the notation $\widetilde{\mathbf{UV}}$ is slightly ambiguous, but we will use it nevertheless. Both $\widetilde{\mathbf{UV}}^j$ and $\widetilde{\mathbf{uv}}^j$ are uniquely determined by the choice of $\widetilde{\mathbf{UV}}^j(0)$.)

Assume that f_1, f_2 are vector fields of class C^r for some positive integer r. Then by repeated integrations by parts we see that the solutions x^j of (20) satisfy the integral equation

$$\begin{aligned} x^j(t) = x_0 \quad &+ \sum_{|I|\le r} \int_0^t v_I^j(s) f_I\Big(x^j(s)\Big)ds - \sum_{|I|<r} \Big[\widetilde{UV}_I^j(\cdot) f_I\Big(x^j(\cdot)\Big)\Big]_0^t \\ &- \sum_{|I|=r} \int_0^t \widetilde{uv}_I^j(s) f_I\Big(x^j(s)\Big)ds \,, \end{aligned} \tag{28}$$

where we use the notation $[\psi(\cdot)]_a^b$ for $\psi(b)-\psi(a)$ and, for a multi-index $I=(\ell_1,\ldots,\ell_k)$, $f_I(x)$ denotes the vector

$$f_I(x) \stackrel{\text{def}}{=} D(D\cdots D((Df_{\ell_k})f_{\ell_{k-1}})\ldots,f_{\ell_2})f_{\ell_1}(x) \ . \tag{29}$$

For I as above, let us also define

$$[X_I] \stackrel{\text{def}}{=} [X_{\ell_1},[X_{\ell_2},\ldots,[X_{\ell_{k-1}},X_{\ell_k}]\ldots]] \ , \tag{30}$$

$$[f_I] \stackrel{\text{def}}{=} [f_{\ell_1},[f_{\ell_2},\ldots,[f_{\ell_{k-1}},f_{\ell_k}]\ldots]] \ . \tag{31}$$

With these notations, we can now state our convergence theorem.

Theorem 2 *Let r be a positive integer. Let $\{\mathbf{u}^j\}\subset L^1([0,T],\mathbb{R}^2)$ be a sequence of ordinary inputs. Let $\mathbf{v}$ be an integrable $\hat{A}(\mathbf{X})$-valued function on $[0,T]$, given by*

$$\mathbf{v}(t) = \sum_{0<|I|\le r} v_I(t) X_I \ .$$

Let $V_I(t)=\int_0^t v_I(s)ds$. Assume that there exists a sequence $\{\mathbf{V}^j\}$ of $\hat{A}(\mathbf{X})$-valued functions on $[0,T]$ with absolutely continuous components, given by $\mathbf{V}^j(t)=\sum_{0<|I|\le r} V_I^j(t)X_I$, and a sequence $\{\widetilde{\mathbf{UV}}^j\}$ of IED's of $\mathbf{u}^j$ and $\mathbf{v}^j$, such that

A1. *$\{V_I^j\}_{j=1}^\infty$ converges uniformly to V_I for $0<|I|\le r$;*

A2. *$\{\widetilde{UV}_I^j\}_{j=1}^\infty$ converges uniformly to 0 for $0<|I|\le r$;*

A3. *the L^1 norms of the derivatives v_I^j for $0<|I|\le r$ and $\widetilde{uv}_I^j$ for $|I|=r$ are uniformly bounded.*

Then:

C1. $\mathbf{v}$ *is an extended input of order* $\leq r$*;*

C2. $\mathbf{v}$ *is given by the formula*

$$\mathbf{v} = \sum_{0<|I|\leq r} \frac{v_I(t)}{|I|}[X_I] \; ; \tag{32}$$

C3. *the* $\mathbf{u}^j$ *EI-converge to* $\mathbf{v}$ *for* C^r *systems of the form (1).*

Remark 6. In simpler terms, the conclusion of Theorem 2 says that, for every choice of C^r vector fields f_1, f_2, and every initial condition x_0, if the unique solution $x^\infty(t)$ of the initial value problem

$$\dot{x}(t) = \sum_{|I|\leq r} \frac{v_I(t)}{|I|}[f_I](x) \;\; , \;\; x(0) = x_0 \tag{33}$$

is defined on the whole interval $[0, T]$, and x^j is the solution of (20), then the x^j are defined on the whole interval $[0, T]$ for j large enough, and converge to x^∞ uniformly as $j \to \infty$. Notice that the initial value problem (33) satisfies the conditions of the Carathéodory theorem on existence and uniqueness of solutions of ordinary differential equations, because the right-hand side is of class C^1 in x for each fixed t, measurable in t for each x, and is bounded, together with its first partial derivatives with respect to x, by an integrable function of t, as long as x stays in a compact set. ∎

PROOF OF THEOREM 2. We will first establish the property stated in Remark 6 for the initial value problem

$$\dot{x}(t) = \sum_{|I|\leq r} v_I(t) f_I(x) \;\; , \;\; x(0) = x_0 \; , \tag{34}$$

where the f_I were defined in (29). We will then show that the right-hand side of the first equation of (34) is actually equal to that of the first equation of (33).

Assume first that f_1 and f_2 are compactly supported, so all the trajectories x^j and x^∞ exist on the whole interval $[0, T]$ and are contained in a fixed compact set K. If we could prove that the sequence $\{x^j\}$ is equicontinuous, then our conclusion would follow from (28), using the boundedness of the L^1 norms of the v_I^j and $\widetilde{uv}_I^j$, and the fact that the $\widetilde{UV}_I^j$ converge uniformly to zero. (By equicontinuity, every subsequence $\{x^{j(\ell)}\}$ of $\{x^j\}$ has a subsequence $\{x^{j(\ell(k))}\} \stackrel{\rm def}{=} \{\sigma^k\}$ that converges uniformly to a limit σ. Then the integrals $\int_0^t \tilde{v}_I^k(s) f_I\Big(\sigma^k(s)\Big)ds - \int_0^t \tilde{v}_I^k(s) f_I\Big(\sigma(s)\Big)ds$, where $\tilde{v}_I^k = v_I^{j(\ell(k))}$, converge uniformly to zero, because the L^1 norms of the $\tilde{v}_I^k$ are bounded. The integrals $\int_0^t \tilde{v}_I^k(s) f_I\Big(\sigma(s)\Big)ds$ converge to $\int_0^t v_I(s) f_I\Big(\sigma(s)\Big)ds$ because the continuous function $s \to f_I\Big(\sigma(s)\Big)$ can be uniformly approximated by piecewise constant functions, and Assumption **A1** implies that $\int_0^t \tilde{v}_I^k(s)\psi(s)ds \to \int_0^t v_I(s)\psi(s)ds$ if ψ is piecewise constant. So $\int_0^t \tilde{v}_I^k(s) f_I\Big(\sigma^k(s)\Big)ds \to \int_0^t v_I(s) f_I\Big(\sigma(s)\Big)ds$. A similar argument shows that $\int_0^t \widetilde{uv}_I^{j(\ell(k))}(s) f_I\Big(\sigma^k(s)\Big)ds \to 0$. So $\sigma = x^\infty$ by uniqueness. In particular σ is independent of the subsequence chosen.)

