Series Expansions for Analytic Systems Linear in Control [1]

Francesco Bullo [a]

[a] *Coordinated Science Laboratory and General Engineering Department*
University of Illinois at Urbana-Champaign, 1308 W. Main St, Urbana, IL 61801
Tel: (+1) 217 333 0656, Fax: (+1) 217 244 1653, Email: bullo@uiuc.edu

Abstract

This paper presents a series expansion for the evolution of a class of nonlinear systems characterized by constant input vector fields. We present a series expansion that can be computed via explicit recursive expressions, and we derive sufficient conditions for uniform convergence over the finite and infinite time horizon. Furthermore, we present a simplified series and convergence analysis for the setting of second order polynomial vector fields. The treatment only relies on elementary notions on analytic functions, number theory, and operator norms.

Key words: nonlinear systems, perturbation theory, Volterra series

1 Introduction

This papers studies series expansions for the evolution of a class of nonlinear control systems. Series expansions are an enabling tool in trajectory generation and optimization problems, e.g., see Lafferriere and Sussmann [18], Leonard and Krishnaprasad [19], Rui et al. [24], and the author's work [5]. Furthermore, series expansions play a key role in the study of sufficient and necessary conditions for local nonlinear controllability; see Kawski [12], and in other areas such as geometric integration; see McLachlan et al. [22], and realization theory; see Isidori [10, Section 3.4 and 3.5].

Volterra series and other types of expansions have received much attention in the literature. Some early work includes the Magnus [21] and Chen [6] series. Volterra series were then studied in Brockett [2], Bruni et al. [3], Gilbert [9], Lesiak and Krener [20]. Fliess [7] later provided a comprehensive treatment of what is now known as the Chen-Fliess series. Motivated by controllability and normal form theory, Kawski and Sussmann [13] obtained increasingly sophisticated versions of the Chen-Fliess series. In a related line of research, the two textbooks Rugh [23] and Schetzen [27] focus on the input/output representation of nonlinear systems via series expansions. Some advanced results within this context are found in Sandberg [25] and Boyd and Chua [1]. These works and this work's contribution are compared below.

The main contributions of this paper are novel series expansions for nonlinear control systems described by smooth ordinary differential equations linear in control. The model system we consider is described by the differential equation

$$\dot{x}(t) = f(x(t)) + Bu(t),$$

where the vector field f is analytic, the matrix B belongs to $\mathbb{R}^{n\times m}$, and the components $(u_1, \dots, u_m)$ of the control u are piecewise continuous functions of time. We call such analytic systems *linear in control*. For this class of nonlinear systems we present a series expansion that can be computed via explicit recursive expressions. Additionally, we present a simplified series assuming the components of f are polynomial of first and second degree. The presentation and derivation rely only on elementary tools and the final series appear in a format similar but not identical to the classic Volterra format.

The convergence properties for the series are characterized via asymptotic bounds on the truncation error. The series expansion converges uniformly over the infinite time horizon provided the linearized system is exponentially stable (i.e., the Jacobian linearization of the drift vector field at the origin is a Hurwitz matrix) and the input norm is bounded by a computable constant. Alternatively, for an arbitrary input, the series is guaranteed to converge over a computable finite time interval.

[1] Submitted to Automatica on May 7, 2000. Revised on December 1, 2001. This version: 14 February 2003.

These limitations are similar to the ones present in perturbation and averaging methods in dynamical systems theory, e.g., see Khalil [14, Section 8.2] and Sanders and Verhulst [26, page 71].

The presentation is organized as follows. We start by comparing the novel series expansions with several previous approaches. Section 2 reviews some basic facts on operator norms. Section 3 discusses the setting of analytic systems linear in control while Section 4 discusses the special case of polynomial systems of second order. A final discussion is presented in Section 5.

Comparison with Volterra and Chen-Fliess expansions

For a review of Chen-Fliess and Volterra series we refer to Chapter 3, and to the Bibliographical Notes in Isidori [10], and to the survey in Fliess et al. [8]. Many of the following comments are inspired by these works.

Volterra series It is known that bilinear systems admit explicit expressions for the Volterra kernels. Because the *Carleman linearization* procedure transforms a generic nonlinear control system into a bilinear system, see Rugh [23, Section 3.3] and Krener [17], explicit expressions for the kernels of any nonlinear system are theoretically available. However, the Carleman linearization technique has various disadvantages. First of all, the procedure generates a high dimensional system which only approximates the correct dynamics; see the discussion in Rugh [23, Section 3.2]. Furthermore, only weaker convergence properties are presented in the treatments [17, 23] than the ones provided in this work, see Theorem 3.1 below.

For generic analytical systems, Lesiak and Krener [20] give explicit (not recursive) expression for the Volterra kernels under the assumption that the flow map of the drift vector field is available. From manipulations of the Chen-Fliess series (see below), one can derive the series expansions for these kernels. To the best of the author's knowledge, no explicit or recursive expressions are currently available.

Chen-Fliess series This is arguably the most successful approach to writing the evolution of a nonlinear analytic control system, see Fliess [7] and Fliess et al. [8]. Remarkably, the terms of the Chen-Fliess expansion are explicitly known. Modified versions of the Chen-Fliess expansion are proposed by Sussmann [28] and later by Kawski and Sussmann [13].

With respect to the work presented here, the Chen-Fliess expansion is more explicit (i.e., we only provide recursive expressions), and more general (i.e., we require systems linear in control). On the other hand, our series has a similar interpretation to the Volterra series: the kth term has order k with respect to the input magnitude, e.g., the first order term has the conventional interpretation of being the response of the linearized system. Additionally, the series has stronger convergence properties than the Chen-Fliess series, e.g., convergence over the infinite time horizon.

