On expansions for nonlinear systems, error estimates and convergence issues

Karine Beauchard, Jérémy Le Borgne, Frédéric Marbach

On expansions for nonlinear systems, error estimates and convergence issues

Karine Beauchard*, Jérémy Le Borgne*, Frédéric Marbach*

March 10, 2023

Abstract

Explicit formulas expressing the solution to non-autonomous differential equations are of great importance in many application domains such as control theory or numerical operator splitting. In particular, intrinsic formulas allowing to decouple time-dependent features from geometry-dependent features of the solution have been extensively studied.

First, we give a didactic review of classical expansions for formal linear differential equations, including the celebrated Magnus expansion (associated with coordinates of the first kind) and Sussmann's infinite product expansion (associated with coordinates of the second kind). Inspired by quantum mechanics, we introduce a new mixed expansion, designed to isolate the role of a time-invariant drift from the role of a time-varying perturbation.

Second, in the context of nonlinear ordinary differential equations driven by regular vector fields, we give rigorous proofs of error estimates between the exact solution and finite approximations of the formal expansions. In particular, we derive new estimates focusing on the role of time-varying perturbations. For scalar-input systems, we derive new estimates involving only a weak Sobolev norm of the input.

Third, we investigate the local convergence of these expansions. We recall known positive results for nilpotent dynamics and for linear dynamics. Nevertheless, we also exhibit arbitrarily small analytic vector fields for which the convergence of the Magnus expansion fails, even in very weak senses. We state an open problem concerning the convergence of Sussmann's infinite product expansion.

Eventually, we derive approximate direct intrinsic representations for the state and discuss their link with the choice of an appropriate change of coordinates.

Contents

*Univ Rennes, CNRS, IRMAR - UMR 6625, F-35000 Rennes, France

1 Introduction

1.1 Motivations

There are multiple situations in which one desires to compute the solution to a differential equation whose dynamics depend on time. One often looks for explicit formulas, depending preferentially on intrinsic quantities, which describe the composition of flows, or even the continuous composition of flows. Some important applications are listed below.

- **Control theory.** Here, the dynamics depend on time mostly through the choice of time-varying controls. One looks for explicit formulas of the continuous product of flows in order to be able to construct controls for which this resulting flow drives a given initial state to a desired target state. In order to establish necessary and sufficient conditions for controllability, one is interested in intrinsic formulas. It is our main motivation.

- **Numerical splitting methods.** Here, the splitting algorithm applies sequentially a succession of basic flows, composed with appropriate time steps. One is interested in choosing correctly the base flows and the time steps in order to approximate the most precisely possible the solution to the true complex flow. Formulas concerning the composition of flows are essential to compute the order of the resulting numerical scheme. We refer to the survey [20] and the introduction books [19, 58]. Composition of flows formulas are also very useful in particular settings like Hamiltonian systems [22] or in the presence of a small perturbation of a reference flow [81]. Concerning numerical methods, more generally, we refer to [82] (respectively [37]) for a survey on Butcher series (resp. post-Lie algebras), algebraic tools related to the algebras manipulated in the sequel.

- **Stochastic differential equations.** Here, the dynamics depend on time through the sources of randomness, say Brownian motions. One wishes to investigate the influence of the randomness on the final state and thus looks for explicit formulas involving iterated Stratanovich integrals to construct a representation of the flow, see e.g. [12, 15, 28, 32].

- **Differential equations on Lie groups.** Sometimes, the state itself of the differential equation belongs to a Lie group, as in [64]. Then, looking for an intrinsic approximation of the state helps to preserve structure which would be lost otherwise. In particular, writing the product of multiple flows as a single flow is important. There are also control problems for differential equations set on Lie groups, as in [67]. Some works, e.g. [30], also tackle the hard question of obtaining Magnus-type expansions, which are intrinsically linear, for nonlinear equations within matrix Lie groups.

- **Analysis of time-periodic systems.** When investigating the behavior of time-periodic systems, some authors borrow tools from "chronological calculus" or expressions of the "logarithm of the flow" (described below). For example, such techniques are used to study stability and asymptotic stability of time-periodic systems of ODEs; see the non linear Floquet Theorem 3.2 and the high-order averaging procedure Theorem 7.1 in [93], or the recent higher-order averaging results of [79].

1.2 Short historical survey

We start with a short survey of some of the many approaches related with the computation of solutions to formal linear differential equations, say

$$\dot{x}(t) = X(t)x(t), \tag{1}$$

together with some initial condition $x(0)$. We recall in Section 1.2.4 the consequences of such results for nonlinear ordinary differential equations.

1.2.1 Iterated integration and Chen-Fliess expansion

A straightforward approach to solving (1) consists in what can be seen as a Picard iteration. For small times, starting from the initial approximation $x(t) \approx x(0)$, one then enhances the approximation iteratively by substituting it in the equation and obtains successively $x(t) \approx x(0) + \int_0^t X(s)x(0)\,\mathrm{d}s$, then $x(t) \approx x(0) + \int_0^t X(s)x(0)\,\mathrm{d}s + \int_0^t X(s)\int_0^s X(s')x(0)\,\mathrm{d}s'\,\mathrm{d}s$ and so on.

In the context of control theory, this expansion is known as the Chen-Fliess expansion, after being popularized by the works [33, 45]. Its main advantages are its simplicity and nice convergence

properties (see Section 5.1). However, it also has some strong drawbacks, which we detail in Remark 17 and Remark 167 and motivate the investigation of other expansions.

1.2.2 Magnus expansion

When $X(t)$ is piecewise constant, for example with values X_1 for $t \in [0,1]$ and X_2 for $t \in [1,2]$, one has formally, $x(2) = e^{X_2} e^{X_1} x(0)$. Hence, the computation of solutions to (1) has a deep link with the famous Campbell [29], Baker [11], Hausdorff [61], Dynkin [40] formula ("CBHD formula" in the sequel).

This formula has a long and rich history which involves forgotten contributions of other authors such as Schur, Poincaré, Pascal or Yosida. As noted by Bourbaki in [25], "*chacun considère que les démonstrations de ses prédécesseurs ne sont pas convaincantes*" (each one considers that the proofs of his predecessors are not convincing). We therefore encourage the reader to dive into the fascinating retrospectives [2] and [23] to understand the progressive construction of its proof throughout the decades. This formula is a formal identity expressing the product of the exponentials of two (non-commutative) indeterminates X_1 and X_2 as the single exponential of a series of Lie brackets (i.e. nested commutators) of these indeterminates, of which the first terms are well-known:

$$e^{X_2} e^{X_1} = \exp\left(X_1 + X_2 + \frac{1}{2}[X_2, X_1] + \dots \right). \tag{2}$$

When more than two exponentials are multiplied, say e^{X_1} through e^{X_n}, one can of course iterate the formula (2) with itself to formally express the product of n exponentials as the single exponential of a complicated series. Letting $n \to +\infty$, one is lead to computing a continuous product of exponentials, which corresponds, heuristically, to solving (1).

Magnus performed a breakthrough by deriving in [80] the first formal representation of the solution to (1) as the exponential of a series, of which the first terms are

$$x(t) = \exp\left(\int_0^t X(\tau_1)\,\mathrm{d}\tau_1 + \frac{1}{2}\int_0^t \int_0^{\tau_1} [X(\tau_1), X(\tau_2)]\,\mathrm{d}\tau_2\,\mathrm{d}\tau_1 + \cdots \right) x(0). \tag{3}$$

This formula can be seen as the continuous counterpart of the CBHD formula and highlights important structural properties of the solutions to (1) (see Section 2.3).

1.2.3 Infinite products

The CBHD formula and the Magnus formula share the goal of expressing the desired quantity as the exponential of a single, although complicated, object. Other approaches go the other way around and try to express the desired quantity as a long (infinite) product of exponentials of very simple objects.

A well-known example is the Lie-Trotter product formula (see e.g. [101]), often used for numerical splitting methods which attempts to give a meaning to the equality

$$e^{X_1 + X_2} = \lim_{n \to +\infty} \left(e^{\frac{X_1}{n}} e^{\frac{X_2}{n}} \right)^n, \tag{4}$$

the interest relying on the fact that the exponentials of X_1 and X_2 are assumed to be easier to compute in some sense than the direct exponential of $X_1 + X_2$.

Another related formula is the Zassenhaus expansion, described by Magnus in [80], which allows to decompose the same quantity $e^{X_1+X_2}$ as an infinite product of exponentials of linear combinations of nested commutators of strictly increasing lengths, whose first terms are

$$e^{X_1+X_2} = e^{X_1} e^{X_2} \exp\left(-\frac{1}{2}[X_1, X_2] \right) \exp\left(\frac{1}{3}[X_2, [X_1, X_2]] + \frac{1}{6}[X_1, [X_1, X_2]] \right) \cdots \tag{5}$$

In the context of differential equations such as (1), a nice formula is Sussmann's infinite product expansion, introduced in [100]. When $X(t)$ is given as a linear combination of elementary generators, e.g. $X(t) = a_1(t)X_1 + a_2(t)X_2$, Sussmann's infinite product expansion is given by a product of exponentials of Lie monomials, such as

$$x(t) = e^{\xi_1 X_1} e^{\xi_2 X_2} e^{\xi_{12}[X_1,X_2]} e^{\xi_{112}[X_1,[X_1,X_2]]} e^{\xi_{212}[X_2,[X_1,X_2]]} \cdots x(0), \tag{6}$$

where the ξ_i are scalar functions of time given by explicit formulas from the functions a_1 and a_2. Compared to other expansions, this formula is both intrinsic (such as the Magnus expansion) and involves coefficients which are easily computed by induction (such as the Chen-Fliess expansion).

1.2.4 Consequences for nonlinear ordinary differential equations

Although the expansions mentioned above concern linear formal differential equations, they can be adapted to ordinary nonlinear differential equations on smooth manifolds governed by smooth vector fields. Indeed, one can identify vector fields with linear operators acting on smooth functions, and points of the manifold with the linear operator on smooth functions corresponding to evaluation at this point. This method allows to recast the nonlinear equation into a linear equation set on a larger space, for which the formal linear expansions can be used (see Section 4.1).

This linearization technique has been used by Sussmann in [99, Proposition 4.3] to prove the convergence of the Chen-Fliess expansion for nonlinear ordinary differential equations driven by analytic vector fields, by Agrachev and Gamkrelidze in the context of control theory (see [3, 4, 49] in which they derive an exponential representation of flows, very similar to Magnus' expansion, using the *chronological calculus* framework) and by Strichartz (see [97] and his derivation of the *generalized CBHD formula*, with applications related to sub-Riemannian geometry).

At a formal level, all identities mentioned above (almost) always make sense. However, if the indeterminates are replaced by objects coming from analysis (say vector fields, matrices or differential operators), convergence issues arise. Generally speaking, convergence often requires that one either assumes that the objects are small enough or that the generated Lie algebra has additional structure, like nilpotence.

1.3 Main goals and organization of this article

This article is both a survey on some classical expansions for nonlinear systems, a research article containing new results and counter-examples and a toolbox for future works. In particular, we aim at the following goals.

- In Section 2 we give a **didactic review of classical expansions for formal linear differential equations**. Our introduction to this algebraic topic is written with a view to making it understandable by readers with minimal algebraic background. We review the following classical expansions:
  1. the Chen-Fliess formula,
  2. the Magnus or generalized CBHD formula (associated with coordinates of the first kind),
  3. Sussmann's infinite product formula (associated with coordinates of the second kind).
- We introduce a **new formal mixed expansion**, inspired by quantum mechanics, designed to isolate the role of a time-invariant drift from the role of a time-varying perturbation (see Theorem 41), which we name *Magnus expansion in the interaction picture* and for which we define coordinates of the *pseudo-first kind* by analogy with first and second kind coordinates.
- We recall in Section 3 **classical well-posedness results and estimates** for products and Lie brackets of analytic vector fields, which are used throughout the article.

- In the context of nonlinear ordinary differential equations driven by regular vector fields, we give in Section 4 **rigorous proofs of error estimates** between the exact solution and finite approximations of each of these four formal expansions. These estimates are part of the mathematical folklore for the Chen-Fliess and Magnus expansions, but are new for our mixed expansion (see Proposition 99) and for Sussmann's infinite product expansion (see Proposition 104). We strive towards providing estimates with similar structures for the four expansions and which are valid under parsimonious regularity assumptions.

- We investigate the convergence of these expansions in Section 5. We recall **known positive convergence results** for smooth vector fields generating nilpotent Lie algebras and for small linear dynamics (matrices). For our new expansion, we investigate the subtle convergence under a natural partial nilpotency assumption (see Corollary 122). In this case, convergence requires analyticity, contrary to the proofs we give for the other expansions under a full nilpotency assumption.

- For analytic vector fields, only the Chen-Fliess expansion is known to converge. We give in Section 5.2 **new strong counter examples** to the convergence of CBHD and Magnus expansions, which **disprove the convergence** of these expansions even for analytic vector fields and in very weak senses (see Proposition 114). We state an open problem concerning the convergence of Sussmann's infinite product for analytic vector fields (see Open problem 135).

- When the system involves a time-invariant drift and a time-varying perturbation, we show in Section 6 that only the Magnus expansion fails to provide **well-behaved estimates with respect to the perturbation size**. For the three other expansions, it turns out to be possible to obtain such estimates by summing well-defined infinite partial series which converge for analytical vector fields (see mostly Propositions 139 and 143).

- In the particular case of scalar-input systems, we prove in Section 7 **new errors estimates involving a negative Sobolev norm of the time-varying input** (see mostly Propositions 149 and 157). Such estimates are the best compatible with the regularity of the input-to-state map and can be helpful for specific applications.

- Eventually, we derive in Section 8 **approximate direct intrinsic representations of the state** for nonlinear systems, which don't require the computation of flows (see Proposition 161). Our formulas can be viewed as almost-diffeomorphisms and might be useful for applications in control theory. Unfortunately, we also study a counter-example which demonstrates that one cannot obtain an exact representation through a diffeomorphism.

2 Formal expansions for linear dynamics

In this section, we consider formal linear differential equations, recall classical expansions valid in this formal setting (for which there is no convergence issue; see nevertheless Remark 12) and introduce a new mixed expansion which isolates the role of a perturbation in the dynamics. Here and in the sequel, the adjective *formal* denotes situations in which we work within the realm of formal power series (see Definition 11).

2.1 Notations

We recall classical definitions and notations for usual algebraic objects. In the sequel, $\mathbb{K}$ denotes the field $\mathbb{R}$ or $\mathbb{C}$. All statements and proofs hold for both base fields. It will be implicit that all vector spaces and algebras are constructed from the base field $\mathbb{K}$.

2.1.1 Free algebras

We refer to the books [65, 90] for thorough introductions to Lie algebras and free Lie algebras.

Definition 1 (Indeterminates). *Let I be a finite set. At the formal level, we consider a set $X := \{X_i;\ i \in I\}$ of* indeterminates, *indexed by I. For applications, we will substitute in their place matrices or vector fields. Most often, we will write $I = [\![1, q]\!]$ for some $q \in \mathbb{N}^*$, or $I = [\![0, q]\!]$ when we want to isolate the role of the indeterminate X_0.*

Definition 2 (Free monoid). *For I as above, we denote by I^* the* free monoid *over I, i.e. the set of finite sequences of elements of I endowed with the concatenation operation. More precisely, if $\sigma = (\sigma_i)_{1\leq i\leq \ell}$ and $\sigma' = (\sigma'_i)_{1\leq i\leq \ell'}$ are elements of I^*, then the concatenation of σ and σ' is the sequence $\sigma \cdot \sigma' = (\sigma''_i)_{1\leq i\leq \ell+\ell'}$ where $\sigma''_i = \sigma_i$ if $1 \leq i \leq \ell$ and $\sigma''_i = \sigma'_{i-\ell}$ if $\ell + 1 \leq i \leq \ell + \ell'$. It is common to write the elements of I^* as words whose letters are elements of I, by juxtaposition of the elements of the sequence. With this point of view, the concatenation operation is the juxtaposition of words. For a more detailed exposition, see [26, §7.2].*

For $\sigma = (\sigma_1, \dots \sigma_k) \in I^*$, where k is the length of σ also denoted by $|\sigma|$, we let $X_\sigma := X_{\sigma_1} \cdots X_{\sigma_k}$. This operation defines an homomorphism from I^* to X^*, the free monoid over X (monomials over X).

Definition 3 (Free algebra). *For X as above, we consider $\mathcal{A}(X)$ the* free associative algebra *generated by X over the field $\mathbb{K}$, i.e. the unital associative algebra of polynomials of the non commutative indeterminates X (see also [26, Chapter 3, Section 2.7, Definition 2]). $\mathcal{A}(X)$ can be seen as a graded algebra:*

$$\mathcal{A}(X) = \bigoplus_{n\in\mathbb{N}} \mathcal{A}_n(X), \tag{7}$$

where $\mathcal{A}_n(X)$ is the finite-dimensional $\mathbb{K}$-vector space spanned by monomials of degree n over X. In particular $\mathcal{A}_0(X) = \mathbb{K}$ and $\mathcal{A}_1(X) = \mathrm{span}_{\mathbb{K}}(X)$.

Definition 4 (Free Lie algebra). *For X as above, $\mathcal{A}(X)$ is endowed with a natural structure of Lie algebra, the Lie bracket operation being defined by $[a, b] = ab - ba$. This operation satisfies $[a, a] = 0$ and the Jacobi identity $[a, [b, c]] + [c, [a, b]] + [b, [c, a]] = 0$. We also write $[a, b]$ as $\mathrm{ad}_a(b)$ (respectively $\underline{\mathrm{ad}}_b(a)$) which allows for iterated left (resp. right) bracketing. We consider $\mathcal{L}(X)$, the* free Lie algebra *generated by X over the field $\mathbb{K}$, which is defined as the Lie subalgebra generated by X in $\mathcal{A}(X)$. It can be seen as the smallest linear subspace of $\mathcal{A}(X)$ containing all elements of X and stable by the Lie bracket (see also [90, Theorem 0.4]). $\mathcal{L}(X)$ is a graded Lie algebra:*

$$\mathcal{L}(X) = \bigoplus_{n\in\mathbb{N}} \mathcal{L}_n(X), \qquad [\mathcal{L}_m(X), \mathcal{L}_n(X)] \subset \mathcal{L}_{m+n}(X) \tag{8}$$

where, for each $n \in \mathbb{N}$, we define $\mathcal{L}_n(X) := \mathcal{L}(X) \cap \mathcal{A}_n(X)$.

Definition 5 (Nilpotent Lie algebra). *Let L be a Lie algebra. We define recursively the following two-sided Lie ideals: $L^1 := L$ and, for $k \geq 1$, $L^{k+1} := [L, L^k]$ i.e. L^{k+1} is the linear subspace of L generated by brackets of the form $[a, b]$ with $a \in L$ and $b \in L^k$. Let $m \in \mathbb{N}^*$. We say that L is a* nilpotent Lie algebra of index m *when $L^m = \{0\}$ and m is the smallest integer for which this property holds.*

Definition 6 (Free nilpotent Lie algebra). *Let $m \in \mathbb{N}^*$. The* free m-nilpotent Lie algebra *over X is the quotient $\mathcal{N}_m(X) := \mathcal{L}(X)/\mathcal{L}(X)^m$ (with the notation of Definition 5). Then the canonical surjection $\sigma_m : \mathcal{L}(X) \to \mathcal{N}_m(X)$ is an homomorphism of Lie algebras.*

The universal properties of the various free algebras constructed above allow to transport on algebras relations proved at the free level.

Lemma 7. *The following universal properties hold.*

- *For each unital associative algebra A and map $\Lambda : X \to A$, there exists a unique homomorphism of algebras $\mathcal{A}(X) \to A$ that extends Λ.*
- *For each Lie algebra L and map $\Lambda : X \to L$, there exists a unique homomorphism of Lie algebras $\mathcal{L}(X) \to L$ that extends Λ.*
- *Let $m \in \mathbb{N}^*$. For each nilpotent Lie algebra L of index m and map $\Lambda : X \to L$, there exists a unique homomorphism of Lie algebras $\mathcal{N}_m(X) \to L$ that extends Λ.*

2.1.2 Iterated brackets and evaluation

Definition 8 (Iterated brackets)**.** *For X as above, we consider* $\mathrm{Br}(X)$ *the* set of iterated brackets *of elements of X. This set can be defined by induction: for $X_i \in X$, $X_i \in \mathrm{Br}(X)$ and if $b_1, b_2 \in \mathrm{Br}(X)$, then the ordered pair (b_1, b_2) belongs to* $\mathrm{Br}(X)$*. More rigorously, one can define* $\mathrm{Br}(X)$ *as the free magma over X or as the set of rooted full binary trees, with leaves labeled by X.*

For $b \in \mathrm{Br}(X)$, we will use the following notations:

- $|b|$ will denote the length of b (i.e. the number of leaves of the tree).
- If $|b| > 1$, there exists a unique pair $b_1 \in \mathrm{Br}(X)$ and $b_2 \in \mathrm{Br}(X)$ such that $b = (b_1, b_2)$ (left and right factors) which are denoted as $\lambda(b) = b_1$ and $\mu(b) = b_2$. We also write (b_1, b_2) as $\mathrm{ad}_{b_1}(b_2)$ (respectively $\underline{\mathrm{ad}}_{b_2}(b_1)$) which allows iterated left (resp. right) bracketing.
- For $i \in I$, $n_i(b)$ denotes the number of occurrences of the indeterminate X_i in b. When $I = [\![0, q]\!]$ we will also write $n(b) = n_1(b) + \cdots + n_q(b) = |b| - n_0(b)$.

Remark 9. *There is a natural* evaluation *mapping* E *from* $\mathrm{Br}(X)$ *to* $\mathcal{L}(X)$ *defined by induction by $\mathrm{E}(X_i) := X_i$ for $X_i \in X$ and $\mathrm{E}((b_1, b_2)) := [\mathrm{E}(b_1), \mathrm{E}(b_2)]$. Through this mapping,* $\mathrm{Br}(X)$ *spans $\mathcal{L}(X)$ over $\mathbb{K}$, i.e. $\mathcal{L}(X) = \mathrm{span}_{\mathbb{K}} \mathrm{E}(\mathrm{Br}(X))$. This mapping is however not injective: for example, (X_1, X_1) and $(X_2, (X_1, X_1))$ are two different elements of* $\mathrm{Br}(X)$*, both evaluated to zero in $\mathcal{L}(X)$.*

More precisely, the E *map extends to a surjective homomorphism of algebras from the nonassociative free algebra over X (which is the free vector space over* $\mathrm{Br}(X)$*, whose elements are (finite) linear combinations of elements of* $\mathrm{Br}(X)$*, endowed with the natural product map induced by the product in* $\mathrm{Br}(X)$*). Moreover the kernel of the extended* E *is precisely the ideal generated by the relations that define anticommutativity and the Jacobi identity in $\mathcal{L}(X)$. This gives an alternative description of $\mathcal{L}(X)$ as a quotient of the free vector space over* $\mathrm{Br}(X)$*.*

Definition 10 (Subspaces of brackets)**.** *When $I = [\![0, q]\!]$ and $M \in \mathbb{N}$, S_M denotes the vector subspace $\mathcal{L}(X)$ defined by*

$$S_M := \mathrm{span}_{\mathbb{K}} \{\mathrm{E}(b);\ \ b \in \mathrm{Br}(X), n(b) \leq M\}\,. \tag{9}$$

2.1.3 Formal power series, exponential and logarithms

Definition 11 (Formal power series)**.** *We consider the (unital associative) algebra $\widehat{\mathcal{A}}(X)$ of* formal power series *generated by $\mathcal{A}(X)$. An element $a \in \widehat{\mathcal{A}}(X)$ is a sequence $a = (a_n)_{n\in\mathbb{N}}$ written $a = \sum_{n\in\mathbb{N}} a_n$, where $a_n \in \mathcal{A}_n(X)$ with, in particular, $a_0 \in \mathbb{K}$ being its constant term. We also define the Lie algebra of formal Lie series $\widehat{\mathcal{L}}(X)$ as the Lie algebra of formal power series $a \in \widehat{\mathcal{A}}(X)$ for which $a_n \in \mathcal{L}(X)$ for each $n \in \mathbb{N}$. For $a \in \widehat{\mathcal{A}}(X)$ and $\sigma \in I^*$, $\langle a, X_\sigma \rangle$ denotes the coefficient of X_σ in a: $a = \sum_{\sigma\in I^*} \langle a, X_\sigma \rangle X_\sigma$.*

Remark 12. *The definition of $\widehat{\mathcal{A}}(X)$ can be made more rigorous by considering* $\mathrm{val} : \mathcal{A}(X) \to \mathbb{N} \cup \{\infty\}$ *defined by* $\mathrm{val}(a) := \inf\{n \in \mathbb{N}; a \in \bigoplus_{k\geq n} \mathcal{A}_k(X)\}$*. Then $\delta(a, b) := e^{-\mathrm{val}(b-a)}$ is a distance on $\mathcal{A}(X)$, that induces the discrete topology on each $\mathcal{A}_n(X)$, and $\widehat{\mathcal{A}}(X)$ is defined as*

the completion of the metric space $\mathcal{A}(X)$, to which the operations on $\mathcal{A}(X)$ naturally extend as continuous operations, endowing it with a structure of topological algebra. This distance verifies a stronger triangular inequality: $\delta(a,b) \leq \max\{\delta(a,c), \delta(b,c)\}$ (usually referred to as the ultrametric inequality*). This construction allows to write, for $a \in \widehat{\mathcal{A}}(X)$ with components $a_n \in \mathcal{A}_n(X)$,*

$$a = \lim_{N \to +\infty} \sum_{n \leq N} a_n, \tag{10}$$

where the convergence holds with respect to the topology described above. This justifies the notation $a = \sum_{n \in \mathbb{N}} a_n$ used in Definition 11. To avoid confusion with convergence issues associated with the evaluation of formal power series when substituting the indeterminates by objects coming from analysis we shall however not use the term convergence *in this context.*

If $a \in \widehat{\mathcal{A}}(X)$ has zero constant term, we define $\exp(a) \in \widehat{\mathcal{A}}(X)$ and $\log(1+a) \in \widehat{\mathcal{A}}(X)$ as

$$\exp(a) := \sum_{m \geq 0} \frac{a^m}{m!}, \tag{11}$$

$$\log(1+a) := \sum_{m \geq 1} \frac{(-1)^{m-1}}{m} a^m. \tag{12}$$

Since a has zero constant term, one checks that the right-hand sides of (11) and (12) indeed define formal power series of $\widehat{\mathcal{A}}(X)$ (and the sums converge in the sense of the topology constructed in Remark 12). In particular, $\log(\exp(a)) = a$ and $\exp(\log(1+a)) = 1+a$.

Lemma 13. *Let $a, b \in \widehat{\mathcal{A}}(X)$ with zero constant term. Then $a = b$ if and only if $\exp(a) = \exp(b)$.*

Proof. The forward implication is obvious. Conversely, if $\exp(a) = \exp(b)$ in $\widehat{\mathcal{A}}(X)$, then, for every $r \geq 1$, their components in $\mathcal{A}_r$ are equal. Moreover, from (11), one has:

$$(\exp(a))_r = \sum_{k=1}^{r} \sum_{r_1 + \ldots r_k = r} \frac{a_{r_1} \ldots a_{r_k}}{k!} = a_r + \Theta_r\left(a_1, \ldots a_{r-1}\right), \tag{13}$$

for some function Θ_r depending only on the $a_{r'}$ for $r' < r$. Hence, we obtain by induction on $r \geq 1$ that $a_r = b_r$ from the equalities $(\exp(a))_r = (\exp(b))_r$. □

2.2 Formal differential equations and iterated integrals

Using the notations of Section 2.1, for $i \in I$, let $a_i \in L^1(\mathbb{R}_+; \mathbb{K})$ and define a by

$$a(t) := \sum_{i \in I} a_i(t) X_i. \tag{14}$$

In this section, we consider the following formal ordinary differential equation set on $\widehat{A}(X)$, driven by a and associated with some initial data $x^\star$,

$$\begin{cases} \dot{x}(t) = x(t) a(t), \\ x(0) = x^\star, \end{cases} \tag{15}$$

whose solutions are precisely defined in the following way.

Definition 14 (Solution to a formal differential equation)**.** *Let $a_i \in L^1(\mathbb{R}_+; \mathbb{K})$ for $i \in I$ and define a by* (14)*. Let $x^\star \in \widehat{\mathcal{A}}(X)$ with homogeneous components $x_n^\star \in \mathcal{A}_n(X)$. The* solution to the formal differential equation (15) *is the formal-series valued function $x : \mathbb{R}_+ \to \widehat{\mathcal{A}}(X)$, whose homogeneous*

components $x_n : \mathbb{R}_+ \to \mathcal{A}_n(X)$ *are the unique continuous functions that satisfy, for every* $t \geq 0$, $x_0(t) = x_0^\star$ *and, for every* $n \in \mathbb{N}^*$,

$$x_n(t) = x_n^\star + \int_0^t x_{n-1}(\tau) a(\tau) \, \mathrm{d}\tau. \tag{16}$$

Definition 15 (Ordered simplex)**.** *For* $r \in \mathbb{N}^*$ *and* $t > 0$, *we introduce*

$$\Delta^r(t) := \{(\tau_1, \ldots, \tau_r) \in (0,t)^r; \;\; 0 < \tau_1 < \cdots < \tau_r < t\}. \tag{17}$$

Iterating this integral formula yields the following power series expansion, which is the most direct way to compute the solution to (15) and was introduced in [33, 34] and popularized in control theory by [45]. In the "chronological calculus" terminology (not used in the present article), it is called "(right) formal Volterra chronological series" [3, Section 1.5].

Lemma 16 (Chen series)**.** *In the context of Definition 14, the solution to* (15) *with initial data* $x^\star = 1$ *can be expanded as*

$$x(t) = \sum_{\sigma \in I^*} \left(\int_0^t a_\sigma \right) X_\sigma, \tag{18}$$

where $\int_0^t a_\emptyset = 1$ *by convention and, for* $\sigma \in I^*$ *with* $|\sigma| \geq 1$, *we introduce the notation*

$$\int_0^t a_\sigma := \int_{\Delta^n(t)} a_{\sigma_1}(\tau_1) \cdots a_{\sigma_n}(\tau_n) \, \mathrm{d}\tau. \tag{19}$$

Proof. Expansion (18) is a direct consequence of the iterated application of (16) and of the definition of X_σ in Definition 2 and can be proved by induction on the length of σ. □

Remark 17. *Despite its simplicity, the Chen series suffers from a major drawback: it involves non intrinsic quantities and is redundant. As an illustration, this has the following consequences:*

- *The functionals* $\int_0^t a_\sigma$ *for* $\sigma \in I^*$ *are not algebraically independent. For example, for every solution to* (15) *and every* $t \geq 0$, *one has the identity*

$$\langle x(t), X_1 X_2 \rangle + \langle x(t), X_2 X_1 \rangle - \langle x(t), X_1 \rangle \langle x(t), X_2 \rangle = 0 \tag{20}$$

- *In the context of nonlinear ordinary differential equations, the representation* (18) *can fail to converge for smooth vector fields despite strong structural assumptions (see Section 5.1.1).*

- *In the context of nonlinear ordinary differential equations, the representation* (18) *will not be invariant by diffeomorphism (see Remark 167), which would be a desirable invariance.*

This drawback motivates the search for more intrinsic representations of the solutions, which will turn out to involve Lie algebras.

The Chen series give rise to *Fliess operators* (stemming from [45, 46]) which can be defined, given some $c \in \widehat{\mathcal{A}}(X)$, as $a \mapsto \sum_{\sigma \in I^*} \langle c, X_\sigma \rangle \int_0^t a_\sigma$. Such operators are well-defined (converge) provided that the coefficients $\langle c, X_\sigma \rangle$ satisfy an appropriate asymptotic behavior. Fliess operators can be used to model input-output systems and feedback groups. For manipulations of such operators thanks to an underlying Hopf algebra structure, we refer to [53, 54, 55, 56], which investigate the question of whether an interconnection of such operators remains a Fliess operator, and its convergence, both in scalar and multivariate settings. See also [57] for the investigation of global convergence issues, and realization of such formal operators on concrete systems.

2.3 Logarithm of flows, coordinates of the first kind

In the particular case where $a(t)$ is a constant element $a \in \mathcal{A}_1(X)$, evaluating the iterated integrals in (19) yields the elegant formula $x(t) = x^\star \exp(ta)$, with the notation of (11). Of course, it is no longer valid for a time-varying dynamic (because the indeterminates do not commute), but one can wish to find an object of which the flow is the exponential, the *logarithm of the flow*.

In this section, we recall and prove Theorem 27, which states that the logarithm of flows of formal linear differential equations is given by explicit Lie brackets. The key argument is the structure result Theorem 20, which states that the logarithm of the flow is a Lie series, and of which we give an elementary proof based on the differential equation satisfied by the logarithm of the flow. We rely on well-known algebraic results, which we recall, for the sake of giving a self-contained presentation.

2.3.1 A differential equation for the logarithm of the flow

We start by deriving the formal differential equation (23) satisfied by the logarithm of the flow. This equation is well-known (see e.g. [80, Theorem III], [34, Theorem 4.1] or [3, formula (5.2)]). We provide an elementary derivation (see Remark 19).

Proposition 18. *The following statements hold.*

1. *Let $T > 0$ and $z \in \mathcal{C}^1([0,T]; \widehat{\mathcal{A}}(X))$. Then, for every $t \in [0,T]$,*

$$\frac{\mathrm{d}}{\mathrm{d}t} \exp(z(t)) = \exp(z(t)) \sum_{n=0}^{+\infty} \frac{(-1)^n}{(n+1)!} \operatorname{ad}^n_{z(t)}(\dot{z}(t)). \tag{21}$$

2. *Let a be given by (14) and x denote the solution to (15) with initial data $x^\star = 1$. Then $z := \log x$ satisfies, for almost every $t \in \mathbb{R}_+$,*

$$\sum_{n=0}^{+\infty} \frac{(-1)^n}{(n+1)!} \operatorname{ad}^n_{z(t)}(\dot{z}(t)) = a(t), \tag{22}$$

$$\dot{z}(t) = \sum_{n=0}^{+\infty} \frac{(-1)^n B_n}{n!} \operatorname{ad}^n_{z(t)}(a(t)), \tag{23}$$

where the Bernoulli numbers $(B_n)_{n\in\mathbb{N}}$ are defined in (527).

Proof. We prove the two claims successively.

1. The regularity assumption $z \in \mathcal{C}^1([0,T]; \widehat{\mathcal{A}}(X))$ is to be understood component by component, i.e. means that for each $\sigma \in I^*$, $t \mapsto \langle z(t), X_\sigma \rangle$ belongs to $\mathcal{C}^1([0,T];\mathbb{K})$. We have

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} \exp(z(t)) &= \frac{\mathrm{d}}{\mathrm{d}t} \left(\sum_{k=0}^{+\infty} \frac{z^k(t)}{k!} \right) = \sum_{k=0}^{+\infty} \frac{1}{(k+1)!} \sum_{j=0}^{k} z^j(t) \dot{z}(t) z^{k-j}(t) \\ &= \exp(z(t)) \left(\sum_{l=0}^{+\infty} \frac{(-1)^l}{l!} z^l(t) \right) \left(\sum_{k=0}^{+\infty} \frac{1}{(k+1)!} \sum_{j=0}^{k} z^j(t) \dot{z}(t) z^{k-j}(t) \right). \end{aligned} \tag{24}$$

   Letting $n := k + l$ and $i := l + j$, we obtain that

$$\frac{\mathrm{d}}{\mathrm{d}t} \exp(z(t)) = \exp(z(t)) \sum_{n=0}^{+\infty} \frac{1}{(n+1)!} \sum_{i=0}^{n} z^i(t) \dot{z}(t) z^{n-i}(t) \sum_{l=0}^{i} (-1)^l \binom{n+1}{l} \tag{25}$$

The following formulas, which can be proved by induction using Pascal's rule,

$$\sum_{l=0}^{i}(-1)^l\binom{n+1}{l} = (-1)^i\binom{n}{i}, \tag{26}$$

$$\sum_{i=0}^{n}(-1)^i\binom{n}{i}z^i y z^{n-i} = (-1)^n \operatorname{ad}_z^n(y) \tag{27}$$

give the conclusion. Of course, if $z \in W^{1,1}((0,T);\widehat{\mathcal{A}}(X))$ (i.e. absolutely continuous), equations (21) remains true as an equality in $L^1((0,T);\widehat{\mathcal{A}}(X))$, i.e. holding for almost every $t \in (0,T)$.

2. Since $z = \log x$ and $\dot{x} = xa$, (22) is an immediate consequence of (21), using the preceding comment since both x and z have $W^{1,1}$ regularity in time when a has L^1 regularity in time. Starting from (22) and applying $\sum_k (-1)^k B_k/k! \operatorname{ad}_{z(t)}^k$ to both sides yields (23) because

$$\sum_{k=0}^{+\infty}\frac{(-1)^k B_k}{k!}\sum_{\ell=0}^{+\infty}\frac{(-1)^\ell}{(\ell+1)!}\operatorname{ad}_{z(t)}^{k+\ell}(\dot{z}(t)) = \dot{z}(t). \tag{28}$$

This follows from the change of index $n := k+\ell$ and the combinatorial relation (529). □

Remark 19. *The historical proofs of Proposition 18 are written using the Poisson bracket notation* $\{\cdot, z^k\} := (-1)^k \operatorname{ad}_z^k(\cdot)$ *which allows to write* (23) *as the nice equality*

$$\dot{z} = \left\{a, \frac{z}{e^z-1}\right\}, \tag{29}$$

using the generating series (527) *of the Bernoulli numbers. This approach allows elegant computations, but requires some setup (see [80, Section III] or [34, Section 1]), which is why we prefer here the elementary computations used in the preceding proof.*

2.3.2 The logarithm of the flow is a Lie series

The fundamental result concerning the logarithm of the flow is that it is a Lie series. We repeat here the proof given in [34, Theorem 4.2] for the sake of completeness. At least two other approaches can be used: one relying on shuffle relations and Ree's theorem (see Appendix A.1) and another one relying on Friedrich's criterion (see Appendix A.2).

Theorem 20. *Let a be given by* (14) *and x be the solution to* (15) *with initial data $x^\star = 1$. Then, for every $t > 0$, $\log x(t) \in \widehat{\mathcal{L}}(X)$.*

Proof. The proof relies on an iterated integration of (23), where $z = \log x$. More precisely, writing $z = \sum z_n$ where $z_n \in \mathcal{A}_n(X)$, we prove by induction on n that, for every t, $z_n(t) \in \mathcal{L}(X)$. First, for every $t > 0$, since $x_0(t) = 1$, one has $z_0(t) = 0$ so $z_0(t) \in \mathcal{L}(X)$. Then, for every $n \geq 1$, by (23),

$$\dot{z}_n(t) = \sum_{k=0}^{n-1}\frac{(-1)^k B_k}{k!}\sum_{n_1+\cdots+n_k=n-1}[z_{n_1}(t),[z_{n_2}(t),\dots[z_{n_k}(t),a(t)]\cdots]], \tag{30}$$

where the sum ranges over indexes $n_i \geq 1$. Moreover, for every $T > 0$ and $y \in L^1((0,T);\mathcal{L}(X))$ one checks that, for every $t \in [0,T]$, $\int_0^t y \in \mathcal{L}(X)$. By the induction assumption, $z_{n_i}(t) \in \mathcal{L}(X)$ for each $n_i \leq n-1$ and every t. By the previous comment, this property is preserved by the time-integration of (30), so $z_n(t) \in \mathcal{L}(X)$ for every t. □

2.3.3 Notations

We start with an abstract definition of the truncated logarithm of a time-dependent dynamic.

Definition 21. *For $m, r \in \mathbb{N}^*$, we define the set of ordered positive partitions of size m of r,*

$$\mathbb{N}_r^m := \{\mathbf{r} = (\mathbf{r}_1, \dots, \mathbf{r}_m) \in (\mathbb{N}^*)^m;\ \ \mathbf{r}_1 + \cdots + \mathbf{r}_m = r\}\,, \tag{31}$$

where $\mathbb{N}_r^m = \emptyset$ when $r < m$. For each $\mathbf{r} \in \mathbb{N}_r^m$ and $t > 0$, we also define the product of simplexes

$$\Delta^{\mathbf{r}}(t) := \Delta^{\mathbf{r}_1}(t) \times \cdots \times \Delta^{\mathbf{r}_m}(t). \tag{32}$$

Example 22. *The sets $\Delta^{\mathbf{r}}(t)$ will be used as integration domains. As easy examples, one has*

$$\Delta^{(3)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3;\ \ 0 < \tau_1 < \tau_2 < \tau_3 < t\}, \tag{33}$$

$$\Delta^{(1,1,1)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3) \in (0,t)^3\}. \tag{34}$$

A more complex example for $r = 4$, $m = 2$ and $\mathbf{r} = (2,2) \in \mathbb{N}_4^2$ is

$$\Delta^{(2,2)}(t) = \{\tau = (\tau_1, \tau_2, \tau_3, \tau_4) \in (0,t)^4;\ \ 0 < \tau_1 < \tau_2 < t \text{ and } 0 < \tau_3 < \tau_4 < t\}. \tag{35}$$

We now give a notation for the (truncated or complete) logarithm of a time-dependent dynamic. We will see in the sequel why this quantity indeed corresponds to a logarithm.

Definition 23 (Abstract logarithm of a time-varying field)**.** *Let $M \in \mathbb{N}$ or $M = +\infty$, $t > 0$ and F be a map from $[0,t]$ with values in some algebra. We introduce the notation*

$$\mathrm{Log}_M\{F\}(t) := \sum_{r=1}^{M} \frac{1}{r} \sum_{m=1}^{r} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [F(\tau_1), F(\tau_2)], \dots F(\tau_r)]\, \mathrm{d}\tau. \tag{36}$$

Remark 24. *In such an abstract setting, the right-hand side of* (36) *does not make sense since we are not able to define an integral over an abstract algebra (without topology on the algebra and without time-regularity on F). At this stage, we see* (36) *as an abstract formula or notation. We will check, each time we use it, that we can give a meaning to the integrals.*

2.3.4 A preliminary algebraic result

Since the monomials form a basis of $\mathcal{A}(X)$, one can define the following linear map β from $\mathcal{A}(X)$ to $\mathcal{L}(X)$ by setting its values on the monomials by $\beta(1) := 0$, $\beta(X_i) := X_i$ for $1 \le i \le q$, and, for $1 \le i_1, \dots, i_k \le q$ with $k \in \mathbb{N}^*$,

$$\beta(X_{i_1} X_{i_2} \cdots X_{i_k}) := [\cdots [X_{i_1}, X_{i_2}], \dots, X_{i_k}]. \tag{37}$$

This process defines a standard way, the "left to right" or "left normed" bracketing, to associate a Lie bracket to each monomial. The following important result, proved successively by Dynkin [39], Specht [95] and Wever [104] states that, if a polynomial is a Lie element, then it is equal to its left normed bracketing.

Lemma 25 (Dynkin's theorem)**.** *For $a \in \mathcal{A}_n(X)$, $a \in \mathcal{L}(X)$ if and only if $\beta(a) = na$.*

Proof. This statement is contained in the equivalence between *(i)* and *(v)* of [90, Theorem 1.4]. □

Example 26. *The element X_1X_2 does not belong to $\mathcal{L}(X)$. And indeed, $\beta(X_1X_2) = X_1X_2 - X_2X_1 \neq 2X_1X_2$. On the contrary, the element $[X_1, X_2] = X_1X_2 - X_2X_1$ belongs to $\mathcal{L}(X)$. And indeed, $\beta([X_1, X_2]) = (X_1X_2 - X_2X_1) - (X_2X_1 - X_1X_2) = 2[X_1, X_2]$.*

2.3.5 An explicit formula for the logarithm of the flow

We now state an explicit expansion of the state as the exponential of the logarithm of the flow. It is the continuous analogue of the well-known CBHD formula (which we recall in Section 2.3.7 as a corollary). It was originally derived by Magnus in [80, Theorem III] and is thus often referred to as the "Magnus expansion".

Theorem 27. *For $t \in \mathbb{R}_+$ and $x^\star \in \widehat{\mathcal{A}}(X)$, the solution x to (15) satisfies*

$$x(t) = x^\star \exp\left(\mathrm{Log}_\infty\{a\}(t)\right), \tag{38}$$

with the notation of Definition 23.

Proof. First, by linearity, it suffices to prove (38) for $x^\star = 1$. Repeated integration of (16) yields the following formula (which is a slightly different form of the Chen series of Lemma 16),

$$x(t) = 1 + \sum_{r\geq 1} \int_{\Delta^r(t)} a(\tau_1)\cdots a(\tau_r)\,\mathrm{d}\tau. \tag{39}$$

Hence, recalling the definitions (31) of $\mathbb{N}_r^m$ and (32) of $\Delta^{\mathbf{r}}(t)$, one has

$$\log(x(t)) = \sum_{r=1}^{+\infty} \sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} a(\tau_1)a(\tau_2)\cdots a(\tau_r)\,\mathrm{d}\tau. \tag{40}$$

By Theorem 20, for each $t \geq 0$, $\log(x(t)) \in \widehat{\mathcal{L}}(X)$. Hence, applying Lemma 25 to each of its homogeneous components in $\mathcal{A}_r$ proves that

$$\log(x(t)) = \sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots[a(\tau_1), a(\tau_2)], \ldots a(\tau_r)]\,\mathrm{d}\tau. \tag{41}$$

Recalling the notation (23) and taking the exponential concludes the proof of (38). □

Magnus expansions (also called BCH expansions) have been extended to more general structures than Lie algebras, for instance to pre-Lie (another name for "chronological algebras") and post-Lie algebras [47], Rota-Baxter algebras [6, 43] and dendriform algebras [41, 42].

2.3.6 Coordinates of the first kind

Although the expansion (41) already has some interest by itself, it is not written on a basis of $\mathcal{L}(X)$, which has some drawbacks. In this paragraph, we define canonical representations for this expansion, in appropriate bases of $\mathcal{L}(X)$.

Definition 28 (Monomial basis)**.** *Let $\mathcal{B} \subset \mathcal{L}(X)$. We say that $\mathcal{B}$ is a* basis *of $\mathcal{L}(X)$ when each element $a \in \mathcal{L}(X)$ can be written as a unique finite linear combination of elements of $\mathcal{B}$. We say that $\mathcal{B}$ is a* monomial *basis of $\mathcal{L}(X)$ when moreover $\mathcal{B} \subset \mathrm{E}(\mathrm{Br}(X))$. In particular, for such bases, if $b \in \mathcal{B}$, one can define $|b|$, $n_i(b)$ for $i \in I$ and $n(b)$ as in Section 2.1.2 by importing these notions from $\mathrm{Br}(X)$. In particular, for $n \in \mathbb{N}^*$, we use the notations $\mathcal{B}_n := \{b \in \mathcal{B}; |b| = n\}$ and $\mathcal{B}_{[\![1,n]\!]} := \{b \in \mathcal{B}; |b| \leq n\}$.*

Proposition 29. *Let $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$. There exists a unique set of functionals $(\zeta_b)_{b\in\mathcal{B}}$, with $\zeta_b \in \mathcal{C}^0\left(\mathbb{R}_+ \times L^1(\mathbb{R}_+;\mathbb{K})^{|I|};\mathbb{K}\right)$, such that, for every $a_i \in L^1(\mathbb{R}_+;\mathbb{K})$, $x^\star \in \widehat{\mathcal{A}}(X)$ and $t \geq 0$, the solution to (15) satisfies*

$$x(t) = x^\star \exp\left(\sum_{b\in\mathcal{B}} \zeta_b(t,a)b\right). \tag{42}$$

Moreover, the functionals ζ_b are "causal" in the sense that, for every $t \geq 0$, $\zeta_b(t,a)$ only depends on the restrictions of the functions a_i to $[0,t]$.

Proof. For each $b \in \mathcal{B}$, since $\mathcal{B}$ is monomial, only a finite number of summands of the right-hand side of (41) have a non vanishing component along b (indeed, only terms sharing the same homogeneity can be involved). Hence, it is clear that the functionals thereby defined are continuous on $\mathbb{R}_+ \times L^1(\mathbb{R}_+;\mathbb{K})^q$, due to their explicit expression. The sum in (42) is understood in the sense of a well-defined formal power series. Indeed, for each word $\sigma \in I^*$, only a finite number of elements $b \in \mathcal{B}$ have a non-vanishing component $\langle b, X_\sigma \rangle$. □

Definition 30 (Coordinates of the first kind)**.** *The functionals ζ_b are usually called* coordinates of the first kind *associated to the (monomial) basis $\mathcal{B}$ of $\mathcal{L}(X)$.*

The terminology *coordinates of the first kind* or *first species* and the opposition with *coordinates of the second kind* (see Section 2.5.3) is classical, see e.g. [24, III.4.3]. See also Section 2.3.8 for references concerning the computation of such coordinates in the context of control theory.

Remark 31. *Thanks to the monomial nature of the basis, one does not need to specify the full basis in order to define a given functional. For example, if $\lambda \in \mathbb{N}^I$ is a given homogeneity, let*

$$\mathrm{Br}_\lambda(X) := \{b \in \mathrm{Br}(X);\ \ \forall i \in I, n_i(b) = \lambda_i\}. \tag{43}$$

Then the coordinates of the first kind ζ_b for $b \in \mathcal{B} \cap \mathrm{Br}_\lambda(X)$ only depend on $\mathcal{B} \cap \mathrm{Br}_\lambda(X)$.

Remark 32. *An important particular case for applications to control theory is the case $X = \{X_0, X_1\}$, with $a_0(t) = 1$ and $a_1(t) = u(t)$. This corresponds to formal scalar-input control-affine systems $\dot{x}(t) = x(t)(X_0 + u(t)X_1)$. One often writes $\zeta_b(t,u)$ (omitting the dependency on $a_0 \equiv 1$) to denote the coordinates of the first kind in this particular context.*

2.3.7 Campbell Baker Hausdorff Dynkin formula

As a corollary, we obtain the classical finite CBHD formula.

Corollary 33. *Let X be a finite set, $n \in \mathbb{N}^*$ and $y_1, \dots, y_n \in \widehat{\mathcal{L}}(X)$ without constant term. There exists a unique $w \in \widehat{\mathcal{L}}(X)$ such that*

$$e^{y_1} \cdots e^{y_n} = e^w. \tag{44}$$

We will use the notation $w = \mathrm{CBHD}_\infty(y_1, \dots, y_n)$. Moreover, for each monomial basis $\mathcal{B}$ of $\mathcal{L}(\{Y_1, \dots, Y_n\})$, there exists a unique sequence $(\alpha_b)_{b\in\mathcal{B}} \subset \mathbb{K}^{\mathcal{B}}$ such that, for every finite set X and $y_1, \dots, y_n \in \widehat{\mathcal{L}}(X)$

$$\mathrm{CBHD}_\infty(y_1, \dots, y_n) = \sum_{b\in\mathcal{B}} \alpha_b y_b \tag{45}$$

where $y_b := \Lambda(b)$ and $\Lambda : \mathcal{L}(\{Y_1, \dots, Y_n\}) \to \widehat{\mathcal{L}}(X)$ is the homomorphism of Lie algebras such that $\Lambda(Y_j) = y_j$ for $j \in [\![1, n]\!]$.

Proof. We prove that (44) holds with

$$w := \mathrm{Log}_\infty \left\{ \sum_{j=1}^{n} y_j 1_{[j-1,j]} \right\} (n) \tag{46}$$

in the sense of Definition 23.

Step 1: Proof when $X = \{X_1, \dots, X_n\}$ and $y_j = X_j$ for $j \in [\![1, n]\!]$. The solution to (15) with $a(t) = \sum_{j=1}^n X_j 1_{[j-1,j]}(t)$ is $x(t) = x^\star e^{X_1} \cdots e^{X_n}$. By Theorem 27, w satisfies (44). By injectivity of the exponential (see Lemma 13), it is the unique solution. By Proposition 29, the equality (45) holds with $\alpha_b := \zeta_b(n, 1_{[0,1]}, \dots, 1_{[n-1,n]})$.

Step 2: Proof in the general case. Let X be a finite set, $n \in \mathbb{N}^*$, $y_1, \dots, y_n \in \widehat{\mathcal{L}}(X)$. Let $Y := \{Y_1, \dots, Y_n\}$ be another set of indeterminates.

The map $\Lambda : Y \to \widehat{\mathcal{L}}(X)$ defined by $\Lambda(Y_j) = y_j$ for $j \in [\![1, n]\!]$ extends into an homomorphism of algebras $\widehat{\mathcal{A}}(Y) \to \widehat{\mathcal{A}}(X)$, which is also an homomorphism of Lie algebras $\widehat{\mathcal{L}}(Y) \to \widehat{\mathcal{L}}(X)$, that we still denote Λ. Indeed Lemma 7 ensures the extension as an homomorphism of algebras $\mathcal{A}(Y) \to \widehat{\mathcal{A}}(X)$ (resp. an homomorphism of Lie algebras $\mathcal{L}(Y) \to \widehat{\mathcal{L}}(X)$). The extension can be done on $\widehat{\mathcal{A}}(Y)$ (resp. $\widehat{\mathcal{L}}(Y)$) because $y_1, \dots, y_n$ do not have constant terms and the target space $\widehat{\mathcal{A}}(X)$ (resp. $\widehat{\mathcal{L}}(X)$) is a space of formal power series.

Let $W := \mathrm{Log}_\infty\{\sum_{j=1}^n Y_j 1_{[j-1,j]}\}(n) \in \widehat{\mathcal{L}}(Y)$. Then $\Lambda(W) = w$. By applying the homomorphism of algebras Λ to the relation $e^{Y_1} \cdots e^{Y_n} = e^W$ we get (44). By applying the homomorphism of Lie algebras Λ to the relation $W = \sum_{b \in \mathcal{B}} \alpha_b b$ we get (45). □

Despite the fact that the product $e^{y_1} \cdots e^{y_n}$ is of course non-commutative, there is some structure and symmetry inside its logarithm, which we highlight for future use in the following result.