To prove that $\{x^j\}$ is equicontinuous, we let

$$\xi^j(t) = x_0 + \sum_{|I|\le r} \int_0^t v_I^j(s) f_I\Big(x^j(s)\Big)ds - \sum_{|I|=r} \int_0^t \widetilde{uv}_I^j(s) f_I\Big(x^j(s)\Big)ds \ .$$

Since the $\widetilde{UV}_I^j$ converge uniformly to zero, we have $\xi^j - x^j \to 0$ uniformly, so equicontinuity of $\{x^j\}$ will follow if we prove that $\{\xi^j\}$ is equicontinuous. Let

$$\zeta^j(t) = x_0 + \sum_{|I|\le r} \int_0^t v_I^j(s) f_I\Big(\xi^j(s)\Big)ds - \sum_{|I|=r} \int_0^t \widetilde{uv}_I^j(s) f_I\Big(\xi^j(s)\Big)ds \ .$$

Since $f_I(\xi^j) - f_I(x^j) \to 0$ uniformly, the boundedness of the L^1 norms of the v_I^j and $\widetilde{uv}_I^j$ implies that $\zeta^j - \xi^j \to 0$ uniformly, so it suffices to prove equicontinuity of $\{\zeta^j\}$. We can write

$$\begin{aligned}&\textstyle\int_0^t v_I^j(s) f_I\Big(\xi^j(s)\Big)ds = \\ &\qquad \textstyle\int_0^t \Big(v_I^j(s) - v_I(s)\Big) f_I\Big(\xi^j(s)\Big)ds + \int_0^t v_I(s) f_I\Big(\xi^j(s)\Big)ds \ .\end{aligned}$$

The second integral in the right-hand side has an integrand which is the product of a fixed L^1 function times a uniformly bounded sequence, so it is equicontinuous. The first integral can be rewritten via integration by parts as

$$\left[\Big(V_I^j(\cdot)-V_I(\cdot)\Big)f_I\Big(\xi^j(\cdot)\Big)\right]_0^t - \int_0^t \Big(V_I^j(s)-V_I(s)\Big)Df_I\Big(\xi^j(s)\Big)\dot{\xi}^j(s)\,ds\,.$$

The definition of the ξ^j as integrals shows that the $\dot{\xi}^j$ are uniformly bounded in L^1. Since $V_I^j - V_I \to 0$ uniformly, and the functions $s \to Df_I\big(\xi^j(s)\big)$ are uniformly bounded, we conclude that $\int_0^t\big(V_I^j(s) - V_I(s)\big)Df_I\big(\xi^j(s)\big)\dot{\xi}^j(s)\,ds \to 0$ uniformly. Also, $\left[\Big(V_I^j(\cdot)-V_I(\cdot)\Big)f_I\Big(\xi^j(\cdot)\Big)\right]_0^t \to 0$ uniformly, because $V_I^j(t) - V_I(t) \to 0$ uniformly and the $f_I\big(\xi^j(t)\big)$ are uniformly bounded. So the integral $\int_0^t\big(v_I^j(s) - v_I(s)\big)f_I\big(\xi^j(s)\big)ds$ goes to 0 uniformly. Therefore $\{\int_0^t v_I^j(s)f_I\big(\xi^j(s)\big)ds\}$ is equicontinuous. A similar argument shows that the sequence of integrals $\{\int_0^t \widetilde{uv}_I^j(s)f_I\big(\xi^j(s)\big)ds\}$ is equicontinuous for $|I| = r$. So $\{\zeta^j\}$ is equicontinuous, and the equicontinuity of $\{x^j\}$ follows. This completes the proof for the case when f_1 and f_2 are compactly supported.

In the general case, we multiply f_1 and f_2 by a smooth compactly supported function $\varphi(x)$ which is equal to one on a neighborhood U of the set $\{x^\infty(t) : 0 \le t \le T\}$. We apply the conclusion to the resulting system, and get trajectories $\hat{x}^j$ of this new system that converge to x^∞. Then for large j the $\hat{x}^j$ will lie in U, so they will be trajectories of the original system. Our conclusion now follows in full generality, except that we have obtained the limiting initial value problem (34) rather than (33).

To complete our proof, we have to establish the formula

$$\sum_{|I|\le r} v_I(t)f_I(x) = \sum_{|I|\le r}\frac{v_I(t)}{|I|}[f_I](x)\ . \tag{35}$$

This can be done as follows. First observe that the definition of the f_I given in (29) implies the identity

$$f_{\ell_1,\ldots,\ell_k} = L_{f_{\ell_1}}L_{f_{\ell_2}}\ldots L_{f_{\ell_k}}\Phi\ .$$

Here $\Phi : \mathbb{R}^n \to \mathbb{R}^n$ is the identity map (i.e. $\Phi(x) = x$) and, for a vector field $g : \mathbb{R}^n \to \mathbb{R}^n$, L_g is the operator of "Lie differentiation in the direction of g," defined by $L_g\psi(x) = D\psi(x).g(x)$, if ψ is a (scalar- or vector-valued) function on $\mathbb{R}^n$. Next observe that the convergence theorem holds for all sufficiently smooth systems, so in particular it holds for systems in which the vector fields are linear, i.e. $f_k(x) = A_k(x)$, where $A_k : \mathbb{R}^n \to \mathbb{R}^n$ is some linear map. In this case, one can easily check that all the vector functions f_I are also linear, and given by $f_{\ell_1} \dots f_{\ell_k}(x) = A_{\ell_k} \dots A_{\ell_1}(x)$. A particular example of this is the *r-th order nilpotent formal system*