Wiener/Volterra systems representation Finally, it is worth mentioning the input/output approach. In this context, models of nonlinear systems are based on input-output operators, see for example the treatment in Boyd and Chua [1] and the textbooks Rugh [23], Schetzen [27]. Under a "fading memory" assumption, Boyd and Chua [1] obtain convergence properties over the infinite time horizon. This fading memory assumption is consistent with the stability requirements of this work and of the classic perturbation methods from dynamical systems theory, see [14, 26].

2 Preliminaries

Elementary number theory concepts

We refer the reader to [15, 29] for a basic introduction into the subject of generating function. We will only quickly review some specific notions when needed here.

Let $\mathbb{N}$ be the set of positive integer numbers, $\mathbb{R}$ the set of real numbers, and $\mathbb{C}$ the set of complex numbers. Let $k \in \mathbb{N}$, and let $P(k)$ be the set of ordered partitions of k. For example, $P(3)$ is the collection $\{\{3\}, \{2,1\}, \{1,2\}, \{1,1,1\}\}$. The set $P(k)$ contains 2^{k-1} elements. [2] Let $P(i,j)$ be the set of ordered sequences of j integers that sums up to i, and let $P(k) - \{k\}$ be the set $P(k)$ minus the element $\{k\}$.

The initial value problem and various Taylor expansions

Let x take value in $\mathbb{R}^n$ and let t belong to an interval I: the finite time case, i.e., $I = [0,T]$, as well as the infinite time horizon case, i.e., $I = [0,\infty)$, are of interest. Consider the initial value problem

$$\begin{aligned} \dot{x}(t) &= f(x(t)) + g(t) \\ x(0) &= 0, \end{aligned} \tag{1}$$

where f and g are vector fields on $\mathbb{R}^n$. The components $\{f_1, \ldots, f_n\}$ of f are analytic functions in a neighborhood of the origin $0 \in \mathbb{R}^n$, and the components $\{g_1, \ldots, g_n\}$ of g are independent of x and piecewise continuous, uniformly bounded functions of time t. The initial value problem (1) is thought of as a control system by setting $g(t) = Bu(t)$, where $B \in \mathbb{R}^{n \times m}$, and $u = (u_1, \ldots, u_m)$. Since the input vector field does not depend on x, we refer to this system as *linear in control*.

[2] Consider the $(k-1)$ possible vertical lines separating k aligned points. There are 2^{k-1} possible vertical line configurations and each corresponds to an ordered partition.

Let $f(0)=0$, and for $i=1,\dots,n$, develop the functions f_i in a Taylor expansion about the origin via

$$f_i(x) = \sum_{m=1}^{+\infty} \sum_{\substack{j_1+\dots+j_n=m \\ j_1,\dots,j_n\geq 0}} \frac{1}{j_1!\cdots j_n!}\left(\frac{\partial^m}{\partial x_1^{j_1}\dots\partial x_n^{j_n}} f_i(0)\right) x_1^{j_1}\cdots x_n^{j_n}.$$

Equivalently, let $f_i(x) = \sum_{m=1}^{+\infty} f_i^{[m]}(x,\dots,x)$, where the tensors $f_i^{[m]} : \underbrace{\mathbb{R}^n\times\dots\times\mathbb{R}^n}_{m} \to \mathbb{R}$, for $m\in\mathbb{N}$, are computed according to

$$\begin{aligned} &f_i^{[m]}(y_1,\dots,y_m) \\ &= \sum_{k_1,\dots,k_m=1}^{n} \frac{1}{m!}\left(\frac{\partial^m}{\partial x_{k_1}\dots\partial x_{k_m}} f_i(0)\right)(y_1)_{k_1}\cdots(y_m)_{k_m} \\ &= \sum_{k_1,\dots,k_m=1}^{n} \left(f_i^{[m]}\right)^{k_1\dots k_m}(y_1)_{k_1}\cdots(y_m)_{k_m}. \end{aligned}$$

In the last equation, the vectors $y_1,\dots,y_m$ belong to $\mathbb{R}^n$, and the symbol $(y_j)_k$ denotes the kth component of y_j. These definitions are readily repeated in vector format for the vector field f by neglecting the subscript i. It will be convenient to adopt the notation $Ax = f^{[1]}(x)$.

Remark 2.1 *(i) The setting of systems linear in control as in equation* (1) *is not overly restrictive. If the forcing term g is a function of both time and state, i.e., $g=g(t,x)$, a change of coordinates might remove the dependency on x; the necessary and sufficient condition for the existence of this transformation is the involutivity of the distribution $\{g(t,\cdot), t\in\mathbb{R}_+\}$. It is also possible to extend the state space by adding integrators to the control inputs via $\dot{u}=v$. If $x(0)\neq 0$ or if $f(0)\neq 0$, it might be possible to redefine f and g to match the system definition in equation* (1).