Proposition 34. *There exists a family of elements $F_{q,h}(Y_1, \dots, Y_q) \in \mathcal{L}(\{Y_1, \dots, Y_q\})$ for $q \in \mathbb{N}^*$ and $h = (h_1, \dots, h_q) \in (\mathbb{N}^*)^q$, such that*

- *for each $i \in [\![1, q]\!]$, $F_{q,h}(Y_1, \dots, Y_q)$ is of homogeneity h_i with respect to Y_i,*
- *for every $n \geq 2$, $y_1, \dots, y_n \in \widehat{\mathcal{L}}(X)$ with zero constant term,*

$$\mathrm{CBHD}_\infty(y_1, \dots, y_n) = \sum_{\substack{q \in [\![1,n]\!], h \in (\mathbb{N}^*)^q \\ j_1 < \cdots < j_q \in [\![1,n]\!]}} F_{q,h}(y_{j_1}, \dots, y_{j_q}), \tag{47}$$

  where $F_{q,h}(y_{j_1}, \dots, y_{j_q})$ denotes the image of $F_{q,h}(Y_1, \dots, Y_q)$ by the homomorphism of algebras from $\mathcal{L}(\{Y_1, \dots, Y_q\})$ to $\widehat{\mathcal{L}}(X)$ which sends Y_i to y_{j_i} for each $i \in [\![1, q]\!]$.

For $q = 1$, $F_{1,(1)}(Y) = Y_1$ and $F_{1,(h_1)} = 0$ for $h_1 \geq 2$. For $q = 2$ and $h_1 + h_2 \leq 4$,

$$\begin{aligned}
F_{2,(1,1)}(Y) &= \frac{1}{2}[Y_1, Y_2] & F_{2,(2,2)}(Y) &= -\frac{1}{24}[Y_2, [Y_1, [Y_1, Y_2]]] \\
F_{2,(2,1)}(Y) &= \frac{1}{12}[Y_1, [Y_1, Y_2]] & F_{2,(3,1)}(Y) &= 0 \\
F_{2,(1,2)}(Y) &= \frac{1}{12}[Y_2, [Y_2, Y_1]] & F_{2,(1,3)}(Y) &= 0.
\end{aligned} \tag{48}$$

For higher order terms, we state below a recursive formula.

Proof. Using the same Lie algebra homomorphism arguments as in the proof of Corollary 33, it is sufficient to consider the case where $y_i = Y_i$ is an indeterminate.

For $n = 2$, the statement is merely a rewriting of (45) where the terms are grouped by their homogeneity with respect to y_1 and y_2. This defines the elements $F_{1,(1)}(Y_1) = Y_1$ and $F_{1,(h)}(Y_1) = 0$ for $h \geq 2$ and $F_{2,h}(Y_1, Y_2)$ for $h \in (\mathbb{N}^*)^2$ according to the usual two-variables formula, of which the well-known low-order terms are recalled in (48).

We define by induction on $n \geq 3$ the functions $F_{n,h}$ by the relations

$$F_{n,h}(Y_1, \dots, Y_n) := \sum_{m | h_1, \dots, h_{n-1}} F_{2,(m,h_n)}\big(F_{n-1,(\frac{h_1}{m}, \dots, \frac{h_{n-1}}{m})}(Y_1, \dots, Y_{n-1}), Y_n\big). \tag{49}$$

We now prove the result by induction on n. Let $n \geq 3$. By associativity of the product, the formula

for two indeterminates and the induction hypothesis at step $n-1$, we obtain

$$\begin{aligned}
&\mathrm{CBHD}_\infty(Y_1,\dots,Y_n)\\
&=\mathrm{CBHD}_\infty(\mathrm{CBHD}_\infty(Y_1,\dots,Y_{n-1}),Y_n)\\
&=\mathrm{CBHD}_\infty(Y_1,\dots,Y_{n-1})+Y_n+\sum_{g\in(\mathbb{N}^*)^2}F_{2,g}(\mathrm{CBHD}_\infty(Y_1,\dots,Y_{n-1}),Y_n)\\
&=Y_n+\sum_{\substack{q\in[\![1,n-1]\!],h'\in(\mathbb{N}^*)^q\\ j_1<\cdots<j_q\in[\![1,n-1]\!]}}\left(F_{q,h'}(Y_{j_1},\dots,Y_{j_q})+\sum_{g\in(\mathbb{N}^*)^2}F_{2,g}(F_{q,h'}(Y_{j_1},\dots,Y_{j_q}),Y_n)\right).
\end{aligned}\tag{50}$$

We now check that the right-hand side of (50) is the same as the right-hand side of (47). Since we are working in the free Lie algebra over $Y_1,\dots Y_n$, we can proceed by homogeneity.

- The terms not involving Y_n are equal, since they have the same expression.
- The term involving only Y_n on both sides is Y_n itself, so they are equal.
- Now, let $q\in[\![1,n-1]\!]$, $j_1<\cdots<j_q\in[\![1,n-1]\!]$ and $h\in(\mathbb{N}^*)^{q+1}$. We look for the term involving h_i times Y_{j_i} for $i\in[\![1,q]\!]$ and h_{q+1} times Y_n, which is $F_{q+1,h}(Y_{j_1},\dots,Y_{j_q},Y_n)$ in (47). In (50), it is

$$\sum_{h'\in(\mathbb{N}^*)^q}\sum_{g\in(\mathbb{N}^*)^2}F_{2,g}(F_{q,h'}(Y_{j_1},\dots,Y_{j_q}),Y_n),\tag{51}$$

  where the sum is restricted to $g_1h'_i=h_i$ and $g_2=h_{q+1}$. Hence, both terms are equal thanks to the definition (49).

This concludes the proof and gives a way to compute the elements $F_{q,h}$ iteratively. □

Remark 35. *In particular, the component of* $\mathrm{CBHD}_\infty(y_1,\dots,y_n)$ *homogeneous with degree* $h=(h_1,\dots,h_q)$ *with respect to* $(y_{j_1},\dots,y_{j_q})$ *is* $F_{q,h}(y_{j_1},\dots,y_{j_q})$*. It depends neither on the total number* n *of arguments in the initial product, nor on the selection of indexes* $(j_1,\dots,j_q)$*. This is the natural symmetry that we wish to highlight.*

Algorithms to compute iteratively the terms in the CBHD formula are investigated for instance in [85, Section 4.a] or in [18, 31] within Hall bases, or in [10] for an expansion on right-nested brackets, which uses fewer terms.

2.3.8 Computation of some coordinates of the first kind

In this paragraph, we focus on the case $X=\{X_0,X_1\}$. Computing the coordinates of the first kind is of paramount interest for applications (see e.g. [70] where the first 14 such coordinates are computed, and [31, 86] for efficient algorithms and explicit formulas obtained by an approach relying on rooted binary labeled trees).

Here, we calculate as an illustration (and because they will be used later) all coordinates of the first kind on a basis of

$$\dot{S}_1:=\mathrm{span}_{\mathbb{K}}\left\{\mathrm{E}(b);\ \ b\in\mathrm{Br}(X),n_1(b)=1\right\}\subset\mathcal{L}(X),\tag{52}$$

where this notation is chosen so that $S_1=\mathbb{K}X_0\oplus\dot{S}_1$ (see Definition 10). We define moreover

$$S_2^+:=\mathrm{span}_{\mathbb{K}}\left\{\mathrm{E}(b);\ \ b\in\mathrm{Br}(X),n_1(b)\geq 2\right\}\subset\mathcal{L}(X),\tag{53}$$

thanks to which we can write the direct sum decomposition $\mathcal{L}(X)=\mathbb{K}X_0\oplus\dot{S}_1\oplus S_2^+$.

Lemma 36. *The family* $(\mathrm{ad}^k_{X_0}(X_1))_{k\in\mathbb{N}}$ *is a (monomial) basis of* $\dot{S}_1$.

Proof. From (52), $\dot{S}_1$ is spanned by the evaluations in $\mathcal{L}(X)$ of the iterated brackets $b \in \mathrm{Br}(X)$ involving X_1 exactly once. Let $b \in \mathrm{Br}(X)$ be such *an iterated bracket*. We assume $\mathrm{E}(b) \neq 0$ in $\mathcal{L}(X)$ and $b \neq X_1$. Then $\mathrm{E}(b) = [\mathrm{E}(\lambda(b)), \mathrm{E}(\mu(b))]$ thus $\mathrm{E}(\lambda(b))$ and $\mathrm{E}(\mu(b))$ are non null in $\mathcal{L}(X)$. Moreover, either $\lambda(b)$ or $\mu(b)$ does not involve X_1 and is thus equal to X_0. Therefore $\mathrm{E}(b) = \pm[X_0, \mathrm{E}(\bar{b})]$ where $\bar{b} \in \mathrm{Br}(X)$ involves X_1 exactly once and $\mathrm{E}(\bar{b}) \neq 0$. Working by induction on the number k of occurrences of X_0 in b, we obtain $\mathrm{E}(b) = \pm\,\mathrm{ad}^k_{X_0}(X_1)$.

The previous argument proves that the given family spans $\dot{S}_1$. Moreover, this family is linearly independent in $\mathcal{L}(X)$ because two different elements have different lengths. □

We now compute the coordinates of the first kind associated with these elements. Up to our knowledge, the following explicit expression is new.

Proposition 37. *Let* $\mathcal{B}$ *a monomial basis of* $\mathcal{L}(X)$ *containing* X_0 *and the family* $(\mathrm{ad}^k_{X_0}(X_1))_{k\in\mathbb{N}}$. *The associated coordinates of the first kind satisfy, for each* $t > 0$, $a_0, a_1 \in L^1((0,t);\mathbb{K})$ *and* $k \in \mathbb{N}$,

$$\zeta_{\mathrm{ad}^k_{X_0}(X_1)}(t, a_0, a_1) = (-1)^k \sum_{\ell=0}^{k} A_0(t)^{k-\ell} \frac{B_{k-\ell}}{(k-\ell)!} \int_{\Delta^{\ell+1}(t)} a_1(\tau_1)a_0(\tau_2)\cdots a_0(\tau_{\ell+1})\,\mathrm{d}\tau, \tag{54}$$

where $A_0(t) := \int_0^t a_0$ *and the Bernoulli numbers* $(B_n)_{n\in\mathbb{N}}$ *are defined in* (527).

Proof. First, the considered coordinates are well-defined independently on the exact choice of $\mathcal{B}$ (see Remark 31). Let x be the solution to (15) starting from $x^\star = 1$. To simplify the notations in this proof, we write $x(t)$, $\zeta_k(t)$ and $Z(t)$ instead of $x(t,a)$, $\zeta_{\mathrm{ad}^k_{X_0}(X_1)}(t, a_0, a_1)$ and $\mathrm{Log}_\infty\{a\}(t)$. From (42),

$$Z(t) = \sum_{b\in\mathcal{B}} \zeta_b(t,a)b = \zeta_{X_0}(t,a)X_0 + Z_1(t) + Z_2(t), \tag{55}$$

where $Z_2(t) \in S_2^+$ and

$$Z_1(t) := \sum_{k=0}^{+\infty} \zeta_k(t)\,\mathrm{ad}^k_{X_0}(X_1). \tag{56}$$

First, a straightforward identification in (36) yields $\zeta_{X_0} = A_0$ and $\zeta_{X_1}(t) = \int_0^t a_1$. Let $k \in \mathbb{N}^*$. The proof consists in computing $\langle x(t), X_1X_0^k\rangle$ in two ways: first by the differential equation (15), then by the formula $x(t) = e^{Z(t)}$. By definition of the solution to (15), we have, for every word $\sigma \in I^*$ and $t > 0$

$$\langle x(t), X_\sigma X_0\rangle = \int_0^t \langle x(\tau), X_\sigma\rangle a_0(\tau)\,\mathrm{d}\tau. \tag{57}$$

Taking into account that $\langle x(t), X_1\rangle = \int_0^t a_1$, we obtain

$$\langle x(t), X_1X_0^k\rangle = \int_{\Delta^{k+1}(t)} a_1(\tau_1)a_0(\tau_2)\cdots a_0(\tau_{k+1})\,\mathrm{d}\tau. \tag{58}$$

On the other hand, we deduce from the expansion of $x(t) = e^{Z(t)}$ that

$$\left\langle x(t), X_1X_0^k\right\rangle = \left\langle Z(t), X_1X_0^k\right\rangle + \sum_{\ell=2}^{k+1} \frac{1}{\ell!}\left\langle Z(t)^\ell, X_1X_0^k\right\rangle \tag{59}$$

because, for $\ell \geq (k+2)$, $Z(t)^\ell$ is a sum of words with length at least $(k+2)$. For $\ell \in [\![2, k+1]\!]$,

$$Z(t)^\ell = \sum_{j=0}^{\ell-1} (A_0(t)X_0)^j Z_1(t) (A_0(t)X_0)^{\ell-1-j} + Z_{2,\ell}(t), \quad \text{where} \quad Z_{2,\ell}(t) \in S_2^+. \tag{60}$$

Thus

$$\left\langle Z(t)^\ell, X_1X_0^k\right\rangle = \left\langle Z_1(t)(A_0(t)X_0)^{\ell-1}, X_1X_0^k\right\rangle = A_0(t)^{\ell-1}(-1)^{k-\ell+1}\zeta_{k-\ell+1}(t), \tag{61}$$

because the word $X_1X_0^{k-\ell+1}$ appears in the decomposition of $\mathrm{ad}^n_{X_0}(X_1)$ iff $k-\ell+1=n$ and then it appears with coefficient $(-1)^n$. We deduce from (59) and (61) that

$$\left\langle x(t), X_1X_0^k\right\rangle = (-1)^k\zeta_k(t) + \sum_{\ell=2}^{k+1}\frac{(-1)^{k+1-\ell}}{\ell!}A_0(t)^{\ell-1}\zeta_{k+1-\ell}(t). \tag{62}$$

Using (58) and the index change $j=k+1-\ell\in[\![0,k-1]\!]$, we obtain

$$\int_{\Delta^{k+1}(t)} a_1(\tau_1)a_0(\tau_2)\cdots a_0(\tau_{k+1})\,\mathrm{d}\tau = (-1)^k\zeta_k(t) + \sum_{j=0}^{k-1}\frac{(-1)^jA_0(t)^{k-j}}{(k+1-j)!}\zeta_j(t), \tag{63}$$

When $A_0(t)=0$, this formula yields (54) immediately. When $A_0(t)\neq 0$, let, for $j\in\mathbb{N}$,

$$\alpha_j := \frac{\langle x(t), X_1X_0^j\rangle}{A_0(t)^{j+1}} \qquad \text{and} \qquad \beta_j := \frac{(-1)^j\zeta_j(t)}{A_0(t)^{j+1}} \tag{64}$$

we deduce from (63) that

$$\alpha_k = \sum_{j=0}^{k}\frac{\beta_j}{(k+1-j)!}. \tag{65}$$

We have

$$z\left(\sum_{k\geq 0}\alpha_kz^k\right) = \sum_{k\geq 0}\sum_{j=0}^{k}\beta_jz^j\frac{z^{k+1-j}}{(k+1-j)!} = \left(\sum_{j\geq 0}\beta_jz^j\right)(e^z-1) \tag{66}$$

or equivalently

$$\sum_{j\geq 0}\beta_jz^j = \frac{z}{e^z-1}\left(\sum_{k\geq 0}\alpha_kz^k\right) = \sum_{n\geq 0}\sum_{k\geq 0}B_n\frac{z^n}{n!}\alpha_kz^k. \tag{67}$$

Thus, for every $j\in\mathbb{N}^*$

$$\beta_j = \sum_{k=0}^{j}\frac{B_{j-k}}{(j-k)!}\alpha_k. \tag{68}$$

Finally (64) and (58) give (54). □

Remark 38. *Formula* (54) *bears a strong similarity with the differential equation* (23) *satisfied by* $z(t)$*, which also involves the Bernoulli numbers. Unfortunately, we have not been able to obtain a shorter proof using this equation.*

In particular, using Proposition 37, we recover the following very classical formula for the partial coefficients of the CBHD formula (see e.g. [103, equation (2)] or [90, Corollary 3.24]).

Corollary 39. *One has* $e^{X_1}e^{X_0}=e^Z$ *where* $Z=X_0+Z_1+Z_2$, $Z_2\in S_2^+$ *(see* (53)*) and*

$$Z_1 := \sum_{n=0}^{+\infty}\frac{B_n}{n!}\mathrm{ad}^n_{X_0}(X_1) = X_1 - \frac{1}{2}[X_0,X_1] + \sum_{n=1}^{+\infty}\frac{B_{2n}}{(2n)!}\mathrm{ad}^{2n}_{X_0}(X_1). \tag{69}$$

Proof. We apply the previous result to the controls $a_0(t)=\mathbf{1}_{(1,2)}(t)$ and $a_1(t)=\mathbf{1}_{(0,1)}(t)$, for which the solution to (15) with $x^\star=1$ satisfies $x(2)=e^{X_1}e^{X_0}$. For $\ell\in\mathbb{N}^*$ and $0<\tau_1<\cdots<\tau_{\ell+1}<2$,

the real number $a_1(\tau_1)a_0(\tau_1)\cdots a_0(\tau_{\ell+1})$ does not vanish iff $0<\tau_1<1$ and $1<\tau_2<\cdots<\tau_{\ell+1}<2$ and then it equals 1. Thus, for every $k\geq 2$, using (54) and (528),

$$(-1)^k\zeta_k(2)=\sum_{\ell=0}^{k}\frac{B_{k-\ell}}{(k-\ell)!}\frac{1}{\ell!}=\sum_{j=0}^{k}\frac{B_j}{j!(k-j)!}=\frac{B_k}{k!} \tag{70}$$

We conclude by noticing, thanks to (54), that $\zeta_0(2)=1=B_0$ and $\zeta_1(2)=-\frac{1}{2}=B_1$. □

Example 40. *As an example and for later use in the sequel, we compute the coordinates of the first kind for the particular choice $a_0(t):=1$ and $a_1(t):=t$. Let $k\in\mathbb{N}$. Using formula* (54) *of Proposition 37 and the identity* (530) *we obtain*

$$\begin{aligned}\zeta_{\mathrm{ad}^k_{X_0}(X_1)}(t,a)&=(-1)^k\sum_{\ell=0}^{k}t^{k-\ell}\frac{B_{k-\ell}}{(k-\ell)!}\frac{t^{\ell+2}}{(\ell+2)!}\\&=(-1)^kt^{k+2}\sum_{\ell=0}^{k}\frac{B_{k-\ell}}{(k-\ell)!(\ell+2)!}\\&=(-1)^{k+1}t^{k+2}\frac{B_{k+1}}{(k+1)!}.\end{aligned} \tag{71}$$

2.4 Interaction picture, coordinates of the pseudo-first kind

In quantum mechanics, the *interaction picture* is an intermediate representation between the *Schrödinger picture* (in which the state vectors are time-dependent and the operators are time-independent) and the *Heisenberg picture* (in which the state vectors are time-independent and the operators are time-dependent). The interaction picture is particularly useful when the dynamics can be written as the sum of a time-independent part, which can be solved exactly, and a time-dependent perturbation. In this section, we introduce and study a formal counterpart of this situation, that can be useful for applications.

2.4.1 A new formal expansion

In this paragraph, we therefore consider $I=[\![0,q]\!]$ to isolate the role of X_0. For some given $a_i\in L^1(\mathbb{R}_+;\mathbb{K})$ for $i\in[\![1,q]\!]$, we assume that a takes the form

$$a(t)=X_0+\sum_{i=1}^{q}a_i(t)X_i. \tag{72}$$

Theorem 41. *For $t\in\mathbb{R}_+$, $x^\star\in\widehat{\mathcal{A}}(X)$ and a of the form* (72)*, the solution x to* (15) *satisfies*

$$x(t)=x^\star\exp(tX_0)\exp\left(\mathcal{Z}_\infty(t,X,a)\right), \tag{73}$$

where $\mathcal{Z}_\infty(t,X,a):=\mathrm{Log}_\infty\{b_t\}(t)$ with the notation of Definition 23 and

$$b_t(s):=e^{-(t-s)X_0}\left(\sum_{i=1}^{q}a_i(s)X_i\right)e^{(t-s)X_0}=\sum_{i=1}^{q}\sum_{k=0}^{+\infty}\frac{(-1)^k}{k!}(t-s)^ka_i(s)\,\mathrm{ad}^k_{X_0}(X_i) \tag{74}$$

i.e.

$$\begin{aligned}\mathcal{Z}_\infty(t,X,a)=\sum\frac{(-1)^{m-1}}{mr}\int_{\Delta^{\mathbf{r}}(t)}\frac{(\tau_1-t)^{k_1}}{k_1!}\cdots\frac{(\tau_r-t)^{k_r}}{k_r!}a_{i_1}(\tau_1)\cdots a_{i_r}(\tau_r)\,\mathrm{d}\tau\\ [\cdots[\mathrm{ad}^{k_1}_{X_0}(X_{i_1}),\mathrm{ad}^{k_2}_{X_0}(X_{i_2})],\ldots\mathrm{ad}^{k_r}_{X_0}(X_{i_r})],\end{aligned} \tag{75}$$

where the sum is taken over $r\in[\![1,\infty]\!]$, $m\in[\![1,r]\!]$, $\mathbf{r}\in\mathbb{N}^m_r$, $k_1,\ldots,k_r\in\mathbb{N}$ and $i_1,\ldots,i_r\in[\![1,q]\!]$.

Proof. First, note that the second equality in (74) stems from the fact that both functions $g_1(\tau) := e^{-\tau X_0} X_i e^{\tau X_0}$ and $g_2(\tau) := \sum_{k=0}^{+\infty} \frac{(-\tau)^k}{k!} \mathrm{ad}^k_{X_0}(X_i)$ solve the Cauchy problem $\dot{g}(\tau) = [g(\tau), X_0]$ and $g(0) = X_i$, so they are equal.

Let $t > 0$. A key point is to remark that all the definitions and results from the previous paragraphs which are stated for a finite set I of indeterminates are still valid if I is an infinite set. For mathematicians with a background in analysis, all equalities can be understood "in the weak sense" as equalities holding along each monomial. Therefore, for a set of indeterminates $\{Y_{k,i}\}_{k\in\mathbb{N}, i\in[\![1,q]\!]}$, the solution to

$$\dot{z}(s) = z(s)\gamma_t(s) \quad \text{where} \quad \gamma_t(s) := \sum_{k,i} \frac{(-1)^k}{k!}(t-s)^k a_i(s) Y_{k,i}, \tag{76}$$

with initial data $z(0) = 1$ satisfies, thanks to Theorem 27,

$$z(t) = \exp\left(\mathrm{Log}_\infty\{\gamma_t\}(t)\right). \tag{77}$$

Let Θ be the unique homomorphism of algebras from $\widehat{\mathcal{A}}(\{Y_{k,i}\}_{k\in\mathbb{N}, i\in[\![1,q]\!]})$ to $\widehat{\mathcal{A}}(X)$ defined by

$$\Theta(Y_{k,i}) := \mathrm{ad}^k_{X_0}(X_i). \tag{78}$$

Then $z_\Theta(s) = \Theta(z(s))$ satisfies on the one hand $z_\Theta(0) = 1$ and $\dot{z}_\Theta(s) = z_\Theta(s) b_t(s)$, and on the other hand $z_\Theta(t) = \exp\left(\mathrm{Log}_\infty\{b_t\}(t)\right)$.

We introduce the change of variables $y(s) := x(s)e^{(t-s)X_0}$. Then,

$$\dot{y}(s) = \dot{x}(s)e^{(t-s)X_0} - x(s)X_0 e^{(t-s)X_0} = x(s)\left(\sum_{i=1}^{q} a_i(s)X_i\right) e^{(t-s)X_0} = y(s)b_t(s). \tag{79}$$

Hence

$$x(t) = y(t) = y(0)z_\Theta(t) = x^\star e^{tX_0} \exp\left(\mathrm{Log}_\infty\{b_t\}(t)\right), \tag{80}$$

which concludes the proof of (73). □

Remark 42. *In the above proof, $\mathcal{Z}_\infty(t, X, a)$ is defined by the logarithm of the product of two flows: the one associated with $-X_0$ and the one associated with $a(t)$. It is a particular case of the construction of the chronological logarithm of the product of two flows associated with two non-autonomous vector fields, see [4, Section 2.2] or [72, p. 92].*

Remark 43. *In expansion (73), the choice to write $\exp(tX_0)$ to the left of the formal logarithm is arbitrary. One could obtain a similar formula with $\exp(tX_0)$ to the right. Depending on the application one has in mind, both choices can be helpful.*

2.4.2 Coordinates of the pseudo-first kind

Proposition 44. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \ldots, X_q\}$ and $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$. There exists a unique set of functionals $(\eta_b)_{b\in\mathcal{B}}$, with $\eta_b \in \mathcal{C}^0\left(\mathbb{R}_+ \times L^1(\mathbb{R}_+;\mathbb{K})^q;\mathbb{K}\right)$, such that, for every $a_i \in L^1(\mathbb{R}_+;\mathbb{K})$ and $t \geq 0$*

$$\mathcal{Z}_\infty(t, X, a) = \sum_{b\in\mathcal{B}} \eta_b(t, a) b \quad \text{in} \quad \widehat{\mathcal{L}}(X). \tag{81}$$

Moreover, $\eta_{X_0} = 0$ and the functionals η_b are "causal" in the sense that, for every $t \geq 0$, $\eta_b(t, a)$ only depends on the restrictions of the functions a_i to $[0, t]$.

Proof. For every $r \in \mathbb{N}^*$ and $\nu \in \mathbb{N}$ we introduce the finite sum of brackets

$$\mathcal{Z}_\infty^{r,\nu}(t,X,a) = \sum \frac{(-1)^{m-1}}{mr} \int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1 - t)^{k_1}}{k_1!} \cdots \frac{(\tau_r - t)^{k_r}}{k_r!} a_{i_1}(\tau_1) \cdots a_{i_r}(\tau_r)\, \mathrm{d}\tau \\ [\cdots [\mathrm{ad}_{X_0}^{k_1}(X_{i_1}), \mathrm{ad}_{X_0}^{k_2}(X_{i_2})], \ldots \mathrm{ad}_{X_0}^{k_r}(X_{i_r})], \tag{82}$$

where the sum is taken over $m \in [\![1,r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $k_1, \ldots, k_r \in \mathbb{N}$ such that $k_1 + \cdots + k_r = \nu$ and $i_1, \ldots, i_r \in [\![1,q]\!]$. For each term in this sum, the bracket

$$[\cdots [\mathrm{ad}_{X_0}^{k_1}(X_{i_1}), \mathrm{ad}_{X_0}^{k_2}(X_{i_2})], \ldots \mathrm{ad}_{X_0}^{k_r}(X_{i_r})] \tag{83}$$

has a unique expansion on the basis $\mathcal{B}_{r,\nu} = \{b \in \mathcal{B};\ n(b) = r \text{ and } n_0(b) = \nu\}$. By summing these expansions we obtain causal functions $(\eta_b)_{b \in \mathcal{B}_{r,\nu}}$ in $\mathcal{C}^0\left(\mathbb{R}_+ \times L^1(\mathbb{R}_+;\mathbb{K})^q;\mathbb{K}\right)$ such that the following equality holds in $\mathcal{L}(X)$

$$\mathcal{Z}_\infty^{r,\nu}(t,X,a) = \sum_{b \in \mathcal{B}_{r,\nu}} \eta_b(t,a) b. \tag{84}$$

By summing these relations, we get (81). □

Definition 45 (Coordinates of the pseudo-first kind)**.** *We call the functionals η_b* coordinates of the pseudo-first kind *associated to the (monomial) basis $\mathcal{B}$ of $\mathcal{L}(X)$, by analogy with coordinates of the first kind.*

2.4.3 Structure constants and estimates for the coordinates

At the formal level, series such as (81) make sense. However, in the sequel, we will need to give a meaning to such series where the indeterminates are replaced by true objects. To make sure that the resulting series converge, it will be necessary to have estimates on the coordinates of the pseudo-first kind. In this paragraph, we suggest a criterion based on the structure constants of $\mathcal{L}(X)$ relative to the underlying monomial basis to obtain such estimates.

Definition 46 (Structure constants)**.** *Let $\mathcal{B}$ be a basis of $\mathcal{L}(X)$. For every $a, b \in \mathcal{B}$, since $[a,b] \in \mathcal{L}(X)$, it can be written as a finite linear combination of basis elements, say*

$$[a,b] = \sum_{c \in \mathcal{B}} \gamma_{a,b}^c c, \tag{85}$$

where the coefficients $\gamma_{a,b}^c \in \mathbb{K}$ and only a finite number of them are non-zero. The set of these coefficients are called the structure constants of $\mathcal{L}(X)$ relative to the basis $\mathcal{B}$.

Definition 47 (Geometric growth)**.** *Let X be a finite set and $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$. We say that $\mathcal{B}$ has* geometric growth *when there exists $\Gamma \geq 1$ such that, for every $b_1, b_2 \in \mathcal{B}$,*

$$\sum_{c \in \mathcal{B}} |\gamma_{b_1,b_2}^c| \leq \Gamma^{|b_1|+|b_2|}. \tag{86}$$

Definition 48 (Asymmetric geometric growth)**.** *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \ldots, X_q\}$ and $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$. We say that $\mathcal{B}$ has* geometric growth with respect to X_0 *when, for every $k \in \mathbb{N}$, there exists $\Gamma(k) \geq 1$ such that, for every $b_1, b_2 \in \mathcal{B}$ with $n(b_1) + n(b_2) \leq k$,*

$$\sum_{c \in \mathcal{B}} |\gamma_{b_1,b_2}^c| \leq \Gamma(k)^{|b_1|+|b_2|}. \tag{87}$$

Asymmetric geometric growth is a weaker notion than geometric growth (which can be seen as asymmetric geometric growth with a constant Γ independent of k). These definitions therefore lead to the following algebraic open problem:.

Open problem 49. *Which monomial bases $\mathcal{B}$ of $\mathcal{L}(X)$ have (asymmetric) geometric growth?*

Remark 50. *A family of examples of monomial bases of $\mathcal{L}(X)$ is given by Hall bases (see Section 2.5.1, and in particular Definition 56). In the paper [13] dedicated to studying the growth of structure constants for Hall bases of $\mathcal{L}(X)$, we provide examples of Hall bases of $\mathcal{L}(X)$ that have geometric growth (in particular, the classical examples of length-compatible Hall bases and the Lyndon basis have geometric growth, see Theorem 1.5 and Theorem 1.6 in the cited work). More importantly, we show that every Hall basis has asymmetric geometric growth (see Theorem 1.9 in the cited work).*

For such bases, we can prove nice estimates for the coordinates of the pseudo-first kind. We start with an estimate concerning the decomposition of the Lie brackets involved in (82).

Lemma 51. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$ with geometric growth with respect to X_0. For every $r \geq 1$, there exists $C(r) \geq 1$ such that, for every $\ell \geq 2$, $b_1, \dots, b_\ell \in \mathcal{B} \setminus \{X_0\}$ with $n(b_1) + \cdots + n(b_\ell) \leq r$ and $b \in \mathcal{B}$,*

$$|\langle [\cdots [b_1, b_2], \dots b_\ell], b \rangle_{\mathcal{B}}| \leq C(r)^{|b|}, \tag{88}$$

where the bra-ket denotes the component of the Lie bracket along b in its decomposition on $\mathcal{B}$.

Proof. Any $a \in \mathcal{L}(X)$ can be written as a linear combination of basis elements, say $a = \sum_{c \in \mathcal{B}} \alpha_c^a c$, where the coefficients $\alpha_c^a \in \mathbb{K}$ and only a finite number of them are non-zero. We endow $\mathcal{L}(X)$ with the norm $\|a\|_{\mathcal{B}} := \sum_{c \in \mathcal{B}} |\alpha_c^a|$. Then, Definition 48 gives, for every $b_1, b_2 \in \mathcal{B}$, $\|[b_1, b_2]\|_{\mathcal{B}} \leq \Gamma(n(b_1) + n(b_2))^{|b_1|+|b_2|}$. We prove by induction on $\ell \geq 2$ that, for every $b_1, \dots, b_\ell \in \mathcal{B} \setminus \{X_0\}$,

$$\|[\cdots [b_1, b_2], \dots b_\ell]\|_{\mathcal{B}} \leq \Gamma\left(n(b_1) + \cdots + n(b_\ell)\right)^{(\ell-1)(|b_1|+\cdots+|b_\ell|)} \tag{89}$$

which implies Lemma 51 with $C(r) = \Gamma(r)^{r-1}$. The result for $\ell = 2$ is already known. Let $\ell \geq 2$ and $b_1, \dots, b_{\ell+1} \in \mathcal{B} \setminus \{X_0\}$. Then $[\cdots [b_1, b_2], \dots b_\ell] = \sum_{d \in \mathcal{B}} \alpha_d d$ where the sum is finite and $\sum_{d \in \mathcal{B}} |\alpha_d| \leq \Gamma(n(b_1) + \cdots + n(b_\ell))^{(\ell-1)(|b_1|+\cdots+|b_\ell|)}$. Then

$$[[\cdots [b_1, b_2], \dots b_\ell], b_{\ell+1}] = \sum_{d \in \mathcal{B}} \alpha_d [d, b_{\ell+1}] = \sum_{d \in \mathcal{B}} \alpha_d \sum_{c \in \mathcal{B}} \gamma_{d, b_{\ell+1}}^c c = \sum_{c \in \mathcal{B}} \left(\sum_{d \in \mathcal{B}} \alpha_d \gamma_{d, b_{\ell+1}}^c \right) c, \tag{90}$$

where the sums are finite and indexed by $d \in \mathcal{B}$ such that $n(d) + n(b_{\ell+1}) = n(b_1) + \cdots + n(b_{\ell+1})$ and $|d| + |b_{\ell+1}| = |b_1| + \cdots + |b_{\ell+1}|$ thus

$$\begin{aligned} \|[[\cdots [b_1, b_2], \dots b_\ell], b_{\ell+1}]\|_{\mathcal{B}} &= \sum_{c \in \mathcal{B}} \left| \sum_{d \in \mathcal{B}} \alpha_d \gamma_{d, b_{\ell+1}}^c \right| \leq \sum_{d \in \mathcal{B}} |\alpha_d| \sum_{c \in \mathcal{B}} |\gamma_{d, b_{\ell+1}}^c| \\ &\leq \Gamma(n(b_1) + \cdots + n(b_\ell))^{(\ell-1)(|b_1|+\cdots+|b_\ell|)} \Gamma(n(b_1) + \cdots + n(b_{\ell+1}))^{|b_1|+\cdots+|b_{\ell+1}|} \end{aligned} \tag{91}$$

which gives the conclusion. □

Proposition 52. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$ with geometric growth with respect to X_0. Then, for every $M \in \mathbb{N}^*$, there exists $C_M > 0$ such that, for every $T \geq 0$, $u \in L^1((0,T); \mathbb{K}^q)$, $b \in \mathcal{B}$ with $n(b) \leq M$ and $t \in [0,T]$,*

$$|\eta_b(t,u)| \leq \frac{C_M^{|b|}}{|b|!} t^{n_0(b)} \|u\|_{L^1(0,t)}^{n(b)}. \tag{92}$$

Proof. We may assume that $(C(r))_{r \in \mathbb{N}^*}$ given by Lemma 51 is non-decreasing. Then, for every $i_1, \dots, i_r \in [\![1, q]\!]$ and $k_1, \dots, k_r \in \mathbb{N}$, for every $b \in \mathcal{B}$,

$$\left| \left\langle [\cdots [\mathrm{ad}_{X_0}^{k_1}(X_{i_1}), \mathrm{ad}_{X_0}^{k_2}(X_{i_2})], \dots \mathrm{ad}_{X_0}^{k_r}(X_{i_r})], b \right\rangle_{\mathcal{B}} \right| \leq C(r)^{|b|}. \tag{93}$$

Indeed, for each $j \in [\![1,r]\!]$, there exists $b_j \in \mathcal{B}$ with $n(b_j) = 1$ and $|b_j| = k_j + 1$ such that $\mathrm{ad}_{X_0}^{k_j}(X_{i_j}) = \pm b_j$ in $\mathcal{L}(X)$. Indeed, the homogeneous part of $\mathcal{L}(X)$ containing k_j times X_0 and X_{i_j} once is of dimension one.

Let $M \in \mathbb{N}^*$ and $b \in \mathcal{B}$ be such that $n(b) \leq M$. We deduce from (75) that

$$\eta_b(t,u) = \sum \frac{(-1)^{m-1}}{mr} \int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1 - t)^{k_1}}{k_1!} \cdots \frac{(\tau_r - t)^{k_r}}{k_r!} u_{i_1}(\tau_1) \cdots u_{i_r}(\tau_r)\, \mathrm{d}\tau \\ \left\langle [\cdots [\mathrm{ad}_{X_0}^{k_1}(X_{i_1}), \mathrm{ad}_{X_0}^{k_2}(X_{i_2})], \ldots \mathrm{ad}_{X_0}^{k_r}(X_{i_r})], b \right\rangle \tag{94}$$

where the sum is taken over $r \in [\![1,\infty]\!]$, $m \in [\![1,r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $k_1, \ldots, k_r \in \mathbb{N}$ and $i_1, \ldots, i_r \in [\![1,q]\!]$. If the summand bra-ket in (94) does not vanish, then $r = n(b)$ and $k_1 + \cdots + k_r = n_0(b)$. Thus the sum in (94) is taken over the finite set $r = n(b)$, $m \in [\![1, n(b)]\!]$, $k_1, \ldots, k_r \in \mathbb{N}$ such that $k_1 + \cdots + k_r = n_0(b)$ and $i_1, \ldots, i_r \in [\![1,q]\!]$, whose cardinal is bounded by $M 2^{|b|} q^M$. Moreover, for every $r, m, k_1, \ldots, k_r, i_1, \ldots, i_r$ in this set, the associated term in (94) is bounded, thanks to (93), by

$$\frac{t^{k_1}}{k_1!} \cdots \frac{t^{k_r}}{k_r!} \|u\|_{L^1(0,t)}^r C(r)^{|b|} \leq t^{n_0(b)} \|u\|_{L^1}^r \left(2^r C(r)\right)^{|b|} \frac{n(b)!}{|b|!} \tag{95}$$

thanks to (117). Thus

$$|\eta_b(t,u)| \leq \frac{1}{|b|!} M! M q^M \left(2^{M+1} C(M)\right)^{|b|} t^{n_0(b)} \|u\|_{L^1}^{n(b)} \tag{96}$$

which gives the conclusion with, for instance, $C_M := M! M q^M 2^{M+1} C(M)$. □

2.5 Infinite product, coordinates of the second kind

In this section, we present an expansion for the formal power series $x(t)$ solution to (15) as a product of exponentials of the members of a Hall basis of $\mathcal{L}(X)$, multiplied by coefficients that have simple expressions as iterated integrals, called *coordinates of the second kind*. This infinite product is an extension, suggested in [69], of Sussmann's infinite product on length-compatible Hall bases [100] to all Hall bases (understood in the generalized sense of Viennot [102, Theorem 1.2] or Shirshov [94, Definition 1]).

2.5.1 Lazard sets, Hall sets and Hall bases

We start by defining Lazard sets and Hall sets, which are two equivalent notions, as proved by Viennot in [102, Corollary 1.1]. They lead to the essential notion of Hall bases (see Definition 56).

Definition 53 (Lazard set)**.** *A* Lazard set *is a subset B of $\mathrm{Br}(X)$, totally ordered by a relation $<$ and such that, for every $M \in \mathbb{N}^*$, the set $B_{[\![1,M]\!]}$ of elements of B with length at most M, labeled as $B_{[\![1,M]\!]} = \{b_1, \ldots, b_{k+1}\}$ with $k \in \mathbb{N}$ and $b_1 < \cdots < b_{k+1}$ satisfies*

$$\begin{cases} b_1 \in Y_0 := X, \\ b_2 \in Y_1 := \{\mathrm{ad}_{b_1}^j(v); j \in \mathbb{N}, v \in Y_0 \setminus \{b_1\}\}, \\ \cdots \\ b_{k+1} \in Y_k := \{\mathrm{ad}_{b_k}^j(v); j \in \mathbb{N}, v \in Y_{k-1} \setminus \{b_k\}\} \end{cases} \tag{97}$$

and

$$B_{[\![1,M]\!]} \cap Y_k = \{b_{k+1}\}, \tag{98}$$

where condition (98) can equivalently be written

$$B_{[\![1,M]\!]} \cap Y_{k+1} = \emptyset, \tag{99}$$

where $Y_{k+1} := \{\mathrm{ad}_{b_{k+1}}^j(v); j \in \mathbb{N}, v \in Y_k \setminus \{b_{k+1}\}\}$.

The elements $\mathrm{ad}^j_{b_\ell}(v)$ for $\ell \in \{0, \dots, k+1\}$, $j \in \mathbb{N}$ and $v \in Y_{\ell-1} \setminus \{b_\ell\}$ are all different in $\mathrm{Br}(X)$ (identify their left and right factors iteratively) and all belong to B.

Definition 54 (Hall set)**.** *A* Hall set *is a subset B of $\mathrm{Br}(X)$, totally ordered by a relation $<$ and such that*

- $X \subset B$,
- *for $b = (b_1, b_2) \in \mathrm{Br}(X)$, $b \in B$ iff $b_1, b_2 \in B$, $b_1 < b_2$ and either $b_2 \in X$ or $\lambda(b_2) \leq b_1$,*
- *for every $b_1, b_2 \in B$ such that $(b_1, b_2) \in B$ then $b_1 < (b_1, b_2)$.*

When $b = (b_1, (b_3, b_4)) \in B$ then b_1 is "sandwiched" in between b_3 and b, since $b_3 \leq b_1 < b$.

Remark 55. *A Hall set can be built by induction on the length. One starts with the set X as well as an order on it. To find all Hall monomials with length n given those of smaller length, one adds first all (b_1, b_2) with $b_1 \in B$, $|b_1| = n-1$, $b_2 \in X$ and $b_1 < b_2$. Then for each bracket $b_2 = (b_2', b_2'') \in B$ with length $|b_2| < n$ one adds all the (b_1, b_2) with $b_1 \in B$ with $|b_1| = n - |b_2|$ and $b_2' \leq b_1 < b_2$. Finally, one inserts the newly generated monomials of degree n into an ordering, maintaining the condition that $b_1 < (b_1, b_2)$.*

Viennot proves in [102, Corollary 1.1] that a subset B of $\mathrm{Br}(X)$ is a Lazard set iff it is a Hall set. He also proves in [102, Proposition 1.1 and Theorem 1.1] that properties (97) and (98) ensure that $\mathrm{E}(B)$ is a linearly independent and generating family of $\mathcal{L}(X)$. This leads to the following definition.

Definition 56 (Hall basis)**.** *Given B a Hall set (or equivalently a Lazard set), $\mathcal{B} := \mathrm{E}(B)$ is a basis of $\mathcal{L}(X)$. Such bases are call* Hall bases.

Remark 57. *Historically, such bases where introduced by Marshall Hall in [59], based on ideas of Philip Hall in [60]. In his historical narrower definition, the third condition in Definition 54 was replaced by the stronger condition: for every $b_1, b_2 \in B$, $b_1 < b_2 \Rightarrow |b_1| \leq |b_2|$. To avoid confusion with the generalized definition, we name them* length-compatible Hall bases *in the sequel.*

Given a Hall set B, the evaluation mapping E is one to one from B to the associated Hall basis $\mathcal{B}$, so that the elements of the basis (belonging to $\mathcal{L}(X)$) can be identified with the bracket of $b \in B \subset \mathrm{Br}(X)$ of which they are the evaluation. We will use this identification in the sequel when no confusion is possible.

Two famous families of Hall bases of $\mathcal{L}(X)$ are the Chen-Fox-Lyndon basis (see [102, Chapter 1]) and the historical length-compatible Hall bases, for which $b_1 < b_2 \Rightarrow |b_1| \leq |b_2|$.

Example 58. *For instance, with $X = \{X_1, X_2\}$, the elements with length at most 4 of each Hall set B of $\mathcal{L}(X)$ with a length-compatible order $<$ such that $X_1 < X_2$ are: X_1, X_2, (X_1, X_2), $\mathrm{ad}^2_{X_1}(X_2)$, $(X_2, (X_1, X_2))$, $\mathrm{ad}^3_{X_1}(X_2)$ and $(X_2, \mathrm{ad}^2_{X_1}(X_2))$, $\mathrm{ad}^2_{X_2}((X_1, X_2))$. Note that, however, $(X_1, (X_2, (X_1, X_2)))$ does not belong to B because $\lambda((X_2, (X_1, X_2))) = X_2$ is not smaller than X_1, and the following equality holds in $\mathcal{L}(X)$*

$$[X_1, [X_2, [X_1, X_2]]] = [[X_1, X_2], [X_1, X_2]] + [X_2, [X_1, [X_1, X_2]]] = [X_2, \mathrm{ad}^2_{X_1}(X_2)] \tag{100}$$

This illustrates how Definition 54 prevents elements from $\mathrm{Br}(X)$, whose evaluations in $\mathcal{L}(X)$ are linked by Jacobi relations, to appear simultaneously in B.

Remark 59. *Let $X := \{X_0, X_1\}$ and $B \subset \mathrm{Br}(X)$ be a Hall set with an order such that $X_0 < X_1$. The definition of a Hall set implies that, for every $k \in \mathbb{N}$, $\mathrm{ad}^k_{X_0}(X_1) \in B$. Moreover, these are all the elements of B containing X_1 exactly once. Since $\mathrm{E}(B)$ is a basis of $\mathcal{L}(X)$, $\mathrm{E}(B) \cap \dot{S}_1$ is a basis of $\dot{S}_1$ and this provides an alternative proof of Lemma 36.*

2.5.2 Infinite product on a Hall basis

Definition 60 (Infinite product)**.** *Let J be a totally ordered set and $(S^j)_{j\in J}$ be a family of $\widehat{\mathcal{A}}(X)$ such that*

- *for every $j \in J$, $\langle S^j, 1\rangle = 1$*
- *for every $\sigma \in I^*$ with $\sigma \neq \emptyset$, the set $\{j \in J; \langle S^j, X_\sigma\rangle \neq 0\}$ is finite.*

The infinite product $\overleftarrow{\prod_{j\in J}} S^j$ is the element of $\widehat{\mathcal{A}}(X)$ defined by

$$\overleftarrow{\prod_{j\in J}} S^j = \sum_{\sigma\in I^*} P_\sigma X_\sigma, \tag{101}$$

where $P_\emptyset = 1$ and P_σ is the finite sum

$$P_\sigma := \sum_{n=0}^{|\sigma|} \sum_{\substack{\sigma_1,\dots,\sigma_n\in I^*,\\ X_{\sigma_1}\cdots X_{\sigma_n}=X_\sigma}} \sum_{\substack{j_1,\dots,j_n\in J,\\ j_1>\cdots>j_n}} \langle S^{j_1}, X_{\sigma_1}\rangle\cdots\langle S^{j_n}, X_{\sigma_n}\rangle. \tag{102}$$

The following lemma is the key point to generalize rigorously Sussmann's infinite product on length-compatible Hall bases, to all Hall bases.

Lemma 61. *Let $\mathcal{B}$ be a Hall basis and $(\alpha_b)_{b\in\mathcal{B}}$ be a family of $\mathbb{K}$. The infinite product $\overleftarrow{\prod_{b\in\mathcal{B}}} e^{\alpha_b b}$ is well defined in $\widehat{\mathcal{A}}(X)$. Moreover, for every $\sigma \in I^*$,*

$$\left\langle \overleftarrow{\prod_{b\in\mathcal{B}}} e^{\alpha_b b}, X_\sigma \right\rangle = \left\langle \overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,|\sigma|]\!]}}} e^{\alpha_b b}, X_\sigma \right\rangle \tag{103}$$

where $\mathcal{B}_{[\![1,|\sigma|]\!]}$ is ordered by the induced order of $\mathcal{B}$.

Proof. $\mathcal{B}$ is a totally ordered set and, for every $b \in \mathcal{B}$, $\langle e^{\alpha_b b}, 1\rangle = 1$. Let $\sigma \in I^*$ with $|\sigma| \geq 1$. For $\alpha \in \mathbb{K}$ and $b \in \mathcal{B}$, the property $\langle e^{\alpha b}, X_\sigma\rangle \neq 0$ requires $|b| \leq |\sigma|$. Indeed

$$e^{\alpha b} - 1 = \sum_{k=1}^{+\infty} \frac{\alpha^k}{k!} b^k \tag{104}$$

has non vanishing coefficients only on monomials $X_{\sigma'}$ with length $|\sigma'| \geq |b|$. Thus the set $\{b \in \mathcal{B}, \langle e^{\alpha_b b}, X_\sigma\rangle \neq 0\}$ is finite. This proves that the infinite product is well defined in $\widehat{A}(X)$ and, by (102), the formula (103) holds. □

2.5.3 Coordinates of the second kind

Definition 62. *Let $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$. The coordinates of the second kind associated to $\mathcal{B}$ is the unique family $(\xi_b)_{b\in\mathcal{B}}$ of functionals $\mathbb{R}_+ \times L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{K}^I) \to \mathbb{K}$ defined by induction in the following way: for every $t > 0$ and $a \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{K}^I)$*

- *$\xi_{X_i}(t,a) := \int_0^t a_i$, for $i \in I$,*
- *for $b \in \mathcal{B}\setminus X$, there exists a unique pair (b_1,b_2) of elements of $\mathcal{B}$ such that $b_1 < b_2$ and a unique maximal integer $m \in \mathbb{N}^*$ such that $b = \mathrm{ad}^m_{b_1}(b_2)$ and then*

$$\xi_b(t,a) := \frac{1}{m!}\int_0^t \xi^m_{b_1}(\tau,a)\dot{\xi}_{b_2}(\tau,a)\,\mathrm{d}\tau. \tag{105}$$

Formula (105) indeed defines continuous functionals on L^1 and the following estimates hold.

Lemma 63. *Let $a_i \in L^1_{\mathrm{loc}}(\mathbb{R}_+;\mathbb{K})$ for $i \in I$. For every $b \in \mathcal{B}$ and $t \geq 0$,*

$$|\dot{\xi}_b(t,a)| \leq |b||a(t)|\|a\|^{|b|-1}_{L^1(0,t)}, \tag{106}$$

$$|\xi_b(t,a)| \leq \|a\|^{|b|}_{L^1(0,t)}. \tag{107}$$

Proof. Estimate (106) is valid for $b \in X$ because $\dot{\xi}_{X_i}(t) = a_i(t)$ for $i \in I$ and propagated by induction on b using the recursive definition (105). Estimate (107) is obtained by time-integration of (106) for each b. □

2.5.4 Infinite product expansion of the solution to the formal differential equation

The following result is due to Sussmann in [100]. The proof below follows Sussmann's argument. It is recalled for sake of giving a self consistent presentation and also to treat the extension from length-compatible Hall bases to all Hall bases (which were not included in Sussmann's original statement).

Theorem 64. *Let $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$. Let $T > 0$ and $a_i \in L^1((0,T);\mathbb{K})$ for $i \in I$. For every $x^\star \in \widehat{\mathcal{A}}(X)$, the solution to the formal diff (15) satisfies, for every $t \in [0,T]$,*

$$x(t) = x^\star \overleftarrow{\prod_{b\in\mathcal{B}}} e^{\xi_b(t,a)b}. \tag{108}$$

Proof. It is sufficient to prove the formula with $x^\star = 1$. To simplify the notations in this proof, we write $\xi_b(t)$ instead of $\xi_b(t,a)$. By Lemma 61 it is sufficient to prove that, for every $t \in [0,T]$ and $\sigma \in I^*$

$$\langle x(t), X_\sigma\rangle = \left\langle \overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,|\sigma|]\!]}}} e^{\xi_b(t)b}, X_\sigma \right\rangle. \tag{109}$$

Let $\sigma \in I^*$, $M := |\sigma|$, $k \in \mathbb{N}$ and $b_1,\dots,b_{k+1}$ and $Y_0,\dots,Y_{k+1}$ be as in (97). The equality (109) can equivalently we written

$$\langle x(t), X_\sigma\rangle = \left\langle e^{\xi_{b_{k+1}}(t)b_{k+1}}\cdots e^{\xi_{b_1}(t)b_1}, X_\sigma \right\rangle. \tag{110}$$

We define $x_0(t) := x(t)$ and, for $j \in [\![1,k+1]\!]$,

$$x_j(t) := x(t)e^{-\xi_{b_1}(t)b_1}\cdots e^{-\xi_{b_j}(t)b_j}. \tag{111}$$

We prove by induction on $j \in [\![0,k+1]\!]$ that

$$\dot{x}_j(t) = x_j(t)\left(\sum_{b\in Y_j}\dot{\xi}_b(t)b\right) \quad \text{and} \quad x_j(0) = 1. \tag{112}$$

It is clear for $j = 0$ because $x_0(t) = x(t)$, $Y_0 = X$ and $\dot{\xi}_{X_i}(t) = a_i(t)$ for $i \in I$. Let $j \in [\![1,k+1]\!]$. We assume (112) holds at step $j-1$. We deduce from the definition of $x_j(t)$ that

$$x_j(t) = x_{j-1}(t)e^{-\xi_{b_j}(t)b_j}. \tag{113}$$

Since $\xi_{b_j}(0) = 0$, $x_j(0) = 1$. Moreover,

$$\begin{aligned}\dot{x}_j(t) &= x_{j-1}(t)\left(\sum_{b\in Y_{j-1}} \dot{\xi}_b(t) b\right) e^{-\xi_{b_j}(t)b_j} - x_{j-1}(t)\dot{\xi}_{b_j}(t) b_j e^{-\xi_{b_j}(t)b_j} \\ &= x_j(t) e^{\xi_{b_j}(t)b_j}\left(\sum_{b\in Y_{j-1}\setminus\{b_j\}} \dot{\xi}_b(t) b\right) e^{-\xi_{b_j}(t)b_j} \\ &= x_j(t) \sum_{m\in\mathbb{N}} \sum_{b\in Y_{j-1}\setminus\{b_j\}} \frac{\xi_{b_j}^m(t)}{m!} \dot{\xi}_b(t) \operatorname{ad}_{b_j}^m(b)\end{aligned} \tag{114}$$

which ends the proof by induction.