$$\dot{S} = S(u_1 X_1 + u_2 X_2) \ , \ \ S(0) = 1 \,, \tag{36}$$

which evolves in the algebra $A_r(\mathbf{X})$ obtained by "truncating" $A(\mathbf{X})$ at degree r, i.e. declaring all monomials of degree $> r$ to be equal to zero. In this case, the linear map A_k is right multiplication by X_k, so $f_{\ell_1} \dots f_{\ell_k}(S) = S X_{\ell_1} \dots X_{\ell_k}$. The limiting trajectory S^∞ will satisfy $\dot{S}^\infty(t) = S^\infty(t)(\sum_{|I|\le r} v_I(t) X_I)$, i.e. $\dot{S}^\infty(t) = S^\infty(t)\mathbf{v}(t)$. On the other hand, the series $S^\infty(t)$ is an exponential of a Lie series because of the Campbell-Hausdorff Formula (CHF)[3]. But then we can write $\mathbf{v}(t) = \lim_{h\to 0} \frac{1}{h}\Big(S^\infty(t)\Big)^{-1}\Big(S^\infty(t+h) - S^\infty(t)\Big) = \lim_{h\to 0}\frac{1}{h}\left(\Big(S^\infty(t)\Big)^{-1} S^\infty(t+h) - 1\right)$. Using CHF again, we conclude that $\Big(S^\infty(t)\Big)^{-1} S^\infty(t+h) = e^{\Lambda(h)}$, where $\Lambda(h)$ is a Lie series that goes to zero as $h \to 0$. So $\mathbf{v}(t) = \lim_{h\to 0}\frac{1}{h}(\Lambda(h) + \frac{1}{2}\Lambda(h)^2 + \dots) = \lim_{h\to 0}\frac{\Lambda(h)}{h}$ for almost every t. So $\mathbf{v}$ is Lie-series-valued, i.e. an extended input. We then use a general fact about free Lie algebras, according to which, if a noncommutative series $\sum_I a_I X_I$ is known to be a Lie series, then $\sum_I a_I X_I = \sum_I \frac{a_I}{|I|}[X_I]$, where $[X_I]$ is as defined in (30). In particular,

$$\sum_{|I|\le r} v_I(t) X_I = \sum_{|I|\le r} \frac{v_I(t)}{|I|}[X_I] \,.$$

[3]The *exponential* of a series is defined by its usual power series expression. Since we are working in a finite dimensional algebra, there are no convergence problems. The CHF says that a product of exponentials of Lie series is also an exponential of a Lie series. Applying this to a piecewise constant control (u_1, u_2) we conclude that for every trajectory $S(t)$ for such a control, starting with $S(0) = 1$, the series $S(t)$ are exponential Lie series. Taking limits, this remains true for trajectories of arbitrary ordinary controls, and for limits of such trajectories, such as S^∞.

Since this identity is true for purely formal indeterminates X_ℓ, it must remain true if we plug for the X_ℓ any objects in an associative algebra. In particular, we can plug in the operators L_{f_ℓ}, and use repeatedly the fact that $[L_f, L_g] = L_{[f,g]}$, to conclude that $\sum_{|I|\le r} v_I(t)(L_f)_I = \sum_{|I|\le r} \frac{v_I(t)}{I} L_{[f_I]}$. If we then apply both sides to the map Φ, we get (35). ∎

Remark 7. Formula (35) contains as a particular case the observation made in Remark 4. To see this, notice that, according to (35), the expression $\sum_{\ell_1,\ell_2=1}^{2} v_{\ell_1,\ell_2}(Df_{\ell_2})f_{\ell_1}$, which by definition is equal to $\sum_{\ell_1,\ell_2=1}^{2} v_{\ell_1,\ell_2} f_{\ell_1,\ell_2}$, has the value $\frac{1}{2}(v_{1,1}[f_1, f_1] + v_{1,2}[f_1, f_2] + v_{2,1}[f_2, f_1] + v_{2,2}[f_2, f_2])$. Since $[f_1, f_1] = [f_2, f_2] = 0$, $[f_1, f_2] = -[f_2, f_1]$, and $v_{1,2} = -v_{2,1}$, we see that this is in fact equal to $v_{1,2}[f_1, f_2]$, as shown in Remark 4. ∎

Remark 8. If the vector fields f_1, f_2 are only assumed to be of class C^{r-1}, the conclusion still follows, if one assumes uniqueness of solutions for (33). The proof is almost exactly the same, except that an extra step is required in which the vector-valued functions f_I are approximated by functions g_I of class C^∞. ∎

PROOF OF THEOREM 1

We will apply Theorem 2. For this purpose, we try to define sequences $\{V_I^j\}_{j=1}^\infty$, $\{\widetilde{UV}_I^j\}_{j=1}^\infty$ satisfying the conditions of that theorem. As before, the derivatives of these functions will be denoted by v_I^j, $\widetilde{uv}_I^j$.

For any multiindex $I = (\ell_1, \ldots, \ell_k)$, let $\delta_1(I)$ and $\delta_2(I)$ denote, respectively, the number of occurrences of the indices 1 and 2 in I. Then $|I| = \delta_1(I) + \delta_2(I)$. We use induction to define $V_I^j, \widetilde{UV}_I^j$ for $|I| \le 5$ such that

A. The V_I^j have the form $V_I^j = V_I + R_I^j$, where

(i) V_1 and V_2 are given by $V_\ell(t) = \int_0^t \eta_{\ell,0}(s)\, ds$.

(ii) If $2 \le |I| \le 5$, then $V_I = 0$, unless $\delta_1(I) = 3$ and $\delta_2(I) = 2$.

(iii) If I has length 5 and $\delta_1(I) = 3$, $\delta_2(I) = 2$,

$$V_I(t) = \int_0^t \Big(\Theta_I \xi(s) + \tilde{\Theta}_I \tilde{\xi}(s)\Big)\, ds \tag{37}$$

where $\xi, \tilde{\xi}$ are the functions defined in (15) and (16), and the constants Θ_I, $\tilde{\Theta}_I$ will be evaluated below.

(iv) The R_I^j converge to 0 uniformly as $j \to \infty$, and the $\|\dot{R}_I^j\|_{L^1}$ are uniformly bounded for all $|I| \le 5$.

B. For $0 < |I| \le 4$, the $\widetilde{UV}_I^j$ can be written as

$$\widetilde{UV}_I^j(t) = (-1)^{|I|} j^{-\frac{|I|}{5}} \sum_{\hat{\omega} \in \Omega(I)} \frac{\eta_{\hat{\omega}}(t) e^{ij(\sum \hat{\omega})t}}{i^{|I|} h(\hat{\omega})}, \tag{38}$$

where, for $I = (\ell_1, \ldots, \ell_k)$, if $\hat{\omega} = (\omega_1, \ldots, \omega_k)$ belongs to $\Omega(\ell_1) \times \cdots \times \Omega(\ell_k)$, we let

$$h(\hat{\omega}) = \omega_k(\omega_k + \omega_{k-1}) \ldots (\omega_k + \cdots + \omega_1)$$

$$\eta_{\hat{\omega}}(t) = \eta_{\omega_1}(t) \cdots \eta_{\omega_k}(t) \ ,$$

$$\sum \hat{\omega} = \omega_1 + \ldots + \omega_k$$

$$\Omega(I) = \{\hat{\omega} \in \Omega(\ell_1) \times \cdots \times \Omega(\ell_k) : h(\hat{\omega}) \neq 0\}\,.$$

C. $\widetilde{UV}_I^j = 0$ for $|I| = 5$.

Notice that **A**, **B** and **C** imply that Assumptions **A1**, **A2** and **A3** of Theorem 2 hold.