(ii) The sequence of tensors $\{f^{[m]}, m\in\mathbb{N}\}$ uniquely determines the Lie algebraic structure at the origin of the control system in equation (1)*. For example, one can see that*

$$\frac{1}{2}[g(t_2),[g(t_1),f(x)]]\Big|_{x=0} = f^{[2]}\left(g(t_1),g(t_2)\right).$$

Operator norms and their estimates

In defining mapping and norms we follow the notation in [14, Chapter 6]. Consider the normed linear space $\mathcal{L}_\infty^n$ of piecewise continuous, uniformly bounded functions over the interval I

$$x : I\subset\mathbb{R}_+\to\mathbb{R}^n; \quad t\mapsto x(t),$$

with norm

$$\|x\|_{\mathcal{L}_\infty} = \sup_{t\in I}\|x(t)\|_\infty = \sup_{t\in I}\max_{i=1,\dots,n}|x_i(t)| < \infty.$$

Assume the matrix A is Hurwitz or that the interval I is finite, and let H_A be the mapping

$$H_A : \mathcal{L}_\infty^n \to \mathcal{L}_\infty^n; \quad x(t)\mapsto \int_0^t \mathrm{e}^{A(t-\tau)}\,x(\tau)d\tau.$$

The $\mathcal{L}_\infty^n$ induced norm for H_A is

$$\|H_A\|_{\mathcal{L}_\infty} = \left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} = \max_{i=1,\dots,n}\sum_{j=1}^{n}\int_{t\in I}|(\mathrm{e}^{At})_{ij}|dt.$$

Next, we consider the vector field f and its derived tensors $f^{[m]}$. For simplicity we start by considering the 2-tensor $f^{[2]} : \mathcal{L}_\infty^n\times\mathcal{L}_\infty^n\to\mathcal{L}_\infty^n$ defined via

$$(x(t),y(t))\mapsto f^{[2]}(x(t),y(t)).$$

and defining its induced norm $\left\|f^{[2]}\right\|_{\mathcal{L}_\infty}$ via

$$\begin{aligned} \left\|f^{[2]}\right\|_{\mathcal{L}_\infty} &= \max_{\substack{\|y_1\|_{\mathcal{L}_\infty}=1\\ \|y_2\|_{\mathcal{L}_\infty}=1}} \left\|f^{[2]}(y_1,y_2)\right\|_{\mathcal{L}_\infty} \\ &\leq \max_{i=1,\dots,n}\sum_{k_1,k_2=1}^{n}|(f_i^{[2]})^{k_1k_2}|. \end{aligned} \tag{2}$$

More generally, we examine the m-tensor $f^{[m]}$ and define its induced norm via

$$\begin{aligned} \left\|f^{[m]}\right\|_{\mathcal{L}_\infty} &= \max_{\substack{\|y_j\|_{\mathcal{L}_\infty}=1\\ j=1,\dots,m}} \left\|f^{[m]}(y_1,\dots,y_j)\right\|_{\mathcal{L}_\infty} \\ &\leq \max_{i=1,\dots,n}\sum_{k_1,\dots,k_m=1}^{n}\left|\left(f_i^{[m]}\right)^{k_1\dots k_m}\right| \\ &= \max_{i=1,\dots,n}\sum_{k_1,\dots,k_m=1}^{n}\frac{1}{m!}\left|\frac{\partial^m}{\partial x_{k_1}\dots\partial x_{k_m}}f_i(0)\right|. \end{aligned}$$

Note that the estimates on the norm of the mappings are only upper bounds whenever $m\geq 2$. For example, consider the 2-tensor $f^{[2]}\big((x_1,x_2),(y_1,y_2)\big) = (-x_1y_1 + x_2y_2 + x_1y_2 + x_2y_1, 0)$ defined over $\mathcal{L}_\infty^2$. Its $\mathcal{L}_\infty$-norm is 2, whereas the estimate presented in equation (2) is 4.

An upper bound on the operator norm $\left\|f^{[m]}\right\|_{\mathcal{L}_\infty}$ is provided by the *Cauchy estimates* for the Taylor series coefficient of an analytic function, see [16, Section 2.3]. Since the vector field f is analytic at the origin, there exists a $\rho\in\mathbb{R}_+$ such that f is analytic over the domain

$$D_\rho = \{(z_1,\dots,z_n)\in\mathbb{C}^n \mid |z_i|<\rho\}.$$

We let $\|f\|_\rho$ denote the maximum value attained by the magnitude of all components of f over D_ρ. For any collection of nonnegative indices $j_1+\cdots+j_n=m$, we have

$$\left|\frac{\partial^m}{\partial x_1^{j_1}\dots\partial x_n^{j_n}} f_i(0)\right| \le \frac{j_1!\cdots j_n!}{\rho^m}\ \|f\|_\rho .$$

Conservatively bounding the right hand side by $(m!/\rho^m)\|f\|_\rho$, we have

$$\begin{aligned}\left\|f^{[m]}\right\|_{\mathcal{L}_\infty} &\le \sum_{k_1,\dots,k_m=1}^{n} \frac{1}{m!}\frac{j_1!\cdots j_n!}{\rho^m}\|f\|_\rho \\ &\le \frac{1}{\rho^m}\|f\|_\rho \sum_{k_1,\dots,k_m=1}^{n} 1 = \left(\frac{n}{\rho}\right)^m \|f\|_\rho .\end{aligned}$$

Finally, given any scalar analytic function h of a scalar variable η, we let $\text{Remainder}_{M+1}(h)(\eta)$ be its Taylor remainder of order $(M+1)$ about $\eta=0$, i.e., we write

$$h(\eta)=\sum_{m=1}^{M}\frac{h^{(m)}(0)}{m!}\eta^m+\text{Remainder}_{M+1}(h)(\eta).$$

3 A series expansion

Let $\epsilon\in\mathbb{R}_+$ and consider the initial value problem

$$\begin{aligned}\dot{x}(t,\epsilon)&=f(x(t,\epsilon))+\epsilon g(t) \\ x(0,\epsilon)&=0,\end{aligned} \tag{3}$$

where the solution x is a function of both $t\in I$ and $\epsilon\in\mathbb{R}_+$. There is no loss of generality in assuming $\|g\|_{\mathcal{L}_\infty}=\|f\|_\rho$, since the constant ϵ can be redefined. Recall the notation $Ax=f^{[1]}(x)=\partial f/\partial x(0)$.