We deduce from (99) and (112) for $j = (k+1)$ that $x_{k+1}(t) - 1$ has non vanishing coefficients only on monomials $X_{\sigma'}$ with $|\sigma'| > |\sigma|$. Therefore, by (102),

$$\langle x(t), X_\sigma\rangle = \left\langle x_{k+1}(t) e^{\xi_{b_{k+1}}(t)b_{k+1}} \cdots e^{\xi_{b_1}(t)b_1}, X_\sigma \right\rangle = \left\langle e^{\xi_{b_{k+1}}(t)b_{k+1}} \cdots e^{\xi_{b_1}(t)b_1}, X_\sigma \right\rangle, \tag{115}$$

which concludes the proof. □

3 Technical tools about functions and vector fields

In this section, we state classical definitions and technical results about functions and vector fields. For the sake of completeness, the proofs, although classical, are provided.

Throughout the whole article, $d \in \mathbb{N}^*$ denotes the dimension of the state space for the considered ordinary differential equations. We work locally, in neighborhoods of the origin $0 \in \mathbb{K}^d$. For $\delta > 0$, B_δ denotes the closed ball of center 0 and radius δ in the state space $\mathbb{K}^d$.

3.1 Functional spaces for finite or analytic regularity

3.1.1 Conventions for multi-indexes

For $a \in \mathbb{N}^*$ and a multi-index $\alpha = (\alpha^1, \ldots, \alpha^a) \in \mathbb{N}^a$, we use the notations $|\alpha| := \alpha^1 + \cdots + \alpha^a$, $\partial^\alpha := \partial_1^{\alpha^1} \cdots \partial_a^{\alpha^a}$ and $\alpha! := \alpha^1! \cdots \alpha^a!$.

Lemma 65. *The following estimates hold:*

$$\forall n \in \mathbb{N}, \quad n^n e^{-n} e \le n! \le (n+1)^{n+1} e^{-(n+1)} e, \tag{116}$$

$$\forall a \in \mathbb{N}^*, \forall \alpha = (\alpha^1, \ldots, \alpha^a) \in \mathbb{N}^a, \quad 2^{-(a-1)|\alpha|}|\alpha|! \le \alpha! \le |\alpha|! \tag{117}$$

Proof. The first inequality can be proved using classical series-integral comparison and the second by iterating $p!q! \ge 2^{-(p+q)}(p+q)!$ for every $p, q \in \mathbb{N}$. □

3.1.2 Finite-regularity norms

Definition 66 (Regular functions)**.** *Let $a, b \in \mathbb{N}^*$ and K a compact subset of $\mathbb{K}^a$. Let $k \in \mathbb{N}$. We endow $\mathcal{C}^k(K; \mathbb{K}^b)$, the space of functions whose real-derivatives are well-defined and continuous up to order k on an open neighborhood of K to $\mathbb{K}^b$ with the norm*

$$\|f\|_{\mathcal{C}^k} := \sum_{j=1}^{b} \sum_{|\alpha|\le k} \frac{1}{\alpha!} \|\partial^\alpha f_j\|_{L^\infty(K)}, \tag{118}$$

where the sum ranges over multi-indexes $\alpha \in \mathbb{N}^a$ whose sum is at most k and $f_1, \ldots, f_b$ are the components of the vector-valued function f. We denote by $\mathcal{C}^\infty(K; \mathbb{K}^b)$ the intersection of these spaces over $k \in \mathbb{N}$.

Lemma 67 (Submultiplicativity). *Let $a \in \mathbb{N}^*$, K a compact subset of $\mathbb{K}^a$ and $k \in \mathbb{N}$. For every $f, g \in \mathcal{C}^k(K;\mathbb{K})$, one has*

$$\|fg\|_{\mathcal{C}^k} \leq \|f\|_{\mathcal{C}^k} \|g\|_{\mathcal{C}^k}. \tag{119}$$

Proof. Using the multivariate Leibniz formula, one has

$$\begin{aligned} \|fg\|_{\mathcal{C}^k} &= \sum_{|\alpha|\leq k} \frac{1}{\alpha!} \|\partial^\alpha (fg)\|_{L^\infty(K)} \\ &\leq \sum_{|\alpha|\leq k} \frac{1}{\alpha!} \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} \|\partial^\beta f\|_{L^\infty(K)} \|\partial^{\alpha-\beta} g\|_{L^\infty(K)} \leq \|f\|_{\mathcal{C}^k} \|g\|_{\mathcal{C}^k}, \end{aligned} \tag{120}$$

where the sum ranges over all multi-indexes $\beta \in \mathbb{N}^a$ such that $\beta_i \leq \alpha_i$ for each $i \in [\![1, a]\!]$. □

Lemma 68 (Control of gradients). *Let $a \in \mathbb{N}^*$, K a compact subset of $\mathbb{K}^a$ and $k \in \mathbb{N}$. For every $f \in \mathcal{C}^{k+1}(K;\mathbb{K})$ and $j \in [\![1, a]\!]$,*

$$\|\partial_j f\|_{\mathcal{C}^k} \leq (k+1) \|f\|_{\mathcal{C}^{k+1}}. \tag{121}$$

Proof. By (118),

$$\|\partial_j f\|_{\mathcal{C}^k} = \sum_{|\alpha|\leq k} \frac{1}{\alpha!} \|\partial^{\alpha+e_j} f\|_{L^\infty(K)} = \sum_{|\alpha|\leq k} \frac{\alpha_j + 1}{(\alpha+e_j)!} \|\partial^{\alpha+e_j} f\|_{L^\infty(K)} \leq (k+1)\|f\|_{\mathcal{C}^{k+1}} \tag{122}$$

since $\alpha_j \leq |\alpha| \leq k$. □

3.1.3 Analytic norms

Definition 69 (Analytic norms). *Let $a, b \in \mathbb{N}^*$ and K a compact subset of $\mathbb{K}^a$. We define $\mathcal{C}^\omega(K;\mathbb{K}^b)$ the space of real-analytic functions defined on an open neighborhood of K to $\mathbb{K}^b$, as the union for $r > 0$ of the spaces $\mathcal{C}^{\omega,r}(K;\mathbb{K}^b)$, which are the subsets of $\mathcal{C}^\infty(K;\mathbb{K}^b)$ for which the following norm is finite*

$$|||f|||_r := \sum_{i=1}^{b} \sum_{\alpha\in\mathbb{N}^a} \frac{r^{|\alpha|}}{\alpha!} \|\partial^\alpha f_i\|_{L^\infty(K)}. \tag{123}$$

Lemma 70 (Submultiplicativity). *Let $a \in \mathbb{N}^*$, K a compact subset of $\mathbb{K}^a$, $r > 0$. Then, for every $f, g \in \mathcal{C}^{\omega,r}(K;\mathbb{K})$, one has*

$$|||fg|||_r \leq |||f|||_r \, |||g|||_r. \tag{124}$$

Proof. Using the multivariate Leibniz formula, one has

$$\begin{aligned} |||fg|||_r &= \sum_{\alpha\in\mathbb{N}^a} \frac{r^{|\alpha|}}{\alpha!} \|\partial^\alpha (fg)\|_{L^\infty(K)} \\ &\leq \sum_{\alpha\in\mathbb{N}^a} \frac{r^{|\alpha|}}{\alpha!} \sum_{\beta\leq\alpha} \binom{\alpha}{\beta} \|\partial^\beta f\|_{L^\infty(K)} \|\partial^{\alpha-\beta} g\|_{L^\infty(K)} = |||f|||_r \, |||g|||_r, \end{aligned} \tag{125}$$

where the sum ranges over all multi-indexes $\beta \in \mathbb{N}^a$ such that $\beta_i \leq \alpha_i$ for each $i \in [\![1, a]\!]$. □

Lemma 71 (Control of gradients). *Let $a \in \mathbb{N}^*$, K a compact subset of $\mathbb{K}^a$. For every $r_2 > r_1 > 0$, $f \in \mathcal{C}^{\omega,r_2}(K;\mathbb{K})$ and $j \in [\![1, a]\!]$,*

$$|||\partial_j f|||_{r_1} \leq \frac{1}{r_1} \left(e \ln \frac{r_2}{r_1} \right)^{-1} |||f|||_{r_2}. \tag{126}$$

In particular, if $r_2 \leq e r_1$,

$$|||\partial_j f|||_{r_1} \leq \frac{1}{r_2 - r_1} |||f|||_{r_2}. \tag{127}$$

Proof. We start with the first estimate (126). One has

$$\begin{aligned}
|||\partial_j f|||_{r_1} = \sum_{\alpha\in\mathbb{N}^a} \frac{r_1^{|\alpha|}}{\alpha!}\|\partial^{\alpha+e_j} f\|_{L^\infty(K)} &= \frac{1}{r_1}\sum_{\alpha\in\mathbb{N}^a} \frac{r_1^{|\alpha+e_j|}}{(\alpha+e_j)!}\frac{(\alpha+e_j)!}{\alpha!}\|\partial^{\alpha+e_j} f\|_{L^\infty(K)} \\
&\leq \frac{1}{r_1}|||f|||_{r_2} \sup_{\alpha\in\mathbb{N}^a}\left(\frac{r_1}{r_2}\right)^{|\alpha+e_j|}\frac{(\alpha+e_j)!}{\alpha!} \\
&\leq \frac{1}{r_1}|||f|||_{r_2} \sup_{n\geq 1} n\left(\frac{r_1}{r_2}\right)^{n}.
\end{aligned} \tag{128}$$

For $x\in(0,1)$, let $C(x) := \sup_{n\geq 1} nx^n = \sup_{n\geq 1}\exp(\ln n + n\ln x)$. Considering that x is a fixed parameter and differentiating the argument with respect to $n\in[1,+\infty)$ yields

$$\frac{\mathrm{d}}{\mathrm{d}n}(\ln n + n\ln x) = \frac{1}{n} + \ln x. \tag{129}$$

Since $x<1$, this derivative is negative for n large enough. For $x\geq 1/e$, the global maximum is for $n=-1/\ln x$. So its value yields the bound

$$C(x) \leq (-e\ln x)^{-1}. \tag{130}$$

For $x\leq 1/e$, the supremum over n is achieved for $n=1$ and its value is x. Since $x\leq(-e\ln x)^{-1}$ for $x\in(0,1)$, the bound (130) is looser and valid for every $x\in(0,1)$.

The second inequality is a consequence of the estimate $\ln(1+\sigma)\geq\sigma/(e-1)$ for $\sigma\leq e-1$. □

Remark 72. *The first estimate* (126) *is classical (see e.g. [84]). The second estimate* (126) *is a simplified version, restricted to the case when the relative radius loss is small enough. This is the form under which we will use Lemma 71 in the sequel since we consider small radius losses.*

3.2 Estimates for differential operators and Lie brackets

3.2.1 Vector-valued functions, vector fields and differential operators

As is usual, we will identify each smooth vector-valued function with the associated first-order linear differential operator, and we will refer to both objects as a *vector field*. Let K be a compact subset of $\mathbb{K}^d$.

Definition 73 (Vector field)**.** *Given* coordinates $a_1,\ldots,a_d\in\mathcal{C}^\infty(K;\mathbb{K})$*, we define the associated* vector field f *indifferently as the vector-valued function* $f=(a_1,\ldots,a_d)$ *of* $\mathcal{C}^\infty(K;\mathbb{K}^d)$ *(mapping each point of the state space to a vector of* $\mathbb{K}^d$*), or as the first-order linear differential operator* $f=a_1\partial_1+\cdots+a_d\partial_d$ *(acting on smooth functions* $\phi\in\mathcal{C}^\infty(K;\mathbb{K})$*).*

Remark 74 (Composition of vector fields)**.** *Seen as linear operators, vector fields can be composed, yielding higher-order differential operators. Let* $f,g\in\mathcal{C}^\infty(K;\mathbb{K}^d)$ *(denoting their coordinates by* $a_1,\ldots,a_d$ *and* $b_1,\ldots b_d$*) and* $\phi\in\mathcal{C}^\infty(K;\mathbb{K})$*. We will hence write*

$$f\phi = \sum_{i=1}^{d} a_i\partial_i\phi \tag{131}$$

and

$$fg\phi = \sum_{i=1}^{d}\sum_{j=1}^{d} a_i\left(b_j\partial_i\partial_j\phi + (\partial_i b_j)(\partial_j\phi)\right). \tag{132}$$

Similarly, for $k\in\mathbb{N}$*,* $f^k\phi$ *will denote the composition of the linear differential operator* f *with itself* k *times, applied to* ϕ*. Such formulas still make sense in a finite regularity setting, as long as all derivatives are well-defined.*

3.2.2 Estimates for products

Lemma 75. *Let $k \in \mathbb{N}$, $n, b \in \mathbb{N}^*$, $\delta > 0$, $f_1, \dots f_n \in \mathcal{C}^{k+n-1}(B_\delta; \mathbb{K}^d)$ and $\phi \in \mathcal{C}^{k+n}(B_\delta; \mathbb{K}^b)$. With the notations of Remark 74,*

$$\|f_n \cdots f_1 \phi\|_{\mathcal{C}^k} \leq \frac{(k+n)!}{k!} \|f_1\|_{\mathcal{C}^{k+n-1}} \cdots \|f_n\|_{\mathcal{C}^{k+n-1}} \|\phi\|_{\mathcal{C}^{k+n}}. \tag{133}$$

Proof. For $n = 1$, it is a consequence of (118), Lemma 67 and Lemma 68. For $n > 1$, the estimate follows by induction. □

Lemma 76. *Let $r_2 > 0$, $r_1 \in [r_2/e, r_2)$, $n, b \in \mathbb{N}^*$ and $\delta > 0$. Let $f_1, \dots, f_n \in \mathcal{C}^{\omega, r_2}(B_\delta; \mathbb{K}^d)$ and $\phi \in \mathcal{C}^{\omega, r_2}(B_\delta; \mathbb{K}^b)$. With the notations of Remark 74,*

$$|||f_n \cdots f_1 \phi|||_{r_1} \leq \frac{n!}{e} \left(\frac{e}{r_2 - r_1} \right)^n |||f_n|||_{r_2} \cdots |||f_1|||_{r_2} \, |||\phi|||_{r_2}. \tag{134}$$

In particular, under the same assumptions,

$$\|f_n \cdots f_1 \phi\|_{\mathcal{C}^0} \leq n! \left(\frac{5}{r_2} \right)^n |||f_n|||_{r_2} \cdots |||f_1|||_{r_2} \, |||\phi|||_{r_2}. \tag{135}$$

Proof. For $n = 1$, estimate (134) is a consequence of (123), (124) and (127). For $n > 1$, one applies the $n = 1$ estimate n times with a radius increment $(r_2 - r_1)/n$ at each step. This yields more precisely

$$\begin{aligned} |||f_n \cdots f_1 \phi|||_{r_1} &\leq \left(\frac{n}{r_2 - r_1} \right) |||f_n|||_{r_1} \, |||f_{n-1} \cdots f_1 \phi|||_{r_1 + \frac{r_2 - r_1}{n}} \\ &\leq \left(\frac{n}{r_2 - r_1} \right)^n |||\phi|||_{r_2} \prod_{j=1}^{n} |||f_j|||_{r_1 + (n-j)\frac{r_2 - r_1}{n}}, \end{aligned} \tag{136}$$

which concludes the proof because the norm (123) is non-decreasing with respect to r, and we can bound n^n using (116). Estimate (135) is a direct consequence for the particular choice $r_1 = r_2/e$, because $e^2/(e-1) \leq 5$. □

3.2.3 Lie brackets

Definition 77 (Lie bracket of vector fields)**.** *We define the* Lie bracket *of smooth vector fields f and g as the usual commutator of the associated linear differential operators: $[f, g] := fg - gf$ (with the notations of Remark 74). By (132) and Schwarz's theorem, one checks that $[f, g]$ is also a first-order differential operator, which, as a vector-valued function, can be computed as $[f, g] = (Dg)f - (Df)g$.*

Definition 78 (Evaluated Lie bracket)**.** *Let I be a finite set of indices, $X = \{X_i; i \in I\}$ be indeterminates and $\{f_i; i \in I\}$ be $\mathcal{C}^\infty$ vector fields on a subset Ω of $\mathbb{K}^d$. For an iterated bracket $b \in \mathrm{Br}(X)$, we define $f_b := \Lambda(\mathrm{E}(b))$, where $\Lambda : \mathcal{L}(X) \to \mathcal{C}^\infty(\Omega; \mathbb{K}^d)$ is the unique homomorphism of Lie algebras such that $\Lambda(X_i) = f_i$ for every $i \in I$ (see Remark 9 and Lemma 7).*

The vector field f_b is obtained by replacing the indeterminates X_i with the corresponding vector fields f_i in the iterated bracket b, for instance $f_{(X_1,(X_2,X_3))} = [f_1, [f_2, f_3]]$.

The notation f_b will sometimes denote the same vector field, seen as a vector-valued function, under weaker regularity assumptions, for instance $f_i \in \mathcal{C}^{|b|-1}$ and then $f_b \in \mathcal{C}^0$.

Lemma 79 (Finite regularity estimate)**.** *Let $k \in \mathbb{N}$ and $\delta > 0$. Let $b \in \mathrm{Br}(X)$. For $i \in I$, let $f_i \in \mathcal{C}^{k+|b|-1}(B_\delta; \mathbb{K}^d)$. Then,*

$$\|f_b\|_{\mathcal{C}^k} \leq 2^{|b|-1} \frac{(k + |b| - 1)!}{k!} \prod_{i \in I} \|f_i\|_{\mathcal{C}^{k+|b|-1}}^{n_i(b)}. \tag{137}$$

Proof. This follows from (133) because, as can be checked by induction on $|b|$, f_b is a sum of at most $2^{|b|-1}$ terms of the form studied in Lemma 75, where ϕ is one of the vector fields f_i. □

Lemma 80 (Analytic estimate)**.** *Let $r_2 > 0$, $r_1 \in [r_2/e, r_2)$ and $\delta > 0$. Let $b \in \mathrm{Br}(X)$. For $i \in I$, let $f_i \in \mathcal{C}^{\omega,r_2}(B_\delta; \mathbb{K}^d)$. Then,*

$$|||f_b|||_{r_1} \leq \frac{(|b|-1)!}{e} \left(\frac{2e}{r_2 - r_1}\right)^{|b|-1} \prod_{i \in I} |||f_i|||_{r_2}^{n_i(b)}. \tag{138}$$

In particular, under the same assumptions,

$$\|f_b\|_{\mathcal{C}^0} \leq (|b|-1)! \left(\frac{9}{r_2}\right)^{|b|-1} \prod_{i \in I} |||f_i|||_{r_2}^{n_i(b)}. \tag{139}$$

$$\|f_b\|_{\mathcal{C}^1} \leq \max\left\{1, \frac{1}{r_2}\right\} (|b|-1)! \left(\frac{9}{r_2}\right)^{|b|-1} \prod_{i \in I} |||f_i|||_{r_2}^{n_i(b)}. \tag{140}$$

Proof. Estimate (138) stems from (134) because, as can be checked by induction on $|b|$, f_b is a sum of at most $2^{|b|-1}$ terms of the form studied in Lemma 76, where ϕ is one of the vector fields f_i. Estimates (139) and (140) are direct consequences of (138) for the particular choice $r_1 = r_2/e$ because $2e^2/(e-1) \leq 9$ and, for every $r_1 > 0$, $\|f_b\|_{\mathcal{C}^1} \leq \max\{1, \frac{1}{r_1}\} |||f_b|||_{r_1}$. □

Remark 81. *The fact that estimate* (138) *scales like the factorial of the length of the Lie bracket is optimal, as illustrated by the following vector fields. For $x \in \mathbb{R}^2$ with $|x| < 1$, define*

$$f_0(x) := e_1 \quad \text{and} \quad f_1(x) := \frac{1}{1 - x_1} e_2. \tag{141}$$

Using (123)*, one checks that these vector fields belong in particular to $\mathcal{C}^{\omega,r}(B_\delta; \mathbb{R}^2)$ for $r = \frac{1}{4}$ and $\delta = \frac{1}{2}$, with $|||f_0|||_r = 1$ and $|||f_1|||_r = 2$. For $k \in \mathbb{N}$, one has*

$$\mathrm{ad}_{f_0}^k(f_1)(x) = \partial_1^k \left(\frac{1}{1 - x_1}\right) e_2 = \frac{k!}{(1 - x_1)^{k+1}} e_2. \tag{142}$$

Moreover, since f_0 is constant and f_1 depends only on x_1 but is a multiple of e_2, every Lie bracket involving f_1 at least twice vanishes identically. Since these analytic vector fields "saturate" the bounds and exhibit such a nice structure, we will use them repeatedly in our counter-examples.

3.3 Well-posedness of ordinary differential equations

The nonlinear differential equations

$$\dot{x}(t) = f(t, x(t)) \quad \text{and} \quad x(0) = p \tag{143}$$

will be studied in the following classical frameworks.

Lemma 82. *Let $\delta, T > 0$ and $f \in L^1((0,T); \mathcal{C}^1(B_{2\delta}; \mathbb{K}^d))$ such that $\|f\|_{L^1((0,T);\mathcal{C}^0)} < \delta$.*

1. *For each $p \in B_\delta$, there exists a unique function $x(\cdot; f, p) \in \mathcal{C}^0([0,T]; B_{2\delta})$ such that*

$$\forall t \in [0,T], \quad x(t; f, p) = p + \int_0^t f(\tau, x(\tau; f, p))\, \mathrm{d}\tau. \tag{144}$$

2. *If $f \in \mathcal{C}^0([0,T] \times B_{2\delta}; \mathbb{K}^d)$ then $x(\cdot; f, p) \in \mathcal{C}^1([0,T]; B_{2\delta})$ and satisfies (143) pointwise.*

3. *If $f \in \mathcal{C}^\infty([0,T] \times B_{2\delta}; \mathbb{K}^d)$, the map $p \in B_\delta \mapsto x(\cdot; f, p) \in \mathcal{C}^0([0,T]; B_{2\delta})$ is smooth.*

4. *If g satisfies the same assumptions as f, then, for each $p \in B_\delta$ and $t \in [0,T]$,*

$$|x(t; f, p) - x(t; g, p)| \leq \|f - g\|_{L^1((0,t);\mathcal{C}^0)} \exp\left(\|f\|_{L^1((0,t);\mathcal{C}^1)}\right). \tag{145}$$

Proof. We proceed step by step. Let $E := \mathcal{C}^0([0,T]; B_{2\delta})$.

1. Define $\Theta : E \to E$ by $\Theta(x)(t) := p + \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau$ for $x \in E$. Thanks to the smallness assumption on f, $\Theta(x)(t) \in B_{2\delta}$. Let $n \in \mathbb{N}^*$ be such that $\|f\|^n_{L^1((0,T);\mathcal{C}^1)}/n! < 1$. By the Banach fixed-point theorem, Θ^n has a unique fixed point, which is also a fixed point of Θ.

2. If f is continuous, then $t \mapsto \Theta(x(t; f, p))$ belongs to $\mathcal{C}^1([0,T]; B_{2\delta})$ and its derivative at time t is $f(t, x(t; f, p))$.

3. If f is smooth, let $\bar{p} \in B_\delta$, $\bar{x} := x(\cdot; f, \bar{p})$ and define $F : B_\delta \times E \to E$ by

$$\forall t \in [0,T], \quad F(p,x)(t) := x(t) - p - \int_0^t f(\tau, x(\tau))\,\mathrm{d}\tau \tag{146}$$

   Then F is of class $\mathcal{C}^\infty$, vanishes at $(\bar{p}, \bar{x})$ and $\partial_x F(\bar{p}, \bar{x})$ is a bijection on E. By the implicit function theorem, the map $p \mapsto x(\cdot; f, p)$ is $\mathcal{C}^\infty$ on a neighborhood of $\bar{p}$.

4. This follows from a standard Grönwall's lemma argument. □

Lemma 83. *Let $\delta, \delta_u > 0$, $q \in \mathbb{N}^*$ and $f \in \mathcal{C}^\omega(B_{2\delta} \times B_{\mathbb{K}^q}(0, \delta_u); \mathbb{K}^d)$. Let $T := \delta / \|f\|_{\mathcal{C}^0}$. For each $p \in B_\delta$ and $u \in L^\infty((0,T); \mathbb{K}^q)$ with $\|u\|_{L^\infty} \leq \delta_u$, there exists a unique solution $x \in \mathcal{C}^0([0,T]; B_{2\delta})$ to*

$$\begin{cases} \dot{x}(t) = f(x(t), u(t)), \\ x(0) = p, \end{cases} \tag{147}$$

denoted $x(t; f, u, p)$. Moreover, the map $(u, p) \mapsto x(\cdot; f, u, p) \in \mathcal{C}^0([0,T]; B_{2\delta})$ is real-analytic on $B_\delta \times B_{L^\infty(0,T)}(0, \delta_u)$.

Proof. Existence stems from Lemma 82. Analyticity is a consequence of the implicit function theorem, which yields the analyticity of the implicit function when the direct function is analytic (see e.g. [27, Theorem 4.5.4]). □

3.4 Flows, compositions and pushforwards

Here and in the sequel, when we manipulate *flows* of vector fields, we always make sure that an appropriate smallness assumption ensures that the local flow is well-defined up to the time at which we evaluate it.

3.4.1 Definitions and approximations

By applying Lemma 82 to a time-independent vector field we obtain the following object.

Definition 84 (Flow of time-independent vector fields)**.** *Let $\delta > 0$. Let $f \in \mathcal{C}^1(B_{2\delta}; \mathbb{K}^d)$ such that $\|f\|_{\mathcal{C}^0} < \delta$. We denote by e^f the flow at time one of the vector field f,*

$$e^f : \begin{cases} B_\delta \to B_{2\delta}, \\ \quad p \mapsto x(1; f, p), \end{cases} \tag{148}$$

with the notations of Section 3.3. We write $e^f p$ instead of $e^f(p)$ to allow easier composition of flows. When moreover $f \in \mathcal{C}^\infty(B_{2\delta}; \mathbb{K}^d)$, e^f can also be seen as the zero-order linear operator on $\mathcal{C}^\infty(B_{2\delta}; \mathbb{K})$ defined by $e^f \phi : p \mapsto \phi(e^f p)$.

Lemma 85. *Let $\delta > 0$ and $f \in \mathcal{C}^1(B_\delta; \mathbb{K}^d)$. Assume that $\delta' := \delta - \|f\|_{\mathcal{C}^0(B_\delta)} > 0$. For each $p \in B_{\delta'}$, $e^f p$ is well-defined and $e^f p \in B_\delta$. Moreover,*

$$|e^f p - p| \leq \|f\|_{\mathcal{C}^0(B_\delta)}, \tag{149}$$

and

$$\|D(e^f)\|_{\mathcal{C}^0(B_{\delta'})} \leq e^{\|Df\|_{\mathcal{C}^0(B_\delta)}} \leq e^{\|f\|_{\mathcal{C}^1(B_\delta)}}. \tag{150}$$

Proof. The second estimate comes from the fact that $D(e^f)_{|p} = R(1)$ where

$$\dot{R}(t) = Df(e^{tf}p)R(t) \quad \text{and} \quad R(0) = \mathrm{Id}. \tag{151}$$

Thus, by Grönwall's lemma,

$$\|R(1)\| \leq \|\mathrm{Id}\| e^{\int_0^1 \|Df(e^{tf}p)\|\,\mathrm{d}t} \leq e^{\|Df\|}, \tag{152}$$

which concludes the proof. □

The exponential notation is motivated by the possibility to approximate e^f by partial sums of the exponential series of the linear differential operator f. It is completely legitimate in the analytic setting, as underlined by the following result.

Lemma 86 (Approximation of autonomous flows)**.** *Let $\delta > 0$, $f \in \mathcal{C}^1(B_{2\delta}; \mathbb{K}^d)$ with $\|f\|_{\mathcal{C}^0} < \delta$. Using the notations of Remark 74:*

1. *For each $M \in \mathbb{N}$, if $f \in \mathcal{C}^M(B_{2\delta}; \mathbb{K}^d)$ and $\phi \in \mathcal{C}^{M+1}(B_{2\delta}; \mathbb{K})$, for each $p \in B_\delta$,*

$$\left| \left(e^f - \sum_{k=0}^{M} \frac{f^k}{k!} \right) (\phi)(p) \right| \leq \|f\|_{\mathcal{C}^M}^{M+1} \|\phi\|_{\mathcal{C}^{M+1}}. \tag{153}$$

2. *If $f \in \mathcal{C}^\omega(B_{2\delta}; \mathbb{K}^d)$ and $\phi \in \mathcal{C}^\omega(B_{2\delta}; \mathbb{K})$, for t small enough, for each $p \in B_\delta$,*

$$e^{tf}(\phi)(p) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} f^k \phi(p) \tag{154}$$

and the sum converges absolutely in the sense of analytic functions.

Proof. First statement. By the first point of Lemma 82, $e^{tf}(p)$ is well defined for every $t \in [0,1]$ and takes values in $B_{2\delta}$. For $t \in [0,1]$ and $k \in [\![0, M+1]\!]$, we have

$$\frac{\mathrm{d}^k}{\mathrm{d}t^k} \left[\phi(e^{tf}(p)) \right] = \left(f^k \phi \right)(e^{tf}(p)). \tag{155}$$

Thus, the considered sum is the Taylor expansion of order M of the map $t \mapsto \phi(e^{tf}(p))$ at $t = 0$ and

$$\left(e^f - \sum_{k=0}^{M} \frac{f^k}{k!} \right) (\phi)(p) = \int_0^1 \frac{(1-s)^M}{M!} \left(f^{M+1} \phi \right)(e^{sf}(p))\,\mathrm{d}s. \tag{156}$$

This concludes the proof of (153) thanks to the integration in (156) and Lemma 75.

Second statement. Let $r > 0$ be such that $f \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K}^d)$ and $\phi \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K})$. Let $r' \in [r/e, r)$. By (134), for each $k \in \mathbb{N}$,

$$\left| \left| \left| \frac{t^k}{k!} f^k \phi \right| \right| \right|_{r'} \leq \frac{|t|^k}{k!} \frac{k!}{e} \left(\frac{e}{r - r'} \right)^k |||f|||_r^k \, |||\phi|||_r \,, \tag{157}$$

so that the sum converges absolutely in $\mathcal{C}^{\omega,r'}$ when $|t|e\,|||f|||_r < r - r'$. Moreover, by (156) with $f \leftarrow tf$ and (134),

$$\left\| \left(e^{tf} - \sum_{k=0}^{M} \frac{t^k}{k!} f^k \right) (\phi) \right\|_{\mathcal{C}^0} \leq \frac{|t|^{M+1}}{(M+1)!} \|f^{M+1}\phi\|_{\mathcal{C}^0}, \tag{158}$$

where, using (157), the right-hand side tends to zero as $M \to +\infty$ under the same smallness condition; so that the sum converges towards $e^{tf}\phi$ in $\mathcal{C}^{\omega,r'}$ when $|t|e\,|||f|||_r < r - r'$. □

3.4.2 Pushforwards of vector fields by diffeomorphisms

Definition 87 (Pushforward of a vector field by a diffeomorphism)**.** *Let Ω, Ω' be open subsets of $\mathbb{K}^d$. Let $\theta \in \mathcal{C}^1(\Omega; \Omega')$ be a local diffeomorphism from Ω to Ω'. Let $f \in \mathcal{C}^0(\Omega; \mathbb{K}^d)$ be a vector field. We define $\theta_* f \in \mathcal{C}^0(\Omega'; \mathbb{K}^d)$* the pushforward of f by θ *as*

$$(\theta_* f)(q) := (D\theta)_{|\theta^{-1}(q)} f(\theta^{-1}(q)) = (D\theta^{-1})^{-1}_{|q} f(\theta^{-1}(q)). \tag{159}$$

Lemma 88 (Chain rule for pushforwards)**.** *Let $\Omega, \Omega', \Omega''$ be open subsets of $\mathbb{K}^d$. Let $\theta \in \mathcal{C}^1(\Omega; \Omega')$ be a local diffeomorphism from Ω to Ω' and $\theta' \in \mathcal{C}^1(\Omega'; \Omega'')$ be a local diffeomorphism from Ω' to Ω''. Let $f \in \mathcal{C}^0(\Omega; \mathbb{K}^d)$ be a vector field. Then, on Ω'',*

$$\theta'_*(\theta_* f) = (\theta' \circ \theta)_* f. \tag{160}$$

Proof. This is a consequence of the chain rule for differentiation, see e.g. [77, Problem 12-10]. □

Lemma 89 (Lie brackets of pushforwards)**.** *Let Ω, Ω' be open subsets of $\mathbb{K}^d$. Let $\theta \in \mathcal{C}^2(\Omega; \Omega')$ be a local diffeomorphism from Ω to Ω'. Let $f, g \in \mathcal{C}^1(\Omega; \mathbb{K}^d)$ be two vector fields. Then, on Ω',*

$$[\theta_* f, \theta_* g] = \theta_*[f, g]. \tag{161}$$

Proof. This is a consequence of the chain rule for differentiation, see e.g. [77, Corollary 8.31]. □

3.4.3 Composition of vector fields with flows

Lemma 90. *Let $\delta > 0$, $f_0 \in \mathcal{C}^1(B_{2\delta}; \mathbb{K}^d)$ and $t \in \mathbb{R}$ such that $|t|\|f_0\|_{\mathcal{C}^0} < \delta$. Denote by $\Phi_0(t, p) := e^{tf_0}(p)$ the associated flow for $p \in B_\delta$.*

1. *For each $M \in \mathbb{N}$, if $f_0, f_1 \in \mathcal{C}^{M+1}(B_{2\delta}; \mathbb{K}^d)$, then, for each $p \in B_\delta$,*

$$\left| (\partial_p \Phi_0(t,p))^{-1} f_1 (\Phi_0(t,p)) - \sum_{k=0}^{M-1} \frac{t^k}{k!} \operatorname{ad}^k_{f_0}(f_1)(p) \right| \leq \frac{t^M}{M!} \left\| \operatorname{ad}^M_{f_0}(f_1) \right\|_{\mathcal{C}^0}. \tag{162}$$

2. *For each $M \in \mathbb{N}$, if $f_0, f_1 \in \mathcal{C}^{M+1}(B_{2\delta}; \mathbb{K}^d)$ and $\operatorname{ad}^M_{f_0}(f_1) \equiv 0$, then, for each $p \in B_\delta$,*

$$(\Phi_0(-t)_* f_1)(p) = (\partial_p \Phi_0(t,p))^{-1} f_1 (\Phi_0(t,p)) = \sum_{k=0}^{M-1} \frac{t^k}{k!} \operatorname{ad}^k_{f_0}(f_1)(p). \tag{163}$$

   This holds in particular when $\mathcal{L}(\{f_0, f_1\})$ is nilpotent with index at most $(M+1)$.

3. *If $r > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K}^d)$, then, for $|t| < \frac{r}{9|||f_0|||_r}$, for each $p \in B_\delta$,*

$$(\Phi_0(-t)_* f_1)(p) = (\partial_p \Phi_0(t,p))^{-1} f_1 (\Phi_0(t,p)) = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \operatorname{ad}^k_{f_0}(f_1)(p), \tag{164}$$

   where, for every $r' \in [r/e, r)$ the series converges in $\mathcal{C}^{\omega,r'}(B_{2\delta}; \mathbb{K}^d)$ when $|t| < \frac{r-r'}{6|||f_0|||_r}$.

4. Let $H_0, H_1 \in \mathcal{M}_d(\mathbb{K}^d)$ and $M \in \mathbb{N}^*$. Then

$$\left\| e^{H_0} H_1 e^{-H_0} - \sum_{k=0}^{M-1} \frac{1}{k!} \operatorname{ad}_{H_0}^k(H_1) \right\| \leq \frac{(2\|H_0\|)^M}{M!} \|H_1\| e^{2\|H_0\|} \tag{165}$$

and

$$e^{H_0} H_1 e^{-H_0} = \sum_{k=0}^{+\infty} \frac{1}{k!} \operatorname{ad}_{H_0}^k(H_1), \tag{166}$$

where ad *is the commutator of matrices* $\operatorname{ad}_A(B) := [A, B] = AB - BA$ *and* $\|\cdot\|$ *a sub-multiplicative norm on* $\mathcal{M}_d(\mathbb{K})$ *such that* $\|\mathrm{Id}_d\| = 1$.

Proof. We proceed step by step.

1. First, for each $\tau \in [0, t]$, $\Phi_0(\tau, p)$ is well-defined. Taking into account that

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}\tau} \left[(\partial_p \Phi_0(t,p))^{-1} \right] &= -(\partial_p \Phi_0(t,p))^{-1} \frac{\mathrm{d}}{\mathrm{d}\tau} \left[\partial_p \Phi_0(\tau, p) \right] (\partial_p \Phi_0(t,p))^{-1} \\ &= -(\partial_p \Phi_0(t,p))^{-1} D f_{0|\Phi_0(\tau,p)}, \end{aligned} \tag{167}$$

one obtains by induction on $k \in [\![0, M+1]\!]$ that

$$\frac{\mathrm{d}^k}{\mathrm{d}\tau^k} \left[(\partial_p \Phi_0(t,p))^{-1} f_1 \left(\Phi_0(\tau, p) \right) \right] = (\partial_p \Phi_0(t,p))^{-1} \operatorname{ad}_{f_0}^k(f_1) \left(\Phi_0(\tau, p) \right). \tag{168}$$

The Taylor formula

$$\begin{aligned} (\partial_p \Phi_0(t,p))^{-1} f_1 \left(\Phi_0(t,p) \right) - \sum_{k=0}^{M-1} \frac{t^k}{k!} \operatorname{ad}_{f_0}^k(f_1)(p) \\ = \int_0^t \frac{(t-s)^{M-1}}{(M-1)!} (\partial_p \Phi_0(s,p))^{-1} \operatorname{ad}_{f_0}^M(f_1) \left(\Phi_0(s,p) \right) \mathrm{d}s \end{aligned} \tag{169}$$

proves the first statement.

2. Equation (169) yields the conclusion.

3. Let $r' \in [r/e, r)$. Thanks to (138),

$$\left\| \left\| \frac{t^k}{k!} \operatorname{ad}_{f_0}^k(f_1) \right\| \right\|_{r'} \leq \frac{|t|^k}{k!} \frac{k!}{e} \left(\frac{2e}{r - r'} \right)^k |||f_0|||_r^k \, |||f_1|||_r \,, \tag{170}$$

so the series converges absolutely in $\mathcal{C}^{\omega, r'}$ when $2e|t| \, |||f_0|||_r < r - r'$, which is the case when $6|t| \, |||f_0|||_r < r - r'$ because $2e < 6$. The weakest bound, for $r' = r/e$ is $2e|t| \, |||f_0|||_r < (1 - 1/e)r$ and it holds when $9|t| \, |||f_0|||_r < r$ because $2e/(1 - 1/e) < 9$.

Moreover, thanks to (169) and (170),

$$\begin{aligned} \left| (\Phi_0(-t)_* f_1)(p) - \sum_{k=0}^{M-1} \frac{t^k}{k!} \operatorname{ad}_{f_0}^k(f_1)(p) \right| &\leq \frac{|t|^M}{M!} \| \operatorname{ad}_{f_0}^M(f_1) \|_{\mathcal{C}^0} \sup_{s \in [0,t]} \| (\partial_p \Phi_0(s, \cdot))^{-1} \|_{\mathcal{C}^0} \\ &\leq A_0 \, |||f_1|||_r \left(\frac{2e|t| \, |||f_0|||_r}{r - r'} \right)^M, \end{aligned} \tag{171}$$

where A_0 denotes the supremum in the right-hand side of (171) which is finite. So the sum converges towards the pushforward under the same smallness assumption on time.

4. The last statement is proved similarly, by considering the function $t \mapsto e^{tH_0} H_1 e^{-tH_0}$. □

3.4.4 Partial derivative of a flow with respect to a parameter

In this paragraph, we compute the partial derivative of a flow with respect to a parameter on which the vector field depends, under a particular nilpotency assumption.

Lemma 91. *Let J an interval of $\mathbb{R}$. Let $\delta > 0$ and $f \in \mathcal{C}^\infty(J \times B_{4\delta}; \mathbb{K}^d)$ such that $\|f\|_{\mathcal{C}^0} < \delta$. Let $\lambda_0 \in J$, $M \in \mathbb{N}$ and assume that, for each $\lambda \in J$, $\mathrm{ad}^M_{f(\lambda_0)}(f(\lambda)) \equiv 0$. Then, for each $p \in B_\delta$,*

$$\frac{\mathrm{d}}{\mathrm{d}\lambda}\left(e^{f(\lambda)}p\right)_{|\lambda=\lambda_0} = \sum_{k=0}^{M-1} \frac{(-1)^k}{(k+1)!}\,\mathrm{ad}^k_{f(\lambda_0)}\left(\partial_\lambda f(\lambda_0)\right)\left(e^{f(\lambda_0)}p\right). \tag{172}$$

This holds in particular when $\mathcal{L}(f(J))$ is nilpotent with index at most $M+1$.

Proof. Let $\Theta \in \mathcal{C}^\infty([0,1] \times J \times B_\delta)$ defined by $\Theta(t,\lambda,p) := e^{tf(\lambda)}(p)$. Let $p_0 \in B_\delta$ and $\lambda_0 \in J$. Let $x_0(t) := e^{tf(\lambda_0)}(p_0)$ for $t \in [0,1]$. Then, the desired derivative is $\partial_\lambda\Theta(1,\lambda_0,p_0) = z(1)$ where z is the solution to $z(0) = 0$ and

$$\dot{z}(t) = \partial_x f(\lambda_0, x_0(t))z(t) + \partial_\lambda f(\lambda_0, x_0(t)). \tag{173}$$

Let $R : (t,s) \in [0,1]^2 \to \mathcal{M}_d(\mathbb{K})$ be the resolvent associated with the linearized system at p_0, which is the solution to $R(s,s) = \mathrm{Id}$ and

$$\partial_t R(t,s) = \partial_x f(\lambda_0, x_0(t))R(t,s), \tag{174}$$

i.e. $R(t,s) = \partial_p\Theta(t-s,\lambda_0,x_0(s))$. Then by Duhamel's principle

$$\begin{aligned} z(1) &= \int_0^1 R(\tau,1)^{-1}\partial_\lambda f(\lambda_0, x_0(\tau))\,\mathrm{d}\tau \\ &= \int_0^1 \left(\partial_p\Theta(\tau-1,\lambda_0,x_0(1))\right)^{-1}\partial_\lambda f(\lambda_0, \Theta(\tau-1,\lambda_0,x_0(1)))\,\mathrm{d}\tau. \end{aligned} \tag{175}$$

By (163) of Lemma 90 with $t \leftarrow \tau - 1$, $f_0 \leftarrow f(\lambda_0,\cdot)$, $f_1 \leftarrow \partial_\lambda f(\lambda_0,\cdot)$ and $p \leftarrow x_0(1)$,

$$z(1) = \int_0^1 \sum_{k=0}^{M-1} \frac{(\tau-1)^k}{k!}\,\mathrm{ad}^k_{f(\lambda_0)}\left(\partial_\lambda f(\lambda_0)\right)\left(x_0(1)\right)\mathrm{d}\tau, \tag{176}$$

which gives the conclusion. □

4 Error estimates in time for nonlinear vector fields

Using a classical linearization principle (see Section 4.1), we show that the formal expansions for linear equations of Section 2 can yield approximate formulas in the context of nonlinear ordinary differential equations. We derive rigorous error bounds at every fixed order with respect to time, involving finite sums or products.

4.1 Linearization principle for nonlinear vector fields

We explain how, by seeing vector fields as first-order differential operators and points on the manifold as the operator of evaluation at this point, one classically recasts a nonlinear ODE driven by smooth vector fields to a linear equation set on a larger space of operators on smooth functions. This approach is notably used in [3, 4] (replacing nonlinear objects by infinite-dimensional linear ones is the foundation of the "chronological calculus") and in [97]. More generally, the idea of replacing the study of a space by the study of the ring of functions on that space is reminiscent of the representation results of [50]. For readers with a background in PDE analysis, this linearization principle can be seen as a "reversed method of characteristics": it transforms a nonlinear ODE into a linear transport PDE, considered at the level of evolution operators.

4.1.1 Definition of an operator acting on smooth functions

When $T > 0$ and $f \in \mathcal{C}_c^\infty([0,T] \times \mathbb{K}^d; \mathbb{K}^d)$ satisfies $\|f\|_{L^1((0,T);\mathcal{C}^0)} < 1$, we take the nonlinear ODE (143) back to a linear framework by considering, for every $t \in [0,T]$ the linear operator $L(t)$ on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$ defined, for $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$, by

$$L(t)\varphi : p \mapsto \varphi\left(x(t;f,p)\right). \tag{177}$$

$L(t)\varphi$ is of class $\mathcal{C}^\infty$ as a composition of $\mathcal{C}^\infty$ functions, by the third statement of Lemma 82. $L(t)\varphi$ is compactly supported in $\mathbb{K}^d$ because φ is and $|x(t;f,p) - p| \leq 1$ for every $p \in \mathbb{K}^d$, by the first statement of Lemma 82 (which is of course invariant by translation of the origin). We don't specify the dependence of $L(t)$ with respect to f to simplify the notations.

For every $p \in \mathbb{K}^d$, the map $t \in [0,T] \mapsto \big(L(t)\varphi\big)(p)$ belongs to $\mathcal{C}^1([0,T];\mathbb{K})$ and satisfies, for every $t \in [0,T]$, using the notations of Remark 74,

$$\frac{\mathrm{d}}{\mathrm{d}t}\big(L(t)\varphi\big)(p) = D\varphi\Big(x(t;f,p)\Big)f\Big(t,x(t;f,p)\Big) = \big(L(t)f(t)\varphi\big)(p). \tag{178}$$

Thus, L satisfies the following linear equation

$$\frac{\mathrm{d}}{\mathrm{d}t}L(t) = L(t)f(t) \tag{179}$$

in the weak sense explicited above. For every fixed $t \in [0,T]$,

$$\forall \varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}), \forall p \in \mathbb{K}^d, \quad \big(L(t)\varphi\big)(p) = \varphi(p) + \int_0^t \big(L(\tau)f(\tau)\varphi\big)(p)\,\mathrm{d}\tau, \tag{180}$$

where the symbol $\int_0^t$ is the Lebesgue integral on $L^1((0,t);\mathbb{K})$. We will use the following notation to refer to this property:

$$L(t) = \mathrm{Id} + \int_0^t L(\tau)f(\tau)\,\mathrm{d}\tau. \tag{181}$$

In the sequel, all integral equalities between operators on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$ should be understood in this weak sense (after evaluation on a test function and at a point). The right-hand side refers to the composition of two operators on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$: $L(\tau)$ and $f(\tau)$, seen as a first-order differential operator on smooth functions.

Equation (179) is now a linear differential equation satisfied by the object $L(t)$ (in a much larger space), so one can hope to apply the linear results of the previous sections.

4.1.2 Approximating sequence

In order to approximate the operator $L(t)$, it is natural to introduce the sequence $(L_j)_{j\in\mathbb{N}}$ of time-dependent operators on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$ defined, for every $t \in [0,T]$, by $L_0(t) := \mathrm{Id}$ and, for $j \in \mathbb{N}$,

$$L_{j+1}(t) := \int_0^t L_j(\tau)f(\tau)\,\mathrm{d}\tau, \tag{182}$$

where this definition should be understood in the weak sense. Hence, recalling Remark 74,

$$L_j(t) = \int_{\Delta^j(t)} f(\tau_1)\cdots f(\tau_j)\,\mathrm{d}\tau, \tag{183}$$

where the integration domain is the ordered simplex of Definition 15. Then, for every $j \in \mathbb{N}$, L_j is "of order j with respect to f", and a differential operator of order at most j (with respect to x) on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$. And this sequence indeed allows to approximate $L(t)$ in the sense exposed in Proposition 92 below, in a finite regularity setting.

4.2 Chen-Fliess expansion

The approximating sequence for the operator $L(t)$ yields the following straight-forward estimate for the Chen-Fliess expansion of the state, which can also be found in [3, Proposition 2.3].

Proposition 92. *For every $M \in \mathbb{N}$, $\delta > 0$, $T > 0$, $f \in L^1((0,T);\mathcal{C}^{\max(M,1)}(B_{2\delta};\mathbb{K}^d))$, with $\|f\|_{L^1((0,T);\mathcal{C}^0)} < \delta$ and $\varphi \in \mathcal{C}^{M+1}(B_{2\delta};\mathbb{K})$, for each $t \in [0,T]$, with the notations of Remark 74,*

$$\left|\varphi\left(x(t;f,p)\right) - \sum_{j=0}^{M} \int_{\Delta^j(t)} \big(f(\tau_1)\cdots f(\tau_j)\varphi\big)(p)\,\mathrm{d}\tau\right| \leq (M+1)!\|f\|^{M+1}_{L^1((0,t);\mathcal{C}^M)}\|\varphi\|_{\mathcal{C}^{M+1}}. \tag{184}$$

In particular, for each $p \in B_\delta$,

$$\left|x(t;f,p) - \sum_{j=0}^{M} \int_{\Delta^j(t)} \big(f(\tau_1)\cdots f(\tau_j)\mathrm{Id}_d\big)(p)\,\mathrm{d}\tau\right| \leq (M+1)!\|f\|^{M+1}_{L^1((0,t);\mathcal{C}^M)}. \tag{185}$$

Hence, if $f \in L^\infty((0,T);\mathcal{C}^M)$, both estimates correspond to a bound scaling like t^{M+1}.

Proof. Let $p \in B_\delta$. Thanks to Lemma 82, $x(\tau;f,p)$ is well-defined for $\tau \in [0,T]$ and $x(\cdot;f,p) \in \mathcal{C}^1([0,T];\mathbb{K}^d)$. Thus, for each $\tau \in [0,T]$,

$$\varphi(x(\tau;f,p)) = \varphi(p) + \int_0^\tau \big(f(\tau_1)\varphi\big)(x(\tau_1;f,p))\,\mathrm{d}\tau_1. \tag{186}$$

By iterating this formula, we obtain for $t \in [0,T]$,

$$\begin{aligned}\varphi\left(x(t;f,p)\right) - \varphi(p) - \sum_{j=1}^{M} \int_{\Delta^j(t)} \big(f(\tau_1)\cdots f(\tau_j)\varphi\big)(p)\,\mathrm{d}\tau \\ = \int_{\Delta^{M+1}(t)} \big(f(\tau_1)\cdots f(\tau_{M+1})\varphi\big)\left(x(\tau_{M+1};f,p)\right)\,\mathrm{d}\tau,\end{aligned} \tag{187}$$

which concludes the proof of (184) using Lemma 75. Then (185) follows by applying (184) to coordinate functions. □

4.3 Magnus expansion in the usual setting

In Section 4.3.1, we state a precise estimate of the difference between the exact flow and the exponential of its truncated logarithm. In Section 4.3.2, we show that this estimate implies a similar estimate for the CBHD formula. Section 4.3.3 is devoted to a technical result used in the proof, which transposes to vector fields a formal integral identity.

4.3.1 Standard error estimate in time

The following estimate can be viewed as a refined version of classical time-focused estimates (see e.g. [87, Proposition 4.3]). It bears a lot of similarity with [38, Theorem 1.32], but is both easier to state and to prove in our flat setting since [38] is concerned with the truncated logarithm of flows in general Riemannian manifolds. We propose a proof for sake of completeness, and because this precise estimate is the founding principle of the new estimate, proved in the next section. It relies on the usual arguments, used for instance in [97] and [3, Proposition 4.1] (which states a slightly tighter estimate).

Proposition 93. *For every $M \in \mathbb{N}$, there exists $\delta_M, C_M > 0$ such that, for every $\delta > 0$, $T > 0$, $f \in L^1((0,T);\mathcal{C}^{\max(M^2,1)}(B_{2\delta};\mathbb{K}^d))$ with $\|f\|_{L^1((0,T);\mathcal{C}^{M^2})} \leq \delta_M \min\{1;\delta\}$, $p \in B_\delta$ and $t \in [0,T]$,*

$$\left|x(t;f,p) - e^{Z_M(t,f)}p\right| \leq C_M \|f\|^{M+1}_{L^1((0,t);\mathcal{C}^{M^2})}, \tag{188}$$

where $Z_M(t,f) := \mathrm{Log}_M\{f\}(t)$ is the vector field introduced in Definition 23.

Hence, if $f \in L^\infty((0,T);\mathcal{C}^{M^2}(B_{2\delta};\mathbb{K}^d))$ this estimate corresponds to a bound scaling like t^{M+1}.

Moreover, if $f(t,x) = \sum_{i\in I} u_i(t) f_i(x)$ with $u_i \in L^1((0,T);\mathbb{K})$ and $f_i \in \mathcal{C}^{M^2}(B_{2\delta};\mathbb{K}^d)$, then, for each monomial basis $\mathcal{B}$ of $\mathcal{L}(X)$,

$$Z_M(t,f) = \sum_{b\in\mathcal{B}_{[\![1,M]\!]}} \zeta_b(t,u) f_b \tag{189}$$

where the functionals ζ_b are the associated coordinates of the first kind and f_b are the evaluated Lie brackets (see Definitions 28, 30 and 78).