We now define $V_I^j, \widetilde{UV}_I^j$. Integration by parts gives

$$\int_0^t u_\ell^j(s)\, ds = \int_0^t \eta_{\ell,0}(s) ds + \int_0^t \sum_{\omega \in \Omega(\ell)} j^{\frac{4}{5}} \eta_\omega(s) e^{ij\omega s}\, ds$$

$$= \int_0^t \eta_{\ell,0}(s) ds + j^{-\frac{1}{5}} \sum_{\omega \in \Omega(\ell)} \{\frac{\eta_\omega(t) e^{ij\omega t}}{i\omega} - \frac{\eta_\omega(0)}{i\omega} - \int_0^t \frac{\eta'_\omega(s) e^{ij\omega s}}{i\omega} ds\}\,.$$

We let

$$
\begin{aligned}
V_\ell(t) &= \int_0^t \eta_{\ell,0}(s)ds, \\
R_\ell^j(t) &= -j^{-\frac{1}{5}} \sum_{\omega\in\Omega(\ell)} \left\{ \frac{\eta_\omega(0)}{i\omega} + \int_0^t \frac{\eta'_\omega(s)}{i\omega} e^{ij\omega s}\, ds \right\}, \\
V_\ell^j(t) &= V_\ell(t) + R_\ell^j(t).
\end{aligned}
$$

Then

$$
\widetilde{UV}_\ell^j(t) = V_\ell^j(t) - \int_0^t u_\ell^j(s)ds = -j^{-\frac{1}{5}} \sum_{\omega\in\Omega(\ell)} \frac{\eta_\omega(t)}{i\omega} e^{ij\omega t}.
$$

If we multiply $\widetilde{UV}_{\ell_2}^j(s)$ by $u_{\ell_1}^j(s)$ and integrate, we get

$$
\int_0^t u_{\ell_1}^j(s)\widetilde{UV}_{\ell_2}^j(s)\, ds = A_{\ell_1,\ell_2}^j + B_{\ell_1,\ell_2}^j,
$$

where

$$
A_{\ell_1,\ell_2}^j = -j^{-\frac{1}{5}} \int_0^t \eta_{\ell_1,0}(s) \sum_{\omega\in\Omega(\ell_2)} \frac{\eta_\omega(s)}{i\omega} e^{ij\omega s} ds
$$

$$
B_{\ell_1,\ell_2}^j = -j^{\frac{3}{5}} \sum_{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)} \int_0^t \frac{\eta_{\omega_1}(s)\eta_{\omega_2}(s)}{i\omega_2} e^{ij(\omega_1+\omega_2)s}\, ds.
$$

We can split B_{ℓ_1,ℓ_2}^j further into the sum C_{ℓ_1,ℓ_2}^j of those terms for which $\omega_1 + \omega_2 = 0$ and the sum D_{ℓ_1,ℓ_2}^j of those terms for which $\omega_1 + \omega_2 \neq 0$. Then

$$
C_{\ell_1,\ell_2}^j = -j^{\frac{3}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2) \\ \omega_1+\omega_2=0}} \int_0^t \frac{\eta_{\omega_1}(s)\eta_{\omega_2}(s)}{i\omega_2}\, ds.
$$

Since $\eta_{\omega_2} = \overline{\eta_{\omega_1}}$ whenever $\omega_1 + \omega_2 = 0$, we see that each unordered pair $\{\omega_1, \omega_2\}$ such that $\omega_1 + \omega_2 = 0$ makes a contribution to the sum equal to $(\frac{1}{i\omega_1} + \frac{1}{i\omega_2})|\eta_{\omega_1}|^2$, i.e. to zero. Therefore $C_{\ell_1,\ell_2}^j = 0$. So we can rewrite B_{ℓ_1,ℓ_2}^j as a sum of terms $B_{\ell_1,\ell_2}^j(\omega_1, \omega_2)$, ranging over $\omega_1 + \omega_2 \neq 0$. We then use integration by parts to rewrite each of the $B_{\ell_1,\ell_2}^j(\omega_1, \omega_2)$ as a sum of three terms, namely,

$$
B_{\ell_1,\ell_2}^j(\omega_1, \omega_2)_1 = -j^{-\frac{2}{5}} \frac{\eta_{\omega_1}(t)\eta_{\omega_2}(t)e^{ij(\omega_1+\omega_2)t}}{i^2\omega_2(\omega_2 + \omega_1)},
$$

$$B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_2 = j^{-\frac{2}{5}}\frac{\eta_{\omega_1}(0)\eta_{\omega_2}(0)}{i^2\omega_2(\omega_2+\omega_1)}\,,$$

$$B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_3 = j^{-\frac{2}{5}}\int_0^t \frac{\left(\eta_{\omega_1}(s)\eta_{\omega_2}(s)\right)'}{i^2\omega_2(\omega_2+\omega_1)}e^{ij(\omega_1+\omega_2)s}\,ds\,.$$

It is clear that all the $A^j_{\ell_1,\ell_2}$, $B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_2$, $B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_3$ go to zero uniformly and have uniformly bounded derivatives, so we can absorb them into the remainders. So we let

$$R^j_{\ell_1,\ell_2}(t) = A^j_{\ell_1,\ell_2} + \sum B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_2 + \sum B^j_{\ell_1,\ell_2}(\omega_1,\omega_2)_3\,,$$

where the sums are over $(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)$, $\omega_1+\omega_2\neq 0$. We then set $V_{\ell_1,\ell_2}(t)=0$, $V^j_{\ell_1,\ell_2}(t) = V_{\ell_1,\ell_2}(t) + R^j_{\ell_1,\ell_2}(t)$. We then have

$$\begin{aligned}\widetilde{UV}^j_{\ell_1,\ell_2}(t) &= V^j_{\ell_1,\ell_2}(t) - \int_0^t u^j_{\ell_1}(s)\widetilde{UV}^j_{\ell_2}(s)\,ds \\ &= j^{-\frac{2}{5}} \sum_{\substack{(\omega_1,\omega_2)\in\Omega(\ell_1)\times\Omega(\ell_2)\\ \omega_1+\omega_2\neq 0}} \frac{\eta_{\omega_1}(t)\eta_{\omega_2}(t)}{i^2\omega_2(\omega_2+\omega_1)}e^{ij(\omega_1+\omega_2)t}\,.\end{aligned}$$