Theorem 3.1 *Consider the initial value problem in equation* (3). *The solution $x: I\times\mathbb{R}_+\mapsto\mathbb{R}^n$ satisfies the formal expansion*

$$x(t,\epsilon)=\sum_{k=1}^{+\infty}\epsilon^k x_k(t) \tag{4}$$

$$x_1(t)=\int_0^t \mathrm{e}^{A(t-\tau)}\, g(\tau)d\tau \tag{5}$$

$$x_k(t)=\sum_{\{i_1\dots i_m\}\in P(k)-\{k\}} \int_0^t \mathrm{e}^{A(t-\tau)}\, f^{[m]}(x_{i_1}(\tau),\dots,x_{i_m}(\tau))d\tau. \tag{6}$$

Assume f analytic over the domain D_ρ. Without loss of generality let $\|g\|_{\mathcal{L}_\infty}=\|f\|_\rho$, and compute

$$\beta=\left(\frac{n}{\rho}\right)\left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1}\|f\|_\rho .$$

If $\beta\epsilon\le 1+2\beta-2\sqrt{\beta+\beta^2}$, the series in equation (4) *converges absolutely and uniformly in $t\in I$, and for all integers M the truncation error is bounded by*

$$\left\|x-\sum_{k=1}^{M}\epsilon^k x_k\right\|_{\mathcal{L}_\infty}\le\left(\frac{\rho}{n}\right)\text{Remainder}_{M+1}(h_\beta)(\beta\epsilon),$$

where $$h_\beta(\eta)=\frac{1+\eta-\sqrt{1-2(1+2\beta)\eta+\eta^2}}{2(\beta+1)}.$$

Comments

We start by computing some terms of the series. Dropping the argument τ inside the integral, the first few terms of equation (6) read

$$\begin{aligned}x_2(t)&=\int_0^t \mathrm{e}^{A(t-\tau)}\, f^{[2]}(x_1,x_1)d\tau \\ x_3(t)&=\int_0^t \mathrm{e}^{A(t-\tau)}\left\{2f^{[2]}(x_2,x_1)+f^{[3]}(x_1,x_1,x_1)\right\}d\tau \\ x_4(t)&=\int_0^t \mathrm{e}^{A(t-\tau)}\left\{2f^{[2]}(x_3,x_1)+f^{[2]}(x_2,x_2)\right. \\ &\qquad \left.+3f^{[3]}(x_2,x_1,x_1)+f^{[4]}(x_1,x_1,x_1,x_1)\right\}d\tau.\end{aligned}$$

A second remark concerns the truncation error estimate. At $M=0$, the estimate turns into an upper bound on the solution $\|x\|_{\mathcal{L}_\infty}$. In other words, whenever convergence is guaranteed we have

$$\|x\|_{\mathcal{L}_\infty}<\left(\frac{\rho}{n}\right)\left(1-\sqrt{\frac{\beta}{1+\beta}}\right).$$

The convergence properties are similar to the ones discussed in [14, Chapter 8]. The condition $\beta\epsilon\le 1+2\beta-2\sqrt{\beta+\beta^2}$, with $\beta=\left(\frac{n}{\rho}\right)\left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1}\|f\|_\rho$, implies the following statement: for any stable system there exists a small enough $\epsilon^\star=\epsilon^\star(\beta)$ such that for all $\epsilon<\epsilon^\star$ the series converges. Alternatively, given a specific value of ϵ, convergence is assured by finding a small enough β, that is, a small enough $\left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1}$. This is always possible since this norm goes to zero as the length of interval I vanishes. A lower bound to $1+2\beta-2\sqrt{\beta+\beta^2}$ is $1/(4\beta+2)$. This bound is asymptotically correct as β goes to $+\infty$. Accordingly, the convergence criteria can be restated in terms of $\epsilon\le 1/(4\beta^2+2\beta)$.

Proof of Theorem 3.1

The proof entails two parts. We first derive the power series in a formal matter, then show its absolute convergence. Following the perturbation methodology in [14, Section 8.1], we proceed as follows: we let $x(t, \epsilon)$ be the solution to the initial value problem (3) (necessarily a function of ϵ), substitute $x(t, \epsilon)$ into both sides of (3), develop both sides in their power series and equate same powers of ϵ. In computing power series expansions of a generic function $y(\epsilon)$, it is convenient to use the notation $[\epsilon^k]\ y(t, \epsilon)$ for the coefficient of ϵ^k in $y(t, \epsilon)$; this notation is taken from [15, Section 1.2.9] and [29, Section 1.2]. For the left side of equation (3) we easily have

$$[\epsilon^k]\ \dot{x}(t, \epsilon) = [\epsilon^k]\ \sum_{j=1}^{+\infty} \epsilon^j \dot{x}_j(t) = \dot{x}_k(t).$$

For the right hand side we compute

$$\begin{aligned} & f(x)\Big|_{x=\sum_{j=1}^{+\infty} \epsilon^j x_j} + \epsilon g(t) \\ & \quad = \sum_{m=1}^{+\infty} f^{[m]} \left(\sum_{i_1=1}^{+\infty} \epsilon^{i_1} x_{i_1}, \ldots, \sum_{i_m=1}^{+\infty} \epsilon^{i_m} x_{i_m} \right) + \epsilon g(t). \end{aligned}$$