Proof. For $M = 0$, $Z_0(t,f) = 0$ thus (188) holds with $C_0 = 1$ because $|x(t;f,p) - p| \leq \|f\|_{L^1(\mathcal{C}^0)}$. From now on $M \in \mathbb{N}^*$ is fixed. By Definition 23, there exists $C'_M > 0$ such that, for every $\delta > 0$, $T > 0$, $f \in L^1((0,T);\mathcal{C}^{M-1})$ with $\|f\|_{L^1(\mathcal{C}^{M-1})} \leq 1$ and $t \in [0,T]$,

$$\|\mathrm{Log}_M\{f\}(t)\|_{\mathcal{C}^0} \leq C'_M \|f\|_{L^1(\mathcal{C}^{M-1})}. \tag{190}$$

In particular, for every $\delta > 0$, $T > 0$, $f \in L^1((0,T);\mathcal{C}^{M-1})$ with $\|f\|_{L^1(\mathcal{C}^{M-1})} \leq \min\{1;\delta;\delta/C'_M\}$, for every $p \in B_\delta$ and $t \in [0,T]$,

- $x(t;f,p)$ is well defined and belongs to $B_{2\delta}$,
- for every $s \in [0,1]$, $e^{s\mathrm{Log}_M\{f\}(t)}p$ is well defined belongs to $B_{2\delta}$.

This happens, in particular, when $\|f\|_{L^1(\mathcal{C}^{M-1})} \leq \delta_M \min\{1;\delta\}$ with $\delta_M := \min\{1; 1/C'_M\}$.

From now on, we fix $\delta, T > 0$ and $f \in L^1((0,T);\mathcal{C}^{M^2})$ with $\|f\|_{L^1(\mathcal{C}^{M^2})} \leq \delta_M \min\{1;\delta\}$.

In order to use the operators $L(t)$ defined in Section 4.1, we assume that $f \in \mathcal{C}^\infty_c([0,T]\times\mathbb{K}^d;\mathbb{K}^d)$. This is not restrictive because this space is dense in $L^1((0,T);\mathcal{C}^{M^2}(B_{2\delta};\mathbb{K}^d))$ and both sides of (188) are continuous for the $L^1((0,T);\mathcal{C}^{M^2})$ topology on f (see the fourth item of Lemma 82). Moreover, this regularization procedure is merely an heuristical convenience, since all the computations performed below make perfect sense even in our finite regularity setting.

Step 1: Construction of the formal logarithm. We introduce $Z_M(t,f)$ the finite sum of terms "of order at most M with respect to f" in the following formal power series (recall the formal power series for $\log(1+x)$):

$$\log L(t) = \sum_{m\in\mathbb{N}^*} \frac{(-1)^{m-1}}{m} \left(\sum_{j\in\mathbb{N}^*} L_j(t)\right)^m, \tag{191}$$

with the notation of (182), i.e. we define

$$Z_M(t,f) := \sum_{r=1}^{M}\sum_{m=1}^{M} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r}\in\mathbb{N}_r^m} L_{\mathbf{r}_m}(t)\cdots L_{\mathbf{r}_1}(t), \tag{192}$$

where $\mathbb{N}_r^m$ is defined in (31). For instance,

$$Z_3 = L_1 + \left(L_2 - \frac{1}{2}L_1^2\right) + \left(L_3 - \frac{1}{2}\left(L_1L_2 + L_2L_1\right) + \frac{1}{3}L_1^3\right). \tag{193}$$

Then, by (183),

$$Z_M(t,f) = \sum_{r=1}^{M} \sum_{m=1}^{M} \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \int_{\Delta^{\mathbf{r}}(t)} f(\tau_1) \cdots f(\tau_r) \, \mathrm{d}\tau, \tag{194}$$

A priori, $Z_M(t,f)$ is thus an inhomogeneous differential operator on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$, of order at most M. Using Lemma 96 (see below in the next paragraph) and Definition 23, $Z_M(t,f) = \mathrm{Log}_M\{f\}(t)$ and satisfies (189). Thus $Z_M(t,f)$ is a smooth vector field, i.e. both a vector-valued function and a first-order differential operator.

Step 2: Strategy for the proof of the estimate. The key observation is that it is sufficient to prove that there exists $C_M > 0$ (independent of δ, T, f) such that, for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$,

$$\left| \left(L(t) - e^{Z_M(t,f)} \right)(\varphi)(p) \right| \leq C_M \|f\|_{L^1(\mathcal{C}^{M^2})}^{M+1} \|\varphi\|_{\mathcal{C}^{M^2+1}}. \tag{195}$$

Then, the conclusion follows by considering an appropriate $\mathcal{C}_c^\infty$ truncation of the coordinate functions $\varphi_j : x \in \mathbb{K}^d \mapsto x_j \in \mathbb{K}$. To prove (195), we will decompose the difference in three terms

$$L - e^{Z_M} = \left(L - \sum_{j=0}^{M} L_j \right) + \left(\sum_{j=0}^{M} L_j - \sum_{k=0}^{M} \frac{Z_M^k}{k!} \right) + \left(\sum_{k=0}^{M} \frac{Z_M^k}{k!} - e^{Z_M} \right), \tag{196}$$

with the notation of (182). The first term is estimated in Proposition 92.

Step 3: Bound for $\sum L_j - \sum \frac{Z_M^k}{k!}$. By (192), this operator is a (finite) linear combination of terms of the form $L_{j_1}(t) \cdots L_{j_p}(t)$ where $p \in \mathbb{N}^*$, $j_1, \ldots, j_p \in [\![1, M]\!]$ and $M+1 \leq j_1 + \ldots + j_p \leq M^2$. Indeed, $Z_M(t,f)$ is also the finite sum of terms "of order at most M with respect to f" in the formal power series (191). Thus, there exists $C_M'' > 0$ (independent of δ, T, f) such that, for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$,

$$\left| \left(\sum_{j=0}^{M} L_j(t) - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!} \right)(\varphi)(p) \right| \leq C_M'' \|f\|_{L^1(\mathcal{C}^{M^2-1})}^{M+1} \|\varphi\|_{\mathcal{C}^{M^2}}. \tag{197}$$

Step 4: Bound for $\sum \frac{Z_M^k}{k!} - e^{Z_M}$. Using Lemma 86 for the time-independent vector field $Z_M(t,f)$ (where $t \in [0,T]$ has been fixed), estimate (153) yields for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$,

$$\left| \left(e^{Z_M(t,f)} - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!} \right)(\varphi)(p) \right| \leq \|Z_M(t,f)\|_{\mathcal{C}^M}^{M+1} \|\varphi\|_{\mathcal{C}^{M+1}}. \tag{198}$$

We deduce from (36) the existence of $C_M''' > 0$ (independent of δ, T, f) such that for every $t \in [0,T]$

$$\|Z_M(t,f)\|_{\mathcal{C}^M} \leq C_M''' \|f\|_{L^1((0,t);\mathcal{C}^{2M-1})}. \tag{199}$$

Hence, for every $p \in B_\delta$, $t \in [0,T]$ and $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$

$$\left| \left(e^{Z_M(t,f)} - \sum_{k=0}^{M} \frac{Z_M(t,f)^k}{k!} \right)(\varphi)(p) \right| \leq (C_M''')^{M+1} \|f\|_{L^1(\mathcal{C}^{2M-1})}^{M+1} \|\varphi\|_{\mathcal{C}^{M+1}}. \tag{200}$$

Gathering (184), (197) and (200) concludes the proof of (188). □

4.3.2 Campbell Baker Hausdorff Dynkin formula

We deduce from Proposition 93 the following estimate for the classical CBHD formula with q time-independent vector fields.

Corollary 94. *For every $M \in \mathbb{N}^*$, there exists $\delta_M, C_M > 0$ such that, for every $\delta > 0$, $q \in \mathbb{N}^*$, $f_1, \dots, f_q \in \mathcal{C}^{M^2}(B_{2\delta}; \mathbb{K}^d)$ with $\sum_{1 \le j \le q} \|f_j\|_{\mathcal{C}^{M^2}} \le \delta_M \min\{1; \delta\}$,*

$$\left\| e^{f_q} \cdots e^{f_1} - e^{\mathrm{CBHD}_M(f_1, \dots, f_q)} \right\|_{\mathcal{C}^0} \le C_M \|f\|^{M+1} \tag{201}$$

where $\mathrm{CBHD}_M(f_1, \dots, f_q) = \mathrm{Log}_M\{f\}(q)$, where the time-dependent vector field f is defined by $f : (t, x) \in [0, q] \times B_{2\delta} \mapsto \sum_{j=1}^{q} 1_{[j-1,j]}(t) f_j(x)$ and $\|f\| := \|f\|_{L^1(\mathcal{C}^{M^2})} = \sum_{1 \le j \le q} \|f_j\|_{\mathcal{C}^{M^2}}$.

Moreover, for each monomial basis $\mathcal{B}$ of $\mathcal{L}(\{X_1, \dots, X_q\})$

$$\mathrm{CBHD}_M(f_1, \dots, f_q) = \sum_{b \in \mathcal{B}_{[\![1, M]\!]}} \alpha_b f_b \tag{202}$$

where $(\alpha_b)_{b \in \mathcal{B}} \subset \mathbb{K}^{\mathcal{B}}$ is given by Corollary 33.

Proof. Because of the particular form of f, we have $x(t; f, p) = e^{f_q} \cdots e^{f_1} p$. Thus the estimate (201) is an application of Proposition 93. Let $\Lambda : \mathcal{L}(\{X_1, \dots, X_q\}) \to \mathcal{L}(\{f_1, \dots, f_q\})$ be the homomorphism of Lie algebras such that $\Lambda(X_j) = f_j$. The map CBHD_M is defined by a finite sum of Lie brackets, thus it commutes with Λ

$$\mathrm{CBHD}_M(f_1, \dots, f_q) = \Lambda(\mathrm{CBHD}_M(X_1, \dots, X_q)) = \Lambda\left(\sum_{b \in \mathcal{B}_{[\![1, M]\!]}} \alpha_b b \right) = \sum_{b \in \mathcal{B}_{[\![1, M]\!]}} \alpha_b \Lambda(b), \tag{203}$$

which proves (202). □

4.3.3 Replacing products with brackets in logarithm integrals

The goal of this section is to prove Lemma 96, which is a key point in the proof of Proposition 93, as it allows to replace products of differential operators with Lie brackets in the integrals involved in the computation of the logarithm of the flow.

We first state and prove a corollary of Theorem 27 in algebras. Indeed, Theorem 27 is a statement about formal differential equations, but it has consequences for concrete realizations, e.g. for systems governed by vector fields or matrices (this will be used in Section 5.2.2).

Corollary 95. *Let A be a unital associative algebra over $\mathbb{K}$ and A_1 be a finite dimensional linear subspace of A. Then, for every $r \in \mathbb{N}^*$, $t > 0$ and $\mathrm{a} \in L^1((0, t); A_1)$, one has*

$$\begin{aligned} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} \mathrm{a}(\tau_1)\mathrm{a}(\tau_2) \cdots \mathrm{a}(\tau_r)\, \mathrm{d}\tau = \\ \frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [\mathrm{a}(\tau_1), \mathrm{a}(\tau_2)], \dots \mathrm{a}(\tau_r)]\, \mathrm{d}\tau, \end{aligned} \tag{204}$$

where the equality should be seen as an equality between elements of a finite dimensional linear subspace of A (generated by monomials of terms in A_1 of degree r), so that one can give a meaning to the integrals without introducing any topology on A.

Moreover, if $\mathrm{a}(\tau) = \sum_{i \in I} \alpha_i(\tau) \mathrm{y}_i$ with $\alpha_i \in L^1((0, t); \mathbb{K})$ and $\mathrm{y}_i \in \mathrm{A}$ then, for each monomial basis $\mathcal{B}_r$ of $\mathcal{L}_r(X)$,

$$\frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [\mathrm{a}(\tau_1), \mathrm{a}(\tau_2)], \dots \mathrm{a}(\tau_r)]\, \mathrm{d}\tau = \sum_{b \in \mathcal{B}_r} \zeta_b(t, \alpha) \mathrm{y}_b, \tag{205}$$

where the functionals ζ_b are the associated coordinates of the first kind and $\mathrm{y}_b = \Upsilon(b)$ where $\Upsilon : \mathcal{A}(X) \to A$ is the homomorphism of algebras such that $\Upsilon(X_i) = \mathrm{y}_i$ (see Definition 30 and Lemma 7).

Proof. Let $q \in \mathbb{N}^*$ be the dimension of A_1 (as a linear subspace) and $\mathrm{y}_1, \dots \mathrm{y}_q$ be a linear basis of A_1. Let $\alpha_i \in L^1((0,t);\mathbb{K})$ denote the components of $\mathrm{a}(\cdot)$ in the basis $\mathrm{y}_1, \dots \mathrm{y}_q$, i.e. $\mathrm{a}(\tau) = \alpha_1(\tau)\mathrm{y}_1 + \ldots + \alpha_q(\tau)\mathrm{y}_q$ for almost every $\tau \in [0,t]$. Then $\mathrm{a}(t) = \Upsilon(a(t))$ where $a(\tau) := \alpha_1(\tau)X_1 + \ldots + \alpha_q(\tau)X_q \in \mathcal{A}_1(X)$. From (40) and (41), one obtains that (204) holds for $a(\cdot)$. Applying the homomorphism Υ of algebras to both sides proves (204) for $\mathrm{a}(\cdot)$. The same strategy proves (205). □

Lemma 96. *For every $r \in \mathbb{N}^*$, $t > 0$ and $f \in \mathcal{C}_c^\infty([0,t] \times \mathbb{K}^d; \mathbb{K}^d)$,*

$$\sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} f(\tau_1) f(\tau_2) \cdots f(\tau_r) \, \mathrm{d}\tau = \frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [f(\tau_1), f(\tau_2)], \dots f(\tau_r)] \, \mathrm{d}\tau, \tag{206}$$

which should be seen as an equality between linear operators on $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$, hence only valid after evaluation at a function φ at a point p, so that the integrals are integrals of scalar-valued functions.

Moreover, if $f(\tau, x) = \sum_{i \in I} u_i(\tau) f_i(x)$ with $u_i \in L^1((0,t);\mathbb{K})$ and $f_i \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}^d)$ then

$$\frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [f(\tau_1), f(\tau_2)], \dots f(\tau_r)] \, \mathrm{d}\tau = \sum_{b \in \mathcal{B}_r} \zeta_b(t,u) f_b, \tag{207}$$

where $\mathcal{B}_r$ is a monomial basis of $\mathcal{L}_r(X)$, the functionals ζ_b are the associated coordinates of the first kind and f_b are the evaluated Lie brackets (see Definitions 4, 30 and 78).

Proof. Let $(f_n)_{n \in \mathbb{N}^*}$ be a sequence of functions in $\mathcal{C}_c^\infty([0,t] \times \mathbb{K}^d; \mathbb{K}^d)$ such that f_n takes values in an at-most n-dimensional vector subspace E_n of $\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}^d)$ and $\|f_n - f\|_{L^1((0,t);\mathcal{C}^r)} \to 0$ when $n \to \infty$. For example, one can choose an n-points trapezoidal approximation of f. For each fixed n, applying Corollary 95 with $A = \mathrm{Op}(\mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K}))$ and $A_1 = E_n$ proves (206) for f_n. Let $\varphi \in \mathcal{C}_c^\infty(\mathbb{K}^d;\mathbb{K})$ and $p \in \mathbb{K}^d$. For each $n \in \mathbb{N}^*$, we deduce that

$$\sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} (f_n(\tau_1) \cdots f_n(\tau_r)\varphi)(p) \, \mathrm{d}\tau = \frac{1}{r} \sum_{m=1}^{r} \sum_{\mathbf{r} \in \mathbb{N}_r^m} \frac{(-1)^{m-1}}{m} \int_{\Delta^{\mathbf{r}}(t)} ([\cdots [f_n(\tau_1), f_n(\tau_2)], \dots f_n(\tau_r)]\varphi)(p) \, \mathrm{d}\tau. \tag{208}$$

For each fixed φ and p, both sides converge as $n \to +\infty$ towards the same quantities for f. This proves that (206) holds as an equality between linear operators. Applying (205) gives (207). □

Remark 97. *Although most algebraic results of Section 2 remain valid for infinite alphabets (sets of indeterminates), there is a difficulty when one wishes to "evaluate" equalities in the free algebra over an infinite alphabet towards some target algebra (one must somehow introduce compatible topologies on both sides). Our approach to prove Lemma 96, where f is allowed to take values in the infinite-dimensional space $\mathcal{C}_c^\infty$, therefore relies on a discretization scheme to return to a finite alphabet, and the convergence of the involved integrals in a weak sense. Another approach, followed in [91, 92], consists in introducing definitions allowing an infinite (continuous) number of generators and proving analogous algebraic results in such a setting.*

4.4 Magnus expansion in the interaction picture

In this section, we consider the nonlinear ordinary differential equation

$$\dot{x}(t) = f_0(x) + f_\sharp(t,x) \tag{209}$$

We show how the formal expansion introduced in Section 2.4 allows to obtain error bounds at every order in the size of the time-varying perturbation $f_\sharp$, provided that the flow of f_0 is known. Such estimates can be useful for example to design splitting methods in the case of a small perturbation (see e.g. [20, Section 3.6] or [21, Section 2]).

The results of this section can be seen as quite natural, but we are not aware of references containing the same statements. We adopt the notations specified in the following definition.

Definition 98. *Let $M \in \mathbb{N}$, $\delta > 0$, $T > 0$ and $f_0 \in \mathcal{C}^{M^2+1}(B_{5\delta};\mathbb{K}^d)$ such that $T\|f_0\|_{\mathcal{C}^0} < \delta$. Let $f_\sharp \in L^1((0,T);\mathcal{C}^{M^2}(B_{5\delta};\mathbb{K}^d))$ and $t \in [0,T]$. We consider*

- *$\Phi_0 \in \mathcal{C}^{M^2+1}([0,T]\times B_{4\delta};B_{5\delta})$ the flow associated with f_0 i.e. $\Phi_0(\tau;p) = e^{\tau f_0}(p)$,*
- *$g_t \in L^1((0,T);\mathcal{C}^{M^2}(B_{4\delta};\mathbb{K}^d))$ defined by*

$$g_t(\tau,y) := (\Phi_0(t-\tau)_* f_\sharp(\tau))(y) = (\partial_p \Phi_0(\tau - t, y))^{-1} f_\sharp\big(\tau, \Phi_0(\tau-t,y)\big), \tag{210}$$

- *$\mathcal{Z}_M(t,f_0,f_\sharp) := \mathrm{Log}_M\{g_t\}(t) \in \mathcal{C}^{M^2-M+1}(B_{4\delta};\mathbb{K}^d)$ in the sense of Definition 23.*

4.4.1 Error bound

Proposition 99. *Let $M,\delta,T,f_0,f_\sharp$ as in Definition 98. There exists $\gamma = \gamma(M,\delta,\|f_0\|_{\mathcal{C}^{M^2+1}}) > 0$ such that, if*

$$\|f_\sharp\|_{L^1((0,T);\mathcal{C}^{M^2})} \le \gamma \tag{211}$$

then, for every $p \in B_\delta$ and $t \in [0,\gamma]$,

$$\left| x(t; f_0+f_\sharp, p) - e^{\mathcal{Z}_M(t,f_0,f_\sharp)} e^{tf_0} p \right| \le C_M \|g_t\|^{M+1}_{L^1((0,t);\mathcal{C}^{M^2})} \tag{212}$$

where $C_M > 0$ is the constant of Proposition 93.

Hence, if $f_\sharp \in L^\infty((0,T);\mathcal{C}^{M^2}(B_{5\delta};\mathbb{K}^d))$, estimate (212) scales like t^{M+1}.

Proof. Let $\eta_M, C_M > 0$ be as in Proposition 93. There exists $T^* = T^*(\delta, \|f_0\|_{\mathcal{C}^{M^2+1}}) \in (0,T]$ such that, for every $f_\sharp \in L^1((0,T);\mathcal{C}^{M^2}(B_{5\delta};\mathbb{K}^d))$ and $t \in [0,T^*]$

$$\|g_t\|_{L^1((0,t),\mathcal{C}^{M^2})} \le 2\|f_\sharp\|_{L^1((0,t),\mathcal{C}^{M^2})}. \tag{213}$$

Let $\gamma := \min\{T^*, \delta, \frac{\delta_M}{2}\min\{1,\delta\}\}$. Let $f_\sharp \in L^1((0,T);\mathcal{C}^{M^2}(B_{5\delta};\mathbb{K}^d))$ with $\|f_\sharp\|_{L^1(\mathcal{C}^{M^2})} < \gamma$. Then, for every $p \in B_\delta$ and $\tau \in [0,T]$, $x(\tau; f_0+f_\sharp, p)$ is well defined and belongs to $B_{3\delta}$. To simplify the notations in this proof, we write $x(\tau)$ instead of $x(\tau; f_0+f_\sharp, p)$. Let $t \in [0,\gamma]$. The function $y : [0,t] \to \mathbb{K}^d$ defined by

$$y(\tau) := \Phi_0\big(t-\tau; x(\tau)\big) \tag{214}$$

takes values in $B_{4\delta}$ and satisfies, for every $\tau \in [0,t]$,

$$\dot{y}(\tau) = g_t\left(\tau, y(\tau)\right). \tag{215}$$

By (213), $\|g_t\|_{L^1(\mathcal{C}^{M^2})} < 2\eta \le \delta_M \min\{1,\delta\}$ thus, by Proposition 93

$$|y(t) - e^{\mathcal{Z}_M(t,f_0,f_\sharp)} y(0)| \le C_M \|g_t\|^{M+1}_{L^1((0,t);\mathcal{C}^{M^2})} \tag{216}$$

which is exactly (212) because $y(t) = x(t)$ and $y(0) = e^{tf_0}p$. □

4.4.2 Expansions of $\mathcal{Z}_M$

Proposition 100. *Let $M, \delta, T, f_0, f_\sharp$ as in Definition 98. Let $r > 0$. If $f_0 \in \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$ and $f_\sharp \in \mathcal{C}^0([0,T]; \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d))$ then, for $0 \leq \tau \leq t \leq \min\{T; \frac{r}{9|||f_0|||_r}\}$*

$$g_t(\tau, \cdot) = e^{(\tau - t)\,\mathrm{ad}_{f_0}}(f_\sharp(\tau)) = \sum_{k=0}^{+\infty} \frac{(\tau - t)^k}{k!}\,\mathrm{ad}^k_{f_0}(f_\sharp(\tau)) \tag{217}$$

and

$$\begin{aligned}\mathcal{Z}_M(t, f_0, f_\sharp) = \sum \frac{(-1)^{m-1}}{rm} \int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1 - t)^{k_1}}{k_1!} \cdots \frac{(\tau_r - t)^{k_r}}{k_r!} \\ \left[\cdots \left[\mathrm{ad}^{k_1}_{f_0}(f_\sharp(\tau_1)), \mathrm{ad}^{k_2}_{f_0}(f_\sharp(\tau_2))\right], \ldots, \mathrm{ad}^{k_r}_{f_0}(f_\sharp(\tau_r))\right] \mathrm{d}\tau,\end{aligned} \tag{218}$$

where the sum is taken over $r \in [\![1, M]\!]$, $m \in [\![1, r]\!]$, $\mathbf{r} \in \mathbb{N}^m_r$, and $k_1, \ldots, k_r \in \mathbb{N}$. Moreover, for every $r' \in [r/e, r)$ and $0 \leq \tau \leq t \leq \min\{T; \frac{r-r'}{6|||f_0|||_r}\}$, the series (217) and (218) converge absolutely in $\mathcal{C}^{\omega,r'}(B_{5\delta}; \mathbb{K}^d)$.

Proof. We apply the third statement of Lemma 90 to f_0 and $f_\sharp(\tau)$ to get (217). The absolute convergence in this series allows to interchange the sums and the integrals. □

When the perturbation $f_\sharp(t, x)$ is affine, i.e. of the form $\sum_{i=1}^q u_i(t) f_i(x)$, by analogy with Theorem 41, we use the notation $\mathcal{Z}_M(t, f, u)$ instead of $\mathcal{Z}_M\left(t, f_0, \sum_{i=1}^q u_i f_i\right)$, with $f = (f_0, f_1, \ldots, f_q)$ and $u = (u_1, \ldots, u_q)$. In this context, we have the following result, that emphasizes that $\mathcal{Z}_M$ is a truncated version of $\mathcal{Z}_\infty$.

Proposition 101. *Let $M, \delta, T, f_0, f_\sharp$ as in Definition 98. Let $r > 0$. If $f_0 \in \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$ and $f_\sharp(t, x) = \sum_{i=1}^q u_i(t) f_i(x)$ where $u_i \in L^1(0, T)$ and $f_i \in \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$. Then*

$$\mathcal{Z}_M(t, f, u) = \lim_{N \to \infty} \sum_{\substack{b \in \mathcal{B} \\ n(b) \leq M \\ n_0(b) \leq N}} \eta_b(t, u) f_b \tag{219}$$

where, for every $r' \in [r/e, r]$ the limit holds in $\mathcal{C}^{\omega,r'}(B_{5\delta}; \mathbb{K}^d)$ when $0 \leq t \leq \min\{T; \frac{r-r'}{6|||f_0|||_r}\}$.

Proof. Let $X = \{X_0, X_1, \ldots X_q\}$ and $\Lambda : \mathcal{L}(X) \to \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$ be the homomorphism of Lie algebras such that $\Lambda(X_i) = f_i$ for $i \in [\![0, q]\!]$ (see Lemma 7). By applying Λ to each term in the equality (84) (where $\mathcal{Z}^{r,\nu}_\infty(t, X, a)$ is the finite sum defined in (82)), we obtain for every $r \in \mathbb{N}^*$ and $\nu \in \mathbb{N}$

$$\mathcal{Z}^{r,\nu}_\infty(t, f, u) = \sum_{b \in \mathcal{B}_{r,\nu}} \eta_b(t, u) f_b. \tag{220}$$

By Proposition 100

$$\mathcal{Z}_M(t, f, u) = \lim_{N \to \infty} \sum_{\nu=0}^{N} \sum_{r=1}^{M} \mathcal{Z}^{r,\nu}_\infty(t, f, u) \tag{221}$$

where for every $r' \in [r/e, r]$ the limit holds in $\mathcal{C}^{\omega,r'}(B_{5\delta}; \mathbb{K}^d)$ when $0 \leq t \leq \min\{T; \frac{r-r'}{6|||f_0|||_r}\}$. This proves (219). □

Remark 102. *Although the family $\eta_b(t, u) f_b$ for $b \in \mathcal{B} \cap S_M = \{b \in \mathcal{B}; n(b) \leq M\}$ (using Definition 10) is not proved to be absolutely summable, equality (219) gives a sense to the expression*

$$\mathcal{Z}_M(t, f, u) = \sum_{b \in \mathcal{B} \cap S_M} \eta_b(t, u) f_b. \tag{222}$$

Indeed, the proof above justifies the absolute summability of appropriate packages $\mathcal{Z}^{r,\nu}_\infty(t, f, u)$ for $r \in [\![1, M]\!]$ and $\nu \in \mathbb{N}$ of this family. The full absolute summability over $\mathcal{B} \cap S_M$ is investigated in the next subsection.

4.4.3 Absolute convergence for coordinates of the pseudo-first kind

Continuing the discussion started in Section 2.4.3 we state a criterion on the basis $\mathcal{B}$ which entails the absolute summability for analytic vector fields of the family $\eta_b(t,u)f_b$ for $b \in \mathcal{B} \cap S_M = \{b \in \mathcal{B}; n(b) \leq M\}$ (using Definition 10).

Proposition 103. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$ and $\mathcal{B}$ a Hall basis of $\mathcal{L}(X)$; or more generally a monomial basis of $\mathcal{L}(X)$ with geometric growth with respect to X_0 (see Definition 48 and Remark 50).*

Let $M, \delta, T, f_0, f_\sharp$ as in Definition 98. Let $r > 0$. We assume $f_0 \in \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$ and $f_\sharp(t,x) = \sum_{i=1}^q u_i(t) f_i(x)$ where $u_i \in L^1(0,T)$ and $f_i \in \mathcal{C}^{\omega,r}(B_{5\delta}; \mathbb{K}^d)$.

Let $r' \in [r/e, r)$. There exists $T^* = T^*(M, q, r, r', |||f_0|||_r) > 0$ such that, for every $t \in (0, T^*)$ and $u \in L^1((0,t), \mathbb{K}^q)$

$$\mathcal{Z}_M(t, f, u) = \sum_{b \in \mathcal{B} \cap S_M} \eta_b(t,u) f_b \tag{223}$$

where the series converges absolutely in $\mathcal{C}^{\omega,r'}(B_\delta; \mathbb{K}^d)$.

Proof. By (92) of Proposition 52 and (138), for every $b \in \mathcal{B} \cap S_M$ and $t \in [0,T]$

$$|\eta_b(t,u)| \, |||f_b|||_{r'} \leq \frac{r-r'}{2e^2} \left(\frac{2eC_M t |||f_0|||_r}{r-r'} \right)^{n_0(b)} \left(\frac{2eC_M}{r-r'} \|u\|_{L^1(0,t)} |||f|||_r \right)^{n(b)} \tag{224}$$

where $|||f|||_r := \max\{|||f_j|||_r \, ; j \in [\![0,q]\!]\}$. In particular, if $|t| < T^*(M,r,r') := \frac{r-r'}{4(q+1)eC_M |||f_0|||_r}$ then the series $\sum \eta_b(t,a) f_b$ converges absolutely in $\mathcal{C}^{\omega,r'}$ because

$$\sum_{b \in \mathcal{B} \cap S_M} (2(q+1))^{-n_0(b)} \leq \sum_{n=1}^{M} \sum_{n_0=0}^{+\infty} (q+1)^{n_0+n} (2(q+1))^{-n_0} \leq M(q+1)^M. \tag{225}$$

□

4.5 Sussmann's infinite product expansion

Let $T > 0$. In this section, we consider affine systems of the form

$$\dot{x}(t) = \sum_{i \in I} u_i(t) f_i(x(t)) \quad \text{and} \quad x(0) = p, \tag{226}$$

where, for $i \in I$, f_i is a vector field and $u_i \in L^1((0,T); \mathbb{K})$. When well-defined, its solution is denoted $x(t; f, u, p)$. For every norm $\|\cdot\|$ on vector fields, $\|f\|$ denotes $\sum_{i \in I} \|f_i\|$.

Proposition 104. *Let $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$ and $(\xi_b)_{b \in \mathcal{B}}$ be the associated coordinates of the second kind. For every $M \in \mathbb{N}^*$, there exist $C_M, \eta_M > 0$ such that the following property holds. Let $T, \delta > 0$, $f_i \in \mathcal{C}^{2M}(B_{3\delta}; \mathbb{K}^d)$ and $u_i \in L^1((0,T); \mathbb{K})$ for $i \in I$. Assume that*

$$\|u\|_{L^1(0,T)} \|f\|_{\mathcal{C}^M} \leq \eta_M \min\{1, \delta\}. \tag{227}$$

Then, for each $t \in [0,T]$ and $p \in B_\delta$,

$$\left| x(t; f, u, p) - \overrightarrow{\prod_{b \in \mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t,u) f_b} p \right| \leq C_M \|u\|_{L^1(0,t)}^{M+1} \|f\|_{\mathcal{C}^{2M}}^{M+1} \left(1 + \|f\|_{\mathcal{C}^{2M}}^{M-1}\right), \tag{228}$$

where the arrow above the product symbol designates the order for the product, i.e. with the notations of Definition 53

$$\overrightarrow{\prod_{b \in \mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t,u) f_b} = e^{\xi_{b_1}(t,u) f_{b_1}} \cdots e^{\xi_{b_{k+1}}(t,u) f_{b_{k+1}}}. \tag{229}$$

Proof. Let $M \in \mathbb{N}^*$. We adopt the notations $b_1, \dots, b_{k+1}$ and $Y_0, \dots, Y_{k+1}$ of Definition 53. For $j \in [\![1, k+1]\!]$, we denote by Φ_j the flow associated with f_{b_j}, i.e. $\Phi_j(t,p) := e^{tf_{b_j}}(p)$. To simplify the notations in this proof, we write $x(t)$ and $\xi_b(t)$ instead of $x(t; f, u, p)$ and $\xi_b(t, u)$. Let $\eta_M := 1/(4|I|M!)$. For brevity, we use the shorthand notation $F := \|f\|_{\mathcal{C}^{2M-1}}$.

Step 1: Well-definition of the flows. Using (227),

$$\left\| \sum_{i \in I} u_i f_i \right\|_{L^1((0,T);\mathcal{C}^0)} \leq \eta_M \min\{1, \delta\} \leq \delta. \tag{230}$$

Thus, for $t \in [0,T]$, $x(t)$ is well-defined and $x(t) \in B_{2\delta}$. For $b \in \mathcal{B}$, using (107) and Lemma 79, we obtain, for each $t \in [0,T]$,

$$\|\xi_b(t) f_b\|_{\mathcal{C}^1} \leq \ell! 2^{\ell-1} \|u\|_{L^1(0,t)}^{\ell} \|f\|_{\mathcal{C}^\ell}^{\ell}, \tag{231}$$

where $\ell := |b|$. Hence, using the crude estimate $|\mathcal{B}_\ell| \leq |I|^\ell$, we obtain, for each $t \in [0,T]$,

$$\begin{aligned} \sum_{b \in \mathcal{B}_{[\![1,M]\!]}} \|\xi_b(t) f_b\|_{\mathcal{C}^1} &\leq \sum_{\ell=1}^{M} |\mathcal{B}_\ell| \ell! 2^{\ell-1} \|u\|_{L^1}^{\ell} \|f\|_{\mathcal{C}^\ell}^{\ell} \\ &\leq M! \sum_{\ell=1}^{+\infty} \left(2|I| \|u\|_{L^1} \|f\|_{\mathcal{C}^\ell}\right)^{\ell} \\ &\leq \frac{2M!|I|\|u\|_{L^1}\|f\|_{\mathcal{C}^M}}{1 - 2|I|\|u\|_{L^1}\|f\|_{\mathcal{C}^M}} \leq \min\{1, \delta\}. \end{aligned} \tag{232}$$

Thus, for every $j \in [\![1, k+1]\!]$,

$$x_j(t) := e^{-\xi_{b_j}(t) f_{b_j}} \cdots e^{-\xi_{b_1}(t) f_{b_1}}(x(t)) \tag{233}$$

is well-defined and belongs to $B_{3\delta}$.

Step 2: Estimates along a Lazard elimination. We prove by induction on $j \in [\![0, k+1]\!]$ the existence of a numerical constant $C_j > 0$ such that

$$(\mathcal{H}_j) : \begin{cases} \dot{x}_j(t) = \sum_{b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_j} \dot{\xi}_b(t) f_b(x_j(t)) + \varepsilon_j(t), \\ x_j(0) = p, \end{cases} \tag{234}$$

where

$$|\varepsilon_j(t)| \leq C_j |u(t)| \|u\|_{L^1}^{M} F^{M+1}(1 + F^{M-1}). \tag{235}$$

First, letting $x_0(t) := x(t)$ by convention, $(\mathcal{H}_0)$ holds with $\varepsilon_0 = 0$, $C_0 = 0$ because $\dot{\xi}_{X_i}(t) = u_i(t)$ for $i \in I$. Let $j \in [\![1, k+1]\!]$ and assume that $(\mathcal{H}_{j-1})$ holds. We deduce from the definition of $x_j(t)$ that

$$x_j(t) = e^{-\xi_{b_j}(t) f_{b_j}}(x_{j-1}(t)) = \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) \tag{236}$$

and, using $(\mathcal{H}_{j-1})$, that

$$\dot{x}_j(t) = -\dot{\xi}_{b_j}(t) f_{b_j}(x_j(t)) + \sum_{b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_{j-1}} \dot{\xi}_b(t) \partial_p \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) f_b(x_{j-1}(t)) + \widetilde{\varepsilon}_j(t) \tag{237}$$

where $\widetilde{\varepsilon}_{j-1}(t) := \partial_p \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) \varepsilon_{j-1}(t)$. By (232), $\|\xi_{b_j}(t) f_{b_j}\|_{\mathcal{C}^1} \leq 1$, so, using (150),

$$|\widetilde{\varepsilon}_{j-1}(t)| \leq e |\varepsilon_{j-1}(t)|. \tag{238}$$

Moreover, for each $b \in \mathcal{B}$

$$\partial_p \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) f_b(x_{j-1}(t)) = \left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)), \tag{239}$$

thus,

$$\dot{x}_j(t) = \sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}\setminus\{b_j\}} \dot{\xi}_b(t)\left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) + \widetilde{\varepsilon}_j(t). \tag{240}$$

For $b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_j\setminus\{b_j\}$, we introduce the maximal integer $h(b)\in\mathbb{N}^*$ such that

$$|b| + (h(b)-1)|b_j| \leq M. \tag{241}$$

Then, by the first statement of Lemma 90 and Definition 62

$$\begin{aligned}\dot{\xi}_b(t)\left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) &= \sum_{m=1}^{h(b)-1} \frac{\xi_{b_j}^m(t)}{m!}\dot{\xi}_b(t) f_{\mathrm{ad}_{b_j}^m(b)}(x_j(t)) + \overline{\varepsilon}_b^j(t)\\ &= \sum_{m=1}^{h(b)-1} \dot{\xi}_{\mathrm{ad}_{b_j}^m(b)}(t) f_{\mathrm{ad}_{b_j}^m(b)}(x_j(t)) + \overline{\varepsilon}_b^j(t)\end{aligned} \tag{242}$$

where

$$|\overline{\varepsilon}_b^j(t)| \leq |\dot{\xi}_b(t)|\frac{|\xi_{b_j}(t)|^{h(b)}}{h(b)!}\|f_{\mathrm{ad}_{b_j}^{h(b)}(b)}\|_{C^0}. \tag{243}$$

By definition of $h(b)$ we have $M+1\leq |b|+h(b)|b_j|\leq M+|b_j|\leq 2M$. Thus, using Lemma 79, (106) and (107), we get

$$\begin{aligned}|\overline{\varepsilon}_b^j(t)| &\leq |u(t)|\|u\|_{L^1}^{|b|+h(b)|b_j|-1}\frac{|b|}{h(b)!}2^{2M}(2M-1)!F^{M+1}(1+F^{M-1})\\ &\leq |u(t)|\|u\|_{L^1}^{M} M 2^{2M}(2M-1)!F^{M+1}(1+F^{M-1}).\end{aligned} \tag{244}$$

By definition of Y_j in Definition 53, we obtain $(\mathcal{H}_j)$ with

$$\varepsilon_j(t) := \widetilde{\varepsilon}_{j-1}(t) + \sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}\setminus\{b_j\}} \overline{\varepsilon}_b^j(t). \tag{245}$$

that satisfies (235) with, for instance $C_{j+1} := eC_j + |I|^{M+1}M2^{2M}(2M-1)!$.

Step 3: Conclusion. Taking into account that $\mathcal{B}_{[\![1,M]\!]}\cap Y_{k+1}=\{0\}$, we get $\dot{x}_{k+1}(t)=\varepsilon_{k+1}(t)$ thus $|x_{k+1}(t)-p|\leq C_{k+1}\|u\|_{L^1}^{M+1}F^{M+1}(1+F^{M-1})$, i.e.

$$\left|\overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{-\xi_b(t,u)f_b}x(t) - p\right| \leq C_{k+1}\|u\|_{L^1}^{M+1}F^{M+1}(1+F^{M-1}). \tag{246}$$

Applying the locally Lipschitz map $e^{\xi_{b_1}(t,u)f_{b_1}}\cdots e^{\xi_{b_{k+1}}(t,u)f_{b_{k+1}}}$ to the two terms in the left-hand side, we get another constant $C_M>0$ such that (228) holds. Note that (232) and (150) ensure that $C_M\leq eC_{k+1}$, so that C_M depends indeed only on M. □

5 Convergence results and issues

The formal expansions of Section 2 generally exhibit poor convergence properties for smooth vector fields without any additional assumption. Nevertheless, one can hope to obtain convergence results in the following particular contexts:

- **Nilpotent Lie algebras**. Here, one assumes that the Lie algebra generated by the set of smooth vector fields $\{f(t,\cdot);\ t\in[0,T]\}$ is nilpotent (see Definition 5). This structural assumption turns most of the involved infinite expansions into finite ones, and it is thus reasonable to expect convergence properties.

- **Banach algebras**. Here, one assumes that the vector fields are actually linear in the space variable, e.g. that $f(t,x) = A(t)x$ for some $A(t) \in \mathcal{M}_d(\mathbb{K})$. This assumption yields better estimates for Lie brackets (since products of matrices behave more nicely than differentiation of nonlinear vector fields) and it is thus reasonable to expect convergence properties. In this section, we give statements for matrices for consistence, but similar results can be obtained for linear operators in a Banach algebra.
- **Analytic vector fields**. Here, one assumes that the vector fields are locally real-analytic, i.e. than their k-th derivative grows roughly no more that $k!$. This bound is compatible with the $1/k!$ factors which come out of the corresponding time integrals, and it is thus reasonable to expect convergence properties.

In the following paragraphs, we investigate the convergence properties of each expansion in each of these three reasonable contexts and encounter some surprises. We summarize the results in the following table.

Expansion	Lie-Nilpotent	Banach	Analytic
Chen-Fliess	No (Section 5.1.1)	Global (Section 5.1.2)	Yes (Section 5.1.3)
Magnus in the usual setting	Yes for $\mathcal{C}^\infty$ (Section 5.2.1)	Small time (Section 5.2.2)	No (Section 5.2.3)
Magnus in the interaction picture	Yes for $\mathcal{C}^\omega$ (Section 5.3.1)	Small perturbation (Section 5.3.2)	No (Section 5.3.3)
Sussmann's infinite product	Yes for $\mathcal{C}^\infty$ (Section 5.4.1)	Small time (Section 5.4.2)	Open problem (Section 5.4.3)

5.1 Chen-Fliess expansion

5.1.1 Counter-example for nilpotent vector fields

As already discussed in Remark 17, the Chen-Fliess expansion is not an intrinsic representation of the flow and involves quantities which are not Lie brackets of the dynamics. Therefore, this expansion is not expected to converge under a Lie-nilpotency assumption. The following counter-example (where the dynamic does not depend on time, thereby obviously generating a nilpotent Lie algebra of order 2) proves that this expansion indeed relies on quantities which are not Lie brackets.

Proposition 105. *There exists $f_0 \in \mathcal{C}^\infty(\mathbb{R};\mathbb{R})$ such that, for every $t \in (0,1]$, the solution $x(t;f,0)$ to (143) with $f(t,x) := f_0(x)$ satisfies, with the notations of Remark 74,*

$$\lim_{N\to+\infty} \left| x(t;f,0) - \sum_{n=0}^{N} \frac{t^n}{n!} \left(f_0^n \mathrm{Id}_1\right)(0) \right| = +\infty. \tag{247}$$

Proof. For every sequence $(\alpha_n)_{n\in\mathbb{N}} \in \mathbb{R}^{\mathbb{N}}$, there exists $f_\alpha \in \mathcal{C}^\infty(\mathbb{R};\mathbb{R}) \cap L^\infty(\mathbb{R};\mathbb{R})$ with $f_\alpha(0) = 1$ such that

$$\forall n \geq 2, \quad \left(f_\alpha^n \mathrm{Id}_1\right)(0) = \alpha_n. \tag{248}$$

This is an easy consequence of Borel's lemma. Indeed, for $n \geq 2$ and $f_\alpha(0) = 1$,

$$\left(f_\alpha^n \mathrm{Id}_1\right)(0) = f_\alpha^{(n-1)}(0) + P_n\left(f_\alpha(0), \ldots, f_\alpha^{(n-2)}(0)\right), \tag{249}$$

for some polynomial P_n. Thus, given a sequence $(\alpha_n)_{n\in\mathbb{N}}$, one can prescribe an appropriate value for $f_\alpha^{(n-1)}$ and recursively ensure (248). Let f_0 be a vector field constructed following this process

for $\alpha_n := n!^2$. On the one hand, since $f_0 \in L^\infty(\mathbb{R};\mathbb{R})$, $x(t;f,0)$ is bounded for $t \in [0,1]$. On the other hand, thanks to (248), for each $t > 0$

$$\sum_{n=0}^{N} \frac{t^n}{n!} \left(f_0^n \mathrm{Id}_1\right)(0) = \sum_{n=0}^{N} n! t^n \to +\infty, \tag{250}$$

which proves (247). □

Remark 106. *In this counter-example, the local change of coordinates which transforms $f_0(x)e_1$ into the constant vector field e_1 allows to transform the ODE on x to a new ODE for which the Chen-Fliess expansion is finite (and thus convergent). It would be even more interesting to construct a counter-example, probably in dimension $d \geq 2$, for which no local change of coordinates can restore the convergence of the Chen-Fliess expansion.*

5.1.2 Global convergence for matrices

Let $T > 0$. In this paragraph, we study linear systems of the form

$$\dot{x}(t) = A(t)x(t) \quad \text{and} \quad x(0) = p, \tag{251}$$

where $A \in L^1((0,T);\mathcal{M}_d(\mathbb{K}))$. The solution is denoted $x(t;A,p)$.

Proposition 107. *Let $T > 0$ and $A \in L^1((0,T);\mathcal{M}_d(\mathbb{K}))$. For each $t \in [0,T]$ and $p \in \mathbb{K}^d$,*

$$x(t;A,p) = p + \sum_{j=1}^{+\infty} \int_{\Delta^j(t)} A(\tau_j)\cdots A(\tau_1)p \,\mathrm{d}\tau, \tag{252}$$

where the series converges absolutely.

Proof. To simplify the notations, we write $x(t)$ instead of $x(t;A,p)$. By Grönwall's lemma, we have $|x(\tau)| \leq |p|e^{\|A\|_{L^1(0,\tau)}}$ for every $\tau \in [0,T]$. By iterating the formula

$$x(\tau) = p + \int_0^\tau A(\tau')x(\tau')\,\mathrm{d}\tau' \tag{253}$$

we obtain, for every $M \in \mathbb{N}^*$

$$\begin{aligned} &\left| x(t) - p - \sum_{j=1}^{M-1} \int_{\Delta^j(t)} A(\tau_j)\cdots A(\tau_1)p\,\mathrm{d}\tau \right| = \left| \int_{\Delta^M(t)} A(\tau_M)\cdots A(\tau_1)x(\tau)\,\mathrm{d}\tau \right| \\ &\leq \int_{\Delta^M(t)} \|A(\tau_M)\|\cdots\|A(\tau_1)\|\,\mathrm{d}\tau |p| e^{\|A\|_{L^1(0,t)}} = \frac{\|A\|_{L^1(0,t)}^M}{M!}|p|e^{\|A\|_{L^1(0,t)}}, \end{aligned} \tag{254}$$

which proves the convergence. Similar estimates prove the absolute convergence. □

5.1.3 Local convergence for analytic vector fields

For analytic vector fields, it is well known that the Chen-Fliess series (also called "(right) chronological exponential" in [3, Section 2.1]) converges locally in time (see e.g. [3, Proposition 2.1], or [99, Proposition 4.3] for slightly different assumptions). The analyticity assumption is necessary, as highlighted by the counter-example of Section 5.1.1.

Proposition 108. *Let $T, \delta, r > 0$. There exists $\eta > 0$ such that, for $f \in L^1([0,T]; \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K}^d))$ with $\|f\|_{L^1(\mathcal{C}^{\omega,r})} \leq \eta$, $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K}^d)$, $t \in [0,T]$ and $p \in B_\delta$,*

$$\varphi\left(x(t; f, p)\right) = \varphi(p) + \sum_{j=1}^{+\infty} \int_{\Delta^j(t)} (f(\tau_1) \cdots f(\tau_j))(\varphi)(p) \, \mathrm{d}\tau, \tag{255}$$

where the sum converges absolutely. In particular,

$$x(t; f, p) = p + \sum_{j=1}^{+\infty} \int_{\Delta^j(t)} (f(\tau_1) \cdots f(\tau_j))(\mathrm{Id}_d)(p) \, \mathrm{d}\tau. \tag{256}$$

Proof. Let $\eta := \min\{\delta/2, r/10\}$. By Lemma 82, $x(t; f, p)$ is well defined for $t \in [0,T]$, $p \in B_\delta$ and belongs to $B_{2\delta}$. Moreover, by Lemma 76, we have, for every $j \in \mathbb{N}^*$

$$\int_{\Delta^j(t)} \left|(f(\tau_1) \cdots f(\tau_j))(\varphi)(p)\right| \, \mathrm{d}\tau \leq j! \left(\frac{5}{r}\right)^j \cdot \frac{\|f\|^j}{j!} \left|\!\left|\!\left|\varphi\right|\!\right|\!\right|_r, \tag{257}$$

where $\|f\| := \|f\|_{L^1((0,t);\mathcal{C}^{\omega,r})}$, which proves the absolute convergence because the right-hand side is bounded by $2^{-j} \left|\!\left|\!\left|\varphi\right|\!\right|\!\right|_r$. Eventually, we deduce from (187) and Lemma 76 that

$$\left| \varphi\left(x(t; f, p)\right) - \sum_{j=0}^{M-1} \int_{\Delta^j(t)} (f(\tau_1) \cdots f(\tau_j))(\varphi)(p) \, \mathrm{d}\tau \right| \leq 2^{-M} \left|\!\left|\!\left|\varphi\right|\!\right|\!\right|_r, \tag{258}$$

which proves (255). □

5.2 Magnus expansion in the usual setting

5.2.1 Equality for nilpotent systems

The goal of this section is to prove that the Magnus expansion is an exact expansion for regular vector fields generating a nilpotent Lie algebra (see Proposition 110).

If the vector fields are analytic in space, a simple proof can be given (see e.g. [66, Remark A.1] for the case of the CBHD formula), with the following steps. First, by density, one can assume that the dynamic depends analytically on time. Then, the maps $t \mapsto x(t)$ and $t \mapsto e^{Z_M}(t)$ are analytic. Because of the nilpotency assumption, $Z_M = Z_{M'}$ for every $M' \geq M$ and estimate (188) proves that both functions have the same Taylor expansion at $t = 0$, and are thus equal.

For non-analytic vector fields, the proof is much more intricate. A sketch of proof is briefly suggested in [4, Proposition 2.4]. In this paragraph, we write the proof completely. The difficulty is to formulate the question in the nilpotent Lie algebra generated by the vector fields, in order to conclude with the universal property of free nilpotent Lie algebras (Lemma 7).

To that end, we start with the following statement.

Lemma 109. *Let a be given by (14), $M \in \mathbb{N}^*$, $Z_M(t) := \mathrm{Log}_M\{a\}(t)$ with the notation of Definition 23. Then for every $t \in \mathbb{R}$, the following equality holds in $\mathcal{N}_{M+1}(X)$*

$$\sum_{n=0}^{M-1} \frac{(-1)^n}{(n+1)!} \mathrm{ad}^n_{Z_M(t)}(\dot{Z}_M(t)) = a(t), \tag{259}$$

where $Z_M(t)$ belongs to the space $\bigoplus_{r \in [\![1,M]\!]} \mathcal{L}(X)^r$ which is identified with $\mathcal{N}_{M+1}(X)$ as a vector space.

Proof. The canonical surjection $\sigma_{M+1} : \mathcal{L}(X) \to \mathcal{N}_{M+1}(X)$ is an homomorphism of Lie algebras. Applying this homomorphism to (22), where $z(t) = \mathrm{Log}_\infty\{a\}(t)$ thanks to Theorem 27, proves (259). □

Proposition 110. *Let $M \in \mathbb{N}^*$. There exists $\eta_M > 0$ such that, for every $T, \delta > 0$ and every time-varying vector field $f : [0,T] \to \mathcal{C}^\infty(B_{4\delta}; \mathbb{K}^d)$ such that $\mathcal{L}(f([0,T]))$ is nilpotent with index at most $M+1$ and $f \in L^1((0,T); \mathcal{C}^M(B_{4\delta}; \mathbb{K}^d))$ with*

$$\|f\|_{L^1((0,T);\mathcal{C}^M)} + \|f\|^M_{L^1((0,T);\mathcal{C}^M)} \leq \eta_M \delta, \tag{260}$$

then, for each $p \in B_\delta$ and $t \in [0,T]$, one has $x(t; f, p) = e^{Z_M(t,f)}(p)$ where $Z_M(t,f) := \mathrm{Log}_M\{f\}(t)$ is the vector field defined in Definition 23.

Proof. Let $M \in \mathbb{N}^*$. By Definition 23 and Lemma 79, there exists $\eta_M > 0$ such that,

$$\|Z_M(t,f)\|_{\mathcal{C}^1} \leq \frac{1}{\eta_M} \left(\|f\|_{L^1(\mathcal{C}^M)} + \|f\|^M_{L^1(\mathcal{C}^M)} \right). \tag{261}$$

In particular, for every $p \in B_\delta$, $e^{Z_M(t,f)}(p)$ is well-defined thanks to Eq. (260).

Step 1: Proof for $f(t,x) = \sum_{j=1}^q a_j(t) f_j(x)$ with $q \in \mathbb{N}^*$, $a_j \in \mathcal{C}^\infty([0,T]; \mathbb{K})$ and $f_j \in \mathcal{C}^\infty_c(\mathbb{K}^d; \mathbb{K}^d)$. By uniqueness in the Cauchy-Lipschitz theorem, it is sufficient to prove that for every $t \in [0,T]$ and $p \in B_\delta$,

$$\frac{\mathrm{d}}{\mathrm{d}t}(e^{Z_M(t,f)}(p)) = f\left(t, e^{Z_M(t,f)}(p)\right). \tag{262}$$

By Definition 23, the map $(t,p) \mapsto Z_M(t,f)(p)$ belongs to $\mathcal{C}^\infty([0,T] \times B_{4\delta}; \mathbb{K}^d)$. Thanks to the nilpotency assumption, $\mathrm{ad}^M_{Z_M(t,f)}(Z_M(\tau,f)) = 0$ on $B_{4\delta}$ for every $t, \tau \in [0,T]$. Thus Lemma 91 yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(e^{Z_M(t,f)}(p)\right) = \sum_{k=0}^M \frac{(-1)^k}{(k+1)!} \mathrm{ad}^k_{Z_M(t,f)}\left(\dot{Z}_M(t,f)\right)\left(e^{Z_M(t,f)}(p)\right). \tag{263}$$

Let $\Lambda : \mathcal{N}_{M+1}(X) \to \mathcal{L}(f_1, \ldots, f_q)$ be the homomorphism of nilpotent Lie algebras such that $\Lambda(X_j) = f_j$ for $j = 1, \ldots, q$. By applying Λ to the equality (259), we obtain that the right-hand side of the above equality is $f(t, e^{Z_M(t,f)}(p))$.