Once again, we can multiply $\widetilde{UV}^j_{\ell_2,\ell_3}(s)$ by $u^j_{\ell_1}(s)$, integrate, and split the result into an A part and a B part. The A part goes to zero uniformly and has uniformly bounded derivatives. The B part is a sum over triples $(\omega_1,\omega_2,\omega_3)\in\Omega(\ell_1)\times\Omega(\ell_2)\times\Omega(\ell_3)$. All these triples necessarily satisfy $\omega_1+\omega_2+\omega_3\neq 0$ and $\omega_2+\omega_3\neq 0$. (Notice that there are no triples $(\omega_1,\omega_2,\omega_3)$ whose sum is equal to zero.) We then use integration by parts to rewrite the B part as a sum of three collections of terms, namely,

$$B^j_{\ell_1,\ell_2,\ell_3}(\omega_1,\omega_2,\omega_3)_1 = j^{-\frac{3}{5}}\frac{\eta_{\omega_1}(t)\eta_{\omega_2}(t)\eta_{\omega_3}(t)e^{ij(\omega_1+\omega_2+\omega_3)t}}{i^3\omega_3(\omega_3+\omega_2)(\omega_3+\omega_2+\omega_1)}\,,$$

and terms $B^j_{\ell_1,\ell_2,\ell_3}(\omega_1,\omega_2,\omega_3)_2$, $B^j_{\ell_1,\ell_2,\ell_3}(\omega_1,\omega_2,\omega_3)_3$ that go to zero uniformly and have uniformly bounded derivatives. Again, we incorporate the A, B_2 and B_3 parts into the remainder $R^j_{\ell_1,\ell_2,\ell_3}$, define $V_{\ell_1,\ell_2,\ell_3}=0$, and let $V^j_{\ell_1,\ell_2,\ell_3} = V_{\ell_1,\ell_2,\ell_3} + R^j_{\ell_1,\ell_2,\ell_3}$. Then the $\widetilde{UV}^j_{\ell_1,\ell_2,\ell_3}$ are just the sum of $-j^{-\frac{3}{5}}\frac{\eta_{\omega_1}(t)\eta_{\omega_2}(t)\eta_{\omega_3}(t)e^{ij(\omega_1+\omega_2+\omega_3)t}}{i^3\omega_3(\omega_3+\omega_2)(\omega_3+\omega_2+\omega_1)}$, ranging over $(\omega_1,\omega_2,\omega_3)\in\Omega(\ell_1)\times\Omega(\ell_2)\times\Omega(\ell_3)$, $\omega_1+\omega_2+\omega_3\neq 0$, $\omega_2+\omega_3\neq 0$.

The next step is exactly the same, except that now there are 4-tuples of frequencies that add up to zero. So, before we carry out the integration by parts, the contribution of these terms must be controlled. This is done exactly as in our analysis of the terms of order two: any such 4-tuple $(\omega_1, \ldots, \omega_4)$ must involve two pairs consisting of a frequency and its negative. The contribution of such a 4-tuple is then equal to a numerator that remains unchanged when $(\omega_1, \ldots, \omega_4)$ is replaced by $(-\omega_1, \ldots, -\omega_4)$, divided by a product of three linear functionals of the ω's. The contribution of each pair and that of its negative cancel.

Finally, when we carry out the last step, we obtain an expression for $u^j_{\ell_1} \widetilde{UV}^j_{\ell_2,\ell_3,\ell_4,\ell_5}$ in which the exponent $j^{\frac{4}{5}}$ of the first factor exactly offsets the exponent of the second factor. When we integrate, all the terms whose five frequencies do not add up to zero produce, exactly as before, A and B terms. Moreover, at this point the B_1 terms have a factor $\frac{1}{j}$ in the denominator, so they go to zero uniformly and have uniformly bounded derivatives. So all the terms coming from the 5-tuples of frequencies that do not add up to zero can be incorporated into the remainders. However, at this point there are 5-tuples that add up to zero. These 5-tuples can only occur for multiindices I such that $\delta_1(I) = 3$, $\delta_2(I) = 2$. So all the V_I for $|I| = 5$ will vanish, except possibly when $\delta_1(I) = 3$, $\delta_2(I) = 2$.

Finally, it is easy to compute V_I when $\delta_1(I) = 3$, $\delta_2(I) = 2$, and the result turns out to be precisely of the form given in (37). Indeed,

$$V_I(t) = \sum_{\hat{\omega} \in \tilde{\Omega}(I)} \int_0^t \frac{\eta_{\hat{\omega}}(s)}{\tilde{h}(\hat{\omega})} ds \ ,$$

where, for $I = (\ell_1, \ell_2, \ell_3, \ell_4, \ell_5)$, $\hat{\omega} = (\omega_1, \ldots, \omega_5) \in \Omega(\ell_1) \times \ldots \times \Omega(\ell_5)$, we let $\tilde{h}(\hat{\omega}) = h(\omega_2, \omega_3, \omega_4, \omega_5)$, and then define $\tilde{\Omega}(I)$ to be the set of those $\hat{\omega} = (\omega_1, \ldots, \omega_5) \in \Omega(\ell_1) \times \ldots \times \Omega(\ell_5)$ for which $\tilde{h}(\hat{\omega}) \neq 0$ and $\omega_1 + \ldots + \omega_5 = 0$. The crucial point now is that the only way to choose the five components $\omega_1, \ldots, \omega_5$ of $\hat{\omega}$ so that their sum is equal to zero is to take them to be the five elements of $\bar{\Omega}$, or their negatives, or the five elements of $\bar{\Omega}'$, or their negatives. Each of these four possibilities leads to 12 choices. For instance, the first possibility says that $(\omega_1, \ldots, \omega_5)$ has to be the same as that $(w_1, \ldots, w_5)$ up to a permutation π of $\{1, \ldots, 5\}$, i.e. $\omega_k = w_{\pi(k)}$.

Moreover, π has to be such that $\{\pi(1), \pi(2), \pi(3)\} = \{k : \ell_k = 1\}$. The set P_I of all such permutations obviously contains 12 elements. Moreover, for each such $\hat{\omega}$ the corresponding contribution to V_I is $\frac{\eta_{w_1} \cdots \eta_{w_5}}{\tilde{h}(\hat{\omega})}$. Summing over all the $\pi \in P_I$ gives a contribution

$$\left(\sum_{\pi \in P_I} \frac{1}{\tilde{h}\big(w_{\pi(1)}, \ldots, w_{\pi(5)}\big)} \right) \eta_{w_1} \ldots \eta_{w_5} \, .$$

The second possibility, in which the ω_k are the negatives of the w_k, gives exactly the same contribution, except that $\eta_{w_1} \ldots \eta_{w_5}$ is replaced by its complex conjugate. (Notice that the functions $\tilde{h}$ are even.) The other two possibilities make an identical contribution, except that now the w_ks are replaced by the w'_k.