The coefficient of ϵ is $f^{[1]}(x_1) + g(t) = Ax_1 + g(t)$. Hence

$$\dot{x}_1 = [\epsilon]\ \big(f(x) + \epsilon g(t)\big) = Ax_1 + g(t).$$

Equation (5) in the text follows from noting that the initial condition for x_1, as well as for any other x_k, is zero. We compute the coefficient of ϵ^k as

$$\begin{aligned} & [\epsilon^k]\ f(x(t, \epsilon)) \\ & \quad = [\epsilon^k]\ \sum_{m=1}^{+\infty} f^{[m]} \left(\sum_{i_1=1}^{+\infty} \epsilon^{i_1} x_{i_1}, \ldots, \sum_{i_m=1}^{+\infty} \epsilon^{i_m} x_{i_m} \right) \\ & \quad = \sum_{m=1}^{k} [\epsilon^k]\ f^{[m]} \left(\sum_{i_1=1}^{k} \epsilon^{i_1} x_{i_1}, \ldots, \sum_{i_m=1}^{k} \epsilon^{i_m} x_{i_m} \right) \\ & \quad = \sum_{m=1}^{k} \sum_{\{i_1 \ldots i_m\} \in P(k,m)} f^{[m]}(x_{i_1}, \ldots, x_{i_m}) \qquad (7) \\ & \quad = f^{[1]}(x_k) + \sum_{\{i_1 \ldots i_m\} \in P(k) - \{k\}} f^{[m]}(x_{i_1}, \ldots, x_{i_m}), \end{aligned}$$

where $P(k, m)$ is the set of ordered sequences of m integers summing up to k. This proves equation (6), since the differential equation for the order ϵ^k term is

$$\dot{x}_k = Ax_k + \sum_{\{i_1 \ldots i_j\} \in P(k) - \{k\}} f^{[j]}(x_{i_1}, \ldots, x_{i_j}).$$

In the second part of the proof we seek an upper bound on ϵ to guarantee that the series in equation (4) converges absolutely and uniformly over $t \in I$. Using the operators norms and bounds discussed in Section 2, we have

$$\|x_1\|_{\mathcal{L}_\infty} \le \left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} \|g\|_{\mathcal{L}_\infty} = \left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} \|f\|_\rho ,$$

and since $\left\|f^{[m]}\right\|_{\mathcal{L}_\infty} \le \|f\|_\rho \left(\frac{n}{\rho}\right)^m$,

$$\begin{aligned} \|x_k\|_{\mathcal{L}_\infty} & \le \left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} \|f\|_\rho \\ & \sum_{\{i_1 \ldots i_m\} \in P(k) - \{k\}} \left(\frac{n}{\rho}\right)^m \|x_{i_1}\|_{\mathcal{L}_\infty} \cdots \|x_{i_m}\|_{\mathcal{L}_\infty} . \end{aligned}$$

Let $\beta = \left(\frac{n}{\rho}\right) \left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} \|f\|_\rho$, and define the series of positive numbers $a_1 = 1$, and

$$a_k = \beta \sum_{\{i_1 \ldots i_m\} \in P(k) - \{k\}} a_{i_1} \cdots a_{i_m}$$

or equivalently $a_k = \dfrac{\beta}{1+\beta} \displaystyle\sum_{\{i_1 \ldots i_m\} \in P(k)} a_{i_1} \cdots a_{i_m}$. By induction one can show that

$$\|x_k\|_{\mathcal{L}_\infty} \le \left(\frac{\rho}{n}\right) \beta^k a_k. \qquad (8)$$

To characterize the behavior of the sequence $\{a_k, k \in \mathbb{N}\}$ we resort to the method of generating functions; see [15, 29]. Let $h(\eta) = \sum_{k=1}^{+\infty} a_k \eta^k$ and compute

$$h(\eta) = \eta + \frac{\beta}{\beta+1} \sum_{k=2}^{+\infty} \eta^k \sum_{\{i_1 \ldots i_m\} \in P(k)} a_{i_1} \cdots a_{i_m},$$

where

$$\begin{aligned} & \sum_{k=2}^{+\infty} \eta^k \sum_{\{i_1 \ldots i_m\} \in P(k)} a_{i_1} \cdots a_{i_m} \\ & \qquad = -\eta + \sum_{k=1}^{+\infty} \sum_{\{i_1 \ldots i_m\} \in P(k)} (a_{i_1} \eta^{i_1}) \cdots (a_{i_m} \eta^{i_m}) \end{aligned}$$

In the spirit of the generating function method one performs the simplification

$$\begin{aligned} & \sum_{k=1}^{+\infty} \sum_{\{i_1 \ldots i_m\} \in P(k)} (a_{i_1} \eta^{i_1}) \cdots (a_{i_m} \eta^{i_m}) \\ & \quad = \sum_{j=1}^{+\infty} \left(\sum_{i=1}^{+\infty} a_i \eta^i \right)^j = \sum_{j=1}^{+\infty} (h(\eta))^j = \frac{h(\eta)}{1 - h(\eta)}, \qquad (9) \end{aligned}$$

where the first equality is equivalent to equation (7) and the last equality holds under the assumption $h < 1$. This bound will be established *a posteriori*. The remaining steps only rely on basic algebra. We compute h as a function of η from the equation

$$h = \eta + \frac{\beta}{\beta+1}\left(-\eta + \frac{h}{1-h}\right)$$

to obtain (discarding a second solution)

$$h(\eta) = \frac{1+\eta-\sqrt{1-2(1+2\beta)\eta+\eta^2}}{2(\beta+1)}.$$

The function h is defined real for any $0 \leq \eta \leq 1+2\beta - 2\sqrt{\beta+\beta^2}$ and attains a maximum value of

$$\max_{0\leq\eta\leq 1+2\beta-2\sqrt{\beta+\beta^2}} h = 1 - \sqrt{\frac{\beta}{\beta+1}}.$$

As β increases, the convergence region and the maximum value of h diminishes. The bound on η translates into a (conservative) estimate on how large ϵ can be in order for the series in equation (4) to converge. The bound on $h(\eta)$ translates into an estimate of the corresponding norm of the displacement $\|x(t)\|_{\mathcal{L}_\infty}$ over the domain of guaranteed convergence. In any case h is always less than unity, so that the equality (9) is justified *a posteriori*.