Step 2: Proof for a general time-dependent vector field f. We apply Step 1 to a sequence f_n of simple functions, taking values in $f([0,T])$, converging towards f in $L^1((0,T); \mathcal{C}^M(B_{4\delta}; \mathbb{K}^d))$. We get the conclusion by passing to the limit in both sides, using the fourth item of Lemma 82. □

5.2.2 Convergence for linear systems

In this paragraph, we consider linear systems of the form (251). Since the Magnus expansion was designed for linear systems, its convergence in this context has received much attention. Depending on the exact convergence notion that one considers and on the way one groups terms, different sufficient conditions for the convergence can be derived. In [97], $T\|A\|_{L^\infty(0,T)} \leq 1$ is shown to be a sufficient condition for convergence on $[0,T]$ thanks to a careful estimate of the combinatorial terms. In [83], $\|A\|_{L^1(0,T)} < \pi$ is shown to be a sufficient condition for convergence using complex analysis.

We give below a short elementary proof with a sub-optimal constant, for the sake of completeness and because it will be used in the sequel. Let $\|\cdot\|$ be a sub-multiplicative norm on $\mathcal{M}_d(\mathbb{K})$.

Proposition 111. *Let $T > 0$ and $A \in L^1((0,T); \mathcal{M}_d(\mathbb{K}))$ such that $\|A\|_{L^1(0,T)} < \frac{1}{4}$. For each $t \in [0,T]$,*

$$Z_\infty(t) := \sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^r \frac{(-1)^{m-1}}{m} \sum_{\mathbf{r} \in \mathbb{N}^m_r} \int_{\Delta^{\mathbf{r}}(t)} [\cdots [A(\tau_1), A(\tau_2)], \ldots A(\tau_r)] \, \mathrm{d}\tau \tag{264}$$

is well defined in $\mathcal{M}_d(\mathbb{K})$ and, for every $p \in \mathbb{K}^d$, $x(t; A, p) = e^{-Z_\infty(t)} p$, where the brackets refer to commutators of matrices, i.e. $[A, B] = AB - BA$.

Proof. Step 1: Absolute convergence of $Z_\infty(t)$. Let $r \in \mathbb{N}^*$. For every $m \in [\![1, r]\!]$ and $\mathbf{r} \in \mathbb{N}_r^m$,

$$\begin{aligned}\int_{\Delta^{\mathbf{r}}(t)} \|[\cdots[A(\tau_1), A(\tau_2)], \dots A(\tau_r)]\| \,\mathrm{d}\tau \\ \leq \int_{\Delta^{\mathbf{r}}(t)} 2^r \|A(\tau_1)\| \dots \|A(\tau_r)\| \,\mathrm{d}\tau \leq 2^r \left(\int_0^t \|A(\tau)\| \,\mathrm{d}\tau\right)^r.\end{aligned} \tag{265}$$

Moreover, recalling the definition of (31), $|\mathbb{N}_r^m| = \binom{r-1}{m-1}$ and $\sum_{m=1}^r \binom{r-1}{m-1} = 2^{r-1}$. Thus,

$$\sum_{r=1}^{+\infty} \frac{1}{r} \sum_{m=1}^{r} \frac{1}{m} \sum_{\mathbf{r}\in\mathbb{N}_r^m} \int_{\Delta^{\mathbf{r}}(t)} \|[\cdots[A(\tau_1), A(\tau_2)], \dots A(\tau_r)]\| \,\mathrm{d}\tau \leq \sum_{r=1}^{+\infty} (4\|A\|_{L^1})^r < \infty. \tag{266}$$

Step 2: Formula for the solution $L \in \mathcal{C}^1([0,t]; \mathcal{M}_d(\mathbb{K}))$ of

$$\begin{cases} L'(\tau) = L(\tau)A(\tau) \\ L(0) = \mathrm{Id}_d. \end{cases} \tag{267}$$

By working as in the proof of Proposition 107, we obtain

$$L(t) = \mathrm{Id}_d + \sum_{r=1}^{+\infty} \int_{\Delta^r(t)} A(\tau_1) \cdots A(\tau_r) \,\mathrm{d}\tau \tag{268}$$

where the series converges absolutely. Moreover, we have

$$\left\| \sum_{r=1}^{+\infty} \int_{\Delta^r(t)} A(\tau_1) \cdots A(\tau_r) \,\mathrm{d}\tau \right\| \leq \sum_{r=1}^{+\infty} \frac{\|A\|_{L^1}^r}{r!} < e^{\frac{1}{4}} - 1 < 1. \tag{269}$$

Thus

$$\log(L(t)) = \sum_{m=1}^{+\infty} \frac{(-1)^m}{m} \left(\sum_{r=1}^{+\infty} \int_{\Delta^r(t)} A(\tau_1) \cdots A(\tau_r) \,\mathrm{d}\tau \right)^m \tag{270}$$

is well defined in $\mathcal{M}_d(\mathbb{K})$ and $L(t) = e^{\log(L(t))}$. By applying Corollary 95 with $A = A_1 = \mathcal{M}_d(\mathbb{K})$, we get $\log(L(t)) = Z_\infty(t)$.

Step 3: Conclusion. The resolvent $R(\tau)$ associated to the linear system $\dot{x} = A(\tau)x$ with initial condition at $\tau = 0$ is $R(\tau) = L(\tau)^{-1}$. Thus $x(t) = R(t)p = e^{-Z_\infty(t)}p$. □

Remark 112. *For $X, Y \in \mathcal{M}_d(\mathbb{K})$ such that $\|X\| + \|Y\| < \frac{1}{8}$, the previous statement implies the convergence of the CBHD formula, yielding a matrix Z_∞ such that $e^X e^Y = e^{Z_\infty}$. Some authors have investigated the optimal convergence domain in different contexts for the CBHD formula. Such a domain sometimes depends on the summation process (i.e. the way terms are grouped together) and the exact question one asks (existence of a logarithm, absolute summability of the series, convergence of the remainder, etc.). Better sufficient conditions than ours can be found for instance in [18], for instance, $\|X\| + \|Y\| < \frac{\ln 2}{2}$. We refer to [16] for a nice survey of the convergence questions regarding the CBHD formula.*

Remark 113. *The smallness assumption (on time or on the matrices) is in general necessary, both for the CBHD formula (see [16, Example 2.3] or [103, Section II]) and for the Magnus expansion (see [83], where the authors also prove that, although the condition $\|A\|_{L^1(0,T)} < \pi$ is not necessary for convergence, there exists A with $\|A\|_{L^1(0,T)} = \pi$ for which the Magnus series at time π does not converge).*

5.2.3 Divergence for arbitrarily small analytic vector fields

The convergence of Magnus expansions is deeply linked with the convergence of the CBHD series. For analytic vector fields, it is expected that both series diverge (see e.g. [4, p.1671] or [97, p.335] for statements without examples). Some authors nevertheless suggested that, despite the divergence of the series, the flows could converge for analytic vector fields (see [97, p.335] and [71, p.241]).

In this paragraph, we give explicit counter-examples to the convergence, even in the weak sense of the flows, for arbitrarily small analytic vector fields, of both the CBHD series and the Magnus expansion. Similarly to counter-examples concerning the convergence of the CBHD series for large matrices (see e.g. [16, Theorem 2.5]), our construction relies on the choice of generators for which many brackets vanish thanks to their particular structure, and the remaining non-vanishing brackets are associated with coordinates of the first kind involving Bernoulli numbers.

Proposition 114. *There exists $\delta > 0$ and $f_0, f_1 \in \mathcal{C}^\omega(B_\delta; \mathbb{R}^2)$ such that,*

$$\forall M \in \mathbb{N}, \exists C_M, \varepsilon_M > 0, \forall \varepsilon \in [0, \varepsilon_M], \quad \left| e^{\varepsilon f_0} e^{\varepsilon f_1}(0) - e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) \right| \leq C_M \varepsilon^{M+1}, \tag{271}$$

where $\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)$ is defined in Corollary 94, but, simultaneously, for every $\varepsilon > 0$,

$$\lim_{M \to +\infty} |\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(0)| = +\infty \tag{272}$$

and

$$\lim_{M \to +\infty} \left| e^{\varepsilon f_0} e^{\varepsilon f_1}(0) - e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) \right| = +\infty. \tag{273}$$

Proof. Let f_0, f_1 as in Remark 81. For these vector fields, estimate (271) comes from Corollary 94. Due to their structure, the only non vanishing brackets are those containing f_1 at most once. Therefore, formula (69) of Corollary 39 yields, for $M \geq 1$,

$$\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0) = \varepsilon f_0 + \sum_{k=0}^{M-1} \frac{B_k}{k!} \varepsilon^{k+1} \operatorname{ad}_{f_0}^k(f_1). \tag{274}$$

Hence, using (142),

$$\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(x) = \varepsilon e_1 + \varepsilon \Theta_M^\varepsilon(x_1) e_2, \tag{275}$$

where we introduce, for $q \in \mathbb{R}$,

$$\Theta_M^\varepsilon(q) := \sum_{k=0}^{M-1} B_k \varepsilon^k (1-q)^{-k-1}. \tag{276}$$

In particular,

$$|\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)(0)| \geq |\varepsilon \Theta_M^\varepsilon(0)|. \tag{277}$$

Since the odd Bernoulli numbers except B_1 are zero, when $M = 2M' + 2$ with $M' \geq 1$, $\Theta_{2M'+2}^\varepsilon = \Theta_{2M'+1}^\varepsilon$. Then,

$$\Theta_{2M'+1}^\varepsilon(q) = \frac{1}{1-q} - \frac{\varepsilon}{2(1-q)^2} + \sum_{k=1}^{M'} B_{2k} \varepsilon^{2k} (1-q)^{-2k-1}. \tag{278}$$

In particular, using (531),

$$\Theta_{2M'+1}^\varepsilon(0) = 1 - \frac{\varepsilon}{2} + \sum_{k=1}^{M'} B_{2k} \varepsilon^{2k} = 1 - \frac{\varepsilon}{2} + \sum_{k=1}^{M'} (-1)^{k+1} \frac{2(2k)!}{(2\pi)^{2k}} \zeta(2k) \varepsilon^{2k}. \tag{279}$$

Thus, for every fixed $\varepsilon > 0$, $|\Theta_M^\varepsilon(0)| \to +\infty$ when $M \to +\infty$, because it involves a sum of the form $\sum_{k=1}^{M'} a_k$ where $|a_{k+1}|/|a_k| \to +\infty$ when $k \to +\infty$. Using (277), this proves (272).

For $p \in \mathbb{R}^2$ close enough to the origin, one can also compute the flow $e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(p)$, which is $y(1)$ where y is the solution to the ODE $y(0) = p$ and

$$\dot{y}_1(s) = \varepsilon \quad \text{and} \quad \dot{y}_2(s) = \varepsilon \Theta_M^\varepsilon \left(y_1(s) \right). \tag{280}$$

Solving successively for y_1 then y_2 yields $y_1(s) = p_1 + s\varepsilon$ and

$$y_2(s) = p_2 + \int_{y_1(0)}^{y_1(s)} \Theta_M^\varepsilon(h) \, \mathrm{d}h. \tag{281}$$

Thus,

$$e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(p) = (p_1 + \varepsilon) e_1 + \left(p_2 + \int_{p_1}^{p_1 + \varepsilon} \Theta_M^\varepsilon(h) \, \mathrm{d}h \right) e_2. \tag{282}$$

In particular,

$$e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) = \varepsilon e_1 + \left(\int_0^{\varepsilon} \Theta_M^\varepsilon(h) \, \mathrm{d}h \right) e_2. \tag{283}$$

When $M = 2M' + 2$ with $M' \geq 1$, using (278), we get

$$e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0) = \varepsilon e_1 + \left(-\ln(1-\varepsilon) - \frac{\varepsilon}{2} \left(\frac{1}{1-\varepsilon} - 1 \right) + \sum_{k=1}^{M'} \frac{B_{2k}}{2k} \varepsilon^{2k} \left(\frac{1}{(1-\varepsilon)^{2k}} - 1 \right) \right) e_2. \tag{284}$$

Hence, for the same reason as above, the flow satisfies $|e^{\mathrm{CBHD}_M(\varepsilon f_1, \varepsilon f_0)}(0)| \to +\infty$ when $M \to +\infty$, which proves (273). □

Remark 115. *If one sees (x_1, x_2) as (q, p) in an Hamiltonian setting, one checks that the vector fields defined in* (141) *and used in this counter-example are associated with the Hamiltonians $\mathcal{H}_0(q,p) := p$ and $\mathcal{H}_1(q,p) := \ln(1-q)$. Therefore, assuming an Hamiltonian structure on the considered vector fields does not provide enough structure to yield convergence.*

One could wonder if assuming even more structure on the dynamics, for example assuming that it is time-reversible, prevents the construction of such counter-examples.

Open problem 116. *Do there exist Hamiltonians $\mathcal{H}_0$ and $\mathcal{H}_1$ on $\mathbb{R}^{2d}$, which are time-reversible (i.e. satisfy $\mathcal{H}_i(q,p) = \mathcal{H}_i(q,-p)$ for every $q, p \in \mathbb{R}^d$), locally real-analytic near zero and for which the convergence of the CBHD series fails as in Proposition 114?*

The counter-example of Proposition 114 for the convergence of the CBHD series allows to build counter-examples to the convergence of the Magnus expansion which blow up instantly, despite analytic regularity in both time and space.

Proposition 117. *There exist $T, \delta > 0$ and $f \in \mathcal{C}^\omega([0,T] \times B_\delta; \mathbb{R}^2)$ such that, for every $\varepsilon > 0$ and $t \in (0, T]$,*

$$\lim_{M \to +\infty} |Z_M(t, \varepsilon f)(0)| = +\infty \tag{285}$$

and

$$\lim_{M \to +\infty} \left| x(t; \varepsilon f, 0) - e^{Z_M(t, \varepsilon f)}(0) \right| = +\infty, \tag{286}$$

where x is the solution to $\dot{x}(t) = \varepsilon f(t, x(t))$ with $x(0) = 0$ and $Z_M(t, \varepsilon f) = \mathrm{Log}_M\{\varepsilon f\}(t)$.

Proof. Let $T = 1$. We define $f(t,x) := f_0(x) + t f_1(x)$, where f_0 and f_1 are defined in Remark 81. Similarly as for the previous construction, only Lie brackets involving f_1 at most once are non-vanishing. Moreover, the coordinates of the first kind associated with the controls $a_0(t) = 1$ and $a_1(t) = t$ have been computed in Example 40. Hence, recalling (142), we have

$$Z_M(t, \varepsilon f) = \varepsilon t e_1 + \sum_{k=0}^{M-1} \varepsilon^{k+1} \frac{(-1)^{k+1} t^{k+2}}{(k+1)!} B_{k+1} \frac{k!}{(1-x_1)^{k+1}} e_2. \tag{287}$$

Proceeding along the same lines as in the proof of Proposition 114 allows to conclude that both $Z_M(t,\varepsilon f)(0)$ and $e^{Z_M(t,\varepsilon)}(0)$ diverge when $M \to +\infty$. □

5.3 Magnus expansion in the interaction picture

5.3.1 Nilpotent systems

For ODEs of the form (209), the starting point of the interaction picture is to factorize the flow of f_0. Hence, the roles of f_0 and f_1 are asymmetric. One can expect that, under the assumption that Lie brackets of f_0 and f_1 containing at least $M+1$ times f_1 identically vanish, the Magnus expansion in the interaction picture should yield an equality of the form

$$x(t; f_0 + f_1, p) = e^{\mathcal{Z}_M(t,f_0,f_1)} e^{tf_0} p, \tag{288}$$

where $\mathcal{Z}_M(t,f_0,f_1)$ is defined in Proposition 99. We prove in this paragraph that it is indeed the case, when f_0 and f_1 are analytic. However, contrary to the case of the usual Magnus expansion (see Section 5.2.1), we give examples highlighting the fact that the analyticity assumption cannot be removed, which is quite surprising but stems from the mixing induced by pushforwards.

We therefore start with the following definition.

Definition 118 (Semi-nilpotent family of vector fields)**.** *Let Ω be an open subset of $\mathbb{K}^d$. Let $\mathcal{F} \subset \mathcal{C}^\infty(\Omega;\mathbb{K}^d)$, $f_0 \in \mathcal{C}^\infty(\Omega;\mathbb{K}^d)$ and $M \in \mathbb{N}^*$. We say that the family of vector fields $\mathcal{F}$ is* semi-nilpotent of index M with respect to f_0 *if every bracket of elements of $\mathcal{F} \cup \{f_0\}$ involving M elements of $\mathcal{F}$ vanishes identically on Ω and M is the smallest positive integer for which this property holds.*

Remark 119. *Some authors (see e.g. [62, Section 3]) refer to this situation by saying that $\mathcal{L}(\mathcal{S}_1)$ is nilpotent of index M, where $\mathcal{S}_1 := \{\mathrm{ad}^k_{f_0}(f); f \in \mathcal{F}, k \in \mathbb{N}\}$. Both definitions are equivalent, thanks to the Jacobi identity for Lie brackets.*

Proposition 120. *Let $T, \delta > 0$. Let $M \in \mathbb{N}$. Let $f_0 \in \mathcal{C}^\infty(B_{4\delta};\mathbb{K}^d)$ with $T\|f_0\|_{\mathcal{C}^0} \leq \delta$. There exists $\eta > 0$ such that, for every $f_1 : [0,T] \to \mathcal{C}^\infty(B_{4\delta};\mathbb{K}^d)$ with $f_1 \in L^1([0,T];\mathcal{C}^{M+1}(B_{4\delta};\mathbb{K}^d))$ and $\|f_1\|_{L^1(\mathcal{C}^M)} \leq \eta$, the following family is well-defined*

$$\mathcal{G} := \{\Phi_0(-t)_* f_1(t);\ \ t \in [0,T]\} \subset \mathcal{C}^\infty(B_\delta;\mathbb{K}^d). \tag{289}$$

and, assuming moreover that $\mathcal{G}$ is nilpotent of index $M+1$, then, for each $t \in [0,T]$ and $p \in B_\delta$, the solution to (209) satisfies (288).

Proof. Let $t > 0$. As in the proof of Proposition 99, we introduce the new variable $y(s) := \Phi_0(t-s, x(s))$. Then $\dot{y}(s) = g_t(s, y(s))$, where g_t is defined in (210). Thanks to Lemma 88, $g_t(s) = \Phi_0(t)_* \Phi_0(-s)_* f_1(s)$. Thanks to the assumption and to Lemma 89, the family $\{g_t(s);\ \ s \in [0,t]\}$ is nilpotent of index $M+1$. Thus, by Proposition 110, $y(t) = e^{\mathcal{Z}_M(t,f_0,f_1)} y(0)$. Since $x(t) = y(t)$ and $y(0) = \Phi_0(t,p)$, this concludes the proof of (288). □

Lemma 121. *Let $T, \delta > 0$, $\mathcal{F} \subset \mathcal{C}^\infty(B_{4\delta};\mathbb{K}^d)$, $f_0 \in \mathcal{C}^\infty(B_{4\delta};\mathbb{K}^d)$ such that $T\|f_0\|_{\mathcal{C}^0} \leq \delta$. The following family is well-defined*

$$\mathcal{G} := \{\Phi_0(-t)_* f;\ \ t \in [0,T], f \in \mathcal{F}\} \subset \mathcal{C}^\infty(B_\delta;\mathbb{K}^d). \tag{290}$$

Assume that the family $\mathcal{F}$ is semi-nilpotent of index M with respect to f_0 and that there exists $r > 0$ such that $\mathcal{F} \cup \{f_0\} \subset \mathcal{C}^{\omega,r}(B_\delta;\mathbb{K}^d)$. Then $\mathcal{G}$ is nilpotent of index M.

Proof. For $t \in [0,T]$ and $f \in \mathcal{F}$, equation (164) of Lemma 90 implies that

$$\Phi_0(-t)_* f = \sum_{k=0}^{+\infty} \frac{t^k}{k!} \mathrm{ad}^k_{f_0}(f) \tag{291}$$

and that the series converges absolutely in $\mathcal{C}^M(B_\delta;\mathbb{K}^d)$ (in particular). Hence, if $t_1,\dots,t_M\in[0,T]$ and $f_1,\dots f_M\in\mathcal{F}$, the bracket

$$\begin{aligned}&[\Phi_0(-t_M)_*f_M,[\cdots[\Phi_0(-t_2)_*f_2,\Phi_0(-t_1)_*f_1]\cdots]]\\&\qquad=\sum_{k_1,\dots k_M\in\mathbb{N}}\frac{t_1^{k_1}\cdots t_M^{k_M}}{k_1!\cdots k_M!}[\mathrm{ad}_{f_0}^{k_M}(f_M),[\cdots[\mathrm{ad}_{f_0}^{k_2}(f_2),\mathrm{ad}_{f_0}^{k_1}(f_1)]\cdots]]\end{aligned}\tag{292}$$

vanishes thanks to the assumption and the absolute convergence of the sums. The same is true for every other bracket structure, which proves that $\mathcal{G}$ is nilpotent of index M. □

Corollary 122. *Let $T,\delta,r>0$. Let $f_0\in\mathcal{C}^{\omega,r}(B_{4\delta};\mathbb{K}^d)$ such that $T\|f_0\|_{\mathcal{C}^0}\le\delta$ and $f_1\in L^1([0,T];\mathcal{C}^{\omega,r}(B_{4\delta};\mathbb{K}^d))$. Assume moreover that $\mathcal{F}:=\{f_1(t,\cdot);\ t\in[0,T]\}$ is semi-nilpotent of index $M+1$ with respect to f_0. Then, for each $t\in[0,T]$ and $p\in B_\delta$, the solution to* (209) *satisfies* (288)*, where $\mathcal{Z}_M(t,f_0,f_1)$ is defined in Proposition 99.*

Proof. This corollary is a direct consequence of Proposition 120 and Lemma 121. □

The analyticity assumption in Lemma 121 is necessary, as illustrated by the following counter-example for smooth functions.

Example 123. *We consider smooth vector fields on $\mathbb{R}^3$. Let $\chi\in\mathcal{C}^\infty(\mathbb{R};\mathbb{R})$ with $\chi\equiv0$ on $\mathbb{R}_-$ and $\chi(x)>0$ for $x>0$. Let f_0 and $\mathcal{F}:=\{f_1,f_2\}$ where*

$$f_0(x):=e_2,\tag{293}$$
$$f_1(x):=\chi(x_2)x_1e_3,\tag{294}$$
$$f_2(x):=\chi(-x_2)e_1.\tag{295}$$

Heuristically, f_1 and f_2 commute because they have disjoint (touching) supports, but the flow of f_0 involved in (290) *mixes these supports for every positive time. This is possible only because χ is not analytic.*

First, we check that $\mathcal{F}$ is semi-nilpotent of order 2 with respect to f_0. Indeed, for every $j\in\mathbb{N}$,

$$\mathrm{ad}_{f_0}^j(f_1)(x)=\chi^{(j)}(x_2)x_1e_3,\tag{296}$$
$$\mathrm{ad}_{f_0}^j(f_2)(x)=(-1)^j\chi^{(j)}(-x_2)e_1.\tag{297}$$

Thus, for $j,k\in\mathbb{N}$, $[\mathrm{ad}_{f_0}^j(f_1),\mathrm{ad}_{f_0}^k(f_1)]$ (resp. $[\mathrm{ad}_{f_0}^j(f_2),\mathrm{ad}_{f_0}^k(f_2)]$) vanishes because both vector fields are multiples of e_3 but independent of x_3 (resp. multiples of e_1 but independent of x_1). Moreover,

$$[\mathrm{ad}_{f_0}^j(f_1),\mathrm{ad}_{f_0}^k(f_2)](x)=-(-1)^k\chi^{(k)}(-x_2)\chi^{(j)}(x_2)e_3=0,\tag{298}$$

because the supports of $\chi(\cdot)$ and $\chi(-\cdot)$ only touch at $x_2=0$ where all derivatives vanish.

Second, let us check however that the family $\mathcal{G}$ defined in (290) *is not nilpotent of index 2. Indeed, for $t\ge0$ and $x\in\mathbb{R}^3$, $\Phi_0(t)(x)=x+te_2$. Thus, for $f\in\mathcal{C}^\infty(\mathbb{R}^3;\mathbb{R}^3)$, $(\Phi_0(-t)_*f)(x)=f(x+te_2)$. Therefore, for every $T>0$, $\mathcal{G}$ is well-defined on $\mathbb{R}^3$. Moreover,*

$$[f_2,(\Phi_0(-t)_*f_1)](x)=\chi(-x_2)\chi(x_2+t)e_3.\tag{299}$$

In particular, for every $\varepsilon>0$, $[f_2,(\Phi_0(-2\varepsilon)_*f_1)](-\varepsilon e_2)=\chi(\varepsilon)^2e_3\neq0$, which prevents the family $\mathcal{G}$ from being nilpotent of index 2 (even locally in time and space).

The analyticity assumption in Corollary 122 is also necessary, as illustrated by the following counter-example for smooth functions, inspired by the previous one.

Example 124. *We consider smooth vector fields on $\mathbb{R}^3$. Let $\chi \in \mathcal{C}^\infty(\mathbb{R};\mathbb{R})$ with $\chi \equiv 0$ on $\mathbb{R}_-$ and $\chi(x) > 0$ for $x > 0$. Let $f_0(x) := e_2$ and $f_1(t,x) := f_1(x)$ (independent of time) with*

$$f_1(x) := 2\chi^{(1)}(x_2)x_1e_3 + \chi^{(1)}(-x_2)e_1. \tag{300}$$

For $j \in \mathbb{N}$, one has

$$\mathrm{ad}^j_{f_0}(f_1)(x) = \partial_2^j f_1(x) = 2\chi^{(j+1)}(x_2)x_1e_3 + (-1)^j\chi^{(j+1)}(-x_2)e_1. \tag{301}$$

Thus, for every $j_1, j_2 \in \mathbb{N}$,

$$\begin{aligned}[\mathrm{ad}^{j_1}_{f_0}(f_1), \mathrm{ad}^{j_2}_{f_0}(f_1)](x) = {}& 2(-1)^{j_1}\chi^{(j_1+1)}(-x_2)\chi^{(j_2+1)}(x_2)e_3 \\ & - 2(-1)^{j_2}\chi^{(j_2+1)}(-x_2)\chi^{(j_1+1)}(x_2)e_3 = 0\end{aligned} \tag{302}$$

because the supports of $\chi(\cdot)$ and $\chi(-\cdot)$ only touch at $x_2 = 0$, where all derivatives vanish. Hence each bracket of f_0 and f_1 involving f_1 at least twice vanishes identically on $\mathbb{R}^3$. Thus, for every $T > 0$, the family $\mathcal{F} := \{f_1(t,\cdot);\ t \in [0,T]\} = \{f_1\}$ is semi-nilpotent of index 2 with respect to f_0. Let us prove that, despite this property, equality (288) with $M = 1$ fails.

Computation of the state. *We solve $\dot{x} = f_0(x) + f_1(x)$ for some initial data p. Solving the ODE successively for x_2, x_1 and x_3, we obtain*

$$x_1(t) = p_1 + \chi(-p_2) - \chi(-p_2 - t), \tag{303}$$
$$x_2(t) = p_2 + t, \tag{304}$$
$$x_3(t) = p_3 + 2\left(\chi(p_2 + t) - \chi(p_2)\right)(p_1 + \chi(-p_2)). \tag{305}$$

In particular, with $t = 2\varepsilon$ and $p = -\varepsilon e_2$, $x(2\varepsilon; f_0 + f_1, -\varepsilon e_2) = (\chi(\varepsilon), \varepsilon, 2\chi(\varepsilon)^2)$.

Computation of the flow. *We compute $e^{\mathcal{Z}_1(t,f_0,f_1)}e^{tf_0}(p)$ for some initial data p. One has $\Phi_0(\tau, q) = q + \tau e_2$. Hence, in particular $(\Phi_0(\tau)_* f_1)(q) = f_1(q - \tau e_2)$. Moreover $\mathcal{Z}_1(t, f_0, f_1)(q) = \int_0^t g_t(s,q)\,\mathrm{d}s$ where $g_t(s,q) = (\Phi_0(t-s)_* f_1)(q)$. Hence $g_t(s,q) = f_1(q - (t-s)e_2)$ and*

$$\begin{aligned}\mathcal{Z}_1(t, f_0, f_1)(q) &= \int_0^t f_1(q + (s-t)e_2)\,\mathrm{d}s \\ &= 2q_1(\chi(q_2) - \chi(q_2 - t))e_3 + (\chi(-q_2 + t) - \chi(-q_2))e_1.\end{aligned} \tag{306}$$

Then $e^{\mathcal{Z}_1(t,f_0,f_1)}e^{tf_0}p = e^{\mathcal{Z}_1(t,f_0,f_1)}(p + te_2)$ is $y(1)$ where y is the solution to $y(0) = p + te_2$ and $\dot{y}(s) = \mathcal{Z}_1(t, f_0, f_1)(y(s))$. Solving the ODE successively for y_2, y_1 and y_3, we obtain

$$y_1(s) = p_1 + s(\chi(-p_2) - \chi(-p_2 - t)), \tag{307}$$
$$y_2(s) = p_2 + t, \tag{308}$$
$$y_3(s) = p_3 + (\chi(p_2 + t) - \chi(p_2))\left(2p_1 s + s^2(\chi(-p_2) - \chi(-p_2 - t))\right). \tag{309}$$

In particular, with with $t = 2\varepsilon$ and $p = -\varepsilon e_2$, $e^{\mathcal{Z}_1(2\varepsilon,f_0,f_1)}e^{2\varepsilon f_0}(-\varepsilon e_2) = (\chi(\varepsilon), \varepsilon, \chi(\varepsilon)^2)$. Thus, for every $\varepsilon > 0$,

$$\left|x(2\varepsilon, f_0 + f_1, -\varepsilon e_2) - e^{\mathcal{Z}_1(2\varepsilon,f_0,f_1)}e^{2\varepsilon f_0}(-\varepsilon e_2)\right| = \chi^2(\varepsilon) > 0. \tag{310}$$

5.3.2 Convergence for linear systems

Let $T > 0$. In this paragraph, we study linear systems of the form

$$\dot{x}(t) = (H_0 + H_1(t))\,x(t) \quad \text{and} \quad x(0) = p, \tag{311}$$

where $H_0 \in \mathcal{M}_d(\mathbb{K})$ and $H_1 \in L^1((0,T); \mathcal{M}_d(\mathbb{K}))$. Let $\|\cdot\|$ be a sub-multiplicative norm on $\mathcal{M}_d(\mathbb{K})$.

Proposition 125. *Let $T > 0$, $H_0 \in \mathcal{M}_d(\mathbb{K})$ and $H_1 \in L^1((0,T);\mathcal{M}_d(\mathbb{K}))$ such that $\|H_1\|_{L^1(0,T)} < \frac{e^{-2T\|H_0\|}}{8}$. Then, for each $t \in [0,T]$ and $p \in \mathbb{K}^d$ the solution to (311) satisfies $x(t) = e^{-Z_\infty(t)}e^{tH_0}p$ where $Z_\infty(t)$ is defined by (264) with*

$$A_t(\tau) = e^{(t-\tau)H_0}H_1(\tau)e^{(\tau-t)H_0} = \sum_{k=0}^{+\infty} \frac{(t-\tau)^k}{k!} \operatorname{ad}_{H_0}^k(H_1). \tag{312}$$

Proof. The function $y : \tau \in [0,t] \mapsto e^{(t-\tau)H_0}x(\tau)$ satisfies $y'(\tau) = A(\tau)$, $y(0) = e^{tH_0}p$. Thus, by Proposition 111, $y(t) = e^{-Z_\infty(t)}e^{tH_0}p$, which gives the conclusion because $y(t) = x(t)$. □

Remark 126. *The Magnus expansion in the usual setting (Proposition 111), when applied directly to $A(t) = H_0 + H_1(t)$ requires a smallness assumption on $T\|H_0\|$ (through the condition $\|A\|_{L^1(0,T)} < \frac{1}{8}$), even for small perturbations H_1. On the contrary, the Magnus expansion in the interaction picture (Proposition 125) holds even when $T\|H_0\|$ is large, provided that the perturbation H_1 is small enough.*

Remark 127. *More generally, in [4, p. 1671], the authors consider the formal power series expressing the chronological logarithm of two flows, associated to two non-autonomous vector fields. They explain that, when the vector fields take values in a Banach algebra, and one of them is small enough, then this series converges. Proposition 125 is an illustration.*

5.3.3 Divergence for arbitrary small analytic vector fields

Generally speaking, since, as illustrated in Section 5.2.3, the Magnus expansion does not converge for analytic vector fields, one cannot expect that the Magnus expansion in interaction picture converges for analytic vector fields.

For instance, if $f_0 = 0$, or if, for some $a \in [\![1,d]\!]$, $f_0(x)$ is a linear combination of $e_1,\dots,e_a$ with coefficients depending only on $x_1,\dots,x_a$ and $f_1(t,x)$ is a linear combination of $e_{a+1},\dots e_d$, with coefficients depending only on $x_{a+1},\dots x_d$, then the vector field $g_t(\tau) = \Phi_0(t-\tau)_* f_1(\tau)$ defined in (210) and involved in the Magnus in the interaction picture formula satisfies $g_t(\tau) = f_1(\tau)$.

Hence, each counter-example to the convergence of the usual Magnus expansion also yields counter-examples to the convergence of the Magnus expansion in the interaction picture.

More generally, in [4, p. 1671], the authors consider the formal power series expressing the chronological logarithm of two flows, associated to two non-autonomous vector fields. They claim that, even for analytic vector fields, this series does not converge in general. The counter examples of the present article illustrate this assertion.

5.4 Sussmann's infinite product expansion

5.4.1 Equality for nilpotent systems

In this section, we study affine systems of the form (226).

Proposition 128. *Let $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$ and $(\xi_b)_{b\in\mathcal{B}}$ be the associated coordinates of the second kind. For every $M \in \mathbb{N}^*$, there exist $\eta_M > 0$ such that the following property holds. Let $T, \delta > 0$, $f_i \in \mathcal{C}^\infty(B_{3\delta};\mathbb{K}^d)$ and $u_i \in L^1((0,T);\mathbb{K})$ for $i \in I$. Assume that the Lie algebra generated by the f_i for $i \in I$ is nilpotent of index at most $M+1$. Then, under the smallness assumption (227), for each $t \in [0,T]$ and $p \in B_\delta$,*

$$x(t;f,u,p) = \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t,u)f_b}p. \tag{313}$$

Proof. The proof strategy is the same as for Proposition 104. We apply the second statement of Lemma 90 instead of the first one, which gives $\varepsilon_j = 0$ for each $j \in [\![0,k+1]\!]$. The smallness assumption guarantees that all flows are well-defined. □

5.4.2 Bilinear systems

Let $T > 0$. In this paragraph, we study the convergence of Sussmann's infinite product expansion for bilinear systems of the form

$$\dot{x}(t) = \left(\sum_{i\in I} u_i(t)A_i\right) x(t) \quad \text{and} \quad x(0) = p \tag{314}$$

where $A_i \in \mathcal{M}_d(\mathbb{K})$ are time-invariant and $u_i \in L^1((0,T);\mathbb{K})$. When well-defined, its solution is denoted $x(t;A,u,p)$. Local convergence is proved in Proposition 131 while an example illustrating the lack of global convergence is proposed in Proposition 132.

Local convergence. The main goal of this paragraph is to prove Proposition 131 which asserts that Sussmann's infinite product expansion for system (314) converges locally (i.e. for small matrices, small controls or small time).

Before proving this result, we need a definition for an ordered infinite product (given in Definition 129 below) and a sufficient condition for its convergence (given in Lemma 130 below).

Defining the ordered product of a family of matrices indexed by a length-compatible Hall basis is straightforward, because there exists an indexation of the family by $\mathbb{N}$ which is compatible with the order induced by the Hall basis (since it does not involve infinite segments). Hence, one is brought back to the classical case of a sequence of products and usual definitions and convergence criteria can be used.

For arbitrary Hall bases (in the generalized sense of Definition 54), the situation is more intricate, due to the potential infinite segments which can prevent the order of the basis from being compatible with the order of natural integers. For example, in the Lyndon basis of $\mathcal{L}(X)$ for $X = \{X_0, X_1\}$ with $X_0 < X_1$, $\mathrm{ad}^k_{X_0}(X_1) < (X_0, X_1)$ for all $k \geq 2$, so there is an infinite segment before (X_0, X_1). This problem already appears for a product which would be indexed by $\mathbb{N}^2$ with the lexicographic order

$$(0,0) < (0,1) < (0,2) < \cdots < (1,0) < (1,1) < (1,2) < \cdots < (2,0) < \cdots \tag{315}$$

We therefore propose a natural definition and a basic sufficient condition for convergence based on absolute convergence. In what follows, $\|\cdot\|$ is a submultiplicative norm on $\mathcal{M}_d(\mathbb{K})$ such that $\|\mathrm{Id}\| = 1$, for instance a subordinated norm.

Definition 129. *Let J be a totally ordered set and $(A_j)_{j\in J}$ matrices of $\mathcal{M}_d(\mathbb{K})$. We say that the ordered product of the e^{A_j} over J* converges *when there exists $M \in \mathcal{M}_d(\mathbb{K})$ such that, for every $\varepsilon > 0$, there exists a finite subset J_0 of J such that, for every finite subset J_1 of J containing J_0, one has*

$$\left\| M - \overleftarrow{\prod_{j\in J_1}} e^{A_j} \right\| \leq \varepsilon. \tag{316}$$

When such an M exists, it is unique and we write

$$M = \overleftarrow{\prod_{j\in J}} e^{A_j}. \tag{317}$$

The following natural convergence criteria also appears in [68].

Lemma 130. *Let J be a totally ordered set and $(A_j)_{j\in J}$ matrices of $\mathcal{M}_d(\mathbb{K})$ such that*

$$\sum_{j\in J} \|A_j\| < +\infty. \tag{318}$$

Then the ordered product of the e^{A_j} over J converges in the sense of Definition 129.

Proof. Let α be the left-hand side of (318).

Step 1: Basic claims. We start with straightforward claims. First, for every $j \in J$, one has

$$\|e^{A_j} - \mathrm{Id}\| \leq e^{\|A_j\|} - 1 \leq \|A_j\| e^{\|A_j\|} \leq \|A_j\| e^{\alpha}. \tag{319}$$

Second, for every finite part $J' \subset J$, one has

$$\left\| \overleftarrow{\prod_{j \in J'}} e^{A_j} \right\| \leq \prod_{j \in J'} e^{\|A_j\|} \leq e^{\alpha}. \tag{320}$$

Third, for every finite parts $J_0 \subset J_1 \subset J$, one has

$$\left\| \overleftarrow{\prod_{j \in J_1}} e^{A_j} - \overleftarrow{\prod_{j \in J_0}} e^{A_j} \right\| \leq e^{3\alpha} \sum_{j \in J_1 \setminus J_0} \|A_j\|. \tag{321}$$

Indeed, writing $J_1 \setminus J_0 = \{j_1 > \cdots > j_n\}$, we have the following telescopic decomposition

$$\overleftarrow{\prod_{j \in J_1}} e^{A_j} - \overleftarrow{\prod_{j \in J_0}} e^{A_j} = \sum_{k=1}^{n} \overleftarrow{\prod_{\substack{j \in J_0 \\ j > j_k}}} e^{A_j} \left(e^{A_{j_k}} - \mathrm{Id}\right) \overleftarrow{\prod_{\substack{j \in J_1 \\ j < j_k}}} e^{A_j}, \tag{322}$$

which, together with the two first claims, proves estimate (321).

Step 2: Construction of a limit. We construct a possible limit. For each $n \geq 2$, let

$$J_n := \left\{ j \in J, \ \ \|A_j\| > \frac{1}{n} \right\}. \tag{323}$$

Thanks to assumption (318), the sets J_n are finite and, moreover,

$$\varepsilon_n := \sum_{j \in J \setminus J_n} \|A_j\| \to 0. \tag{324}$$

Now, for each $n \geq 2$, we define the matrix

$$M_n := \overleftarrow{\prod_{j \in J_n}} e^{A_j}. \tag{325}$$

This defines a Cauchy sequence in the complete space $\mathcal{M}_d(\mathbb{K})$. Indeed, for every $n < p$, thanks to estimate (321), one has

$$\|M_n - M_p\| \leq e^{3\alpha} \varepsilon_n. \tag{326}$$

Hence, there exists $M \in \mathcal{M}_d(\mathbb{K})$ towards which the sequence $(M_n)_{n \geq 2}$ converges. By letting $[p \to \infty]$ in the previous inequality we obtain, for every $n \geq 2$

$$\|M_n - M\| \leq e^{3\alpha} \varepsilon_n. \tag{327}$$

Step 3: Proof of convergence. We now prove that the ordered product of the e^{A_j} over J converges to M in the sense of Definition 129. Let $\varepsilon > 0$. Let $n \geq 2$ large enough such that $e^{3\alpha} \varepsilon_n < \varepsilon/2$. For every finite set J_1 containing J_n, condition (316) holds thanks to (327) and (321). □

Proposition 131. *Let $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$, $(\xi_b)_{b \in \mathcal{B}}$ be the coordinates of the second kind associated to $\mathcal{B}$. There exists $\eta > 0$ such that the following property holds. Let $A_i \in \mathcal{M}_d(\mathbb{K})$ for $i \in I$. For $b \in \mathcal{B}$, we define the matrix $A_b := \Lambda(b)$ where $\Lambda : \mathcal{L}(X) \to \mathcal{M}_n(\mathbb{K})$ is the homomorphism of Lie algebras such that $\Lambda(X_i) = A_i$ for $i \in I$ (see Lemma 7). Let $T > 0$ and $u_i \in L^1((0,T);\mathbb{K})$ for $i \in I$. Assume that*

$$\|u\|_{L^1(0,T)} \|A\| \leq \eta. \tag{328}$$

Then, for each $t \in [0,T]$ and $p \in \mathbb{K}^d$, the ordered product of the $e^{\xi_b(t,u)A_b}$ over $b \in \mathcal{B}$ converges. Moreover, for every $p \in \mathbb{K}^d$,

$$x(t; A, u, p) = \overrightarrow{\prod_{b \in \mathcal{B}}} e^{\xi_b(t,u)A_b} p. \tag{329}$$

Proof. Let $\eta := 1/(8|I|^2)$. Let $T > 0$. Below, the variable t implicitly belongs to $[0,T]$. To simplify the notations we write $\xi_b(t)$ instead of $\xi_b(t,u)$.

Step 1: Convergence of the ordered product of the $e^{\xi_b(t)A_b}$ over $b \in \mathcal{B}$. One obtains, by induction on $|b|$, that for every $b \in \mathcal{B}$, $\|A_b\| \le (2\|A\|)^{|b|}$. Thus, recalling (107),

$$\|\xi_b(t)A_b\| \le \left(2\|A\|\|u\|_{L^1(0,t)}\right)^{|b|}. \tag{330}$$

Taking into account that $|\mathcal{B}_\ell| \le |I|^\ell$, we obtain, using (328),

$$\sum_{b\in\mathcal{B}} \|\xi_b(t)A_b\| \le \sum_{\ell=1}^{+\infty} \left(2|I|\|A\|\|u\|_{L^1(0,t)}\right)^\ell \le 1 \tag{331}$$

and Lemma 130 gives the conclusion.

Step 2: Estimates along a Lazard elimination in $\mathcal{B}_{[\![1,M]\!]}$. Let $M \in \mathbb{N}^*$. We adopt the notations $b_1,\dots,b_{k+1}$ and $Y_0,\dots,Y_{k+1}$ of Definition 53 and we define $x_0(t) := x(t)$ and, for $j \in [\![1,k+1]\!]$

$$x_j(t) := e^{-\xi_{b_j}(t,u)A_{b_j}} \cdots e^{-\xi_{b_1}(t,u)A_{b_1}} x(t). \tag{332}$$

We prove by induction on $j \in [\![0,k+1]\!]$ that

$$(\mathcal{H}_j): \begin{cases} \dot{x}_j(t) = \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_j} \dot{\xi}_b(t)A_b + \varepsilon_j(t)\right) x_j(t), \\ x_j(0) = p, \end{cases} \tag{333}$$

where $\varepsilon_0 = 0$ and

$$\|\varepsilon_j(t)\| \le \left(M|I|\|A\||u(t)|(4|I|\|A\|\|u\|_{L^1(0,t)})^M + \|\varepsilon_{j-1}(t)\|\right) e^{2\|\xi_{b_j}(t)A_{b_j}\|}. \tag{334}$$

First, $(\mathcal{H}_0)$ holds with $\varepsilon_0 = 0$ because $x_0(t) = x(t)$ and $\dot{\xi}_{X_i}(t) = u_i(t)$ for $i \in I$. Let $j \in [\![1,k+1]\!]$ and assume that $(\mathcal{H}_{j-1})$ holds. We deduce from the definition of x_j that

$$x_j(t) = e^{-\xi_{b_j}(t)A_{b_j}} x_{j-1}(t) \tag{335}$$

and from $(\mathcal{H}_{j-1})$ that

$$\begin{aligned} \dot{x}_j(t) &= -\dot{\xi}_{b_j}(t)A_{b_j}x_j(t) + e^{-\xi_{b_j}(t)A_{b_j}} \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}} \dot{\xi}_b(t)A_b + \varepsilon_{j-1}(t)\right) e^{\xi_{b_j}(t)A_{b_j}} x_j(t) \\ &= \left(\sum_{b\in\mathcal{B}_{[\![1,M]\!]}\cap Y_{j-1}\setminus\{b_j\}} \dot{\xi}_b(t) e^{-\xi_{b_j}(t)A_{b_j}} A_b e^{\xi_{b_j}(t)A_{b_j}} + \widetilde{\varepsilon}_{j-1}(t)\right) x_j(t) \end{aligned} \tag{336}$$

where $\widetilde{\varepsilon}_{j-1}(t) := e^{-\xi_{b_j}(t)A_{b_j}} \varepsilon_{j-1}(t) e^{\xi_{b_j}(t)A_{b_j}}$ satisfies,

$$\|\widetilde{\varepsilon}_{j-1}(t)\| \le \|\varepsilon_{j-1}(t)\| e^{2\|\xi_{b_j}(t)A_{b_j}\|}. \tag{337}$$

For $b \in \mathcal{B}_{[\![1,M]\!]} \cap Y_{j-1} \setminus \{b_j\}$, let $h(b) \in \mathbb{N}^*$ be the maximal integer such that (241) holds and

$$\overline{\varepsilon}_b^j(t) := \dot{\xi}_b(t) e^{-\xi_{b_j}(t)A_{b_j}} A_b e^{\xi_{b_j}(t)A_{b_j}} - \sum_{m=0}^{h(b)-1} \dot{\xi}_b(t) \frac{\xi_{b_j}^m(t)}{m!} A_{\mathrm{ad}_{b_j}^m(b)}. \tag{338}$$

Then, by definition of Y_j, $(\mathcal{H}_j)$ holds with ε_j defined by (245). Using the fourth statement of Lemma 90, (106), (330), we obtain

$$\begin{aligned}\|\overline{\varepsilon}_b^j(t)\| &\leq |\dot{\xi}_b(t)|\frac{(2\|\xi_{b_j}(t)A_{b_j}\|)^{h(b)}}{h(b)!}\|A_b\|e^{2\|\xi_{b_j}(t)A_{b_j}\|}\\ &\leq |b||u(t)|\|u\|_{L^1}^{|b|-1}\cdot\frac{(2\|A\|\|u\|_{L^1})^{h(b)|b_j|}2^{h(b)}}{h(b)!}\cdot(2\|A\|)^{|b|}\\ &\leq M|u(t)|(4\|A\|\|u\|_{L^1})^M\|A\|,\end{aligned} \tag{339}$$

taking into account $M+1 \leq |b|+h(b)|b_j| \leq 2M$ and $\|A\|\|u\|_{L^1} \leq 1$.

We deduce from (245), (337), (339) and the relation $|\mathcal{B}_{[\![1,M]\!]}| \leq |I|^{M+1}$ that (334) holds.

Step 4: Proof of an estimate on the ordered product of the $e^{\xi_b(t)A_b}$ over $\mathcal{B}_{[\![1,M]\!]}$. We deduce from (334), (331) and the relation $k+1 = |\mathcal{B}_{[\![1,M]\!]}| \leq |I|^{M+1}$ that

$$\|\varepsilon_{k+1}(t)\| \leq eM\|A\||I|^2|u(t)|(4|I|^2\|A\|\|u\|_{L^1(0,t)})^M. \tag{340}$$

Hence, using (328),

$$\|\varepsilon_{k+1}\|_{L^1(0,t)} \leq \frac{e}{4}(4|I|^2\|A\|\|u\|_{L^1(0,t)})^{M+1} \leq 2^{-M}. \tag{341}$$

We deduce from $(\mathcal{H}_{k+1})$, (99) and Grönwall's lemma that

$$\left|\overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{-\xi_b(t,u)A_b}x(t)-p\right| = |x_{k+1}(t)-p| \leq \int_0^t |\varepsilon_{k+1}(\tau)x_{k+1}(\tau)|d\tau \leq 2^{-M}e|p| \tag{342}$$

Multiplying both sides by the finite product $\overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t,u)A_b}$ gives

$$\left|x(t)-\overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,M]\!]}}} e^{\xi_b(t,u)A_b}p\right| \leq e^2 2^{-M}|p| \tag{343}$$

Passing to the limit $[M\to\infty]$ in the previous estimate gives (329). □

Lack of global convergence. The goal of this paragraph is to illustrate that the smallness assumption (328) in Proposition 131 is necessary because the equality does not hold globally.

Proposition 132. *Consider the constant control $\overline{u} : t \in \mathbb{R}_+ \mapsto (1,1) \in \mathbb{R}^2$.*

1. *There exist a Hall basis $\mathcal{B}$ of $\mathcal{L}(\{X_1,X_2\})$ and a subsequence $(b_k)_{k\in\mathbb{N}}$ of $\mathcal{B}$ such that*

$$\exists\gamma>0, \forall k\in\mathbb{N}, t\geq 0, \quad \xi_{b_k}(t,\overline{u}) \geq \left(\frac{t}{\gamma}\right)^{|b_k|} \tag{344}$$

2. *There exists $A_1, A_2 \in \mathcal{M}_3(\mathbb{C})$ and $t>\gamma$ such that $(e^{\xi_{b_k}(t,\overline{u})A_{b_k}})_{k\in\mathbb{N}}$ does not converge to Id_3 in $\mathcal{M}_3(\mathbb{C})$. Thus, the ordered product of the $e^{\xi_b(t,\overline{u})A_b}$ over $\mathcal{B}$ does not converge in $\mathcal{M}_3(\mathbb{C})$.*

Proof. For the first point we adapt an argument due to Sussmann in [100, pages 333-335]. We define by induction two sequences $(b_k^1)_{k\in\mathbb{N}}$ and $(b_k^2)_{k\in\mathbb{N}}$ of $\mathrm{Br}(\{X_1,X_2\})$ by

$$b_0^1 = X_1, \qquad b_0^2 = X_2, \qquad b_{k+1}^1 := [b_k^2,[b_k^1,b_k^2]], \qquad b_{k+1}^2 := [b_k^1,[b_k^1,b_k^2]]. \tag{345}$$

There exists a Hall basis of $\mathcal{L}(\{X_1,X_2\})$, whose order, denoted $<$, is compatible with length and such that, for every $k\in\mathbb{N}$, $b_k^1, b_k^2 \in \mathcal{B}$ and $b_k^1 < b_k^2$. It suffices to choose, on the brackets with length 3^k, some order such that $b_k^1 < b_k^2$. Then, automatically, $[b_k^1,b_k^2] \in \mathcal{B}$ and thus $b_{k+1}^1, b_{k+1}^2 \in \mathcal{B}$. Such

a process indeed allows to construct a Hall basis (see Remark 55), provided that one chooses an arbitrary length-compatible order on all other brackets.

To lighten the notations, we write $\xi_b(t)$, instead of $\xi_b(t,\overline{u})$. We have $\xi_{X_1}(t) = \xi_{X_2}(t) = t$. An easy induction shows that, for every $b \in \mathcal{B}$, $\dot{\xi}_b(t) = t^{|b|-1}/\alpha_b$, where $\alpha_b \in \mathbb{N}^*$. The constants α_b can be computed recursively: $\alpha_{X_1} = \alpha_{X_2} = 1$ and, if $b = \mathrm{ad}_{b_1}^m(b_2)$ with $m \in \mathbb{N}^*$, $b_1 < b_2$ and $\lambda(b_2) < b_1$ then $\alpha_b = \alpha_{b_1}^m |b_1|^m m! \alpha_{b_2}$. In particular, for every $k \in \mathbb{N}$,

$$\alpha_{b_{k+1}^1} = \alpha_{b_k^1}\alpha_{b_k^2}^2 |b_k^2||b_k^1| = \alpha_{b_k^1}\alpha_{b_k^2}^2 3^{2k}, \qquad \alpha_{b_{k+1}^2} = 2\alpha_{b_k^1}^2 \alpha_{b_k^2} |b_k^1|^2 = 2\alpha_{b_k^1}^2 \alpha_{b_k^2} 3^{2k}. \tag{346}$$

Let $\beta_k = \max\{\alpha_{b_k^1}, \alpha_{b_k^2}\}$. Then, $\beta_0 = 1$ and, by the previous relations,

$$\beta_{k+1} \leq 3^{2k+1}\beta_k^3. \tag{347}$$

Thus $\theta_k := 3^{-k}\ln(\beta_k)$ satisfies $\theta_0 = 0$ and

$$\theta_{k+1} \leq \theta_k + (2k+1)3^{-(k+1)}\ln(3), \tag{348}$$

which leads to $\theta_k \leq \eta := \sum_{j=1}^{+\infty}(2j+1)3^{-(j+1)}\ln(3)$ i.e. $\beta_k \leq (\gamma')^{3^k}$ where $\gamma' = e^\eta$. Therefore, for every $k \in \mathbb{N}$ and $j \in \{1,2\}$ we have

$$\left|\xi_{b_k^j}(t)\right| \geq \frac{1}{|b_k^j|}\left(\frac{t}{\gamma'}\right)^{|b_k^j|}. \tag{349}$$

Let $\gamma > \gamma'$ be such that, for every $k \in \mathbb{N}$, $\frac{1}{3^k}\left(\frac{\gamma}{\gamma'}\right)^{3^k} \geq 1$. Then (344) holds, for instance with $b_k = b_k^1$.