So Theorem 2 tells us that the control sequence defined in (12) and (13) EI-converges to

$$\mathbf{u}^\infty(t) = \eta_{1,0}(t) X_1 + \eta_{2,0}(t) X_2 + \sum_{I \in \Lambda} \frac{1}{5} v_I(t) \, [X_I] \, , \tag{39}$$

where Λ consists of all the indices I such that $\delta_1(I) = 3$, $\delta_2(I) = 2$, and

$$v_I(t) = \Theta_I \xi(t) + \tilde{\Theta}_I \tilde{\xi}(t) \, . \tag{40}$$

One can verify that there are six nonvanishing formal brackets $[X_I]$, $I \in \Lambda$. Moreover, to compute each v_I we need to compute Θ_I and $\tilde{\Theta}_I$, each of which is itself a sum over 12 permutations. So the above formula expresses the $(3, 2)$ homogeneous component of u^∞ as a sum of 144 terms. Although this is in principle not too hard to compute, we prefer to present a shortcut that considerably simplifies the calculation.

Clearly, each member of $\{[X_I] : I \in \Lambda\}$ is a linear combination of B_0 and B_1. Moreover, the functions v_I are themselves linear combinations with constant coefficients of ξ and $\tilde{\xi}$. This means that

$$\sum_{I \in \Lambda} \frac{1}{5} v_I(t) \, [X_I] = \Big(\Theta_1 \xi(t) + \Theta_2 \tilde{\xi}(t) \Big) B_0 + \Big(\Theta_3 \xi(t) + \Theta_4 \tilde{\xi}(t) \Big) B_1 \, .$$

So we need to figure out the values of the four coefficients Θ_i.

To do this, it is convenient to use the *product expansion of the Chen-Fliess series* described in [16]. This is a particularly simple way to express the solution of (36) as a product of exponentials $e^{C_B(t)B}$, where the B belong to a *P. Hall basis* of $L(\mathbf{X})$.

For details about the product expansion we refer the reader to [16]. The main purpose of this expansion is to solve a technical problem that arises when we try to express the solution of (36) in terms of iterated integrals. (We can regard (36) as evolving in $\hat{A}(\mathbf{X})$ or in the "truncated" finite-dimensional algebra $A_r(\mathbf{X})$.) Due to the Campbell-Hausdorff Formula, we know *a priori* that the solution $S(t)$ is an exponential $e^{Z(t)}$, where $Z(t)$ is a Lie series. It is clear that $Z(t)$ is uniquely determined by $S(t)$. However, if we try to write $Z(t)$ as a linear combination of brackets such as the $[X_I]$ that we introduced earlier, we have to cope with the difficulty that these brackets are not linearly independent, because there are many linear relations among them, arising from the skew symmetry property $[P,Q]+[Q,P]=0$ and the Jacobi identity $[P,[Q,R]]+[Q,[R,P]]+[R,[P,Q]]=0$. This leads to a "combinatorial explosion," an example of which was already encountered above when we pointed out that there are six nonzero brackets $[X_I]$ with $\delta_1([X_I])=3$, $\delta_2([X_I])=2$.

To reduce the combinatorial explosion one needs an explicit prescription for selecting a basis of $L(\mathbf{X})$. One such choice is provided by the use of a P. Hall basis. The precise definition is as follows. A *P. Hall basis* of $L(\mathbf{X})$ is a set $\mathcal{B}$ of formal brackets $B \in \mathcal{FB}r(\mathbf{X})$ (cf. Remark 3), endowed with a total ordering $\prec$, such that the pair $(\mathcal{B},\prec)$ satisfies:

PH1. If $B, B' \in \mathcal{B}$ and $\delta(B) < \delta(B')$, then $B \prec B'$.

PH2. Every X_k is in $\mathcal{B}$.

PH3. If B is a formal bracket and $\delta(B) > 1$, so that B can be written in a unique way as $[B_1, B_2]$ (cf. Remark 3), then $B \in \mathcal{B}$ if and only if (i) $B_1 \in \mathcal{B}$, (ii) $B_2 \in \mathcal{B}$, (iii) $B_1 \prec B_2$ and (iv) either (iv.a) $\delta(B_2) = 1$ or (iv.b) $B_2 = [B_3, B_4]$ and $B_3 \preceq B_1$.

If $B \in \mathcal{B}$ is such that $\delta(B) > 1$, it is clear the left factor (cf. Remark 3) of B is also in $\mathcal{B}$. Moreover, B can be written in a unique way as $\mathrm{ad}^{\kappa}_{B_1}(B_2)$, where $B_1 \prec B_2$ and either $\delta(B_2) = 1$ or the left factor B_3 of B_2 satisfies $B_3 \prec B_1$. (Here we are using the standard notation ad_B to denote the operator $Z \to [B, Z]$.)

The main theorem of [16] gives an explicit formula for $S(t)$ as a product $S(t) = \prod_{B\in\mathcal{B}} e^{C_B(u)(t)B}$, with the factors ordered from right ot left. The functions $C_B(u)(t)$ are given by $C_B(u)(t) = \int_0^t c_B(u)(s)\,ds$. The $c_B(u)$ and $C_B(u)$ are defined recursively as follows. For $B = X_k$, we let $c_B(u)(t) = u_k(t), 0 \le t \le T$. Assume that $c_B(u)$ and $C_B(u)$ have been defined for all $B \in \mathcal{B}$ with degree $\le n$, and let $B \in \mathcal{B}_{n+1}$. Then B can be written in a unique way as $\mathrm{ad}^{\kappa}_{B_1}(B_2)$, where either $\delta(B_2) = 1$ or the left factor B_3 of B_2 satisfies $B_3 \prec B_1$. We then define

$$c_B(u) = \frac{1}{\kappa!}\Big(C_{B_1}(u)\Big)^{\kappa} c_{B_2}(u)\,.$$

With the help of the product formula, we can easily compute the values of the four coefficients Θ_i. First, notice that the linear combination $2\Theta_1 B_0 + 2\Theta_3 B_1$ is precisely the limiting extended input $\bar{\mathbf{u}}^\infty$ that we would have obtained if we had used, instead of our u^j, the HOS $\{\bar{u}^j\}$ given by a formula similar to (13) but with $\eta_{1,0} = \eta_{2,0} \equiv 0$, $\eta_{w_\ell} \equiv 1$, and $\eta_{w'_\ell} \equiv 0$. So to compute Θ_1 and Θ_3 it suffices to set

$$\begin{aligned} \bar{\Omega}(1) &= \{\pm w_1, \pm w_2, \pm w_3, \}\,, && (41)\\ \bar{\Omega}(2) &= \{\pm w_4, \pm w_5, \}\,, && (42) \end{aligned}$$

$$\begin{aligned} \bar{u}^j_1(t) &= j^{\frac{4}{5}} \sum_{\omega\in\bar{\Omega}(1)} e^{ij\omega t}\,, && (43)\\ \bar{u}^j_2(t) &= j^{\frac{4}{5}} \sum_{\omega\in\bar{\Omega}(2)} e^{ij\omega t}\,, && (44) \end{aligned}$$

compute the product expansion of the corresponding solution $\bar{S}^j$ of (36), and then take the limit as $j \to \infty$. We know that the limit $\bar{S}^\infty$ of the S^j satisfies

$$\frac{d}{dt}\bar{S}^\infty(t) = \bar{S}^\infty(t)\bar{\mathbf{u}}^\infty(t)\ .$$