Finally, from the estimate in equation (8) we obtain

$$\begin{aligned}\left\| x - \sum_{k=1}^{M} \epsilon^k x_k \right\|_{\mathcal{L}_\infty} &\leq \sum_{k>M} \epsilon^k \|x_k\|_{\mathcal{L}_\infty} \\ \leq \left(\frac{\rho}{n}\right) \sum_{k>M} a_k \epsilon^k \beta^k &= \left(\frac{\rho}{n}\right) \mathrm{Remainder}_{M+1}\left(h(\beta\epsilon)\right).\end{aligned}$$

The convergence statement follows by noting that $h(\beta\epsilon)$ can be developed in a convergent Taylor expansion about $\epsilon = 0$ in a radius $\beta\epsilon \leq 1+2\beta-2\sqrt{\beta+\beta^2}$.

4 Second order polynomial systems

Polynomial vector fields are common in example applications, e.g., see [5], and are important in the study of normal forms, e.g., see [11]. Furthermore, control systems may be written in polynomial form via coordinate transformations as well as via dynamic extension of the state space. In this section, we investigate whether simpler expressions or stronger convergence properties are available for this subclass of systems. Specifically, we focus on systems described by a vector field f whose components are first and second order polynomial functions. In other words, we consider a control system with $f^{[1]}$ and $f^{[2]}$ as only non-vanishing tensors:

$$\begin{aligned}&\dot{x}(t,\epsilon) = Ax(t,\epsilon) + f^{[2]}(x(t,\epsilon), x(t,\epsilon)) + \epsilon g(t) \qquad (10)\\ &x(0,\epsilon) = 0.\end{aligned}$$

Theorem 4.1 *Consider the initial value problem in equation* (10)*. The solution $x : I \times \mathbb{R}_+ \mapsto \mathbb{R}^n$ satisfies*

$$\begin{aligned}x(t,\epsilon) &= \sum_{k=1}^{+\infty} \epsilon^k x_k(t)\\ x_1(t) &= \int_0^t \mathrm{e}^{A(t-\tau)}\, g(\tau) d\tau\\ x_k(t) &= \sum_{i=1}^{k-1} \int_0^t \mathrm{e}^{A(t-\tau)}\, f^{[2]}(x_i(\tau), x_{k-i}(\tau)) d\tau, \quad k \geq 2.\end{aligned}$$

Without loss of generality assume $\|g\|_{\mathcal{L}_\infty} = \left\|f^{[2]}\right\|_{\mathcal{L}_\infty}$, and compute

$$\beta = 2\left\|\mathrm{e}^{At}\right\|_{\mathcal{L}_1} \left\|f^{[2]}\right\|_{\mathcal{L}_\infty}.$$

If $\beta^2\epsilon < 1$, the series converges absolutely and uniformly in $t \in I$, and for all integers M the truncation error is bounded by

$$\left\| x - \sum_{k=1}^{M} \epsilon^k x_k \right\|_{\mathcal{L}_\infty} \leq \frac{1}{\beta}\, \mathrm{Remainder}_{M+1}\left(1 - \sqrt{1-\beta^2\epsilon}\right).$$

Comments

The first few terms of the resulting series are:

$$\begin{aligned}x_2(t) &= \int_0^t \mathrm{e}^{A(t-\tau)}\, f^{[2]}(x_1, x_1) d\tau\\ x_3(t) &= \int_0^t \mathrm{e}^{A(t-\tau)} \left\{2f^{[2]}(x_2, x_1)\right\} d\tau\\ x_4(t) &= \int_0^t \mathrm{e}^{A(t-\tau)} \left\{2f^{[2]}(x_3, x_1) + f^{[2]}(x_2, x_2)\right\} d\tau.\end{aligned}$$

Note the agreement with the expressions for the analytic case. The polynomial nature of the control system (10) leads to simplifications in the bound on the solution $\|x\|_{\mathcal{L}_\infty}$ and in the computation of the parameter β: no norms over complex planes are required. Whenever convergence is guaranteed it holds

$$\|x\|_{\mathcal{L}_\infty} < \frac{1}{\beta}.$$

Even though the estimates for polynomial vector fields have a simpler expression, they qualitatively agree with the ones for the more general analytic vector field case: given the parameter β, the series converges for all forcing terms smaller in magnitude than a constant $\epsilon^\star = \epsilon^\star(\beta)$.