For the second point, let, for $j \in \{1,2,3\}$, $F_j \in \mathcal{M}_3(\mathbb{R})$ be the matrix of the linear map $x \in \mathbb{R}^3 \mapsto e_j \wedge x$. Then $[F_1,F_2] = F_3$, $[F_2,F_3] = F_1$ and $[F_3,F_1] = F_2$. In particular

$$[F_2,[F_1,F_2]] = F_1, \quad [F_1,[F_1,F_2]] = -F_2. \tag{350}$$

We consider $A_1 = e^{i\frac{\pi}{6}}F_1$ and $A_2 = e^{i\frac{\pi}{6}}F_2$ in $\mathcal{M}_3(\mathbb{C})$. One easily proves by induction on $k \in \mathbb{N}^*$ that $A_{b_k^1} = (-1)^{k+1}iF_1$ and $A_{b_k^2} = -iF_2$. We have, for every $k \in \mathbb{N}$ and $t \in \mathbb{R}$

$$e^{\xi_{b_k}(t)A_{b_k}} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \cosh(\xi_{b_k}(t)) & i(-1)^k\sinh(\xi_{b_k}(t)) \\ 0 & i(-1)^{k+1}\sinh(\xi_{b_k}(t)) & \cosh(\xi_{b_k}(t)) \end{pmatrix} \tag{351}$$

By (344), this sequence of matrices diverges for every $t > \gamma$. □

5.4.3 Investigation for analytic vector fields

In this paragraph, we study affine systems (226). Our goal is to explain the difficulty of the convergence question for Sussmann's infinite product for arbitrary analytic vector fields. First, we state a definition (Definition 133) and a sufficient condition for the convergence (Lemma 134), in the same spirit as for matrices. Then we show that they do not provide convergence for general analytic vector fields and we formulate an open problem.

Definition 133. *Let J be a totally ordered set, $\delta > 0$ and $(f_j)_{j\in J}$ a family of $\mathcal{C}^1(B_{2\delta};\mathbb{K}^d)$. We say that the ordered product of the e^{f_j} over J converges uniformly on B_δ if there exists $g \in \mathcal{C}^0(B_\delta;\mathbb{K}^d)$*

such that, for every $\varepsilon > 0$, there exists a finite subset J_0 of J such that, for every finite subset J_1 of J containing J_0, and for every $p \in B_\delta$ one has

$$\left\| g(p) - \overleftarrow{\prod_{j\in J_1}} e^{f_j} p \right\| \leq \varepsilon. \tag{352}$$

When such a g exists, it is unique and we write

$$g = \overleftarrow{\prod_{j\in J}} e^{f_j}. \tag{353}$$

Lemma 134. *Let J be a totally ordered set, $\delta > 0$ and $(f_j)_{j\in J}$ a family of $\mathcal{C}^1(B_{2\delta}; \mathbb{K}^d)$ such that*

$$\sum_{j\in J} \|f_j\|_{\mathcal{C}^0} < \delta \quad \text{and} \quad \alpha := \sum_{j\in J} \|f_j\|_{\mathcal{C}^1} < \infty. \tag{354}$$

Then the ordered product of the e^{f_j} over J converges uniformly on B_δ and is e^α-Lipschitz.

Proof. We proceed as in the proof of Lemma 130.

Step 1: Basic claims. First, for every finite subset $J' \subset J$ and $p \in \mathbb{K}^d$ with $|p| \leq 2\delta - \sum_{j\in J'} \|f_j\|_{\mathcal{C}^0}$, then

$$\overleftarrow{\prod_{j\in J'}} e^{f_j} p \in B_{2\delta} \qquad \text{and} \qquad \left\| \partial_p \left[\overleftarrow{\prod_{j\in J'}} e^{f_j} p \right] \right\| \leq \prod_{j\in J'} e^{\|f_j'\|_{\mathcal{C}^0}} \leq e^\alpha \tag{355}$$

because of Lemma 85 and the chain rule.

Second, for every finite parts $J_0 \subset J_1 \subset J$ and $p \in \mathbb{K}^d$ with $|p| \leq 2\delta - \sum_{j\in J_1} \|f_j\|_{\mathcal{C}^0}$ one has

$$\left\| \overleftarrow{\prod_{j\in J_1}} e^{f_j} p - \overleftarrow{\prod_{j\in J_0}} e^{f_j} p \right\| \leq e^\alpha \sum_{j\in J_1\setminus J_0} \|f_j\|_{\mathcal{C}^0}. \tag{356}$$

Indeed, writing $J_1 \setminus J_0 = \{j_1 > \cdots > j_n\}$, we have the following telescopic decomposition

$$\overleftarrow{\prod_{j\in J_1}} e^{f_j} p - \overleftarrow{\prod_{j\in J_0}} e^{f_j} p = \sum_{k=1}^{n} \left\{ \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) e^{f_{j_k}} \left(\overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{f_j} \right) p - \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) \left(\overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{f_j} \right) p \right\}. \tag{357}$$

For $k \in [\![1, n]\!]$, let $x_k := \overleftarrow{\prod_{\substack{j\in J_1 \\ j<j_k}}} e^{f_j} p$ which is a point in $B_{2\delta - \|f_{j_k}\|_{\mathcal{C}^0}}$. By (355) and (149), the term with index k in the previous sum is bounded by

$$\left| \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) e^{f_{j_k}} x_k - \left(\overleftarrow{\prod_{\substack{j\in J_0 \\ j>j_k}}} e^{f_j} \right) x_k \right| \leq e^\alpha \left| e^{f_{j_k}} x_k - x_k \right| \leq e^\alpha \|f_{j_k}\|_{\mathcal{C}^0}. \tag{358}$$

which, together with (357) proves (356).

Step 2: Construction of a limit. We construct a possible limit. For each $n \geq 2$, let

$$J_n := \left\{ j \in J, \ \|f_j\|_{\mathcal{C}^1} > \frac{1}{n} \right\}. \tag{359}$$

Thanks to assumption (354), the sets J_n are finite and, moreover,

$$\varepsilon_n := \sum_{j\in J\setminus J_n} \|f_j\|_{\mathcal{C}^1} \to 0. \tag{360}$$

Now, for each $n \geq 2$, we define $g_n \in \mathcal{C}^0(B_\delta; \mathbb{K}^d)$ by

$$g_n(p) := \overleftarrow{\prod_{j \in J_n}} e^{f_j} p. \tag{361}$$

This defines a Cauchy sequence in the complete space $\mathcal{C}^0(B_\delta; \mathbb{K}^d)$. Indeed, for every $n < n'$ and $p \in B_\delta$, thanks to estimate (356), one has

$$|g_n(p) - g_{n'}(p)| \leq e^\alpha \varepsilon_n. \tag{362}$$

Hence, there exists $g \in \mathcal{C}^0(B_\delta; \mathbb{K}^d)$ towards which the sequence $(g_n)_{n \geq 2}$ uniformly converges on B_δ. By (355), g_n is e^α-Lipschitz on B_δ for every $n \in \mathbb{N}$, thus so is g. By letting $[n' \to \infty]$ in the previous inequality we obtain, for every $n \geq 2$ and $p \in B_\delta$

$$|g_n(p) - g(p)| \leq e^\alpha \varepsilon_n. \tag{363}$$

Step 3: Proof of convergence. We now prove that the ordered product of the e^{f_j} over J converges uniformly to g on B_δ in the sense of Definition 133. Let $\varepsilon > 0$. Let $n \geq 2$ large enough such that $e^\alpha \varepsilon_n < \varepsilon/2$. For every finite set J_1 containing J_n, condition (352) holds thanks to (363) and (356). □

Now, let us emphasize that, by using estimates on $\xi_b(t,u)$ and f_b depending only on the length of the Lie bracket b, it is not possible to prove the convergence of $\sum |\xi_b(t,u)| \|f_b\|_{\mathcal{C}^1}$, where the sum ranges over $b \in \mathcal{B}$, an arbitrary Hall basis of $\mathcal{L}(X)$.

On the one hand, one easily proves by induction on $|b|$ that, for every $b \in \mathcal{B}$ and $u \in L^\infty$ with $\|u\|_{L^\infty} \leq 1$, there holds $|\xi_b(t,u)| \leq t^{|b|}$. However, by the first statement of Proposition 132, when X contains at least two indeterminates, there are Hall bases (even compatible with length) for which one may not expect an upper bound, function of $|b|$ alone, that behaves better than geometrically. Hence, we should consider the $t^{|b|}$ bound to be sharp, when one restricts to bounds depending only on $|b|$.

On the other hand, if the vector fields are locally analytic, there exists $r, \delta > 0$ such that $f_i \in \mathcal{C}^{\omega,r}(B_\delta; \mathbb{K}^d)$ for $i \in I$. By (140) with $r_1 \leftarrow r$ and $r_2 \leftarrow r/e$ for every $b \in \mathcal{B}$,

$$\|f_b\|_{\mathcal{C}^1} \leq \left(1 + \frac{e}{r}\right) (|b| - 1)! \left(\frac{9}{r}\right)^{|b|-1} F^{|b|}, \tag{364}$$

where $F := \max_{i \in I} |||f_i|||_r$. However, by Remark 81, the dependence in $(|b|-1)!$ is optimal (again, if one restricts to bounds involving only $|b|$).

We deduce from the previous estimates that there exists $C > 0$ such that

$$|\xi_b(t,u)| \|f_b\|_{\mathcal{C}^1} \leq (Ct)^{|b|} |b|!. \tag{365}$$

This bound does not provide the convergence of the considered series. Indeed, for every $t > 0$, $(Ct)^{|b|}|b|! \to +\infty$ as $|b| \to +\infty$, so an argument depending on $|b|$ alone doesn't even prove that the general term tends to zero.

To prove the convergence of Sussmann's infinite product expansion, one therefore either needs a better sufficient condition than Lemma 134 or one needs to prove estimates on ξ_b and f_b that take into account the structure of the bracket b, and not only its length.

Open problem 135. *Does Sussmann's infinite product converge for analytic vector fields?*

In Section 6.4, we prove the convergence (for analytic vector fields) of some infinite subproducts, by applying Lemma 134 with estimates on ξ_b that depend on the structure of b.

6 Error estimates for control systems

In this section, we consider control-affine systems with drift, i.e. of the form

$$\dot{x}(t) = f_0(x(t)) + \sum_{i=1}^{q} u_i(t) f_i(x(t)) \quad \text{and} \quad x(0) = p, \tag{366}$$

where $f_0, \dots, f_q$ are vector fields and $u = (u_1, \dots, u_q) \in L^1(\mathbb{R}; \mathbb{K}^q)$. When well-defined, the solution is denoted $x(t; f, u, p)$ where $f = (f_0, \dots, f_q)$ and $u = (u_1, \dots, u_q)$.

We prove error formulas at every order in $\|u\|_{L^1}$ for the Chen-Fliess expansion, the Magnus expansion in the interaction picture and for Sussmann's infinite product expansion. In each case, the error formula involves an infinite sum or an infinite product which turns out to be well-defined. We also propose a counter-example for the validity of such error estimates for the usual Magnus expansion, for which the infinite sum involved is not well-defined.

6.1 Chen-Fliess expansion

The convergence of the Chen-Fliess series, for control affine systems (366) with analytic vector fields, under a smallness assumption on t and a uniform bound on u, is a classical result, see for instance [36, Proposition 3.37] or [99, Proposition 4.3]. In this section we prove the convergence of the Chen-Fliess expansion, (Proposition 136) under a smallness assumption on $\|u\|_{L^1}$. We also generalize the Chen-Fliess expansion to nonlinear systems (not necessarily affine) with scalar input (Proposition 137), because this fact will be used in Section 7.2.

In the following statement $q \in \mathbb{N}^*$, $I = [\![0, q]\!]$. For a word $\sigma = \sigma_1 \cdots \sigma_\ell \in I^*$, with $\ell \in \mathbb{N}^*$, $\sigma_1, \dots, \sigma_\ell \in I$, and vector fields $f_0, f_1, \dots, f_q$, we denote by f_σ the differential operator $f_{\sigma_1} \cdots f_{\sigma_\ell}$ (with the notations of Remark 74). For $t > 0$ and $u = (u_1, \dots, u_q) \in L^1(0, t)$, the quantity $\int_0^t u_\sigma$ is defined in (19), with $u_0 = 1$.

Proposition 136. *Let $\delta, r > 0$ and $f_0, f_1, \dots, f_q \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K}^d)$. There exists $\eta > 0$ such that, for every $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K})$, $t \in [0, \eta]$ and $u \in L^1((0,t); \mathbb{K}^q)$ such that $\|u\|_{L^1} \leq \eta$ and $p \in B_\delta$, then*

$$\varphi(x(t; f, u, p)) = \sum_{\sigma \in I^*} \left(\int_0^t u_\sigma \right) (f_\sigma \varphi)(p) \tag{367}$$

where the sum converges absolutely, uniformly with respect to (t, u, p). Moreover, for every $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta}; \mathbb{K})$, there exists $C > 0$ such that, for every $M \in \mathbb{N}$, $p \in B_\delta$, $t \in [0, \eta]$ and $u \in L^1((0,t); \mathbb{K}^q)$ such that $\|u\|_{L^1} \leq \eta$, then

$$\left| \varphi(x(t; f, u, p)) - \sum_{n(\sigma) \leq M} \left(\int_0^t u_\sigma \right) (f_\sigma \varphi)(p) \right| \leq (C \|u\|_{L^1})^{M+1}, \tag{368}$$

where the sum ranges over words $\sigma \in I^*$ such that the number of non-zero letters is at most M.

Proof. For $\sigma = \sigma_1 \cdots \sigma_\ell \in I^*$, let $n(\sigma)$ be the number of non zero letters in σ, i.e. $n(\sigma) = |\{i \in [\![1, \ell]\!]; \sigma_i \neq 0\}|$ and $n_0(\sigma)$ be the number of occurrences of the letter zero in σ, i.e. $n_0(\sigma) = |\{i \in [\![1, \ell]\!]; \sigma_i = 0\}|$. Then $\ell = n(\sigma) + n_0(\sigma)$. One proves by induction on the length ℓ of $\sigma \in I^*$ the following estimate, for every $t > 0$ and $u \in L^1((0,t); \mathbb{K}^q)$,

$$\left| \left(\int_0^t u_\sigma \right) \right| \leq \frac{\|u\|_{L^1(0,t)}^{n(\sigma)}}{n(\sigma)!} \frac{t^{n_0(\sigma)}}{n_0(\sigma)!}. \tag{369}$$

Let $\|f\| = \sum_{i=0}^{q} |||f_i|||_r$, $\eta = r/(10\|f\|)$, $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta};\mathbb{K})$, $t \in [0,\eta]$ and $u \in L^1((0,t);\mathbb{K}^q)$ such that $\|u\|_{L^1(0,t)} = \sum_{i=1}^{q} \|u_i\|_{L^1(0,t)} \leq \eta$ and $p \in B_\delta$. Using (369) and (135), we get

$$\left| \left(\int_0^t u_\sigma \right) (f_\sigma \varphi)(p) \right| \leq \|u\|_{L^1(0,t)}^{n(\sigma)} t^{n_0(\sigma)} \left(\frac{10}{r} \|f\| \right)^{\ell} |||\varphi|||_r \tag{370}$$

which proves the absolute convergence of the sum in (367), uniformly with respect to (t,u,p)

The proof of the equality in (367) consists in applying (255) to $f(t,x) = f_0(x) + \sum_{i=1}^{q} u_i(t) f_i(x)$. In particular the sum involved in (368) is the Taylor expansion of order M of $u \mapsto \varphi(x(t;f,u,p))$ at $u = 0$. By adapting Lemma 83 to affine systems with L^1 controls, we get the real-analyticity of the map $u \mapsto \varphi(x(t;f,u,p))$ on $B_{L^1(0,t)}(0,\eta)$ uniformly with respect to $(t,p) \in [0,\eta] \times B_\delta$ which ends the proof of (368). □

The last statement of this section focuses on nonlinear control systems with scalar input

$$\dot{x} = f(x,u) \tag{371}$$

where $f : \mathbb{K}^d \times \mathbb{K} \to \mathbb{K}^d$, When well-defined, the solution of this ODE, with initial condition $x(0) = p$ is denoted $x(t;f,u,p)$. We introduce the notation

$$\int_0^t u^k := \int_{\Delta^n(t)} u(\tau_1)^{k_1} \cdots u(\tau_n)^{k_n} \, \mathrm{d}\tau \tag{372}$$

for every $t > 0$, $u \in L^1((0,t);\mathbb{K})$, and every multi-index $k = (k_1,\ldots,k_n) \in \mathbb{N}^n$ with $n \in \mathbb{N}^*$.

Proposition 137. *Let $r,\delta,\delta_u > 0$, $f \in \mathcal{C}^{\omega,r}(B_{2\delta} \times [-\delta_u,\delta_u];\mathbb{K}^d)$ and $f_k := \frac{1}{k!}\partial_u^k f(\cdot,0)$ for every $k \in \mathbb{N}$. There exists $T^*,\eta > 0$ such that, for every $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta};\mathbb{K})$, $t \in [0,T^*]$, $u \in L^\infty((0,t);\mathbb{K})$ with $\|u\|_{L^\infty} \leq \eta$ and $p \in B_\delta$, with the notations of Remark 74,*

$$\varphi\left(x(t;f,u,p)\right) = \sum_{\substack{n \in \mathbb{N} \\ k \in \mathbb{N}^n}} \left(\int_0^t u^k \right) (f_{k_1} \cdots f_{k_n})(\varphi)(p), \tag{373}$$

where the sum converges absolutely, uniformly with respect to (t,u,p). Moreover, for every $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta};\mathbb{K})$, there exists $C > 0$ such that, for every $M \in \mathbb{N}$, $t \in [0,T^*]$, $u \in L^\infty((0,t);\mathbb{K})$ with $\|u\|_{L^\infty} \leq \eta$ and $p \in B_\delta$

$$\left| \varphi\left(x(t;f,u,p)\right) - \sum_{\substack{n \in \mathbb{N} \\ k \in \mathbb{N}^n, |k| \leq M}} \left(\int_0^t u^k \right) (f_{k_1} \cdots f_{k_n})(\varphi)(p) \right| \leq (C\|u\|_{L^\infty})^{M+1} \tag{374}$$

where the sum is taken over $n \in \mathbb{N}$ and $k = (k_1,\ldots,k_n) \in \mathbb{N}^n$ such that $k_1 + \cdots + k_n \leq M$.

Proof. We define $r' = r/e$,

$$T^* := \min\left\{ \frac{r'}{10\,|||f|||_r}, \frac{\delta}{\|f\|_{\mathcal{C}^0}} \right\}, \qquad \eta := \min\left\{ \delta_u, \frac{r}{10} \right\}. \tag{375}$$

Let $\varphi \in \mathcal{C}^{\omega,r}(B_{2\delta};\mathbb{K})$, $t \in [0,T^*]$, $u \in L^\infty((0,t);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \leq \eta$ and $p \in B_\delta$. Then $x(t;f,u,p) \in B_{2\delta}$.

Step 1: Uniform absolute convergence of the sum in (373). Using the iterated version of (127) and (116), we get, for every $k \in \mathbb{N}$,

$$|||f_k|||_{r'} \leq \frac{1}{k!} \left(\frac{k}{r-r'} \right)^k |||f|||_r \leq \left(\frac{e}{r-r'} \right)^k |||f|||_r \leq \left(\frac{5}{r} \right)^k |||f|||_r . \tag{376}$$

For every $n \in \mathbb{N}^*$ and $k_1, \dots, k_n \in \mathbb{N}$, we have, using (135) and (376)

$$\begin{aligned} |(f_{k_1} \cdots f_{k_n})(\varphi)(p)| &\leq n! \left(\frac{5}{r'}\right)^n |||f_{k_n}|||_{r'} \cdots |||f_{k_1}|||_{r'} \, |||\varphi|||_{r'} \\ &\leq n! \left(\frac{5}{r'}\right)^n \left(\frac{5}{r}\right)^{k_1+\cdots+k_n} |||f|||_r^n \, |||\varphi|||_{r'} \end{aligned} \tag{377}$$

and

$$\left| \int_0^t u^k \right| = \left| \int_{\Delta^n(t)} u(\tau_1)^{k_1} \cdots u(\tau_n)^{k_n} \, \mathrm{d}\tau \right| \leq \frac{t^n}{n!} \|u\|_{L^\infty}^{k_1+\cdots+k_n}. \tag{378}$$

By definition of T^* and η we have $\frac{5t}{r'} |||f|||_r \leq \frac{1}{2}$ and $\frac{5}{r}\|u\|_{L^\infty} \leq \frac{1}{2}$, which gives the conclusion.

Step 2: Equality in (373) and error formula (374). We have $f(\cdot, u) = \sum_{j=0}^{+\infty} u^j f_j$ with convergence in $\mathcal{C}^{\omega,r'}(B_{2\delta}; \mathbb{K}^d)$ uniformly with respect to $u \in B_{\mathbb{K}^q}(0, \eta)$. Thus, the equality (373) is a consequence of Fubini theorem and (255) applied to $(t, x) \mapsto f(x, u(t))$. In particular the finite sum involved in (374) is the Talyor expansion of order M of $u \mapsto \varphi(x(t; f, u, p))$ at $u = 0$. By Lemma 83 $u \mapsto \varphi(x(t; f, u, p))$ is analytic on $B_{L^\infty(0,T^*)}(0, \eta)$ uniformly with respect to $(t, p) \in [0, T^*] \times B_\delta$, which ends the proof of (374). □

6.2 Magnus expansion in the usual setting: a counter-example

Contrary to other expansions, the usual Magnus expansion does not yield, in general, error estimates involving the size of the control. Indeed, the infinite segments which would need to be summed do not converge, even for analytic vector fields, arbitrarily small times and even when the drift vector field vanishes at the origin. The following statement illustrates that even the series defining the terms which are linear with respect to the control does not converge.

Proposition 138. *Let $d := 2$. There exists $T, \delta > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\delta}(B_\delta; \mathbb{K}^d)$ with $f_0(0) = 0$ and a control $u \in \mathcal{C}^\infty([0, T])$, such that, if one defines, for $t \in (0, T)$, the sequence of vector fields*

$$F_n(t) := \sum_{k=0}^{n} \zeta_{\mathrm{ad}_{X_0}^k(X_1)}(t, u) \, \mathrm{ad}_{f_0}^k(f_1), \tag{379}$$

then, for each $\delta^* \in (0, \delta)$ and $t \in (0, T)$, $F_n(t) \in \mathcal{C}^\infty(B_\delta; \mathbb{K}^d)$ does not converge in $\mathcal{C}^0(B_{\delta^*}; \mathbb{K}^d)$.

Proof. We define the following vector fields for $x \in \mathbb{R}^2$ with $|x| < 1$,

$$f_0(x) := x_2 e_1 \quad \text{and} \quad f_1(x) := \frac{1}{1 - x_1} e_2. \tag{380}$$

Then,

$$\mathrm{ad}_{f_0}^k(f_1)(x) = x_2^k \partial_1^k \left(\frac{1}{1 - x_1}\right) e_2 = \frac{k! x_2^k}{(1 - x_1)^{k+1}} e_2. \tag{381}$$

We now choose the particular control $u(t) := t$ for $t \in (0, T)$ with $T = 1$ (the simpler choice, $u(t) := 1$, would not produce a diverging counter-example). Using the expression (71) from Example 40 for the coordinates of the first kind along the brackets $\mathrm{ad}_{X_0}^k(X_1)$ for this particular control, we obtain, for $t \in (0, T)$,

$$F_n(t)(x) = \sum_{k=0}^{n} (-1)^{k+1} t^{k+2} \frac{B_{k+1}}{k+1} \frac{x_2^k}{(1 - x_1)^{k+1}}. \tag{382}$$

Thus, for each $t, \delta^* > 0$, the sequence of vector fields $F_n(t)$ does not converge in $\mathcal{C}^0(B_{\delta^*}; \mathbb{K}^d)$, since for every $x_2 \neq 0$, the general term of the series does not tend to zero because of the asymptotic (531) for Bernoulli numbers. □

6.3 Magnus expansion in the interaction picture

The following statement is an immediate consequence of Proposition 99. It illustrates that, contrary to the classical Magnus expansion, our "Magnus in the interaction picture" expansion allows to obtain error estimates involving the size of the control, at any order.

Proposition 139. *Let $M \in \mathbb{N}$, $\delta > 0$, $T > 0$, $f_0 \in \mathcal{C}^{M^2+1}(B_{5\delta}; \mathbb{K}^d)$ with $T\|f_0\|_{\mathcal{C}^0} \leq \delta$ and $f_1, \dots, f_q \in \mathcal{C}^{M^2}(B_{5\delta}; \mathbb{K}^d)$. There exists $\gamma, C > 0$ such that, for every $u = (u_1, \dots, u_q) \in L^1((0,T); \mathbb{K})$ with*

$$\|u\|_{L^1} \leq \gamma \tag{383}$$

$p \in B_\delta$ and $t \in [0, \gamma]$ then

$$\left| x(t; f, u, p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p \right| \leq C \|u\|_{L^1(0,t)}^{M+1}. \tag{384}$$

In (384), $\mathcal{Z}_M(t, f, u)$ (where implicitly $f = (f_0, f_1, \dots, f_q)$ and $u = (u_1, \dots, u_q)$) is a notation for the vector field $\mathcal{Z}_M\left(t, f_0, \sum_{i=1}^q u_i f_i\right)$, defined in Definition 98 for the affine perturbation $f_\sharp(t, x) = \sum_{i=1}^q u_i(t) f_i(x)$. This notation is chosen by analogy with Theorem 41.

6.4 Sussmann's infinite product expansion

The goal of this section is to prove Proposition 143 which states that, despite the difficulties mentioned in Section 5.4.3 concerning the full convergence of Sussmann's infinite product expansion, some (infinite) subproducts of it do converge and yield error estimates at every order in the size of the control for control-affine systems with drift of the form (366).

We start with an elementary remark (Lemma 140) on the structure of brackets of a Hall set which allows to prove nice asymmetric estimates on the associated coordinates of the second kind (see Lemma 141). The following result proves that, when one tries to factorize the lateral X_0 factors outside of a bracket of a Hall set $B \subset \mathrm{Br}(X)$ with $X_0 \in X$, these X_0 factors cannot alternate sides more than once.

Lemma 140. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$ and $B \subset \mathrm{Br}(X)$ be a Hall set. For each $b \in B$, there exist $m, \overline{m} \in \mathbb{N}$ such that*

$$b = \mathrm{ad}_{X_0}^m \underline{\mathrm{ad}}_{X_0}^{\overline{m}}(b^*), \tag{385}$$

where $\underline{\mathrm{ad}}_{X_0}^{\overline{m}}$ denotes the iterated right bracketing $\overline{m}$ times by X_0 and $b^* \in B$ is such that either $b^* \in X$ or $b = (b_1, b_2)$ with $b_1 \neq X_0$ and $b_2 \neq X_0$.

Proof. The key point is that, by the third condition in Definition 54, for each $b \in B \setminus X$, $\lambda(b) < b$. Let $b \in B$. We disjunct cases.

- If $b \in X$ or ($\lambda(b) \neq X_0$ and $\mu(b) \neq X_0$), then (385) holds with $m = \overline{m} = 0$ and $b^* = b$.
- If $\lambda(b) = X_0$, there exists a unique $m \in \mathbb{N}^*$ and $\tilde{b} \in \mathcal{B}$ such that $b = \mathrm{ad}_{X_0}^m(\tilde{b})$ where $\tilde{b} \in X$ or $\lambda(\tilde{b}) \neq X_0$.
  - If $\tilde{b} \in X$ or $\mu(\tilde{b}) \neq X_0$, (385) holds with $\overline{m} = 0$ and $b^* = \tilde{b}$.
  - Otherwise, there exists a unique $\overline{m} \in \mathbb{N}^*$ and $b^* \in B$ such that $\tilde{b} = \underline{\mathrm{ad}}_{X_0}^{\overline{m}}(b^*)$ where $b^* \in X$ or $\mu(b^*) \neq X_0$.
    - If $b^* \in X$, (385) holds.
    - Else $\mu(b^*) \neq X_0$. one has $\lambda(b^*) < b^*$ as recalled. Moreover, since $\overline{m} \geq 1$, $(b^*, X_0) \in B$ so $b^* < X_0$ (by the second point of Definition 54). Hence $\lambda(b^*) < X_0$. So we also have $\lambda(b^*) \neq X_0$ and (385) holds.

- If $\mu(b) = X_0$, there exists a unique $\overline{m} \in \mathbb{N}^*$ and $\tilde{b} \in B$ such that $b = \underline{\mathrm{ad}}^{\overline{m}}_{X_0}(\tilde{b})$ where $\tilde{b} \in X$ or $\mu(\tilde{b}) \neq X_0$.
  - If $\tilde{b} \in X$, (385) holds with $m = 0$ and $b^* = \tilde{b}$.
  - Else $\mu(\tilde{b}) \neq X_0$. Since $\overline{m} \geq 1$, $(\tilde{b}, X_0) \in B$, so $\tilde{b} < X_0$. Since $\lambda(\tilde{b}) < \tilde{b}$, this proves $\lambda(\tilde{b}) \neq X_0$. So (385) holds with $m = 0$ and $b^* = \tilde{b}$.

Hence, the decomposition (385) always holds. □

We now turn to asymmetric estimates for the coordinates of the second kind, which, contrary to Lemma 63, isolate the role of X_0 associated with the implicit control $u_0 = 1$.

Lemma 141. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $B \subset \mathrm{Br}(X)$ a Hall set and $(\xi_b)_{b\in B}$ the associated coordinates of the second kind. For every $k \in \mathbb{N}^*$, there exists $c_k \geq 1$ such that, for each $b \in B$ with $n(b) = k$, $T > 0$, $u \in L^1((0,T);\mathbb{K}^q)$ and $t \in [0,T]$,*

$$|\xi_b(t,1,u)| \leq \|u\|^k_{L^1(0,t)} \frac{(c_k t)^{n_0(b)}}{n_0(b)!} \tag{386}$$

and

$$|\dot{\xi}_b(t;1,u)| \leq \begin{cases} k|u(t)|\|u\|^{k-1}_{L^1(0,t)} & \text{when } n_0(b) = 0, \\ \|u\|^{k-1}_{L^1(0,t)}\Big(kt|u(t)| + n_0(b)\|u\|_{L^1(0,t)}\Big)\frac{c_k(c_kt)^{n_0(b)-1}}{n_0(b)!} & \text{when } n_0(b) > 0. \end{cases} \tag{387}$$

Proof. In this proof, we write $\xi_b(t)$ instead of $\xi_b(t,1,u)$ by concision for the value at time $t \in [0,T]$ of the coordinate of the second kind associated with the control $u_0 = 1$ and u_i for $i \in [\![1,q]\!]$. First, when (387) holds on $[0,T]$, then so does (386) by time-integration (with the same constant). Hence, we only need to prove the bound on the time derivative of the coordinates.

Step 1: Persistence of the estimates by right bracketing with X_0. Let $k \in \mathbb{N}^*$ and $b \in B$ such that $n(b) = k$. We assume that (386) holds and we prove that $\tilde{b} := (b, X_0)$ satisfies both estimates with the same constant. Since $\dot{\xi}_{X_0}(t) = 1$, we have

$$|\dot{\xi}_{\tilde{b}}(t)| = |\xi_b(t)\dot{\xi}_{X_0}(t)| \leq \|u\|^k_{L^1(0,t)} \frac{(c_kt)^{n_0(b)}}{n_0(b)!}. \tag{388}$$

Hence $\tilde{b}$ satisfies (387) (and (386) by integration) because $c_k \geq 1$ and $n_0(\tilde{b}) > 0$.

Step 2: Persistence of the estimates by arbitrary long left bracketing with X_0, up to $c_k \leftarrow 2c_k$. Let $k \in \mathbb{N}^*$ and $b \in B$ with $n(b) = k$. We assume that (387) holds and we prove that, for every $m \in \mathbb{N}^*$, $\tilde{b} := \mathrm{ad}^m_{X_0}(b)$ satisfies both estimates with a constant $c_k \leftarrow 2c_k$. If $n_0(b) = 0$, it is straightforward to check that $\tilde{b}$ satisfies (387) with $c_k \leftarrow 1$. If $n_0(b) = 1$, we have

$$\begin{aligned} |\dot{\xi}_{\tilde{b}}(t)| &= \frac{1}{m!}|\xi^m_{X_0}(t)\dot{\xi}_b(t)| \\ &\leq \frac{t^m}{m!}\|u\|^{k-1}_{L^1}\Big(kt|u(t)| + n_0(b)\|u\|_{L^1}\Big)\frac{c_k(c_kt)^{n_0(b)-1}}{n_0(b)!} \\ &\leq \|u\|^{k-1}_{L^1}\Big(kt|u(t)| + (m+n_0(b))\|u\|_{L^1}\Big)2^{m+n_0(b)}c_k^{n_0(b)}\frac{t^{m+n_0(b)-1}}{(m+n_0(b))!} \\ &\leq \|u\|^{k-1}_{L^1}\Big(kt|u(t)| + n_0(\tilde{b})\|u\|_{L^1}\Big)(2c_k)^{n_0(\tilde{b})}\frac{t^{n_0(\tilde{b})-1}}{n_0(\tilde{b})!} \end{aligned} \tag{389}$$

because $n_0(\tilde{b}) = m + n_0(b)$ and $c_k \geq 1$. So $\tilde{b}$ satisfies (387) with a constant $c_k \leftarrow 2c_k$.

Step 3: Proof of the estimates by induction on $k \in \mathbb{N}^*$.

Initialization for $k = 1$. For $i \in [\![1, q]\!]$, $\dot{\xi}_{X_i}(t) = u_i(t)$ so both estimates are satisfied with constant 1 when $b \in \{X_1, \dots, X_q\}$. By Lemma 140, Step 1 and Step 2, we deduce that (386) and (387) hold for $k = 1$ with $c_1 = 2$.

Induction $(k-1) \to k$. Let $k \geq 2$ and let us assume that the estimates are proved for every $b \in \mathcal{B}$ with $n(b) \leq (k-1)$. Let $b \in \mathcal{B}$ with $n(b) = k$. By Lemma 140, Step 1 and Step 2, we can assume that $b = \mathrm{ad}_{b_1}^m(b_2)$ with $b_1, b_2 \in \mathcal{B}$, $b_1 \neq X_0$ and ($b_2 \in X$ or $\lambda(b_2) < b_1$) and ($b_2 \neq X_0$ or $m > 1$). Assume that $b_2 \neq X_0$. Then the induction assumption applies to both b_1 and b_2. Let $k_1 := n(b_1)$ and $k_2 := n(b_2)$. Then $k = mk_1 + k_2$, $n_0(b) = mn_0(b_1) + n_0(b_2) \geq n_0(b_2)$. Using the induction assumption and (117) with $a \leftarrow (m+1)$, we obtain, when $n_0(b_2) > 0$,

$$\begin{aligned}
\left|\dot{\xi}_b(t)\right| &= \left|\frac{1}{m!}\xi_{b_1}^m(t)\dot{\xi}_{b_2}(t)\right| \\
&\leq \frac{1}{m!}\left(\|u\|_{L^1}^{k_1}\frac{(c_{k_1}t)^{n_0(b_1)}}{n_0(b_1)!}\right)^m \|u\|_{L^1}^{k_2-1}\Big(k_2 t|u(t)| + n_0(b_2)\|u\|_{L^1}\Big)\frac{c_{k_2}(c_{k_2}t)^{n_0(b_2)-1}}{n_0(b_2)!} \qquad (390)\\
&\leq \|u\|_{L^1}^{k-1}\Big(kt|u(t)| + n_0(b)\|u\|_{L^1}\Big)2^{mn_0(b)}c_{k_1}^{mn_0(b_1)}c_{k_2}^{n_0(b_2)}\frac{t^{n_0(b)-1}}{n_0(b)!}.
\end{aligned}$$

Since $m \leq k$, we have the two desired estimates with $c_k := 2 \cdot 2^k \max\{c_j; j \in [\![1, k-1]\!]\}$, where the first factor 2 comes from Step 2. When $n_0(b_2) = 0$, the proof is similar and easier. When $b_2 = X_0$, the induction hypothesis applies because $m > 1$ so $n(b_1) < n(b)$ and the proof is straightforward. □

Remark 142. *The "persistence" of the estimates with respect to left or right bracketing by X_0, as mentioned and derived in Steps 1 and 2 of the proof of Lemma 141 might be linked with sufficient conditions for small-time local controllability which "ignore" the number of leading (or trailing) X_0 factors (see [17], [63, Theorem 6] or [73, Theorem 3.7]).*

These estimates allow to prove the main result of this section.

Proposition 143. *Let $q \in \mathbb{N}^*$, $X = \{X_0, X_1, \dots, X_q\}$, $\mathcal{B}$ a Hall basis of $\mathcal{L}(X)$ and $(\xi_b)_{b\in\mathcal{B}}$ the associated coordinates of the second kind. Let $M \in \mathbb{N}$, $r, \delta > 0$, $f_0, \dots, f_q \in \mathcal{C}^{\omega,r}(B_{4\delta}; \mathbb{K}^d)$. There exists $\eta, C_M > 0$ such that, for every $u \in L^1((0,T); \mathbb{K}^q)$ with $T \leq \eta$ and $\|u\|_{L^1(0,T)} \leq \eta$, the ordered product of the $e^{\xi_b(t,1,u)f_b}$ over the infinite set $\mathcal{B} \cap S_M = \{b \in \mathcal{B}; n(b) \leq M\}$ (using Definition 10) converges uniformly on B_δ and, for each $t \in [0,T]$ and $p \in B_\delta$,*

$$\left|x(t; f, u, p) - \overrightarrow{\prod_{b\in\mathcal{B}\cap S_M}} e^{\xi_b(t,1,u)f_b}p\right| \leq C_M\|u\|_{L^1(0,t)}^{M+1}. \qquad (391)$$

Proof. In this proof, to simplify the notations, we write $x(t)$, $\xi_b(t)$ and $\|u\|$ instead of $x(t; f, u, p)$, $\xi_b(t, 1, u)$ and $\|u\|_{L^1(0,t)}$. Let $(c_k)_{k\in\mathbb{N}^*}$ be the increasing sequence of constants of Lemma 141. We define

$$C_* := \frac{18\,|||f|||_r}{r}\max_{k\in[\![1,2M]\!]} c_k, \qquad (392)$$

$$\eta := \min\left\{\frac{\delta}{2\|f\|_{\mathcal{C}^1}}, \frac{\min\{1,\delta\}}{2C_*(q+1)M!(1+r)}\right\} \qquad (393)$$

$$C_M := e^{2\delta}(1+r)(2M)!(q+1)^{M+1}C_*^{M+1}. \qquad (394)$$

For $t \in [0,T]$ and $u \in L^1((0,T); \mathbb{K}^q)$ with $T \leq \eta$ and $\|u\| \leq \eta$, using (393),

$$t\|f_0\|_{\mathcal{C}^0} + \sum_{i=1}^q \|u_i\|_{L^1(0,t)}\|f_i\|_{\mathcal{C}^0} \leq \eta\|f\|_{\mathcal{C}^0} \leq \delta. \qquad (395)$$

Hence, for each $p \in B_\delta$, $x(t; f, u, p) \in B_{2\delta}$.

Strategy. Since the product involved in (391) is indexed by the infinite set $\mathcal{B} \cap S_M$, the proof strategy consists in considering the sequence of finite products $\mathcal{B}_{[\![1,L]\!]} \cap S_M$ for $L \in \mathbb{N}^*$ and let $L \to +\infty$. The error between the true solution and the finite product contains both a term scaling like $\|u\|^{M+1}$ which will persist in the limit and a transitory error term which vanishes as $L \to +\infty$. Each bracket in $b \in \mathcal{B}$ is either, not involved at all in the process, involved in the final error, involved in the transitory error term, or involved in the finite product, depending on $L, M, n(b)$ and $n_0(b)$ as pictured in Fig. 1. In Steps 2, 3 and 4, $L \geq M+1$ is fixed. In Step 5, we take the limit $L \to +\infty$.

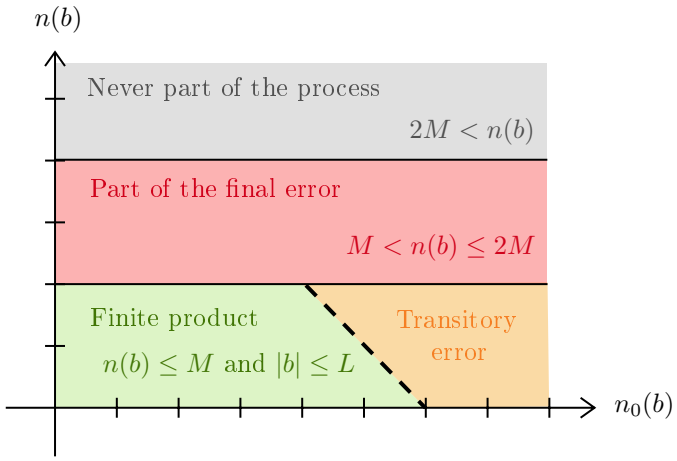

Figure 1: Decomposition of $\mathcal{B}$ along the Lazard elimination process for the product on $\mathcal{B} \cap S_M$.

Step 0: Preliminary estimates. First, using estimate (387) from Lemma 141, for each $b \in \mathcal{B}$ with $n(b) = k$, one has in particular

$$\|\dot{\xi}_b\|_{L^1} \leq \|u\|^k \frac{(c_k t)^{n_0(b)}}{n_0(b)!}. \tag{396}$$

Taking into account that for every $m \in \mathbb{N}^*$, $|\mathcal{B}_m| \leq (q+1)^m$ and using the analytic estimate (140), we obtain the following estimate for the terms which can be part of the final error

$$\begin{aligned}
&\sum_{b \in \mathcal{B} \cap (S_{2M} \setminus S_M)} \|\dot{\xi}_b\|_{L^1} \|f_b\|_{\mathcal{C}^1} \\
&\quad \leq \sum_{k=M+1}^{2M} \sum_{n_0=0}^{+\infty} |\mathcal{B}_{k+n_0}| \|u\|^k \frac{(c_k t)^{n_0}}{n_0!} (1+r) \left(\frac{9 \|\!|f|\!\|_r}{r} \right)^{k+n_0} (k+n_0-1)! \\
&\quad \leq (1+r)(2M-1)! \sum_{k=M+1}^{2M} ((q+1) C_* \|u\|)^k \sum_{n_0=0}^{+\infty} ((q+1) C_* T)^{n_0} \\
&\quad \leq (1+r)(2M)! (q+1)^{M+1} C_*^{M+1} \|u\|^{M+1},
\end{aligned} \tag{397}$$

because $\|u\| \leq \eta$, $T \leq \eta$ and $(q+1)C_*\eta \leq \frac{1}{2}$. For the terms which can be part of the finite product

or of the transitory error, there holds similarly

$$\begin{aligned}
&\sum_{b\in\mathcal{B}\cap S_M} \|\dot\xi_b\|_{L^1}\|f_b\|_{\mathcal{C}^1}\\
&\qquad\le T\|f_0\|_{\mathcal{C}^1}+\sum_{k=1}^{M}\sum_{n_0=0}^{+\infty}|\mathcal{B}_{k+n_0}|\|u\|^k\frac{(c_kt)^{n_0}}{n_0!}(1+r)\left(\frac{9\,|||f|||_r}{r}\right)^{k+n_0}(k+n_0-1)!\\
&\qquad\le T\|f_0\|_{\mathcal{C}^1}+(1+r)(M-1)!\sum_{k=1}^{M}((q+1)C_*\|u\|)^k\sum_{n_0=0}^{+\infty}((q+1)C_*T)^{n_0}\\
&\qquad\le T\|f_0\|_{\mathcal{C}^1}+(1+r)M!(q+1)C_*\|u\|\le\delta.
\end{aligned}\tag{398}$$

Step 1: Convergence of the ordered product of the $e^{\xi_b(t)f_b}$ over $\mathcal{B}\cap S_M$, uniformly on B_δ, towards a Lipschitz map. Thanks to (398), we have

$$\sum_{b\in\mathcal{B}\cap S_M}|\xi_b(t)|\,\|f_b\|_{\mathcal{C}^1}\le\sum_{b\in\mathcal{B}\cap S_M}\|\dot\xi_b\|_{L^1}\|f_b\|_{\mathcal{C}^1}\le\delta\tag{399}$$

and Lemma 134 gives the conclusion of Step 1.

Step 2: Lazard structure on $\mathcal{B}_{[\![1,L]\!]}\cap S_M$. We use the notations of Definition 53 to describe $\mathcal{B}_{[\![1,L]\!]}$. There exists $m\in\mathbb{N}$ and an extraction ϕ such that

$$\mathcal{B}_{[\![1,L]\!]}\cap S_M=\{b_{\phi(1)}<\cdots<b_{\phi(m+1)}\}.\tag{400}$$

Let $i\in[\![1,m+1]\!]$ and $n=\phi(i)$. By Definition 53, there exists a unique factorization

$$b_{\phi(i)}=b_n=\mathrm{ad}_{b_{n-1}}^{j_{n-1}}\cdots\mathrm{ad}_{b_1}^{j_1}(b_0)\tag{401}$$

where $b_0\in X$, $j_1,\dots,j_{n-1}\in\mathbb{N}$ (one just identifies left and right factors in $\mathrm{Br}(X)$). For every $j\in[\![1,n-1]\!]\setminus\phi([\![1,i-1]\!])$, b_j contains at least $(L+1)$ occurrences of the variables $X_1,\dots,X_q$, thus it cannot be involved in the factorization of b_n. This proves that

$$\begin{cases}
b_{\phi(1)}\in\widetilde{Y}_0:=X,\\
b_{\phi(2)}\in\widetilde{Y}_1:=\{\mathrm{ad}^j_{b_{\phi(1)}}(v);j\in\mathbb{N},v\in\widetilde{Y}_0\setminus\{b_{\phi(1)}\}\},\\
\cdots\\
b_{\phi(m+1)}\in\widetilde{Y}_m:=\{\mathrm{ad}^j_{b_{\phi(m)}}(v);j\in\mathbb{N},v\in\widetilde{Y}_{m-1}\setminus\{b_{\phi(m)}\}\},
\end{cases}\tag{402}$$

$$\mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_{m+1}=\emptyset,\tag{403}$$

where $\widetilde{Y}_{m+1}:=\{\mathrm{ad}^j_{b_{\phi(m+1)}}(v);j\in\mathbb{N},v\in\widetilde{Y}_m\setminus\{b_{\phi(m+1)}\}\}$.

Step 3: Proof of estimates along the Lazard elimination on $\mathcal{B}_{[\![1,L]\!]}\cap S_M$. To simplify the notations, from now on, we write $\mathcal{B}_{[\![1,L]\!]}\cap S_M=\{b_1<\cdots<b_{m+1}\}$ and we use (402) and (403) with $\phi=\mathrm{Id}$. Let $x_0(t):=x(t)$. By (399), for every $j\in[\![1,m+1]\!]$,

$$x_j(t):=e^{-\xi_{b_j}(t)f_{b_j}}\cdots e^{-\xi_{b_1}(t)f_{b_1}}x(t)\tag{404}$$

is well-defined and belongs to $B_{3\delta}$. The goal of Step 3 is to prove by induction on $j\in[\![0,m+1]\!]$ that

$$(\mathcal{H}_j):\begin{cases}\dot x_j(t)=\sum_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_j}\dot\xi_b(t)f_b(x_j(t))+\varepsilon_j(t),\\ x_j(0)=p,\end{cases}\tag{405}$$

where

$$\|\varepsilon_j\|_{L^1} \le e^{|\xi_{b_j}(t)|\|f_{b_j}\|_{\mathcal{C}^1}}\|\varepsilon_{j-1}\|_{L^1} + \sum_{\tilde{b}\in Z_j}\|\dot{\xi}_{\tilde{b}}\|_{L^1}\|f_{\tilde{b}}\|_{\mathcal{C}^0}, \tag{406}$$

where $Z_j \subset (\mathcal{B}\cap S_{2M})\setminus(\mathcal{B}_{[\![1,L]\!]}\cap S_M)$ is defined in (413).

First $(\mathcal{H}_0)$ holds with $\varepsilon_0 = 0$ because $\dot{\xi}_{X_0}(t) = 1$ and $\dot{\xi}_{X_i}(t) = u_i(t)$ for $i \in [\![1,q]\!]$. Now, let $j \in [\![1, m+1]\!]$ and assume that $(\mathcal{H}_{j-1})$ holds. We deduce from the definition of x_j that

$$x_j(t) = e^{-\xi_{b_j}(t)f_{b_j}}(x_{j-1}(t)) = \Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right) \tag{407}$$

and thus that

$$\dot{x}_j(t) = \sum_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_{j-1}\setminus\{b_j\}}\dot{\xi}_b(t)\left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) + \widetilde{\varepsilon}_{j-1}(t), \tag{408}$$

where $\widetilde{\varepsilon}_{j-1}(t) = \partial_p\Phi_j\left(-\xi_{b_j}(t), x_{j-1}(t)\right)\varepsilon_{j-1}(t)$. We get $(\mathcal{H}_j)$ with

$$\varepsilon_j(t) := \sum_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_{j-1}\setminus\{b_j\}}\overline{\varepsilon}_b^j(t) + \widetilde{\varepsilon}_{j-1}(t) \tag{409}$$

where, for every $b \in \mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_{j-1}\setminus\{b_j\}$,

$$\overline{\varepsilon}_b^j(t) := \dot{\xi}_b(t)\left(\Phi_j\left(-\xi_{b_j}(t)\right)_* f_b\right)(x_j(t)) - \sum_{k=0}^{h(b)-1}\dot{\xi}_b(t)\frac{\xi_{b_j}^k(t)}{k!}f_{\mathrm{ad}_{b_j}^k(b)}(x_j(t)) \tag{410}$$

where $h(b)\in\mathbb{N}^*$ is the maximal integer such that

$$n(b) + (h(b)-1)\,n(b_j) \le M \qquad \text{and} \qquad |b| + (h(b)-1)|b_j| \le L. \tag{411}$$

By (162),

$$|\overline{\varepsilon}_b^j(t)| \le |\dot{\xi}_b(t)|\frac{|\xi_{b_j}(t)|^{h(b)}}{h(b)!}\|f_{\mathrm{ad}_{b_j}^{h(b)}(b)}\|_{\mathcal{C}^0} = |\dot{\xi}_{\tilde{b}}(t)|\|f_{\tilde{b}}\|_{\mathcal{C}^0}, \tag{412}$$

for $\tilde{b} := \mathrm{ad}_{b_j}^{h(b)}(b)$. Hence, (406) holds with

$$Z_j := \{\mathrm{ad}_{b_j}^{h(b)}(b);\ \ b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M\cap\widetilde{Y}_{j-1}\setminus\{b_j\}\}. \tag{413}$$

This yields $Z_j \subset (\mathcal{B}\cap S_{2M})\setminus(\mathcal{B}_{[\![1,L]\!]}\cap S_M)$ thanks to (411).

Step 4: Proof of an estimate on the finite product over $\mathcal{B}_{[\![1,L]\!]}\cap S_M$. By (406), (399) and (397), we have

$$\begin{aligned}\|\varepsilon_{m+1}\|_{L^1} &\le e^{\delta}\sum_{b\in\mathcal{B}\cap(S_{2M}\setminus S_M)}\|\dot{\xi}_b\|_{L^1}\|f_b\|_{\mathcal{C}^0} + e^{\delta}\sum_{b\in(\mathcal{B}\cap S_M)\setminus\mathcal{B}_{[\![1,L]\!]}}\|\dot{\xi}_b\|_{L^1}\|f_b\|_{\mathcal{C}^0}\\ &\le e^{-\delta}C_M\|u\|^{M+1} + o_{L\to+\infty}(1),\end{aligned} \tag{414}$$

because the series in (398) converges. We deduce from (405) and (403) that

$$\left|\overleftarrow{\prod_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M}}e^{-\xi_b(t,1,u)f_b}x(t) - p\right| = |x_{m+1}(t) - p| \le e^{-\delta}C_M\|u\|^{M+1} + o_{L\to+\infty}(1) \tag{415}$$

By (399), the map $\overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M}}e^{-\xi_b(t,u)f_b}$ is e^{δ} Lipschitz on $B_{3\delta}$. Then, by (415),

$$\left|x(t) - \overrightarrow{\prod_{b\in\mathcal{B}_{[\![1,L]\!]}\cap S_M}}e^{-\xi_b(t,1,u)f_b}p\right| \le C_M\|u\|^{M+1} + o_{L\to+\infty}(1) \tag{416}$$

Step 5: Infinite subproduct limit. By Step 1, the infinite product over $\mathcal{B}\cap S_M$ is well-defined. By letting $L\to+\infty$ in estimate (416), we obtain the conclusion of Proposition 143. □

7 Refined error estimates for scalar-input affine systems

In this section, we consider scalar-input affine systems with drift, i.e. of the form

$$\dot{x}(t) = f_0(x(t)) + u(t) f_1(x(t)) \quad \text{and} \quad x(0) = p, \tag{417}$$

where f_0, f_1 are vector fields on $\mathbb{K}^d$ and $u \in L^1((0,T);\mathbb{K})$. When well-defined, its solution is denoted $x(t; f, u, p)$. Such systems have been extensively studied in control theory, as toy models for more complex situations.

The goal of this section is to improve, in this particular framework, the error estimates of the previous section: the new bound is not expressed in terms of $\|u\|_{L^1}$ but in terms of the L^∞ norm of the time-primitive of the input, which heuristically corresponds to the $W^{-1,\infty}$ norm of u.