Therefore

$$\bar{S}^{\infty}(t) = e^{t(2\Theta_1 B_0 + 2\Theta_3 B_1)}.$$

If we work in the nilpotent algebra $A_5(\mathbf{X})$, then B_0 and B_1 commute, so we have $\bar{S}^{\infty}(t) = e^{2\Theta_1 t B_0} e^{2\Theta_3 t B_1}$. It is easily seen that B_0 and B_1 are different elements of a P. Hall basis. So

$$\begin{aligned} \Theta_1 &= \frac{1}{2t} \lim_{j\to\infty} C_{B_0}(\bar{u}^j)(t), \\ \Theta_3 &= \frac{1}{2t} \lim_{j\to\infty} C_{B_1}(\bar{u}^j)(t). \end{aligned}$$

The product formula gives

$$\begin{aligned} C_{B_0}(\bar{u}^j)(t) = \frac{1}{2!}\int_0^t \Big(\int_0^s \Big(\int_0^{\tau_1} \bar{u}_1^j(\tau_2)\, d\tau_2\Big)\bar{u}_2^j(\tau_1) d\tau_1\Big) \\ \times \Big(\int_0^s \bar{u}_1^j(\tau)\, d\tau\Big)^2 \bar{u}_2^j(s)\, ds, \end{aligned}$$

$$C_{B_1}(\bar{u}^j)(t) = \frac{1}{3!}\int_0^t \Big(\int_0^s \bar{u}_2^j(\tau) d\tau\Big)\Big(\int_0^s \bar{u}_1^j(\tau) d\tau\Big)^3 \bar{u}_2^j(s) ds.$$

To study the asymptotic behavior of these expressions as $j \to \infty$, we just expand each $\bar{u}_\ell^j$ as a sum of six terms (if $\ell = 1$) or four terms (if $\ell = 2$). Then C_{B_0} and C_{B_1} are expressed as sums of $6^3 2^2 = 864$ terms, which may appear to be combinatorially overwhelming. However, it is not hard to see that the only terms that make a nontrivial contribution to the asymptotic behavior are those that correspond to five frequencies that add up to zero. There are exactly 24 such terms. For C_{B_0}, a typical term of this kind is the one that corresponds to the frequencies w_1, w_4, w_2, w_3, w_5. This term contributes $\frac{t}{2w_1 w_2 w_3 (w_1+w_4)}$ to the limit of $C_{B_0}(\bar{u}^j)(t)$. If we symmetrize with respect to $\{w_1, w_2, w_3\}$, and with respect to $\{w_4, w_5\}$, we get 12 terms whose total contribution is precisely $g_{B_0}(w_1, \ldots, w_5)$. Another contribution equal to $g_{B_0}(w_1, \ldots, w_5)$ results from the 12 terms obtained by replacing the w_k by their negatives. So the limit of $C_{B_0}(\bar{u}^j)(t)$ is $2g_{B_0}(w_1, \ldots, w_5)t$. Similarly, the limit of $C_{B_1}(\bar{u}^j)(t)$ is $2g_{B_1}(w_1, \ldots, w_5)t$. From this it follows that $\Theta_1 = g_{B_0}(w_1, \ldots, w_5)$ and $\Theta_3 = g_{B_1}(w_1, \ldots, w_5)$. The formulas for Θ_2 and Θ_4 are exactly the same, except that the w_k are replaced by the w'_k.

The proof of Theorem 1 is now complete. ∎

Remark 9. Using the appropriate algebraic formalism, one can get simple expressions for the functions g_B associated to every P. Hall bracket. One can then prove the following in complete generality. Let ν_1, ν_2 be integers > 0. Write $\nu = \nu_1 + \nu_2$, and let $\mu = \dim L(\mathbf{X})_{\nu_1,\nu_2}$, so that there are exactly μ P. Hall brackets B such that $\delta_1(B) = \nu_1$ and $\delta_2(B) = \nu_2$. Then the corresponding functions g_B are linearly independent on the hyperplane $H = \{(\omega_1, \ldots, \omega_\nu) : \omega_1 + \ldots + \omega_\nu = 0\}$. This fact, that was verified directly here for $\nu_1 = 3$, $\nu_2 = 2$, $\mu = 2$, guarantees that the analogue of the linear system (17), (18) always has a solution. ▮

SOME REMARKS ON THE GENERAL CASE

The analysis carried out in this paper applies almost without change to a general single-bracket extension of a general m-input driftless system. The only new technical difficulty is the problem discussed in Remark 9, that is, the linear independence of the functions g_B for different P. Hall brackets of the same degree with respect to each X_ℓ.

General multibracket extensions can then be handled by superposition, as explained above. (We have already encountered one example of such superposition: the limit input $\mathbf{u}^\infty$ is the sum of two parts, namely, the one arising from the w_ℓ and that resulting from the w'_ℓ. This was a consequence of the independence of the two sets of frequencies. It turns out that, provided one takes care of some algebraic complications, this generalizes. One can produce a desired limiting extended input $\mathbf{v}$ by taking an HOS which is a sum of parts, one for each totally homogeneous component, and making sure that the corresponding sets of frequencies are independent.)

DRIFT AND FEEDBACK

A trivial particular case of the above is the seemingly more general situation where our original system is of the form

$$\dot{x} = f_0(x) + \sum_{k=1}^{m} u_k(t) f_k(x)\,, \tag{45}$$

(i.e. has a drift term), but the extended system

$$\dot{x} = f_0(x) + \sum_{k=1}^{r} v_k(t) f_k(x)\,, \tag{46}$$

has the special property that the new brackets f_k, $k > m$, do not involve f_0. In this case we can apply our general construction thinking of (45) as a system without drift, and of the extended input that we want to approximate as one whose X_0 component happens to be 1. Since the new brackets do not involve f_0, our construction yields approximating inputs u^j whose zeroth component is just 1, so we end up producing inputs for (45) that T-converge to the desired limit.