Finally, we consider the initial value problem

$$\dot{x} = -x + x^2 + u(t), \quad x(0) = 0.$$

We compute $\left\|f^{[2]}\right\|_{\mathcal{L}_\infty} = 1$ and $\left\|e^{At}\right\|_{\mathcal{L}_1} = 1 - e^{-T}$, assuming the interval of interest is $I = [0,T]$. We write $u(t) = \epsilon g(t)$, where we set $\epsilon = \|u\|_{\mathcal{L}_\infty}$ and $g(t) = u(t)/\|u\|_{\mathcal{L}_\infty}$. Accordingly, the assumption $\|g\|_{\mathcal{L}_\infty} = \left\|f^{[2]}\right\|_{\mathcal{L}_\infty}$ in Theorem 4.1 is satisfied. The convergence bound in Theorem 4.1 becomes

$$\|u\|_{\mathcal{L}_\infty} (1 - e^{-T})^2 < 1/4. \tag{11}$$

The guaranteed convergence region is depicted in Figure 1. Since the origin is an exponentially stable equilibrium point for the system, convergence is guaranteed over the infinite time horizon for all $\|u\|_{\mathcal{L}_\infty} < 1/4$. The bound $\|u\|_{\mathcal{L}_\infty} < 1/4$ is tight, since any constant input $u(t) = u > 1/4$ leads to solutions with finite escape time. The solution at $u_{\text{bound}}(t) = 1/4$ is

$$x_{\text{bound}}(t) = \frac{t}{2(2+t)}. \tag{12}$$

Figure 2 presents approximations of increasing order computed according to Theorem 4.1.

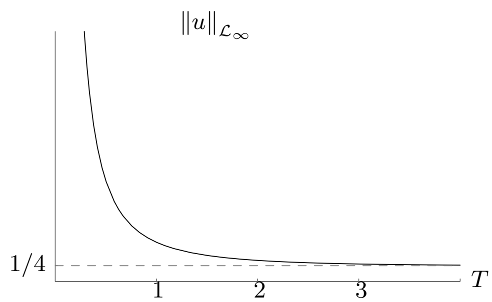

Fig. 1. According to equation (11), convergence for time T is guaranteed provided $\|u\|_{\mathcal{L}_\infty}$ is below the depicted curve.

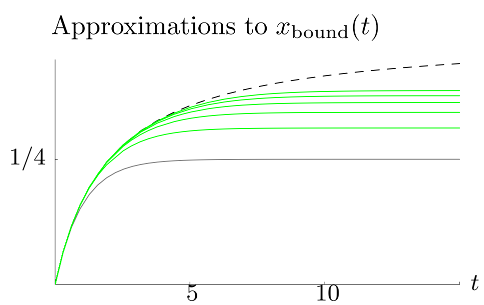

Fig. 2. Comparing the exact evolution in equation (12) (in dashed line) with approximations of increasing order.

Proof of Theorem 4.1

The expressions for the terms of the series are immediately derived from Theorem 3.1. Since $f^{[m]} = 0$ for all $m \geq 3$, we restrict the summation indices $\{i_1 \dots i_m\} \in P(k) - \{k\}$ to the set $P(k,2)$. With regards to the estimates and the convergence properties, we write

$$\|x_1\|_{\mathcal{L}_\infty} \leq \left\|e^{At}\right\|_{\mathcal{L}_1} \|g\|_{\mathcal{L}_\infty} = \left\|e^{At}\right\|_{\mathcal{L}_1} \left\|f^{[2]}\right\|_{\mathcal{L}_\infty}$$

$$\|x_k\|_{\mathcal{L}_\infty} \leq \left\|e^{At}\right\|_{\mathcal{L}_1} \left\|f^{[2]}\right\|_{\mathcal{L}_\infty} \sum_{i=1}^{k-1} \|x_i\|_{\mathcal{L}_\infty} \|x_{k-i}\|_{\mathcal{L}_\infty}.$$

Define the sequence $a_1 = \frac{1}{2}$, $a_k = \frac{1}{2}\sum_{i=1}^{k-1} a_i a_{k-i}$, and prove by induction that

$$\|x_k\|_{\mathcal{L}_\infty} \leq 2^{2k-1} \left\|e^{At}\right\|_{\mathcal{L}_1}^{2k-1} \left\|f^{[2]}\right\|_{\mathcal{L}_\infty}^{2k-1} a_k \equiv \beta^{2k-1} a_k.$$

Modulo a scaling factor, the numbers $\{a_k, k \in \mathbb{N}\}$ are the *Catalan numbers*, see [29, Section 2.3]. Their generating function is

$$h(\eta) = \sum_{k=1}^{+\infty} a_k \eta^k = 1 - \sqrt{1-\eta},$$

which is defined real for any $0 \leq \eta \leq 1$ and attains a maximum value of 1. This observation and the estimate on $\|x_k\|_{\mathcal{L}_\infty}$ readily lead to

$$\left\|x - \sum_{k=1}^{M} \epsilon^k x_k\right\|_{\mathcal{L}_\infty} \leq \sum_{k>M} \epsilon^k \|x_k\|_{\mathcal{L}_\infty} \leq \sum_{k>M} \epsilon^k \beta^{2k-1} a_k$$

$$= \frac{1}{\beta} \sum_{k>M} a_k (\beta^2 \epsilon)^k = \frac{1}{\beta} \text{Remainder}_{M+1} \left(1 - \sqrt{1 - \beta^2 \epsilon}\right).$$

The convergence statement follows by noting that $\sqrt{1-\beta^2\epsilon}$ can be developed in a convergent Taylor expansion about $\epsilon = 0$ in a radius $\epsilon < 1/\beta^2$.

5 Conclusions

Numerous research avenues remain open, including how to characterize the relationship with the Chen-Fliess series and with the series in [4], how to extend Theorem 3.1 to systems with generic forcing term $g = g(t,x)$, and how to bring to bear normal form theory. Furthermore, the simplicity and convergence properties of the novel series might help in areas such as trajectory generation and optimization, controllability, and model reduction.

Acknowledgments

The author thanks Christoforos Hadjicostis, James W. Melody, and Miloš Žefran for a number of exciting conversations. This research was supported in part by the U.S. Army Research Office under grant DAAD 190110716.