This refined estimate is somehow optimal in the scale of Sobolev spaces (as shown by the one dimensional system $\dot{x}(t) = u(t)$) and specific to the scalar-input case (see Section 7.5).

Lowering the Sobolev regularity required on the input is of paramount interest for applications in control theory (see e.g. [14]) and might also be useful for applications to stochastic differential equations where the input is a noise with low regularity (see e.g. [15]).

Definition 144 (Integrated input)**.** *Let $T > 0$ and $u \in L^1((0,T);\mathbb{K})$. In this section, U always denotes the time-primitive of u vanishing at zero, i.e. defined by $U(t) := \int_0^t u(s)\,\mathrm{d}s$ for $t \in [0,T]$.*

7.1 Auxiliary system trick

Enhancing the estimates relies on the following trick which factorizes the dependence of the input and introduces an auxiliary system involving the time-primitive U of the input (and not u itself).

Proposition 145. *Let $\delta > 0$, $f_0, f_1 \in \mathcal{C}^\omega(B_{3\delta};\mathbb{K}^d)$ and $\eta^* > 0$ small enough so that the two following maps are well defined and (globally) analytic*

$$\Phi_1 : \begin{cases} [-\eta^*,\eta^*] \times B_{2\delta} & \to B_{3\delta} \\ (\tau, q) & \mapsto e^{\tau f_1}(q) \end{cases} \quad \text{and} \quad F : \begin{cases} B_{2\delta} \times [-\eta^*,\eta^*] & \to \mathbb{K}^d \\ (q,\tau) & \mapsto (\Phi_1(-\tau)_* f_0)(q). \end{cases} \tag{418}$$

Let $T > 0$ be such that $T\|F\|_{\mathcal{C}^0} \leq \delta$.

1. *For every $p \in B_\delta$ and $U \in \mathcal{C}^0([0,T];\mathbb{K})$ with $\|U\|_{L^\infty} \leq \eta^*$, there exists a unique solution $x_1 \in \mathcal{C}^1([0,T];\mathbb{K}^d)$ to*

$$\begin{cases} \dot{x}_1(t) = F(x_1(t), U(t)), \\ x_1(0) = p, \end{cases} \tag{419}$$

   denoted $x_1(t; F, U, p)$. It takes values in $B_{2\delta}$. Moreover, the map $(p, U) \mapsto x_1(\cdot; F, U, p)$ is analytic from $B_\delta \times B_{\mathcal{C}^0[0,T]}(0, \eta^*)$ to $\mathcal{C}^1([0,T];\mathbb{K}^d)$.

2. *For every $p \in B_\delta$, $t \in [0,T]$ and $u \in L^1((0,T);\mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta^*$,*

$$x(t; f, u, p) = \Phi_1\big(U(t); x_1(t; F, U, p)\big). \tag{420}$$

Proof. The existence of η^* such that Φ_1 and F are well defined and globally analytic results from the third statement of Lemma 90. The analytic dependence of x_1 with respect to (p, U) is given by Lemma 83. By definition of x_1, the right-hand side of (420) satisfies the same Cauchy problem as x thus the two functions are equal. □

7.2 A new formulation of the Chen-Fliess expansion

The goal of this section is to derive of a new formulation of the Chen-Fliess expansion for scalar-input affine systems (417).

Proposition 146. *Let $\delta, r > 0$ and $f_0, f_1 \in \mathcal{C}^{\omega,r}(B_{3\delta}; \mathbb{K}^d)$. There exists $\eta > 0$ such that for every $\varphi \in \mathcal{C}^{\omega,r}(B_{3\delta}; \mathbb{K})$, $t \in [0, \eta]$, $u \in L^1((0,t); \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta$ and $p \in B_\delta$, with the notations of Remark 74,*

$$\varphi(x(t; f, u, p)) = \sum_{\substack{\ell \in \mathbb{N}, n \in \mathbb{N} \\ k \in \mathbb{N}^n}} \frac{U(t)^\ell}{\ell! k!} \left(\int_0^t U^k \right) \Big(f_1^\ell (\mathrm{ad}_{f_1}^{k_1}(f_0)) \cdots (\mathrm{ad}_{f_1}^{k_n}(f_0)) \Big)(\varphi)(p) \tag{421}$$

with the notation (372), where the sum converges absolutely, uniformly with respect to (t, u, p). Moreover, for every $\varphi \in \mathcal{C}^{\omega,r}(B_{3\delta}; \mathbb{K})$, there exists $C > 0$ such that, for every $M \in \mathbb{N}^*$, $t \in [0, \eta]$, $u \in L^1((0,t); \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta$ and $p \in B_\delta$,

$$\begin{aligned} &\left| \varphi(x(t; f, u, p)) - \sum_{\substack{\ell \in \mathbb{N}, n \in \mathbb{N} \\ \ell + |k| \leq M}} \frac{U(t)^\ell}{\ell! k!} \left(\int_0^t U^k \right) \Big(f_1^\ell (\mathrm{ad}_{f_1}^{k_1}(f_0)) \cdots (\mathrm{ad}_{f_1}^{k_n}(f_0)) \Big)(\varphi)(p) \right| \\ &\leq C^{M+1} \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right) \end{aligned} \tag{422}$$

where the sum is taken over $\ell \in \mathbb{N}$, $n \in \mathbb{N}$ and $k = (k_1, \ldots, k_n) \in \mathbb{N}^n$ such that $\ell + k_1 + \cdots + k_n \leq M$.

Proof. Let η^*, T, x_1 be as in Proposition 145, $\|f\| := |||f_0|||_r + |||f_1|||_r$ and

$$\eta := \min \left\{ T, \eta^*, \frac{\delta}{\|f\|}, \frac{r}{28\|f\|} \right\}. \tag{423}$$

Let $\varphi \in \mathcal{C}^{\omega,r}(B_{3\delta}; \mathbb{K})$, $t \in [0, \eta]$, $u \in L^1((0,t); \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta$ and $p \in B_\delta$. Then $x_1(t; F, U, p) \in B_{2\delta}$ and, by (420) and (423), $x(t; f, u, p) \in B_{3\delta}$.

Step 1: Proof of the absolute convergence in (421) uniformly with respect to $p \in B_\delta$. Let $r' := r/e$. Then, by Lemma 80, for every $k \in \mathbb{N}$, $\mathrm{ad}_{f_1}^k(f_0) \in \mathcal{C}^{\omega,r'}(B_{3\delta}; \mathbb{K}^d)$ and

$$\left| \left| \left| \mathrm{ad}_{f_1}^k(f_0) \right| \right| \right|_{r'} \leq \frac{k!}{e} \left(\frac{9}{r} \right)^k \|f\|^{k+1}. \tag{424}$$

Thus, by (135),

$$\begin{aligned} &\left| \Big(f_1^\ell (\mathrm{ad}_{f_1}^{k_1}(f_0)) \cdots (\mathrm{ad}_{f_1}^{k_n}(f_0)) \Big)(\varphi)(p) \right| \\ &\qquad \leq (n+\ell)! \left(\frac{5}{r'} \right)^{n+\ell} |||f_1|||_{r'}^\ell \left| \left| \left| \mathrm{ad}_{f_1}^{k_1}(f_0) \right| \right| \right|_{r'} \cdots \left| \left| \left| \mathrm{ad}_{f_1}^{k_n}(f_0) \right| \right| \right|_{r'} \\ &\qquad \leq (n+\ell)! \left(\frac{14}{r} \right)^{n+\ell} \|f\|^\ell \frac{k_1!}{e} \left(\frac{9}{r} \right)^{k_1} \|f\|^{k_1+1} \cdots \frac{k_n!}{e} \left(\frac{9}{r} \right)^{k_n} \|f\|^{k_n+1} \\ &\qquad \leq e^{-n} (n+\ell)! k_1! \cdots k_n! \left(\frac{14\|f\|}{r} \right)^{n+\ell+k_1+\cdots+k_n}. \end{aligned} \tag{425}$$

Moreover, recalling notation (372),

$$\left| \frac{U(t)^\ell}{\ell! k!} \int_0^t U^k \right| = \left| \frac{U(t)^\ell}{\ell!} \int_{\Delta^n(t)} \frac{U(\tau_1)^{k_1} \cdots U(\tau_n)^{k_n}}{k_1! \cdots k_n!} \, \mathrm{d}\tau \right| \leq \|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n} \frac{t^n}{n!} \frac{1}{\ell! k_1! \cdots k_n!}. \tag{426}$$

Thus it is sufficient to prove the summability over $\ell \in \mathbb{N}, n \in \mathbb{N}^*, k_1, \dots, k_n \in \mathbb{N}$ of the following quantity

$$\begin{aligned}
&\left(\frac{t}{e}\right)^n \frac{(n+\ell)!}{n!\ell!} \left(\frac{14\|f\|}{r}\right)^{n+\ell+k_1+\cdots+k_n} \|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n} \\
&\qquad \leq \left(\frac{t}{e}\right)^n 2^{n+\ell} \left(\frac{14\|f\|}{r}\right)^{n+\ell+k_1+\cdots+k_n} \|U\|_{L^\infty}^{\ell+k_1+\cdots+k_n} \\
&\qquad \leq \left(\frac{28t\|f\|}{er}\right)^n \left(\frac{28\|f\|}{r}\|U\|_{L^\infty}\right)^{\ell+k_1+\cdots+k_n}
\end{aligned} \tag{427}$$

which is ensured by (423).

Step 2: Proof of (421) and (422). Applying Lemma 86 and Proposition 137 we get

$$\begin{aligned}
\varphi(x(t;f,u,p)) &= \varphi\left(e^{U(t)f_1} x_1(t;F,U,p)\right) \\
&= \sum_{\ell=0}^{+\infty} \frac{U(t)^\ell}{\ell!} f_1^\ell \varphi(x_1(t;F,U,p)) \\
&= \sum_{\ell=0}^{+\infty} \frac{U(t)^\ell}{\ell!} f_1^\ell \sum_{\substack{n\in\mathbb{N} \\ k\in\mathbb{N}^n}} \frac{1}{k!} \left(\int_0^t U^k\right) \left((\mathrm{ad}_{f_1}^{k_1}(f_0)) \cdots (\mathrm{ad}_{f_1}^{k_n}(f_0))\right)\Big(\varphi\Big)(p)
\end{aligned} \tag{428}$$

The bound proved in Step 1 allows to exchange the differential operator f_1^ℓ and the second sum, which proves (421). To prove (422), one bounds the queue of the series thanks to (425) and the following consequence of Hölder's inequality, valid when $\ell + |k| \geq (M+1)$

$$\left| U(t)^\ell \int_{\Delta^n(t)} U(\tau_1)^{k_1} \cdots U(\tau_n)^{k_n} \, \mathrm{d}\tau \right| \leq C(\eta) \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right). \tag{429}$$

□

One of the ingredients of the above proof is the Chen-Fliess expansion of the auxiliary system $x_1(t;F,U,p)$, which appears in [3, Section 3] under the denomination "representation of the perturbation flow".

Remark 147. *The bound* (422) *between the exact solution and the truncated Chen-Fliess series (in its' original formulation) is used by Stefani in [96, Lemma 3.1 and Corollary 3.1]. Our proof is both different and shorter.*

Remark 148. *Equality* (421) *where the sum converges absolutely proves that appropriate packages of the Chen-Fliess expansion are absolutely summable under a smallness assumption on $\|U\|_{L^\infty}$, which is weaker than the smallness assumption on $\|u\|_{L^1}$ which is used in Proposition 136 for multi-input systems.*

7.3 Magnus expansion in the interaction picture

In this section, we prove the following enhanced error estimate for the magnus expansion in the interaction picture with scalar input. Our proof relies on an appropriate approximation for the auxiliary system x_1 introduced in Section 7.1.

Proposition 149. *Let $\delta > 0$ and $f_0, f_1 \in \mathcal{C}^\omega(B_{3\delta}; \mathbb{K}^d)$. For every $M \in \mathbb{N}$, there exist $\eta_M, C_M > 0$ such that, for every $T \in [0, \eta_M]$, $u \in L^1((0,T);\mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta_M$, $t \in [0,T]$ and $p \in B_\delta$,*

$$\left| x(t;f,u,p) - e^{\mathcal{Z}_M(t,f,u)} e^{tf_0} p \right| \leq C_M \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right). \tag{430}$$

Proof. In Section 7.3.1, we introduce a vector field $\mathcal{Y}_M(t,f,U)$ such that $e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}(p)$ is a good approximation of the auxiliary state x_1 defined in (419). Since, by (420), $x(t) = e^{U(t)f_1}(x_1(t))$, the desired estimate then relies on the following decomposition

$$\begin{aligned} x(t;f,u,p) - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p = x(t;f,u,p) - e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p \\ + e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p. \end{aligned} \tag{431}$$

Using Proposition 150 and Proposition 154 (see further) for the first and second lines, we get

$$\left|x(t;f,u,p) - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p\right| \leq C_M\left(\|U\|_{L^1}^{M+1} + |U(t)|^{M+1} + \|U\|_{L^{M+1}}^{M+1}\right) \tag{432}$$

which gives the conclusion since $\|U\|_{L^1(0,t)} \leq t^{\frac{M}{M+1}}\|U\|_{L^{M+1}(0,t)}$. □

In Section 7.3.1, we define $\mathcal{Y}_M(t,f,U)$ and prove in Proposition 150 that it indeed provides a good approximation of the auxiliary state. In Section 7.3.2, we explain the link between $e^{U(t)X_1}e^{\mathcal{Y}_M(t,X,U)}$ and $e^{\mathcal{Z}_M(t,X,u)}$ at the formal level. In Section 7.3.3, we show in Proposition 154 that this formal link entails that $e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}$ is close to $e^{\mathcal{Z}_M(t,f,u)}$.

7.3.1 An approximation of the auxiliary state

We use the error formula of Proposition 99 for the Magnus expansion in the interaction picture to obtain an approximation of the auxiliary state.

Proposition 150. *Let $\delta,\rho>0$, $f_0,f_1\in\mathcal{C}^{\omega,\rho}(B_{3\delta};\mathbb{K}^d)$. For every $M\in\mathbb{N}$, there exist $\eta_M, C_M>0$ such that, for every $p\in B_\delta$, $t\in[0,\eta_M]$, $u\in L^1((0,t);\mathbb{K})$ such that $\|U\|_{L^\infty}\leq\eta_M$,*

$$\left|x(t;f,u,p) - e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p\right| \leq C_M\|U\|_{L^1(0,t)}^{M+1} \tag{433}$$

where $\mathcal{Y}_M(t,f,U) := \mathrm{Log}_M\{G_t\}(t)$, and $G_t:[0,t]\times B_{3\delta}\to\mathbb{K}^d$ is defined by

$$G_t(s,y) := \sum_{\substack{k\in\mathbb{N}^*\\ \ell\in\mathbb{N}}} \frac{(s-t)^\ell}{\ell!}\frac{U(s)^k}{k!}\,\mathrm{ad}_{f_0}^\ell\,\mathrm{ad}_{f_1}^k(f_0)(y) \tag{434}$$

and this sum converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta};\mathbb{K}^d)$ with $\rho'=\rho/e$. Moreover,

$$\begin{aligned} \mathcal{Y}_M(t,f,U) = \sum \frac{(-1)^{m-1}}{rm}\int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1-t)^{\ell_1}}{\ell_1!}\frac{U(\tau_1)^{k_1}}{k_1!}\cdots\frac{(\tau_r-t)^{\ell_r}}{\ell_r!}\frac{U(\tau_r)^{k_r}}{k_r!}\,\mathrm{d}\tau \\ \left[\cdots\left[\mathrm{ad}_{f_0}^{\ell_1}(\mathrm{ad}_{f_1}^{k_1}(f_0)),\mathrm{ad}_{f_0}^{\ell_2}(\mathrm{ad}_{f_1}^{k_2}(f_0))\right],\ldots,\mathrm{ad}_{f_0}^{\ell_r}(\mathrm{ad}_{f_1}^{k_r}(f_0))\right], \end{aligned} \tag{435}$$

where the sum is taken over $r\in[\![1,M]\!]$, $m\in[\![1,r]\!]$, $\mathbf{r}\in\mathbb{N}_r^m$, $\ell_1,\ldots,\ell_r\in\mathbb{N}$, $k_1,\ldots,k_r\in\mathbb{N}^*$ and the sum converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta};\mathbb{K}^d)$.

Proof. Step 1: Convergence in (434) and (435). By (138), for every $s\in[0,t]$,

$$\left|\!\left|\!\left|\frac{(s-t)^\ell}{\ell!}\frac{U(s)^k}{k!}\,\mathrm{ad}_{f_0}^\ell\,\mathrm{ad}_{f_1}^k(f_0)\right|\!\right|\!\right|_{\rho'} \leq t^\ell\|U\|_{L^\infty}^k\frac{(k+\ell)!}{\ell!k!}\left(\frac{9}{\rho}\right)^{k+\ell}|||f|||_\rho^{k+\ell+1} \tag{436}$$

thus the sum in (434) converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta};\mathbb{K}^d)$ when t and $\|U\|_{L^\infty}$ are smaller than $\frac{\rho}{18|||f|||_\rho}$.

For every $r \in [\![1, M]\!]$, $m \in [\![1, r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $\ell_1, \dots, \ell_r \in \mathbb{N}$, $k_1, \dots, k_r \in \mathbb{N}^*$, using (117) and the non-decreasing of $q \in [\![1, \infty]\!] \mapsto \|\cdot\|_{L^q(0,t)}$ for $t \in [0,1]$, we get

$$\begin{aligned}\left| \int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1 - t)^{\ell_1}}{\ell_1!} \frac{U(\tau_1)^{k_1}}{k_1!} \cdots \frac{(\tau_r - t)^{\ell_r}}{\ell_r!} \frac{U(\tau_r)^{k_r}}{k_r!} \, \mathrm{d}\tau \right| &(r + |\ell| + |k| - 1)! \left(\frac{9\|f\|}{\rho} \right)^{r+|\ell|+|k|-1} \\ &\leq (2^{r-1}t)^{|\ell|} (2^{r-1} \|U\|_{L^{|k|}(0,t)})^{|k|} \left(\frac{36\|f\|}{\rho} \right)^{r+|\ell|+|k|-1} (r-1)!\end{aligned} \tag{437}$$

Thus, by (138), the sum in (435) converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta}; \mathbb{K}^d)$ when t and $\|U\|_{L^\infty}$ are smaller than $\frac{\rho 2^{-M}}{18 |||f|||_\rho}$.

Step 2: Proof of (433). Let T, η^* and F as in Proposition 145. We introduce the function $F_1 : [0,T] \times B_{2\delta} \to \mathbb{K}^d$ defined by

$$F_1(t,y) := F(y, U(t)) - f_0(y) = \sum_{j=1}^{+\infty} \frac{U(t)^j}{j!} \operatorname{ad}_{f_1}^j(f_0)(y) \tag{438}$$

where the sum converges in $\mathcal{C}^{\omega,\rho'}(B_{2\delta}; \mathbb{K}^d)$ when $\|U\|_{L^\infty} < \frac{\rho}{9|||f|||_\rho}$. Let $M \in \mathbb{N}$. There exists $C > 0$ such that, for every $t \in [0,T]$, $U \in \mathcal{C}^0([0,T]; \mathbb{K})$ with $\|U\|_{L^\infty} \leq \eta^*$, the function F_1 defined by (438) satisfies

$$\|F_1\|_{L^1((0,t);\mathcal{C}^{M^2})} \leq C \|U\|_{L^1(0,t)} \leq CT \|U\|_{L^\infty(0,t)}. \tag{439}$$

Let $C_M > 0$ and $\gamma = \gamma(M, \delta, \|f_0\|_{\mathcal{C}^{M^2+1}}) > 0$ be as in Proposition 99 and

$$\eta_M := \min \left\{ 1, \eta^*, \frac{\rho 2^{-M}}{36 \, |||f|||_\rho}, \frac{\gamma}{CT} \right\}. \tag{440}$$

Let $p \in B_\delta$, $t \in [0, \eta_M]$ and $u \in L^1((0,t); \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta_M$. Then, the convergences of Step 1 hold and $\|F_1\|_{L^1((0,t);\mathcal{C}^{M^2})} \leq \gamma$ thus we can apply Proposition 99 and Proposition 100 to the equation $\dot{x}_1 = f_0(x_1) + F_1(t, x_1)$

$$\left| x_1(t; F, U, p) - e^{\mathcal{Y}_M(t,f,U)} e^{tf_0} p \right| \leq C_M \|G_t\|_{L^1((0,t);\mathcal{C}^{M^2})}^{M+1}. \tag{441}$$

Moreover, there exists C' (depending only on η^*, f_0, f_1) such that

$$\|G_t\|_{L^1((0,t);\mathcal{C}^{M^2})} \leq C' \|U\|_{L^1(0,t)}. \tag{442}$$

Thus, we get (433) by applying the $e^{\eta^* \|f_1\|_{\mathcal{C}^1}}$-Lipschitz map $e^{U(t)f_1}$ to (441). □

In the next paragraphs, we will use the following technical result about $\mathcal{Y}_M(t,f,U)$ and its decomposition in homogeneous components with respect to U.

Lemma 151. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}(B_{3\delta}; \mathbb{K}^d)$. For every $M \in \mathbb{N}^*$, there exists $\eta_M, C_M > 0$ such that, for every $j \in \mathbb{N}^*$, $t \in [0, \eta_M]$, $u \in L^1((0,t), \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta_M$, the sum in the right-hand side of (435) taken over $r \in [\![1, M]\!]$, $m \in [\![1, r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $\ell_1, \dots, \ell_r \in \mathbb{N}$ and $k_1, \dots, k_r \in \mathbb{N}^*$ such that $k_1 + \ldots + k_r = j$, converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta}; \mathbb{K}^d)$ and its sum, denoted $\mathcal{Y}_M^j(t,f,U)$, satisfies*

$$\left| \left| \left| \mathcal{Y}_M^j(t,f,U) \right| \right|_{\rho'} \leq C_M \left(\frac{\|U\|_{L^j(0,t)}}{2\eta_M} \right)^j \tag{443}$$

where $\rho' = \rho/e$. Moreover, $\mathcal{Y}_M(t,f,U) = \sum_{j \in \mathbb{N}^*} \mathcal{Y}_M^j(t,f,U)$ where the sum converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta}; \mathbb{K}^d)$.

Proof. Let $\eta_M > 0$ be as in Proposition 150, $t \in [0, \eta_M]$ and $u \in L^1((0,t);\mathbb{K})$ such that $\|U\|_{L^\infty} < \eta_M$. The sum involved in $\mathcal{Y}_M^j(t,f,U)$ converges absolutely in $\mathcal{C}^{\omega,\rho'}(B_{3\delta};\mathbb{K}^d)$ because it is a subfamily of the one considered in Proposition 150. By (437), there exists $C_M > 0$ (independent of t and U) such that, for every $j \in \mathbb{N}^*$, (443) holds. The non-decreasing of $q \in [\![1,\infty]\!] \mapsto \|\cdot\|_{L^q(0,t)}$ (since $t \leq 1$) gives the last conclusion. □

7.3.2 Identification procedure at the formal level

In this paragraph, we highlight at the formal level the link between $e^{U(t)X_1}e^{\mathcal{Y}_M(t,X,U)}$ and $e^{\mathcal{Z}_M(t,X,u)}$ in $\widehat{\mathcal{L}}(X)$. We start with a new formal factorization, well adapted to estimates with respect to the primitive of the scalar input.

Proposition 152. *Let $X = \{X_0, X_1\}$ and $u \in L^1(\mathbb{R}_+;\mathbb{K})$. For every $x^\star \in \widehat{\mathcal{A}}(X)$, the solution x to the formal differential equation*

$$\begin{cases} \dot{x}(t) = x(t)(X_0 + u(t)X_1), \\ x(0) = x^\star, \end{cases} \tag{444}$$

satisfies, for every $t \in \mathbb{R}_+$,

$$x(t) = x^\star \exp(tX_0) \exp(\mathcal{Y}_\infty(t,X,U)) \exp(U(t)X_1) \tag{445}$$

where $\mathcal{Y}_\infty(t,X,U) \in \widehat{\mathcal{L}}(X)$ is defined by $\mathcal{Y}_\infty(t,X,U) = \mathrm{Log}_\infty\{\beta_t\}(t)$ and $\beta_t : [0,t] \to \widehat{\mathcal{L}}(X)$ is defined by

$$\beta_t(s) = e^{-(t-s)X_0}\left(e^{U(s)X_1}X_0e^{-U(s)X_1} - X_0\right)e^{(t-s)X_0} = \sum_{\substack{k\in\mathbb{N}^* \\ \ell\in\mathbb{N}}} \frac{(s-t)^\ell}{\ell!}\frac{U(s)^k}{k!}\,\mathrm{ad}_{X_0}^\ell\,\mathrm{ad}_{X_1}^k(X_0) \tag{446}$$

i.e.

$$\mathcal{Y}_\infty(t,X,U) = \sum \frac{(-1)^{m-1}}{rm}\int_{\Delta^{\mathbf{r}}(t)} \frac{(\tau_1-t)^{\ell_1}}{\ell_1!}\frac{U(\tau_1)^{k_1}}{k_1!}\cdots\frac{(\tau_r-t)^{\ell_r}}{\ell_r!}\frac{U(\tau_r)^{k_r}}{k_r!}\,\mathrm{d}\tau \\ \left[\cdots\left[\mathrm{ad}_{X_0}^{\ell_1}(\mathrm{ad}_{X_1}^{k_1}(X_0)),\mathrm{ad}_{X_0}^{\ell_2}(\mathrm{ad}_{X_1}^{k_2}(X_0))\right],\ldots,\mathrm{ad}_{X_0}^{\ell_r}(\mathrm{ad}_{X_1}^{k_r}(X_0))\right] \tag{447}$$

where the sum is taken over $r \in \mathbb{N}^*$, $m \in [\![1,r]\!]$, $\mathbf{r} \in \mathbb{N}_r^m$, $\ell_1,\ldots,\ell_r \in \mathbb{N}$, $k_1,\ldots,k_r \in \mathbb{N}^*$.

Proof. First, in the same way as Theorem 27 has been generalized to an infinite alphabet in the proof of Theorem 41, it is possible to generalize Theorem 41 to an infinite alphabet.

The function $x_1 : [0,T] \to \widehat{\mathcal{A}}(X)$ defined by $x_1(t) := x(t)e^{-U(t)X_1}$ satisfies $x_1(0) = x^\star$ and

$$\dot{x}_1 = x_1(t)e^{U(t)X_1}X_0e^{-U(t)X_1} = x_1(t)\left(X_0 + \sum_{k\in\mathbb{N}^*}\frac{U(t)^k}{k!}\,\mathrm{ad}_{X_1}^k(X_0)\right). \tag{448}$$

This equation is of the form $\dot{x}_1(t) = x_1(t)(X_0 + \sum_{k\in\mathbb{N}^*} a_k(t)Y_k)$ for some indeterminates Y_k. Thus, Theorem 41 (adapted to an infinite alphabet) and the homomorphism of algebras sending Y_k to $\mathrm{ad}_{X_1}^k(X_0)$ prove that

$$x_1(t) = x^\star \exp(tX_0)\exp(\mathcal{Y}_\infty(t,X,U)). \tag{449}$$

which gives the conclusion. □

We now use the formal expansion (445) to obtain an alternative formula for $\mathcal{Z}_\infty(t,X,u)$ defined by Theorem 41, in terms of the primitive of the scalar input. For $r,\nu\in\mathbb{N}$, we introduce the finite dimensional subspace of $\mathcal{L}(X)$

$$\mathcal{L}_{r,\nu}(X) := \mathrm{span}\{\mathrm{E}(b);\ \ b\in \mathrm{Br}(X), n_0(b)=\nu, n_1(b)=r\} \tag{450}$$

and $P_{r,\nu}:\widehat{\mathcal{L}}(X)\to\mathcal{L}_{r,\nu}(X)$ the associated canonical projection.

Proposition 153. *Let $X=\{X_0,X_1\}$, $T>0$, $u\in L^1((0,T);\mathbb{K})$, $t\in[0,T]$, $\mathcal{Y}_\infty(t,X,U)$ defined by Proposition 152 and $\mathcal{Z}_\infty(t,X,u)$ defined by Theorem 41. Then, in $\widehat{\mathcal{L}}(X)$,*

$$\mathcal{Z}_\infty(t,X,u) = \mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t,X,U), U(t)X_1\right). \tag{451}$$

In particular, for every $M\in\mathbb{N}^*$, $r\in[\![1,M]\!]$ and $\nu\in\mathbb{N}$,

$$P_{r,\nu}\mathcal{Z}_M(t,X,u) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_M(t,X,U), U(t)X_1\right). \tag{452}$$

In this statement, CBHD_∞ is defined in Corollary 33, CBHD_M is its truncation used in Corollary 94 and $\mathcal{Z}_M(t,X,u)$ is defined in Theorem 41 and used in Proposition 99.

Proof. We deduce from Proposition 152 and Theorem 41 that

$$\exp(\mathcal{Z}_\infty(t,X,u)) = \exp(\mathcal{Y}_\infty(t,X,U))\exp(U(t)X_1). \tag{453}$$

Thus Corollary 33 proves (451). Let $M\in\mathbb{N}^*$, $r\in[\![1,M]\!]$, $\nu\in\mathbb{N}$. We deduce from (451) that

$$P_{r,\nu}\mathcal{Z}_\infty(t,X,u) = P_{r,\nu}\,\mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t,X,U), U(t)X_1\right). \tag{454}$$

By definition, $\mathcal{Z}_\infty(t,X,u)-\mathcal{Z}_M(t,X,u)$ is a linear combination of brackets all involving at least $(M+1)$ occurrences of X_1, thus $P_{r,\nu}\mathcal{Z}_\infty(t,X,u)=P_{r,\nu}\mathcal{Z}_M(t,X,u)$. By definition, $\mathcal{Y}_\infty(t,X,U)$ is a sum of brackets involving all at least one occurrence of X_1, thus

$$P_{r,\nu}\,\mathrm{CBHD}_\infty\left(\mathcal{Y}_\infty(t,X,U), U(t)X_1\right) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_\infty(t,X,U), U(t)X_1\right). \tag{455}$$

Moreover $\mathcal{Y}_\infty(t,X,U)-\mathcal{Y}_M(t,X,U)$ is a linear combination of brackets involving all at least $(M+1)$ occurrences of X_1 thus

$$P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_\infty(t,X,U), U(t)X_1\right) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\mathcal{Y}_M(t,X,U), U(t)X_1\right), \tag{456}$$

which ends the proof of (452). □

7.3.3 Error formula for analytic vector fields

We prove in Proposition 154 an error bound between $e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}$ and $e^{\mathcal{Z}_M(t,f,u)}$.

Proposition 154. *Let $\delta,\rho>0$, $f_0,f_1\in\mathcal{C}^{\omega,\rho}(B_{3\delta};\mathbb{K}^d)$. For every $M\in\mathbb{N}$, there exist $\eta_M,C_M>0$ such that, for every $t\in[0,\eta_M]$, $p\in B_\delta$ and $u\in L^1((0,t);\mathbb{K})$ such that $\|U\|_{L^\infty}\le\eta_M$,*

$$\left|e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p\right| \le C_M\left(|U(t)|^{M+1}+\int_0^t|U|^{M+1}\right). \tag{457}$$

Proof. We split the difference as

$$\begin{aligned} &e^{U(t)f_1}e^{\mathcal{Y}_M(t,f,U)}e^{tf_0}p - e^{\mathrm{CBHD}_M(\mathcal{Y}_M(t,f,U),U(t)f_1)}e^{tf_0}p \\ &\qquad + e^{\mathrm{CBHD}_M(\mathcal{Y}_M(t,f,U),U(t)f_1)}e^{tf_0}p - e^{\mathcal{Z}_M(t,f,u)}e^{tf_0}p. \end{aligned} \tag{458}$$

Taking into account that $\|\mathcal{Y}_M(t,f,U)\|_{\mathcal{C}^{M^2}}\le C\|U\|_{L^1(0,t)}$, the first line is bounded by Corollary 94. Using Grönwall's lemma and Proposition 155 bounds the second line. □

Proposition 155. *Let $\delta, \rho > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,\rho}(B_{3\delta}; \mathbb{K}^d)$ and $\rho' := \rho/e$. For every $\rho'' \in (0, \rho')$, $M \in \mathbb{N}$, there exist $\eta_M, C_M > 0$ such that, for every $t \in [0, \eta_M]$, $u \in L^1((0,t); \mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta_M$,*

$$\left|\left|\left| \mathcal{Z}_M(t,f,u) - \mathrm{CBHD}_M\left(\mathcal{Y}_M(t,f,U), U(t)f_1\right)\right|\right|\right|_{\rho''} \leq C_M \left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} \right). \tag{459}$$

In particular, $\mathcal{Z}_M(t,f,u)$ is the sum of the terms homogeneous with degree at most M with respect to U in $\mathrm{CBHD}_M\left(\mathcal{Y}_M(t,f,U), U(t)f_1\right)$.

Proof. Step 1: Finite approximation of $\mathcal{Y}_M(t,f,U)$. First, by Lemma 151, one can write

$$\mathcal{Y}_M(t,f,U) = \sum_{j=1}^{M} \mathcal{Y}_M^j(t,f,U) + \sum_{j>M} \mathcal{Y}_M^j(t,f,U) =: \overline{\mathcal{Y}}_M(t,f,U) + \mathcal{R}_M(t,f,U), \tag{460}$$

where the remainder satisfies $\left|\left|\left| \mathcal{R}_M(t,f,U) \right|\right|\right|_{\rho'} \leq C\|U\|^{M+1}_{L^{M+1}(0,t)}$. By the triangular and Young inequalities, it is therefore sufficient to prove (459) with $\mathcal{Y}_M$ replaced by the finite truncation $\overline{\mathcal{Y}}_M(t,f,U)$.

Step 2: Identification at the free level. Let $\Lambda : \mathcal{L}(X) \to \mathcal{C}^\omega(B_{3\delta}; \mathbb{K}^d)$ be the homomorphism of Lie algebras such that $\Lambda(X_i) = f_i$. The relation (452) is made of finite linear combinations of brackets of X_0 and X_1. Let $M \in \mathbb{N}$. By applying Λ to this equality, we get, for every $r \in [\![1, M]\!]$, $\nu \in \mathbb{N}$

$$P_{r,\nu}\mathcal{Z}_M(t,f,u) = P_{r,\nu}\,\mathrm{CBHD}_M\left(\overline{\mathcal{Y}}_M(t,f,U), U(t)f_1\right). \tag{461}$$

By definition

$$\mathcal{Z}_M(t,f,u) = \sum_{\nu\in\mathbb{N}} \sum_{r=1}^{M} P_{r,\nu}\mathcal{Z}_M(t,f,u) \tag{462}$$

where the sum converges in $\mathcal{C}^{\omega,\rho'}(B_{3\delta}; \mathbb{K}^d)$ for appropriate $\rho' \in (0,\rho)$, by Proposition 100. Thus, with the notations of (47),

$$\mathcal{Z}_M(t,f,u) - \mathrm{CBHD}_M\left(\overline{\mathcal{Y}}_M(t,f,U), U(t)f_1\right) = \sum_{jh_1+h_2>M} F_{2,h}(\mathcal{Y}_M^j(t,f,U), U(t)f_1), \tag{463}$$

where the sum is taken over $j, h_1, h_2 \in [\![1, M]\!]$.

Step 3: Proof of (459). From now on, $\eta_M > 0$ is given by Proposition 150 and Lemma 151, $t \in [0, \eta_M]$, $u \in L^1((0,t); \mathbb{K})$ is such that $\|U\|_{L^\infty} < \eta_M$ and $\rho'' \in (0, \rho')$. For each term in the finite sum (463), one has, thanks to Lemma 151,

$$\begin{aligned} \left|\left|\left| F_{2,h}(\mathcal{Y}_M^j(t,f,U), U(t)f_1) \right|\right|\right|_{\rho''} &\leq C \left|\left|\left| \mathcal{Y}_M^j(t,f,U) \right|\right|\right|_{\rho'}^{h_1} \left|\left|\left| U(t)f_1 \right|\right|\right|_{\rho'}^{h_2} \\ &\leq C'\|U\|^{jh_1}_{L^j(0,t)}|U(t)|^{h_2} \leq C'\|U\|^{jh_1}_{L^{M+1}(0,t)} t^{h_1 - \frac{jh_1}{M+1}} |U(t)|^{h_2} \end{aligned} \tag{464}$$

which concludes the proof thanks to Young's inequality since $jh_1 + h_2 \geq M + 1$. □

7.4 Sussmann's infinite product expansion

When the input is scalar, the estimates of the coordinates obtained in Lemma 141 can be enhanced to involve only the primitive of the input, at least for Hall bases where X_1 is minimal, which in turn improves the estimate of Proposition 143 (see Proposition 157 below). The hypothesis that X_1 is the minimal element can be seen as the formal counterpart of the auxiliary system trick of Section 7.1.

Lemma 156. *Let $X = \{X_0, X_1\}$, $\mathcal{B}$ be a Hall basis of $\mathcal{L}(X)$ for which X_1 is the minimal element and $(\xi_b)_{b\in\mathcal{B}}$ the associated coordinates of the second kind. For every $k \geq 1$, there exists $c_k \geq 1$ such that, for each $b \in \mathcal{B} \setminus X$ with $n(b) = k$, $T > 0$, $u \in L^1((0,T);\mathbb{K})$ and $t \in [0,T]$,*

$$|\xi_b(t,1,u)| \leq c_k \|U\|^k_{L^k(0,t)} \frac{(c_k t)^{n_0(b)-1}}{(n_0(b)-1)!} \tag{465}$$

and

$$|\dot{\xi}_b(t;1,u)| \leq \begin{cases} c_k |U(t)|^k & \text{when } n_0(b) = 1, \\ c_k |U(t)|^k \frac{(c_k t)^{n_0(b)-1}}{(n_0(b)-1)!} + c_k^2 \|U\|^k_{L^k(0,t)} \frac{(c_k t)^{n_0(b)-2}}{(n_0(b)-2)!} & \text{when } n_0(b) \geq 2. \end{cases} \tag{466}$$

Proof. As for Lemma 141, estimate (465) is obtained, for each b, by time integration of (466). Moreover, still as in Lemma 141, both estimates are invariant by right-bracketing with X_0, and also by arbitrary long left-bracketing with X_0, up to $c_k \leftarrow 2c_k$. Let us prove (465) and (466) by induction on k.

Initialization for $k = 1$. We have $\xi_{X_1}(t) = U(t)$ and $\dot{\xi}_{[X_1,X_0]}(t) = U(t)$. Hence $[X_1, X_0] \in \mathcal{B}$ (because $X_1 < X_0$) satisfies both estimates. By Lemma 140, when $n(b) = 1$, there exist $m, \overline{m} \in \mathbb{N}$ such that $b = \mathrm{ad}^m_{X_0} \underline{\mathrm{ad}}^{\overline{m}}_{X_0}(X_1)$. Since X_1 is minimal, if $b \neq X_1$, $\overline{m} > 0$. Thus, by the previous invariant properties, we get the conclusion with $c_1 := 2$.

Induction $(k-1) \to k$. Let $k \geq 2$ and let us assume that the two estimates are proved for every $b \in \mathcal{B} \setminus X$ with $n(b) \leq (k-1)$. Let $b \in \mathcal{B}$ with $n(b) = k$. By Lemma 140 and the previous invariant properties, we may assume that $b = \mathrm{ad}^m_{b_1}(b_2)$ with $m \in \mathbb{N}^*$, $b_1 < b_2 \in \mathcal{B}$, $b_1 \neq X_0$, ($b_2 \in X$ or $\lambda(b_2) < b_1$) and ($b_2 \neq X_0$ or $m > 1$).

- If $b_1 = X_1$, then $b_2 = X_0$ (otherwise, if $b_2 \notin X$, $\lambda(b_2) < X_1$, which is impossible since X_1 is minimal). Thus
$$|\dot{\xi}_b(t)| = \frac{|U(t)|^m}{m!} \tag{467}$$
so (466) with $c_k = 1$ holds since $n_0(b) = 1$ and $k = m$.

- If $b_1 \neq X_1$, then b_1 satisfies (466) for some $k_1 \in [\![1, k-1]\!]$. Moreover, either $b_2 = X_0$ or $b_2 \notin X$ (because it cannot be X_1). The case ($b_2 = X_0$ and $m > 1$) is easier and left to the reader. Thus we are left with the case where b_2 satisfies (466) for some $k_2 \in [\![1, k-1]\!]$. One has $k = mk_1 + k_2$ and $\nu := n_0(b) = mn_0(b_1) + n_0(b_2) =: m\nu_1 + \nu_2$. Thus,
$$|\dot{\xi}_b(t)| \leq \frac{c_{k_1}^m \|U\|^{mk_1}_{L^{k_1}}}{m!} \frac{(c_{k_1}t)^{m\nu_1 - m}}{(\nu_1-1)!^m} \left(c_{k_2} |U(t)|^{k_2} \frac{(c_{k_2}t)^{\nu_2-1}}{(\nu_2-1)!} + c_{k_2}^2 \|U\|^{k_2}_{L^{k_2}} \frac{(c_{k_2}t)^{\nu_2-2}}{(\nu_2-2)!} \mathbf{1}_{\nu_2 \geq 2} \right). \tag{468}$$
Thanks to Hölders' inequality,
$$\|U\|^{mk_1}_{L^{k_1}} \|U\|^{k_2}_{L^{k_2}} \leq \|U\|^k_{L^k} t^m. \tag{469}$$
Thanks to Hölder's inequality and Young's inequality,
$$\|U\|^{mk_1}_{L^{k_1}} |U(t)|^{k_2} \leq t^m \left(\|U\|^k_{L^k} t^{-1} + |U(t)|^k \right). \tag{470}$$
Moreover, thanks to (117), for $i \in [\![1,2]\!]$,
$$\frac{1}{m!} \frac{1}{(\nu_1-1)!^m} \frac{1}{(\nu_2-i)!} \leq 2^{(m+1)(\nu-i)} \frac{1}{(\nu-i)!}. \tag{471}$$
Combining these inequalities proves (466) with $c_k := 2^{k+2} \max\{c_j; j \in [\![1, k-1]\!]\}$.

□

These enhanced estimates yield the following result.

Proposition 157. *Let $X = \{X_0, X_1\}$, $\mathcal{B}$ a Hall basis of $\mathcal{L}(X)$ for which X_1 is the minimal element and $(\xi_b)_{b\in\mathcal{B}}$ the associated coordinates of the second kind. Let $r, \delta > 0$, $f_0, f_1 \in \mathcal{C}^{\omega,r}(B_{4\delta}; \mathbb{K}^d)$. For each $M \in \mathbb{N}^*$, there exist $\eta_M, C_M > 0$ such that, for every $u \in L^1((0,T);\mathbb{K})$ with $T \leq \eta_M$ and $\|U\|_{L^{M+1}(0,T)} \leq \eta_M$, the ordered product of the $e^{\xi_b(t,1,u)f_b}$ over the infinite set $\mathcal{B} \cap S_M = \{b \in \mathcal{B}; n(b) \leq M\}$ (using Definition 10) converges uniformly on B_δ and, for each $t \in [0,T]$ and $p \in B_\delta$,*

$$\left| x(t; f, u, p) - \overrightarrow{\prod_{b\in\mathcal{B}\cap S_M}} e^{\xi_b(t,1,u)f_b} p \right| \leq C_M \|U\|^{M+1}_{L^{M+1}(0,t)}. \tag{472}$$

Proof. The proof is the same as the proof of Proposition 143. The only difference is that we use estimates of Lemma 156 instead of those of Lemma 141. The fact that these enhanced estimates are not valid for $b \in X$ doesn't come into play. Indeed, neither X_0 nor X_1 are involved in the final error term (397). □

7.5 Failure of the primitive estimate for multiple inputs

Proposition 149 relying only on the primitive of the input is specific to the scalar-input case and fails for multiple inputs. As an illustration, for $\delta > 0$ and $f_0, f_1 \in \mathcal{C}^\infty(B_\delta; \mathbb{K}^d)$, in the degenerate case $M = 0$ and the particular case $f_0(0) = 0$, $p = 0$, estimate (430) implies that, for every $T > 0$, there exists $C_T > 0$ such that, for $t \in [0,T]$ and $u \in L^1(0,T)$ with $\|U\|_{L^\infty} \leq 1$,

$$|x(t; u, 0)| \leq C_T \|U\|_{L^\infty}. \tag{473}$$

As illustrated by the following example, even this very crude estimate fails for multiple inputs, because the $W^{-1,\infty}$ norms are not sufficient to bound the nonlinear terms arising in the dynamic.

Example 158. *Let $T > 0$ and consider the following system on $\mathbb{R}^2$:*

$$\begin{cases} \dot{x}_1 = u, \\ \dot{x}_2 = v x_1, \end{cases} \tag{474}$$

where u and v are two scalar inputs. There exists $u_n, v_n \in L^1(0,T)$ such that

$$\|U_n\|_{L^\infty} + \|V_n\|_{L^\infty} \to 0 \quad \text{and} \quad |x(t; (u_n, v_n), 0)| \not\to 0, \tag{475}$$

where U_n is the primitive of u_n and V_n the primitive of V_n. Indeed, let $n \in \mathbb{N}^*$ and define $u_n(t) := n \cos n^2 t$ and $v_n(t) := n \sin n^2 t$. Then one has

$$\|U_n\|_{L^\infty} + \|V_n\|_{L^\infty} \leq \frac{2}{n}. \tag{476}$$

Moreover, $x_1(t) = U_n(t) = (\sin n^2 t)/n$ and

$$x_2(T) = \int_0^T v_n(t) U_n(t)\, \mathrm{d}t = \int_0^T \sin^2(n^2 t)\, \mathrm{d}t \to \frac{T}{2}, \tag{477}$$

as $n \to +\infty$. This proves (475).

Remark 159. *Although Proposition 149 does not hold for multiple inputs, we expect that the proof method can be adapted to obtain asymmetric estimates, involving for example $\|U\|_{L^\infty} + \|v\|_{L^\infty}$ in the two-inputs case (or the converse). Such asymmetric estimates have been used successfully to obtain sharp results for particular control systems in [51].*

8 On direct intrinsic representations of the state

The expansions studied above in this article unfortunately don't provide a direct intrinsic representation of the state. The Magnus and Sussmann expansions are given with intrinsic quantities (Lie brackets of the vector fields) but they require to compute one or multiple flows in order to recover the state. The Chen-Fliess expansion gives directly a formula for the state, but it depends on non-intrinsic quantities (see Remark 17 and Remark 167). In this section, we investigate the possibility of finding a direct intrinsic formula for the state. We discuss this possibility in the context of affine systems.

8.1 Approximate direct intrinsic representations

We prove in this section approximate direct intrinsic representations which achieve the desired goal up to a small error. We believe that the formulas we derive can be of interest for applications to control theory as they give approximate expressions for the state in terms of the inputs and Lie brackets of the vector fields evaluated at the origin.

We start with an elementary result, which bounds the error when replacing a flow by the value of the vector field.

Lemma 160. *Let $\delta > 0$ and $z \in \mathcal{C}^1(B_\delta; \mathbb{K}^d)$ such that $\|z\|_{\mathcal{C}^0} \leq \delta$. Then*

$$|e^z(0) - z(0)| \leq |z(0)| \|Dz\|_{\mathcal{C}^0} e^{\|Dz\|_{\mathcal{C}^0}}. \tag{478}$$

Proof. Let $x(t) := e^{tz}(0)$ for $t \in [0,1]$. Then, for every $t \in [0,1]$,

$$|x(t) - tz(0)| \leq \int_0^t |z(x(\tau)) - z(0)| \, \mathrm{d}\tau \leq \frac{t^2}{2} \|Dz\|_{\mathcal{C}^0} |z(0)| + \int_0^t \|Dz\|_{\mathcal{C}^0} |x(\tau) - \tau z(0)| \, \mathrm{d}\tau \tag{479}$$

and by Grönwall's lemma, $|x(t) - tz(0)| \leq \frac{t^2}{2} \|Dz\|_{\mathcal{C}^0} |z(0)| e^{t\|Dz\|_{\mathcal{C}^0}}$. □

This elementary estimate allows to obtain approximate direct intrinsic representations from the various Magnus expansions described above.

Proposition 161. *Let $M \in \mathbb{N}^*$, $\delta > 0$ and $q \in \mathbb{N}^*$.*

1. *Let $I = [\![0, q]\!]$ or $I = [\![1, q]\!]$. Let $f_i \in \mathcal{C}^{M^2}(B_\delta; \mathbb{K}^d)$ for $i \in I$. For $T > 0$ and $u \in L^\infty((0,T); \mathbb{K}^q)$, if $x(t; f, u, 0)$ denotes the solution to (226) with $p = 0$ and $Z_M(t, f, u)$ denotes the vector field defined in Proposition 93 (called $Z_M(t, \sum_{i \in I} u_i f_i)$ in this statement), then, as $T \to 0$,*

$$x(t; f, u, 0) = Z_M(t, f, u)(0) + O\left(t^{M+1} + t|x(t; f, u, 0)|\right). \tag{480}$$

in the following sense: there exist $C, \eta > 0$ such that, for every $T \in (0, \eta]$ and $u \in L^\infty((0,T); \mathbb{K}^q)$ with $\|u\|_{L^\infty} \leq 1$, for each $t \in [0,T]$,

$$|x(t; f, u, 0) - Z_M(t, f, u)(0)| \leq C\left(t^{M+1} + t|x(t; f, u, 0)|\right). \tag{481}$$

2. *Let $T > 0$, $f_0, \ldots, f_q \in \mathcal{C}^{M^2+1}(B_{2\delta}; \mathbb{K}^d)$ with $f_0(0) = 0$ and $T\|f_0\|_{\mathcal{C}^0} \leq \delta$. For $u \in L^1((0,T); \mathbb{K}^q)$, if $x(t; f, u, 0)$ denotes the solution to (366) with $p = 0$ and $\mathcal{Z}(t; f, u)$ denotes the vector field defined in Proposition 139, then, as $\|u\|_{L^1} \to 0$,*

$$x(t; f, u, 0) = \mathcal{Z}_M(t, f, u)(0) + O\left(\|u\|_{L^1(0,t)}^{M+1} + |x(t; f, u, 0)|^{1+\frac{1}{M}}\right). \tag{482}$$

3. *Let $f_0, f_1 \in \mathcal{C}^\omega(B_{3\delta};\mathbb{K}^d)$ with $f_0(0) = 0$. Let $T > 0$ as in Proposition 145. For $u \in L^1((0,T);\mathbb{K})$, if $x(t;f,u,0)$ denotes the solution to (417) with $p = 0$ and $\mathcal{Z}(t,f,u)$ denotes the vector field defined in Proposition 139 (with $q = 1$), then, as $(T, \|U\|_{L^\infty}) \to 0$,*

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left(\|U\|^{M+1}_{L^{M+1}(0,t)} + |x(t;f,u,0)|^{1+\frac{1}{M}}\right), \tag{483}$$

Proof. Proof of the first statement. By Proposition 93, there exists $C_1 > 0$ and $T^* > 0$ such that for every $u \in L^\infty((0,T^*);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \le 1$ and $t \in [0,T^*]$,

$$\left|x(t;f,u,0) - e^{Z_M(t,f,u)}(0)\right| \le C_1 t^{M+1}. \tag{484}$$

By the explicit expression of $Z_M(t,f,u)$, there exists $C_2 > 0$ such that for every $u \in L^\infty((0,T^*);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \le 1$ and $t \in [0,T^*]$,

$$\|Z_M(t,f,u)\|_{\mathcal{C}^1} \le C_2 t. \tag{485}$$

Thus, by Lemma 160, there exists $C_3 > 0$ such that, for every for every $u \in L^\infty((0,T^*);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \le 1$ and $t \in [0,T^*]$,

$$\left|e^{Z_M(t,f,u)}(0) - Z_M(t,f,u)(0)\right| \le C_3 t \left|Z_M(t,f,u)(0)\right|. \tag{486}$$

Then, by triangular inequality, for every $u \in L^\infty((0,T^*);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \le 1$ and $t \in [0,T^*]$

$$|x(t;f,u,0) - Z_M(t,f,u)(0)| \le C_1 t^{M+1} + C_3 t |Z_M(t,f,u)(0)| \tag{487}$$

and in particular, for $t \le T \le 1/(2C_3)$

$$|Z_M(t,f,u)(0)| \le 2\,|x(t;f,u,0)| + 2C_1 t^{M+1}. \tag{488}$$

This gives (481) with $C = \max\{2C_1; 2C_3\}$ and $\eta := \min\{T^*, 1/(2C_3)\}$.

Proof of the second statement. The strategy is the same: one starts from the estimate in Proposition 139, then applies Lemma 160 to $\mathcal{Z}_M(t,f,u)$ and concludes thanks to the following estimate, implied by the explicit expressions of the vector field

$$\|\mathcal{Z}_M(t,f,u)\|_{\mathcal{C}^1} = \underset{\|u\|_{L^1}\to 0}{O}\left(\|u\|_{L^1(0,t)}\right), \tag{489}$$

and Young's inequality.