It has been suggested by L. Gurvits (personal communication) that the approximation algorithm can be used to produce *feedback laws*, as follows. Suppose that for a suitable Lie bracket extension (46) of (45) there is a feedback control law $v_k = \varphi_k(x)$, $k = 1, \ldots, r$, that has some desired properties. Then one may try to produce time-dependent feedback laws $u_k = \psi_{k,\varepsilon}(x, t)$ that, as $\varepsilon \to 0$, give rise to trajectories that converge to those of the closed-loop system

$$\dot{x} = f_0(x) + \sum_{k=1}^{r} \varphi_k(x) f_k(x), \tag{47}$$

at least on some fixed time-interval $[0, T]$. Our algorithm makes it possible to produce such an approximation as follows. Assuming that the functions φ_k are sufficiently smooth, then we can regard the control system

$$\dot{x} = f_0(x) + \sum_{k=1}^{r} w_k \varphi_k(x) f_k(x), \tag{48}$$

as arising from a control specialization of a Lie bracket extension of a system

$$\dot{x} = f_0(x) + \sum_{k=1}^{\rho} \hat{w}_k g_k(x)\,, \tag{49}$$

where the vector fields g_k are linear combinations of $f_1, \ldots, f_m$ with sufficiently smooth coefficients. For example: suppose $m = 2$, and (47) is

$$\dot{x} = f_0(x) + \varphi_1(x) f_1(x) + \varphi_2(x) f_2(x) + \varphi_3(x)[f_1, f_2](x)\ .$$

We rewrite $\varphi_3[f_1, f_2](x)$ as $[f_1, \varphi_3 f_2] - (L_{f_1}\varphi_3) f_2$, and let $g_1 = f_1$, $g_2 = \varphi_1 f_1$, $g_3 = \varphi_2 f_2$, $g_4 = (L_{f_1}\varphi_3) f_2$, $g_5 = \varphi_3 f_2$. Then (48) is just given by

$$\dot{x} = f_0(x) + w_1 g_2(x) + w_2 g_3(x) + w_3 g_4(x) + w_3 [g_1, g_5](x)\,,$$

which arises from the extension

$$\Sigma_e\ :\ \dot{x} = f_0(x) + \sum_{k=1}^{5} \bar{w}_k g_k(x) + \bar{w}_6 [g_1, g_5](x)$$

of

$$\Sigma\ :\ \dot{x} = f_0(x) + \sum_{k=1}^{5} \hat{w}_k g_k(x)$$

by specializing the control to $\bar{w}_1 = \bar{w}_5 = 0$, $\bar{w}_2 = w_1$, $\bar{w}_3 = w_2$, $\bar{w}_4 = \bar{w}_6 = w_3$. If we use our algorithm to produce a sequence $\{\hat{w}^j\}$ of inputs for Σ that T-converge to the extended input $X_2 + X_3 + X_4 + [X_1, X_5]$, then the time-dependent feedbacks

$$\begin{aligned} u_1^j(x,t) &= \hat{w}_1^j(t) + \hat{w}_2^j(t)\varphi_1(x)\,, \\ u_2^j(x,t) &= \hat{w}_3^j(t)\varphi_2(x) + \hat{w}_4^j(t)(L_{f_1}\varphi_3)(x) + \hat{w}_5^j(t)\varphi_3(x) \end{aligned}$$

have the desired properties.

References

[1] Bourbaki, N., *Elements of Mathematics: Lie Groups and Lie Algebras*, Hermann (1975) Paris.

[2] Brockett, R. W., "Control Theory and Singular Riemannian Geometry," in *New Directions in Applied Mathematics*, Springer-Verlag, New York (1981), pp. 11-27.

[3] Fernandes, C., L. Gurvits, and Z.X. Li, "Foundations of Nonholonomic Motion Planning," Technical Report, Robotics Research Laboratory, Courant Institute of Mathematical Sciences, 1991.

[4] Gurvits, L., and Z.X. Li, "Theory and Application of Nonholonomic Motion Planning," Technical Report, Courant Institute of Mathematical Sciences, 1990.

[5] Haynes, G. W., and H. Hermes, "Nonlinear Controllability via Lie Theory," *SIAM J. Control and Optimization* **8**, No. 4, Nov. 1970, pp. 450-460.

[6] Hauser, J., S. S. Sastry, and P. V. Kokotovic, "Nonlinear Control via Approximate Input-output Linearization: The Ball and Beam Example," *Proc. 28th IEEE CDC*, Tampa, Florida, Dec. 1989, pp. 1987-1993.

[7] Kurzweil, J., and J. Jarnik, "Limit Process in Ordinary Differential Equations," *Journal of Applied Mathematics and Physics* **38**, March 1987, pp. 241-256.

[8] Kurzweil, J., and J. Jarnik, "Iterated Lie Brackets in Limit Processes in Ordinary Differential Equations," *Results in Mathematics* **14**, 1988, pp. 125-137.

[9] Kurzweil, J., and J. Jarnik, "A Convergence Effect in Ordinary Differential Equations," in *Asymptotic Methods of the Mathematical Physics*, (Russian), 301, 'Naukova Pumka,' Kiev, 1989.

[10] Lafferriere, G., "A General Strategy for Computing Steering Controls of Systems without Drift," Submitted to the 1991 CDC.

[11] Lafferriere, G., and H. J. Sussmann, "Motion Planning for Controllable Systems without Drift: A Preliminary Report," Rutgers University Systems and Control Center Report SYCON-90-04.

[12] Lafferriere, G., and H. J. Sussmann, "Motion Planning for Controllable Systems without Drift," *Proc. IEEE Int. Conf. Robotics and Automation,* Sacramento, CA, April 1991.

[13] Murray, R. M., and S. S. Sastry, "Grasping and Manipulations Using Multifingered Robot Hands," Memorandum No. UCB/ERL M90/24, Electronics Research Laboratory, Univ. of California, Berkeley, March 1990.

[14] Murray, R. M., and S. S. Sastry, "Steering Controllable Systems," *Proc. 29th IEEE CDC*, Honolulu, Hawaii, Dec. 1990.

[15] Sastry, S. S., and Z. Li, "Robot Motion Planning with Nonholonomic Constraints," *Proc. 28th IEEE CDC*, Tampa, Florida, Dec. 1989, pp. 211-216.

[16] Sussmann, H. J., "A Product Expansion for the Chen Series," in *Theory and Applications of Nonlinear Control Systems*, C. Byrnes and A. Lindquist Eds., North-Holland, 1986, pp. 323-335.

[17] Sussmann, H. J., "A General Theorem on Local Controllability," *SIAM J. Control and Optimization* **25** No. 1, January 1987, pp. 158-194.

[18] Sussmann, H. J., and W. Liu, "Limiting Behavior of Trajectories for Highly Oscillating Controls," in preparation.

[19] Sussmann, H. J., and W. Liu, "Limits of Highly Oscillatory Controls and Approximation of General Paths by Admissible Trajectories." To appear in *Proc. 30th IEEE CDC*, Brighton, UK, Dec. 1991.