References

[1] S. Boyd and L. O. Chua. Fading memory and the problem of approximating nonlinear operators with Volterra series. *IEEE Transactions on Circuits and Systems*, 32(11):1150–61, 1985.

[2] R. W. Brockett. Volterra series and geometric control theory. *Automatica*, 12(2):167–176, 1976.

[3] C. Bruni, G. Di Pillo, and G. Koch. On the mathematical models of bilinear systems. *Ricerche di Automatica*, 2:11–26, 1971.

[4] F. Bullo. Series expansions for the evolution of mechanical control systems. *SIAM Journal on Control and Optimization*, 40(1):166–190, 2001.

[5] F. Bullo, N. Ehrich Leonard, and A. D. Lewis. Controllability and motion algorithms for underactuated Lagrangian systems on Lie groups. *IEEE Transactions on Automatic Control*, 45(8):1437–1454, 2000.

[6] K. T. Chen. Integration of paths, geometric invariants, and a generalized Baker-Hausdorff formula. *Annals of Mathematics*, 67:164–178, 1957.

[7] M. Fliess. Fonctionelles causales non linéares et indeterminées non commutatives. *Bull. Soc. Math.*, 109:3–40, 1981.

[8] M. Fliess, M. Lamnabhi, and F. Lamnabhi-Lagarrigue. An algebraic approach to nonlinear functional expansions. *IEEE Transactions on Circuits and Systems*, 30(8):554–70, 1983.

[9] E. G. Gilbert. Functional expansions for the response of nonlinear differential systems. *IEEE Transactions on Automatic Control*, 22(6):909–921, 1977.

[10] A. Isidori. *Nonlinear Control Systems*. Springer Verlag, New York, NY, third edition, 1995. ISBN 3540199160.

[11] W. Kang and A. J. Krener. Extended quadratic controller form and dynamic state feedback linearization of nonlinear systems. *SIAM Journal on Control and Optimization*, 30(6):1319–1337, 1992.

[12] M. Kawski. High-order small-time local controllability. In H. J. Sussmann, editor, *Nonlinear Controllability and Optimal Control*, pages 441–477. Dekker, New York, NY, 1990. ISBN 0824782585.

[13] M. Kawski and H. J. Sussmann. Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory. In U. Helmke, D. Pratzel-Wolters, and E. Zerz, editors, *Operators, Systems, and Linear Algebra*, pages 111–128. Teubner, Stuttgart, Germany, 1997.

[14] H. K. Khalil. *Nonlinear Systems*. Prentice Hall, Englewood Cliffs, NJ, second edition, 1995. ISBN 0132280248.

[15] D. E. Knuth. *The Art of Computer Programming, Volume 1, Fundamental Algorithms*. Addison Wesley, Reading, MA, third edition, 1997. ISBN 0201896834.

[16] S. G. Krantz. *Function Theory of Several Complex Variables*. Pure and Applied Mathematics. John Wiley & Sons, New York, NY, 1982. ISBN 0471093246.

[17] A. J. Krener. Linearization and bilinearization of control systems. In *Allerton Conference on Communication, Control and Computing*, pages 834–843, Monticello, IL, September 1974. University of Illinois at Urbana-Champaign.

[18] G. Lafferriere and H. J. Sussmann. A differential geometric approach to motion planning. In Z. Li and J. F. Canny, editors, *Nonholonomic Motion Planning*, pages 235–270. Kluwer Academic Publishers, Boston, MA, 1993. ISBN 0792392752.

[19] N. Ehrich Leonard and P. S. Krishnaprasad. Motion control of drift-free, left-invariant systems on Lie groups. *IEEE Transactions on Automatic Control*, 40(9):1539–1554, 1995.

[20] C. Lesiak and A. J. Krener. The existence and uniqueness of Volterra series for nonlinear systems. *IEEE Transactions on Automatic Control*, 23(6): 1090–1095, 1978.

[21] W. Magnus. On the exponential solution of differential equations for a linear operator. *Communications on Pure and Applied Mathematics*, VII:649–673, 1954.

[22] R. I. McLachlan, G. R. W. Quispel, and N. Robidoux. Unified approach to Hamiltonian systems, Poisson systems, gradient systems, and systems with Lyapunov functions or first integrals. *Physical Review Letters*, 81(12):2399–403, 1998.

[23] W. J. Rugh. *Nonlinear System Theory: The Volterra/Wiener Approach*. John Hopkins University Press, Baltimore, Maryland, 1981. ISBN 0801825490.

[24] C. Rui, I. V. Kolmanovsky, and N. H. McClamroch. Nonlinear attitude and shape control of spacecraft with articulated appendages and reaction wheels. *IEEE Transactions on Automatic Control*, 45(8): 1455–69, 2000.

[25] I. W. Sandberg. Expansions for nonlinear systems. *The Bell System Technical Journal*, 61(2):159–199, 1982.

[26] J. A. Sanders and F. Verhulst. *Averaging Methods in Nonlinear Dynamical Systems*. Springer Verlag, New York, NY, 1985. ISBN 0387962298.

[27] M. Schetzen. *The Volterra and Wiener Theories of Nonlinear Systems*. Wiley-Interscience, New York, 1989. ISBN 0894643568.

[28] H. J. Sussmann. A product expansion of the Chen series. In C. I. Byrnes and A. Lindquist, editors, *Theory and Applications of Nonlinear Control Systems*, pages 323–335. Elsevier, Oxford, UK, 1986. ISBN 0444700552.

[29] H. S. Wilf. *Generatingfunctionology*. Academic Press, New York, NY, second edition, 1994. ISBN 0127519564.