Proof of the third statement. First, one can assume that $f_1(0) \neq 0$. Indeed, otherwise, both x and $\mathcal{Z}_M$ vanish identically, so the desired estimate is void. Using Proposition 155 and the explicit expression of the vector field $\mathrm{CBHD}_M\left(\mathcal{Y}_M(t,f,U), U(t)f_1\right)$, we obtain in the asymptotics $(t, \|U\|_{L^\infty}) \to 0$

$$\|\mathcal{Z}_M(t,f,u)\|_{\mathcal{C}^1} = O\left(|U(t)| + \|U\|_{L^1(0,t)}\right). \tag{490}$$

Thus, using $f_0(0) = 0$, Proposition 149 and the same strategy as above, we obtain in the asymptotics $(t, \|U\|_{L^\infty}) \to 0$

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left(|U(t)|^{M+1} + \int_0^t |U|^{M+1} + \left(|U(t)| + \|U\|_{L^1}\right)|x(t;f,u,0)|\right). \tag{491}$$

The following proposition and Young's inequality give the conclusion. □

Proposition 162. *Let $\delta > 0$, $f_0, f_1 \in \mathcal{C}^\omega(B_\delta;\mathbb{K}^d)$ with $f_0(0) = 0$ and $f_1(0) \neq 0$. There exists $T, \eta, C > 0$ such that, for every $u \in L^1((0,T),\mathbb{K})$ with $\|U\|_{L^\infty} < \eta$ and $t \in [0,T]$,*

$$|U(t)| \le C\left(|x(t;f,u,0)| + \|U\|_{L^1(0,t)}\right). \tag{492}$$

Proof. With the notations of Proposition 145, $x(t;f,u,0) = e^{U(t)f_1}x_1(t;F,U,0)$ tends to zero when $\|U\|_{L^\infty} \to 0$. A Taylor expansion of order 2 in $x(t;f,u,0) = e^{U(t)f_1}x_1(t;F,U,0)$ provides $C_1 > 0$ such that, for every $t \in [0,T]$ and $u \in L^1((0,T);\mathbb{K})$ such that $\|U\|_{L^\infty} \leq \eta^*$,

$$|x(t;f,u,0) - x_1(t;F,U,0) - U(t)f_1(0)| \leq C_1|U(t)|^2 + C_1|U(t)||x_1(t;F,U,0)|. \tag{493}$$

Moreover, by Grönwall's lemma, there exists $C_2 > 0$ such that

$$|x_1(t;F,U,0)| \leq C_2\|U\|_{L^1(0,t)}. \tag{494}$$

Let $P : \mathbb{K}^d \to \mathbb{K}^d$ defined by $P(y) = \langle y, f_1(0)\rangle/|f_1(0)|^2$. Applying P to the vector in the left-hand side of (493) and using (494), we get the conclusion, when $\|U\|_{L^\infty}$ is small enough. □

Under additional nilpotency assumptions, one can omit the truncation errors in the representation formulas of Proposition 161.

Corollary 163. *Under the same assumptions as in Proposition 161.*

1. *Assume moreover that $\mathcal{L}(\{f_i; i \in I\})$ is nilpotent of index at most $M+1$. Then, as $T \to 0$,*

$$x(t;f,u,0) = Z_M(t,f,u)(0) + O\left(t|x(t;f,u,0)|\right). \tag{495}$$

2. *Assume moreover that $f_i \in \mathcal{C}^\omega(B_{4\delta};\mathbb{K}^d)$ for $i \in I := [\![0,q]\!]$ and that $\{f_i; i \in [\![1,q]\!]\}$ is semi-nilpotent of index at most $M+1$ with respect to f_0. Then, as $\|u\|_{L^1} \to 0$,*

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left(\|u\|_{L^1(0,t)}|x(t;f,u,0)|\right). \tag{496}$$

3. *Assume moreover that $f_0, f_1 \in \mathcal{C}^\omega(B_{4\delta};\mathbb{K}^d)$ and that $\{f_1\}$ is semi-nilpotent of index at most $M+1$ with respect to f_0. Then, as $(T, \|U\|_{L^\infty}) \to 0$,*

$$x(t;f,u,0) = \mathcal{Z}_M(t,f,u)(0) + O\left((\|U\|_{L^1(0,t)} + |U(t)|)|x(t;f,u,0)|\right). \tag{497}$$

Proof. These are straightforward consequences of Proposition 110 (for the first item) and Corollary 122 (for the second and third item, thanks to the analyticity assumption), using the same approach as in the proof of Proposition 161. □

Remark 164. *Estimate* (483) *proves that, for a situation in which $\int_0^t |U|^{M+1}$ is negligible, the state is well approximated by $\mathcal{Z}_M(t,f,u)(0)$, which is a convergent series of iterated Lie brackets of f_0 and f_1 evaluated at 0. We expect that this representation can be useful for applications to control theory, where one tries to relate controllability of the system with geometric relations on the Lie brackets evaluated at zero.*

8.2 Diffeomorphisms and Lie brackets

Lie brackets behave very nicely with respect to local changes of coordinates. Let f_i be smooth vector fields for $i \in I$, $p \in \mathbb{K}^d$ and θ be a smooth local diffeomorphism near p. If $x(t)$ denotes the solution to (226), we define $y(t) := \theta(x(t))$. Then, one checks that y is the solution to

$$\dot{y}(t) = \sum_{i\in I} u_i(t)g_i(y(t)) \quad \text{and} \quad y(0) = p', \tag{498}$$

where $g_i := \theta_* f_i$ and $p' := \theta(p)$. By iterating Lemma 89, Lie brackets of the vector fields defining the dynamics for y can be computed explicitly from those of x. More precisely, for every $b \in \mathrm{Br}(X)$,

$$g_b = \theta_* f_b \tag{499}$$

with the notation of Definition 78. In particular, there exists a linear invertible map $L_p : \mathbb{K}^d \to \mathbb{K}^d$, $L_p := D\theta(p)$, such that, for every $b \in \mathrm{Br}(X)$,

$$g_b(p') = Lf_b(p). \tag{500}$$

Conversely, if the f_i and g_i for $i \in I$ are analytic vector fields, the existence of points p and p' and a linear invertible map L_p such that (500) holds is a sufficient condition for the existence of a local smooth diffeomorphism θ with $\theta(p) = p'$ and such that, for all controls u_i, there holds $y(t) = \theta(x(t))$ where x and y denote the solutions to (226) and (498) for the same set of controls. This nice property is proved in [75, Theorem 1] and was then extended with a more general geometric viewpoint in [98] (see also [5, Theorem 5.5] for a modern presentation).

When (500) only holds for brackets up to some length $M \in \mathbb{N}$ and the controls are uniformly bounded in L^∞, one can prove (see [76]) the existence of a local smooth diffeomorphism θ and a constant C such that

$$|y(t) - \theta(x(t))| \leq Ct^{M+1}. \tag{501}$$

Up to our knowledge, the converse, which is conjectured to be true in [76], is a nice open problem.

Open problem 165. *Let $I = [\![1, q]\!]$ and $X = \{X_1, \dots X_q\}$. Let $p, p' \in \mathbb{K}^d$. Assume that there exists a smooth diffeomorphism θ from a neighborhood of p to a neighborhood of p' and $M \in \mathbb{N}$ such that, for all controls $u_1, \dots u_q \in L^\infty(0, T)$ with $\|u_i\| \leq 1$, estimate* (501) *holds for the trajectories x and y corresponding to the same controls. Does this imply that there exists a linear invertible map such that, for each $b \in \mathrm{Br}(X)$ with $|b| \leq M$,* (500) *holds?*

Open problem 166. *Same question in the context of affine systems with drift, i.e. when $I = [\![0, q]\!]$, $X = \{X_0, X_1, \dots X_q\}$ and the first control u_0 is constrained to be identically equal to 1. This question might be harder because one gets less information from* (501) *as it is valid for less choices of controls since u_0 is heavily constrained.*

Remark 167. *Property* (500) *is specific to Lie brackets and does not hold for products of differential operators. As an illustration, consider the case $\mathbb{K} = \mathbb{R}$, $d = 2$, $X = \{X_0, X_1\}$, $p = p' = 0$ with $L_0 := D\theta(0) = \mathrm{Id}_2$. Then, for every $\phi \in \mathcal{C}^\infty(\mathbb{R}^2; \mathbb{R})$, $(g_b\phi)(0) = (f_b\phi)(0)$, but this relation does not extend to a similar relation between products of f_0 and f_1 and those of g_0 and g_1. For example, with the vector fields $f_0(x) := (0, x_1)$ and $f_1(x) := (1, 0)$ and the smooth diffeomorphism $\theta(x) := (x_1, x_2 + x_1^2)$, one has $g_0(y) = (0, y_1)$ and $g_1(y) = (1, 2y_1)$. In particular, $(f_1^2\phi)(0) = \partial_{11}\phi(0)$ but $(g_1^2\phi)(0) = \partial_{11}\phi(0) + 2\partial_{12}\phi(0)$. This explains why we consider that the Chen-Fliess expansion is not an intrinsic representation of the state, as it depends on quantities which are not invariant through local changes of coordinates.*

8.3 Replacing the Magnus flow by a diffeomorphism

Let f_i for $i \in I$ be smooth vector fields. We consider the solution $x(t; u)$ to (226) with $p = 0$. Let $Z_M(t, u)$ be the vector field defined in Proposition 93 (and called $Z_M(t, \sum_{i \in I} u_i f_i)$ in this statement). By Proposition 93, for each $M \in \mathbb{N}$, $x(t; u)$ is given by the time-one flow of the autonomous vector field $Z_M(t, u)$, up to an error scaling like t^{M+1} when the controls u_i are uniformly bounded in L^∞.

In this paragraph, inspired by the nice properties of Lie brackets with respect to diffeomorphisms recalled above, we attempt to replace the computation of the time-one flow by a diffeomorphism. This can be seen as being related with the converse of the classical question of whether a given diffeomorphism can be represented as the time-one flow of an autonomous vector field (see e.g. [7, 8] for positive answers in particular cases, [7, Section 2] for an elementary necessary condition, and [52] or [88] for statements highlighting that the answer is only rarely positive).

This also corresponds to replacing the terms $x(t; u) + o(|x(t; u)|)$ in Proposition 161 by $\theta(x(t; u))$, where θ is a smooth local diffeomorphism of $\mathbb{K}^d$.

We start with a definition.

Definition 168. *Let $T > 0$ and $n \in \mathbb{N}$. We say that a functional $\beta : [0,T] \times L^\infty((0,T);\mathbb{K}^q) \to \mathbb{K}$ is* homogeneous of degree n with respect to time *when, for every $u \in L^\infty((0,T);\mathbb{K}^q)$, $\lambda \in (0,1]$ and $t \in [0,T]$,*

$$\beta(\lambda t, u^\lambda) = \lambda^n \beta(t,u) \tag{502}$$

where u^λ is defined by $u^\lambda(\lambda t) := u(t)$ for $t \in [0,T]$ and $u^\lambda(\lambda t) := 0$ for $t > T$.

In particular, the product of two homogeneous functionals of degree n and m with respect to time is an homogeneous functional of degree $n+m$. The coordinates of the first kind $\zeta_b(t,u)$, pseudo-first kind $\eta_b(t,u)$ and second kind $\xi_b(t,u)$ are all homogeneous of degree $|b|$ with respect to time. An interesting property of homogeneous functionals is given by the following statement.

Lemma 169. *Let $T > 0$, $n \in \mathbb{N}$ and $\beta : [0,T] \times L^\infty((0,T);\mathbb{K}^q) \to \mathbb{K}$, homogeneous of degree n with respect to time. Assume that there exists $C > 0$ such that, for every $u \in L^\infty((0,T);\mathbb{K}^q)$ with $\|u\|_{L^\infty(0,T)} \leq 1$ and each $t \in [0,T]$,*

$$|\beta(t,u)| \leq C t^{n+1}. \tag{503}$$

Then $\beta \equiv 0$.

Proof. Let $t \in [0,T]$ and $u \in L^\infty((0,T);\mathbb{K}^q)$ such that $\|u\|_{L^\infty(0,T)} \leq 1$. On the one hand, for each $\lambda \in (0,1]$, $\beta(\lambda t, u^\lambda) = \lambda^n \beta(t,u)$. On the other hand, $|\beta(\lambda t, u^\lambda)| \leq C\lambda^{n+1}t^{n+1}$ because $\|u^\lambda\|_{L^\infty} = \|u\|_{L^\infty} \leq 1$. Hence $|\beta(t,u)| \leq C\lambda t^{n+1}$ for each $\lambda \in (0,1]$ so $\beta(t,u) = 0$. □

One could wonder if the following proposition holds.

False proposition 170. *Let $X = \{X_i ; i \in I\}$, $\mathcal{B}$ be a monomial basis of $\mathcal{L}(X)$. Let $T > 0$. There exists a family $(\beta_b)_{b\in\mathcal{B}}$ of functionals from $[0,T] \times L^\infty((0,T);\mathbb{K}^q)$ to $\mathbb{K}$, with β_b homogeneous of degree $|b|$ with respect to time, such that the following statement holds. Let $\delta > 0$ and $f_i \in \mathcal{C}^\infty(B_\delta;\mathbb{K}^d)$ for $i \in I$. There exists a smooth diffeomorphism θ of $\mathbb{K}^d$ near $p = 0$ such that, for each $M > 0$, there exists $C_M, T_M > 0$ such that, for every $u \in L^\infty((0,T);\mathbb{K}^q)$ with $\|u\|_{L^\infty} \leq 1$, for each $t \in [0,T_M]$,*

$$|\theta(x(t;u)) - y_M(t;u)| \leq C_M t^{M+1}, \tag{504}$$

and

$$y_M(t;u) = \theta(0) + \sum_{|b|\leq M} \beta_b(t,u) g_b(\theta(0)), \tag{505}$$

where $g_b = \theta_* f_b$ and $x(t;u)$ is the solution to (226) starting from $p = 0$.

The functionals β_b would be the analog of the coordinates of the first and second kind described earlier. A formula such as (505) would be ideal for applications to control theory for example, since it is expressed on intrinsic quantities (Lie brackets) and allows to compute $x(t;u)$ directly without solving for flows (one recovers $x(t;u) \approx \theta^{-1}(y(t;u))$). In some sense, it corresponds to asking if there exists a local change of coordinates for which the Chen-Fliess expansion only involves Lie bracket terms (and all the non-Lie bracket terms vanish).

Unfortunately, it is impossible in general, as illustrated by the following counter-example.

Proposition 171. *Let $X = \{X_0, X_1\}$. Let $T > 0$ and consider, in $\mathbb{R}^3$, $f_0(x) := (0, x_1 + x_1^2, x_1x_2)$ and $f_1(x) := (1,0,0)$, i.e. the following affine system with drift*

$$\begin{cases} \dot{x}_1 = u, \\ \dot{x}_2 = x_1 + x_1^2, \\ \dot{x}_3 = x_1 x_2, \end{cases} \tag{506}$$

together with the initial data $x(0) = 0$. There exists a monomial basis $\mathcal{B}$ of $\mathcal{L}(X)$, such that, for all functionals $\beta_b : [0,T] \times L^\infty((0,T);\mathbb{R}) \to \mathbb{R}$ for $b \in \mathcal{B}$, homogeneous of degree $|b|$ with respect to time and for every local $\mathcal{C}^6$ diffeomorphism θ of $\mathbb{R}^3$, there exists $M \in [\![1,6]\!]$ and a control $u \in L^\infty((0,T);\mathbb{R})$ with $\|u\|_{L^\infty} \leq 1$ such that (504) does not hold, even for small times.

Proof. Let $\mathcal{B}$ be a length-compatible Hall basis of $\mathcal{L}(X)$ with $X_0 < X_1$.

Step 1: Computation of $y_6(t)$. We define $\dot{\mathcal{B}}_\ell = \{b \in \mathcal{B}; n_1(b) = \ell\}$ for every $\ell \in \mathbb{N}$. Then $\dot{\mathcal{B}}_1 = \{\mathrm{ad}^k_{X_0}(X_1); k \in \mathbb{N}\}$. The computation shows that the only elements $b \in \dot{\mathcal{B}}_1$ such that $f_b \neq 0$ are

$$b_1 = X_1, \qquad b_2 = [X_0, X_1], \qquad c_1 = [X_0, [X_0, X_1]], \tag{507}$$

$$f_{b_1}(x) = e_1, \qquad f_{b_2}(x) = -(1 + 2x_1)e_2 - x_2 e_3, \qquad f_{c_1}(x) = x_1^2 e_3. \tag{508}$$

Thus, the only elements $b \in \dot{\mathcal{B}}_2$ that could satisfy $f_b \neq 0$ are $[b_1, b_2]$, $[b_1, c_1]$, $[b_2, c_1]$. The computation shows that, among them, only the two first ones do satisfy the condition:

$$b_3 = [X_1, [X_0, X_1]], \qquad c_2 = [X_1, \mathrm{ad}^2_{X_0}(X_1)], \tag{509}$$

$$f_{b_3}(x) = -2e_2, \qquad f_{c_2}(x) = 2x_1 e_3. \tag{510}$$

Thus, the only elements $b \in \dot{\mathcal{B}}_3$ with length at most 6 that could satisfy $f_b \neq 0$ are $[b_1, b_3]$, $[b_1, c_2]$, $[b_2, b_3]$, $[b_2, c_2]$, $[c_1, b_3]$. The computation shows that, among them, only the second and the third ones do satisfy the condition:

$$b_4 = \mathrm{ad}^2_{X_1}\,\mathrm{ad}^2_{X_0}(X_1), \qquad b_5 = [[X_0, X_1], [X_1, [X_0, X_1]]], \tag{511}$$

$$f_{b_4}(x) = 2e_3, \qquad f_{b_5}(x) = -2e_3. \tag{512}$$

Thus the only elements $b \in \dot{\mathcal{B}}_4$ with length at most 6 that could satisfy $f_b \neq 0$ are $[b_1, b_4]$ and $[b_1, b_5]$, but the computation shows that they satisfy $f_b = 0$. Therefore, for every $b \in \dot{\mathcal{B}}_4 \cup \dot{\mathcal{B}}_5 \cup \dot{\mathcal{B}}_6$, $f_b = 0$. In conclusion, $b_1, \dots, b_5$ are the only elements $b \in \mathcal{B}$ such that $f_b(0) \neq 0$. In particular, none of them have length 4 or 6, thus

$$y_6(t) = \theta(0) + D\theta(0)\Big(\beta_1(t,u)e_1 - \beta_2(t,u)e_2 - 2\beta_3(t,u)e_2 + 2(\beta_4(t,u) - \beta_5(t,u))e_3\Big) \tag{513}$$

is the sum of 4 homogeneous functionals of degree $1, 2, 3$ and 5. Here and below we write β_j instead of β_{b_j} for brevity.

Step 2: Computation of homogeneous terms with degree 4 and 6 in $\theta(x(t))$. In this step, we consider a local $\mathcal{C}^6$ diffeomorphism θ of $\mathbb{R}^3$ defined on a neighborhood of $p = 0$. For $u \in L^\infty((0,T);\mathbb{R})$, we denote by U the primitive of u such that $U(0) = 0$ and V the primitive of U such that $V(0) = 0$. Straightforward explicit integration of (506) yields

$$x(t;u) = U(t)e_1 + V(t)e_2 + \int_0^t U^2(s)\,\mathrm{d}s e_2 + \frac{1}{2}V^2(t)e_3 + \int_0^t U(s)\int_0^s U^2(s')\,\mathrm{d}s'\,\mathrm{d}s e_3, \tag{514}$$

where the five terms are respectively functionals homogeneous of degree 1 through 5 with respect to time in the sense of Definition 168. Using a Taylor expansion of θ at 0, one obtains (vector-valued) functionals γ_k for $k \in [\![1,6]\!]$, homogeneous of degree k with respect to time such that for every $M \in [\![1,6]\!]$

$$\theta(x(t)) = \theta(0) + \sum_{k=1}^{M} \gamma_k(t,u) + \mathcal{O}(t^{M+1}). \tag{515}$$

In particular

$$\begin{aligned}\gamma_4(t,u) = \frac{1}{2}V^2(t)\partial_3\theta(0) + U(t)\int_0^t U^2\partial_{12}\theta(0) + \frac{1}{2}V^2(t)\partial_{22}\theta(0)\\ + \frac{1}{2}U^2(t)V(t)\partial_{112}\theta(0) + \frac{1}{4!}U^4(t)\partial_1^4\theta(0)\end{aligned} \tag{516}$$

and

$$\begin{aligned}\gamma_6(t,u) &= U(t)\int_0^t U(s)\int_0^s U(s')^2 ds' ds \partial_{13}\theta(0) + \frac{1}{2}V^3(t)\partial_{23}\theta(0) + \frac{1}{2}\left(\int_0^t U^2\right)^2 \partial_{22}\theta(0)\\ &+ \frac{1}{4}U^2(t)V^2(t)\partial_{113}\theta(0) + \frac{1}{2}U(t)V(t)\int_0^t U^2\partial_{122}\theta(0) + \frac{1}{6}V^3(t)\partial_{222}\theta(0)\\ &+ \frac{1}{6}U^3(t)\int_0^t U^2\partial_{1112}\theta(0) + \frac{1}{4}U^2(t)V^2(t)\partial_{1122}\theta(0)\\ &+ \frac{1}{4!}V(t)U^4(t)\partial_1^4\partial_2\theta(0) + \frac{1}{6!}U^6(t)\partial_1^6\theta(0).\end{aligned} \tag{517}$$

Step 3: Denying (504). We proceed by contradiction, assuming that there exists a local $\mathcal{C}^6$ diffeomorphism θ of $\mathbb{R}^3$ such that, for each $M \in [\![1,6]\!]$, there exists $C_M, T_M > 0$ such that (504) holds for every $t \in [0,T_M]$ and $u \in L^\infty((0,T_M);\mathbb{R})$ with $\|u\|_{L^\infty} \leq 1$.

By induction on M, estimate (504), Lemma 169 and (513) imply that $\gamma_1 = \beta_1\partial_1\theta(0)$, $\gamma_2 = -\beta_2\partial_2\theta(0)$, $\gamma_3 = -2\beta_3\partial_2\theta(0)$, $\gamma_4 = 0$, $\gamma_5 = 2(\beta_4-\beta_5)\partial_3\theta(0)$ and $\gamma_6 = 0$.

On the one hand, by choosing u such that $U(t) = 0$ but $V(t) \neq 0$, the relation $\gamma_4(t,u) = 0$ implies that $\partial_{22}\theta(0) = -\partial_3\theta(0) \neq 0$ because θ is a local diffeomorphism. On the other hand, by choosing u such that $U(t) = V(t) = 0$ but $\int_0^t U^2 \neq 0$, the relation $\gamma_6(t,u) = 0$ implies that $\partial_{22}\theta(0) = 0$. This concludes the proof, since we have found incompatible conditions on $\partial_{22}\theta(0)$. □

Remark 172. *This section is written with a focus on time-based estimates. However, a similar "false proposition" could be stated for control-based estimates. The same counter-example also negates this possibility.*

Acknowledgements

The authors would like to thank Matthias Kawski for the time taken to read carefully this long article, many relevant bibliographical and notational suggestions, and his enthusiasm concerning some of the results contained in this work.

The authors also thank Joackim Bernier for discussions about the numerical literature and the importance of the convergence issues discussed in this article for numerical splitting schemes.

The authors benefit from the support of ANR project LEBESGUE, grant ANR-11-LABX-0020. Karine Beauchard and Frédéric Marbach benefit from the support of ANR project TRECOS, grant ANR-20-CE40-0009-01.

A Proving that the logarithm of the flow is a Lie series

A.1 Using shuffle relations and Ree's theorem

In this paragraph, we describe a proof of Theorem 20 relying on Ree's theorem and shuffle relations satisfied by the coefficients of the Chen series. This approach is notably used in [71, 74]. We start with some definitions.

Definition 173 (Shuffle product)**.** *The* shuffle product *is the map from $I^* \times I^*$ to the free vector space over I^* defined by induction on the length of the words by $\emptyset \sqcup\!\sqcup \sigma = \sigma \sqcup\!\sqcup \emptyset := \sigma$ for every $\sigma \in I^*$ (where $\emptyset$ denotes the empty word) and, for every $\sigma, \sigma' \in I^*$ and $\ell, \ell' \in I$,*

$$(\sigma\ell) \sqcup\!\sqcup (\sigma'\ell') := (\sigma \sqcup\!\sqcup (\sigma'\ell'))\ell + ((\sigma\ell) \sqcup\!\sqcup \sigma')\ell' \tag{518}$$

Intuitively, the shuffle product of two words is the sum of all the ways of riffle shuffling these two words together, interleaving their letters (exactly as one would riffle shuffle two packets of a

card deck). For example, the shuffle product of the words ab and cd (over the Latin alphabet) is $abcd + acbd + acdb + cabd + cadb + cdab$.

The following result was introduced in [89, Theorem 2.5] to prove Theorem 20.

Lemma 174 (Ree's theorem)**.** *Let $\gamma : I^* \to \mathbb{K}$ with $\gamma_\emptyset = 1$. We still denote by γ its linear extension to the free vector space over I^*. Consider the formal series $x := \sum_{\sigma \in I^*} \gamma_\sigma X_\sigma$. Then $\log x \in \widehat{\mathcal{L}}(X)$ iff (the linear extension of) γ satisfies the so-called "shuffle relations", i.e. iff for every $\sigma, \sigma' \in I^*$,*

$$\gamma_{\sigma \sqcup\!\sqcup \sigma'} = \gamma_\sigma \gamma_{\sigma'}. \tag{519}$$

Proof. This statement is item *(iii)* in [90, Theorem 3.2]. □

Therefore, to show that $\log x(t)$ is a Lie series, it suffices to check that the coefficients $\int_0^t a_\sigma$ (defined in (19)) of the Chen series satisfy these shuffle relations. We proceed as in [89, Section 2], by induction on $|\sigma| + |\sigma'|$ in (519). Definition (19) can also be written as, for every $\sigma \in I^*$ and $\ell \in I$,

$$\int_0^t a_{\sigma\ell} = \int_0^t a_\sigma(s) a_\ell(s)\,\mathrm{d}s. \tag{520}$$

Since we set $\int_0^t a_\emptyset = 1$ by definition for the empty word $\emptyset$, (519) holds for every $\sigma, \sigma' \in I^*$ when $|\sigma| + |\sigma'| = 1$. Assume now that it holds for $|\sigma| + |\sigma'| \leq n$ for some $n \in \mathbb{N}^*$. Let $\sigma, \sigma' \in I^*$ and $\ell, \ell' \in I$ such that $|\sigma\ell| + |\sigma'\ell'| = n + 1$. Applying successively, (518) and the linearity of the extension of γ, (520), the induction hypothesis and eventually (520) again, we obtain, for every $t \geq 0$,

$$\begin{aligned}
\gamma_{(\sigma\ell) \sqcup\!\sqcup (\sigma'\ell')}(t) &= \gamma_{(\sigma \sqcup\!\sqcup (\sigma'\ell'))\ell}(t) + \gamma_{((\sigma\ell) \sqcup\!\sqcup \sigma')\ell'}(t) \\
&= \int_0^t \gamma_{\sigma \sqcup\!\sqcup (\sigma'\ell')}(s) a_\ell(s)\,\mathrm{d}s + \int_0^t \gamma_{(\sigma\ell) \sqcup\!\sqcup \sigma'}(s) a_{\ell'}(s)\,\mathrm{d}s \\
&= \int_0^t \gamma_\sigma(s) \gamma_{\sigma'\ell'}(s) a_\ell(s)\,\mathrm{d}s + \int_0^t \gamma_{\sigma\ell}(s) \gamma_{\sigma'}(s) a_{\ell'}(s)\,\mathrm{d}s \\
&= \int_0^t \dot{\gamma}_{\sigma\ell}(s) \gamma_{\sigma'\ell'}(s) + \gamma_{\sigma\ell}(s) \dot{\gamma}_{\sigma'\ell'}(s)\,\mathrm{d}s,
\end{aligned} \tag{521}$$

which proves that $\gamma_{(\sigma\ell) \sqcup\!\sqcup (\sigma'\ell')} = \gamma_{\sigma\ell}\gamma_{\sigma'\ell'}$.

A.2 Using Friedrich's criterion

In this paragraph, we describe a proof of Theorem 20 relying on Friedrich's criterion. This approach is notably used in [97, Section 3]. We start with some definitions.

Let $\mathcal{A}(X) \otimes \mathcal{A}(X)$ be the tensor product of the algebra $\mathcal{A}(X)$ with itself (i.e. the tensor product of $\mathcal{A}(X)$ and $\mathcal{A}(X)$, endowed with the product rule $(a \otimes b)(a' \otimes b') := (aa') \otimes (bb')$, see [26, Chapter 3, Section 4.1, Definition 1] for a precise construction). The algebra $\mathcal{A}(X)$ is the universal enveloping algebra of the Lie algebra $\mathcal{L}(X)$, and as such is a Hopf algebra (see [1]). The coproduct homomorphism $\Delta : \mathcal{A}(X) \to \mathcal{A}(X) \otimes \mathcal{A}(X)$ is defined by setting the values $\Delta(1) := 1 \otimes 1$ and $\Delta(X_i) := X_i \otimes 1 + 1 \otimes X_i$ for $1 \leq i \leq q$. This defines a unique homomorphism because $\mathcal{A}(X)$ is freely generated by X as an algebra (see [90, Proposition 1.2] for more detail). The coproduct Δ can then be used to characterize Lie elements, as in the following result, which was proposed by Friedrichs in [48], then proved by multiple authors in the same period [35, 44, 78, 80].

Lemma 175 (Friedrichs' criterion)**.** *For $a \in \mathcal{A}(X)$, $a \in \mathcal{L}(X)$ if and only if the condition $\Delta(a) = a \otimes 1 + 1 \otimes a$ holds.*

Proof. This statement is the equivalence between *(i)* and *(iii)* in [90, Theorem 1.4]. □

Example 176. *The element X_1X_2 does not belong to $\mathcal{L}$. And indeed,*

$$\begin{aligned}\Delta(X_1X_2) = \Delta(X_1)\Delta(X_2) &= (X_1\otimes 1 + 1\otimes X_1)(X_2\otimes 1 + 1\otimes X_2)\\ &= X_1X_2\otimes 1 + X_1\otimes X_2 + X_2\otimes X_1 + 1\otimes X_1X_2\\ &\neq X_1X_2\otimes 1 + 1\otimes X_1X_2.\end{aligned} \tag{522}$$

On the contrary, the element $[X_1, X_2] = X_1X_2 - X_2X_1$ belongs to $\mathcal{L}$. And indeed,

$$\begin{aligned}\Delta([X_1,X_2]) &= \Delta(X_1X_2) - \Delta(X_2X_1)\\ &= (X_1X_2\otimes 1 + X_1\otimes X_2 + X_2\otimes X_1 + 1\otimes X_1X_2)\\ &\quad - (X_2X_1\otimes 1 + X_2\otimes X_1 + X_1\otimes X_2 + 1\otimes X_2X_1)\\ &= [X_1,X_2]\otimes 1 + 1\otimes [X_1,X_2].\end{aligned} \tag{523}$$

The tensor product $\mathcal{A}(X)\otimes\mathcal{A}(X)$ also has a graded structure, with $(\mathcal{A}(X)\otimes\mathcal{A}(X))_n = \bigoplus_{i=0}^n \mathcal{A}_i(X)\otimes\mathcal{A}_{n-i}(X)$. Since the homomorphism Δ is linear and degree preserving, it can be extended as an homomorphism from $\widehat{\mathcal{A}}(X)$ to $\widehat{\mathcal{A}(X)\otimes\mathcal{A}(X)}$, the formal power series over $\mathcal{A}(X)\otimes\mathcal{A}(X)$. For such series with zero constant term, one can define, as in (11), an exponential, say $\exp_\otimes$, which also verifies a uniqueness property such as Lemma 13. One can then derive a criterion to determine whether the logarithm of a formal power series is a Lie element.

Corollary 177. *Let $a\in\widehat{\mathcal{A}}(X)$ with $a_0 = 1$. Then $\log(a)\in\widehat{\mathcal{L}}(X)$ if and only if $\Delta(a) = a\otimes a$.*

Proof. We follow [90, Theorem 3.2]. By linearity and degree preservation, Lemma 175 implies that, for $a\in\widehat{\mathcal{A}}(X)$, $a\in\widehat{\mathcal{L}}(X)$ if and only if $\Delta(a) = a\otimes 1 + 1\otimes a$. For $a\in\widehat{\mathcal{A}}(X)$ with constant term 1,

$$\begin{aligned}\log a\in\widehat{\mathcal{A}}(X) &\iff \Delta(\log(a)) = \log(a)\otimes 1 + 1\otimes\log(a)\\ &\iff \exp_\otimes(\Delta(\log(a))) = \exp_\otimes(\log(a)\otimes 1 + 1\otimes\log(a))\\ &\iff \Delta(\exp(\log(a))) = \exp_\otimes(\log(a)\otimes 1)\exp_\otimes(1\otimes\log(a))\\ &\iff \Delta(a) = ((\exp\log a)\otimes 1)(1\otimes(\exp\log a)) = a\otimes a,\end{aligned} \tag{524}$$

where we used the equality $\Delta(\exp(\cdot)) = \exp_\otimes(\Delta(\cdot))$, because Δ is an homomorphism, and the fact that $\exp_\otimes(b\otimes 1 + 1\otimes c) = \exp_\otimes(b\otimes 1)\exp_\otimes(1\otimes c)$, because $b\otimes 1$ and $1\otimes c$ commute. □

Therefore, to show that $\log(x(t))$ is a Lie series, it suffices to check that $\Delta(x(t)) = x(t)\otimes x(t)$. This can be checked as in [97, Section 3] using the following argument. At the initial time $\Delta(x(0)) = \Delta(1) = 1\otimes 1 = x(0)\otimes x(0)$. On the one hand

$$\frac{\mathrm{d}}{\mathrm{d}t}\Delta(x) = \Delta(\dot{x}) = \Delta(xa) = \Delta(x)\Delta(a) = \Delta(x)(a\otimes 1 + 1\otimes a). \tag{525}$$

On the other hand,

$$\frac{\mathrm{d}}{\mathrm{d}t}(x\otimes x) = \dot{x}\otimes x + x\otimes\dot{x} = (xa)\otimes x + x\otimes(xa) = (x\otimes x)(a\otimes 1 + 1\otimes a). \tag{526}$$

Hence, both quantities satisfy the same formal differential equation with the same initial condition, so they are equal for every $t\geq 0$.

B Elementary numerical identities

B.1 Bernoulli numbers

We use the notation $(B_n)_{n\in\mathbb{N}}$ to denote the Bernoulli numbers, which are defined (using the modern NIST sign and indexing convention) by the identity

$$\forall z\in\mathbb{C}, |z| < 2\pi, \qquad \frac{z}{e^z - 1} = \sum_{n=0}^{+\infty} B_n\frac{z^n}{n!} = 1 - \frac{z}{2} + \sum_{n=1}^{+\infty} B_{2n}\frac{z^{2n}}{(2n)!}. \tag{527}$$

Lemma 178. *The Bernoulli numbers satisfy, for every $n \geq 2$*

$$\sum_{k=0}^{n-1} \binom{n}{k} B_k = 0, \tag{528}$$

$$\sum_{k=0}^{n} \binom{n}{k} \frac{B_k}{n+1-k} = 0, \tag{529}$$

$$\sum_{k=0}^{n} \frac{B_{n-k}}{(n-k)!(k+2)!} = -\frac{B_{n+1}}{(n+1)!}. \tag{530}$$

Moreover, the odd Bernoulli numbers except B_1 vanish and, for every $n \geq 1$,

$$B_{2n} = (-1)^{n+1} \frac{2(2n)!}{(2\pi)^{2n}} \zeta(2n) \sim (-1)^{n+1} 2\sqrt{2\pi n} \left(\frac{n}{\pi}\right)^{2n}, \tag{531}$$

where ζ is the Riemann zeta function.

Proof. The first two identities are classical and can be proved using the generating series of the Bernoulli numbers of (527), respectively by identification in $z = (e^z - 1) \times (z/(e^z - 1))$ for (528) and in $1 = ((e^z - 1)/z) \times (z/(e^z - 1))$ for (529).

The third identity (530) follows from (529) and the computation

$$\begin{aligned} \sum_{k=0}^{n} \frac{B_{n-k}}{(n-k)!(k+2)!} &= \frac{1}{(n+1)!} \sum_{\ell=0}^{n} \binom{n}{n-\ell} \frac{n+1}{n+1-\ell} \frac{B_\ell}{n-\ell+2} \\ &= \frac{1}{(n+1)!} \sum_{\ell=0}^{n} \binom{n+1}{\ell} \frac{B_\ell}{(n+1)-\ell+1} = -\frac{B_{n+1}}{(n+1)!}. \end{aligned} \tag{532}$$

Eventually, the relationship with the Riemann zeta function is proved in [9, equation (12.38)]. The asymptotic is a consequence of the Stirling's approximation and $\zeta(s) \to 1$ as $s \geq 1$ tends to $+\infty$ (which is a direct consequence of the formula $\zeta(s) = \sum n^{-s}$). □

References

[1] Eiichi Abe. *Hopf algebras*, volume 74. Cambridge University Press, Cambridge, 1980.

[2] Rüdiger Achilles and Andrea Bonfiglioli. The early proofs of the theorem of Campbell, Baker, Hausdorff, and Dynkin. *Arch. Hist. Exact Sci.*, 66(3):295–358, 2012.

[3] Andrei Agrachev and Revaz Gamkrelidze. Exponential representation of flows and a chronological enumeration. *Mat. Sb. (N.S.)*, 107(149)(4):467–532, 639, 1978.

[4] Andrei Agrachev and Revaz Gamkrelidze. Chronological algebras and nonstationary vector fields. In *Problems in geometry, Vol. 11 (Russian)*, pages 135–176, 243. Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow, 1980.

[5] Andrei Agrachev and Yuri Sachkov. *Control theory from the geometric viewpoint*. Berlin: Springer, 2004.

[6] Mahdi Jasim Hasan Al-Kaabi, Kurusch Ebrahimi-Fard, and Dominique Manchon. Post-Lie-Magnus expansion and BCH-recursion. *arXiv preprint arXiv:2108.11103*, 2021.

[7] Jaime Arango and Adriana Gómez. Flows and diffeomorphisms. *Rev. Colomb. Mat.*, 32(1):13–27, 1998.

[8] Jaime Arango and Adriana Gómez. Diffeomorphisms as time one maps. *Aequationes Math.*, 64(3):304–314, 2002.

[9] George Arfken, Hans Weber, and Frank Harris. *Mathematical methods for physicists. A comprehensive guide. 7th ed.* Amsterdam: Elsevier/Academic Press, 7th ed. edition, 2013.

[10] Ana Arnal, Fernando Casas, and Cristina Chiralt. A note on the Baker–Campbell–Hausdorff series in terms of right-nested commutators. *Mediterranean Journal of Mathematics*, 18(2):1–16, 2021.

[11] Henry Baker. Alternants and Continuous Groups. *Proc. London Math. Soc. (2)*, 3:24–47, 1905.

[12] Fabrice Baudoin. *An introduction to the geometry of stochastic flows.* Imperial College Press, London, 2004.

[13] Karine Beauchard, Jérémy Le Borgne, and Frédéric Marbach. Growth of structure constants of free Lie algebras relative to Hall bases. *Journal of Algebra*, 612:281–378, 2022.

[14] Karine Beauchard and Frédéric Marbach. Quadratic obstructions to small-time local controllability for scalar-input systems. *J. Differential Equations*, 264(5):3704–3774, 2018.

[15] Gérard Ben Arous. Flots et séries de Taylor stochastiques. *Probab. Theory Related Fields*, 81(1):29–77, 1989.

[16] Stefano Biagi, Andrea Bonfiglioli, and Marco Matone. On the Baker-Campbell-Hausdorff theorem: non-convergence and prolongation issues. *Linear and Multilinear Algebra*, pages 1–19, 2018.

[17] Rosa Maria Bianchini and Gianna Stefani. Sufficient conditions of local controllability. In *1986 25th IEEE Conference on Decision and Control*, pages 967–970. IEEE, 1986.

[18] Sergio Blanes and Fernando Casas. On the convergence and optimization of the Baker-Campbell-Hausdorff formula. *Linear Algebra Appl.*, 378:135–158, 2004.

[19] Sergio Blanes and Fernando Casas. *A concise introduction to geometric numerical integration.* Monographs and Research Notes in Mathematics. CRC Press, Boca Raton, FL, 2016.

[20] Sergio Blanes, Fernando Casas, and Ander Murua. Splitting and composition methods in the numerical integration of differential equations. *Bol. Soc. Esp. Mat. Apl., S$\vec{e}$MA*, 45:89–145, 2008.

[21] Sergio Blanes, Fernando Casas, José Oteo, and Javier Ros. A pedagogical approach to the Magnus expansion. *Eur. J. Phys.*, 31(4):907–918, 2010.

[22] Sergio Blanes and Per Christian Moan. Splitting methods for non-autonomous Hamiltonian equations. *J. Comput. Phys.*, 170(1):205–230, 2001.

[23] Andrea Bonfiglioli and Roberta Fulci. *Topics in noncommutative algebra*, volume 2034 of *Lecture Notes in Mathematics*. Springer, Heidelberg, 2012. The theorem of Campbell, Baker, Hausdorff and Dynkin.

[24] Nicolas Bourbaki. Elements of mathematics. Lie groups and Lie algebras. Part I: Chapters 1-3. English translation. Springer, 1975.

[25] Nicolas Bourbaki. *Éléments d'histoire des mathématiques.* Masson, Paris, 1984. Reprints of the historical notes from the ıt Eléments.

[26] Nicolas Bourbaki. *Algebra I. Chapters 1–3.* Elements of Mathematics (Berlin). Springer-Verlag, Berlin, 1998. Translated from the French, Reprint of the 1989 English translation.

[27] Boris Buffoni and John Toland. *Analytic theory of global bifurcation. An introduction.* Princeton, NJ: Princeton University Press, 2003.

[28] Kevin Burrage and Pamela Burrage. High strong order methods for non-commutative stochastic ordinary differential equation systems and the Magnus formula. *Phys. D*, 133(1-4):34–48, 1999. Predictability: quantifying uncertainty in models of complex phenomena (Los Alamos, NM, 1998).

[29] John Edward Campbell. On a Law of Combination of Operators (Second Paper). *Proc. Lond. Math. Soc.*, 29:14–32, 1897/98.

[30] Fernando Casas and Arieh Iserles. Explicit magnus expansions for nonlinear equations. *Journal of Physics A: Mathematical and General*, 39(19):5445, 2006.

[31] Fernando Casas and Ander Murua. An efficient algorithm for computing the Baker-Campbell-Hausdorff series and some of its applications. *J. Math. Phys.*, 50(3):033513, 23, 2009.

[32] Fabienne Castell. Asymptotic expansion of stochastic flows. *Probab. Theory Related Fields*, 96(2):225–239, 1993.

[33] Kuo-Tsai Chen. Iterated integrals and exponential homomorphisms. *Proc. London Math. Soc. (3)*, 4:502–512, 1954.

[34] Kuo-Tsai Chen. Integration of paths, geometric invariants and a generalized Baker-Hausdorff formula. *Ann. Math. (2)*, 65:163–178, 1957.

[35] Paul Moritz Cohn. Sur le critère de Friedrichs pour les commutateurs dans une algèbre associative libre. *C. R. Acad. Sci. Paris*, 239:743–745, 1954.

[36] Jean-Michel Coron. *Control and nonlinearity*, volume 136. Providence, RI: American Mathematical Society (AMS), 2007.

[37] Charles Curry, Kurusch Ebrahimi-Fard, and Hans Munthe-Kaas. What is a post-lie algebra and why is it useful in geometric integration. In *European Conference on Numerical Mathematics and Advanced Applications*, pages 429–437. Springer, 2017.

[38] Bruce Driver. On truncated logarithms of flows on a riemannian manifold. *arXiv preprint arXiv:1810.02414*, 2018.

[39] Evgenii Dynkin. Calculation of the coefficients in the Campbell-Hausdorff formula. *Doklady Akad. Nauk SSSR (N.S.)*, 57:323–326, 1947.

[40] Evgenii Dynkin. On the representation by means of commutators of the series $\log(e^x e^y)$ for noncommutative x and y. *Mat. Sbornik N.S.*, 25(67):155–162, 1949.

[41] Kurusch Ebrahimi-Fard and Dominique Manchon. A Magnus- and Fer-type formula in dendriform algebras. *Found. Comput. Math.*, 9(3):295–316, 2009.

[42] Kurusch Ebrahimi-Fard and Dominique Manchon. The Magnus expansion, trees and Knuth's rotation correspondence. *Found. Comput. Math.*, 14(1):1–25, 2014.

[43] Kurusch Ebrahimi-Fard and Frédéric Patras. The pre-Lie structure of the time-ordered exponential. *Lett. Math. Phys.*, 104(10):1281–1302, 2014.

[44] David Finkelstein. On relations between commutators. *Comm. Pure Appl. Math.*, 8:245–250, 1955.

[45] Michel Fliess. Fonctionnelles causales non linéaires et indéterminées non commutatives. *Bull. Soc. Math. France*, 109(1):3–40, 1981.

[46] Michel Fliess. Réalisation locale des systemes non linéaires, algebres de lie filtrées transitives et séries génératrices non commutatives. *Inventiones mathematicae*, 71(3):521–537, 1983.

[47] Gunnar Fløystad and Hans Munthe-Kaas. Pre- and post-Lie algebras: the algebro-geometric view. In *Computation and combinatorics in dynamics, stochastics and control. The Abel symposium, Rosendal, Norway, August 16–19, 2016. Selected papers*, pages 321–367. Cham: Springer, 2018.

[48] Kurt Friedrichs. Mathematical aspects of the quantum theory of fields. V. Fields modified by linear homogeneous forces. *Comm. Pure Appl. Math.*, 6:1–72, 1953.

[49] Revaz Gamkrelidze. Exponential representation of solutions of ordinary differential equations. In *Equadiff IV (Proc. Czechoslovak Conf. Differential Equations and their Applications, Prague, 1977)*, volume 703 of *Lecture Notes in Math.*, pages 118–129. Springer, Berlin, 1979.

[50] Israel Gelfand and Mark Naĭmark. On the imbedding of normed rings into the ring of operators in Hilbert space. *Mat. Sb., Nov. Ser.*, 12:197–213, 1943.

[51] Laetitia Giraldi, Pierre Lissy, Clément Moreau, and Jean-Baptiste Pomet. Necessary conditions for local controllability of a particular class of systems with two scalar controls. *arXiv preprint arXiv:1907.04706*, 2019.

[52] Janusz Grabowski. Free subgroups of diffeomorphism groups. *Fundam. Math.*, 131(2):103–121, 1988.

[53] Steven Gray and Luis Duffaut Espinosa. A Faà di Bruno Hopf algebra for a group of Fliess operators with applications to feedback. *Syst. Control Lett.*, 60(7):441–449, 2011.

[54] Steven Gray, Luis Duffaut Espinosa, and Kurusch Ebrahimi-Fard. Faà di Bruno Hopf algebra of the output feedback group for multivariable Fliess operators. *Syst. Control Lett.*, 74:64–73, 2014.

[55] Steven Gray and Kurusch Ebrahimi-Fard. Generating series for networks of Chen-Fliess series. *Syst. Control Lett.*, 147:9, 2021. Id/No 104827.

[56] Steven Gray, Luis Duffaut Espinosa, and Kurusch Ebrahimi-Fard. Additive networks of Chen-Fliess series: Local convergence and relative degree. *arXiv preprint arXiv:2104.08950*, 2021.

[57] Steven Gray and Yuan Wang. Fliess operators on L^p spaces: convergence and continuity. *Systems & Control Letters*, 46(2):67–74, 2002.

[58] Ernst Hairer, Christian Lubich, and Gerhard Wanner. *Geometric numerical integration*, volume 31 of *Springer Series in Computational Mathematics*. Springer, Heidelberg, 2010.

[59] Marshall Hall. A basis for free Lie rings and higher commutators in free groups. *Proc. Am. Math. Soc.*, 1:575–581, 1950.

[60] Philip Hall. A contribution to the theory of groups of prime-power order. *Proc. Lond. Math. Soc. (2)*, 36:29–95, 1933.

[61] Felix Hausdorff. Die symbolische exponentialformel in der gruppentheorie. *Ber. Verh. Kgl. Schs. Ges. Wiss. Leipzig., Math.-phys. Kl.*, 58:19–48, 1906.

[62] Henry Hermes. Local controllability and sufficient conditions in singular problems. ii. *SIAM Journal on Control and Optimization*, 14(6):1049–1062, 1976.

[63] Henry Hermes and Matthias Kawski. Local controllability of a single input, affine system. *Nonlinear analysis and applications*, 109:235–248, 1987.

[64] Arieh Iserles, Hans Munthe-Kaas, Syvert Nørsett, and Antonella Zanna. Lie-group methods. In *Acta numerica, 2000*, volume 9 of *Acta Numer.*, pages 215–365. Cambridge Univ. Press, Cambridge, 2000.

[65] Nathan Jacobson. *Lie algebras*. Dover Publications, Inc., New York, 1979. Republication of the 1962 original.

[66] Frédéric Jean. *Control of nonholonomic systems: from sub-Riemannian geometry to motion planning*. Cham: Springer; Bilbao: BCAM – Basque Center for Applied Mathematics, 2014.

[67] Velimir Jurdjevic and Héctor Sussmann. Control systems on Lie groups. *J. Differential Equations*, 12:313–329, 1972.

[68] Mikhail Karasev and Mikhail Mosolova. Infinite products and t-products of exponents. *Teoreticheskaya i Matematicheskaya Fizika*, 28(2):189–200, 1976.

[69] Matthias Kawski. Controllability via chronological calculus. In *Proceedings of the IEEE Conference on Decision and Control*, volume 3, pages 2920–2925. IEEE, 1999.

[70] Matthias Kawski. Calculating the logarithm of the chen fliess series. In *Proc. MTNS, Perpignan*, 2000.

[71] Matthias Kawski. The combinatorics of nonlinear controllability and noncommuting flows. In *Mathematical control theory, Part 1, 2 (Trieste, 2001)*, ICTP Lect. Notes, VIII, pages 223–311. Abdus Salam Int. Cent. Theoret. Phys., Trieste, 2002.

[72] Matthias Kawski. Chronological calculus in systems and control theory. *Mathematics of Complexity and Dynamical Systems*, page 88, 2011.

[73] Matthias Kawski. High-order small-time local controllability. In *Nonlinear controllability and optimal control*, pages 431–467. Routledge, 2017.

[74] Matthias Kawski and Héctor Sussmann. Noncommutative power series and formal Lie-algebraic techniques in nonlinear control theory. In *Operators, systems, and linear algebra (Kaiserslautern)*, European Consort. Math. Indust., pages 111–128. Teubner, Stuttgart, 1997.

[75] Arthur Krener. On the equivalence of control systems and the linearization of nonlinear systems. *SIAM J. Control*, 11:670–676, 1973.

[76] Arthur Krener. Local approximation of control systems. *J. Differ. Equations*, 19:125–133, 1975.

[77] John Lee. *Introduction to smooth manifolds*, volume 218. New York, NY: Springer, 2nd revised edition, 2013.

[78] Roger Lyndon. A theorem of Friedrichs. *Michigan Math. J.*, 3:27–29, 1955.

[79] Marco Maggia, Sameh A Eisa, and Haithem E Taha. On higher-order averaging of time-periodic systems: reconciliation of two averaging techniques. *Nonlinear Dynamics*, 99(1):813–836, 2020.

[80] Wilhelm Magnus. On the exponential solution of differential equations for a linear operator. *Comm. Pure Appl. Math.*, 7:649–673, 1954.

[81] Robert McLachlan. Composition methods in the presence of small parameters. *BIT*, 35(2):258–268, 1995.

[82] Robert McLachlan, Klas Modin, Hans Munthe-Kaas, and Olivier Verdier. Butcher series: a story of rooted trees and numerical methods for evolution equations. *arXiv preprint arXiv:1512.00906*, 2015.

[83] Per Christian Moan and Jitse Niesen. Convergence of the Magnus series. *Found. Comput. Math.*, 8(3):291–301, 2008.

[84] Clément Mouhot and Cédric Villani. On Landau damping. *Acta Math.*, 207(1):29–201, 2011.

[85] Hans Munthe-Kaas and Brynjulf Owren. Computations in a free Lie algebra. *Philos. Trans. R. Soc. Lond., Ser. A, Math. Phys. Eng. Sci.*, 357(1754):957–981, 1999.

[86] Ander Murua. The Hopf algebra of rooted trees, free Lie algebras, and Lie series. *Found. Comput. Math.*, 6(4):387–426, 2006.

[87] Alexander Nagel, Elias Stein, and Stephen Wainger. Balls and metrics defined by vector fields. I. Basic properties. *Acta Math.*, 155(1-2):103–147, 1985.

[88] Jacob Palis. Vector fields generate few diffeomorphisms. *Bull. Am. Math. Soc.*, 80:503–505, 1974.

[89] Rimhak Ree. Lie elements and an algebra associated with shuffles. *Ann. of Math. (2)*, 68:210–220, 1958.

[90] Christophe Reutenauer. *Free Lie algebras*, volume 7 of *London Mathematical Society Monographs. New Series*. The Clarendon Press, Oxford University Press, New York, 1993.

[91] Leonardo Saenz and Rodolfo Suarez. Lie algebras associated with the exponential solutions of nonautonomous linear differential equations. *J. Math. Phys.*, 42(9):4582–4605, 2001.

[92] Leonardo Saenz and Rodolfo Suarez. A combinatorial approach to the generalized Baker-Campbell-Hausdorff-Dynkin formula. *Systems & Control Letters*, 45(5):357–370, 2002.

[93] Andrei Sarychev. Lie- and chronologico-algebraic tools for studying stability of time-varying systems. *Syst. Control Lett.*, 43(1):59–76, 2001.

[94] Anatolii Shirshov. On the bases of a free Lie algebra. *Algebra Logika*, 1(1):14–19, 1962.

[95] Wilhelm Specht. Die linearen Beziehungen zwischen höheren Kommutatoren. *Math. Z.*, 51:367–376, 1948.

[96] Gianna Stefani. On the local controllability of a scalar-input control system. In *Theory and applications of nonlinear control systems (Stockholm, 1985)*, pages 167–179. North-Holland, Amsterdam, 1986.

[97] Robert Strichartz. The Campbell-Baker-Hausdorff-Dynkin formula and solutions of differential equations. *J. Funct. Anal.*, 72(2):320–345, 1987.

[98] Héctor Sussmann. An extension of a theorem of Nagano on transitive Lie algebras. *Proc. Am. Math. Soc.*, 45:349–356, 1974.

[99] Héctor Sussmann. Lie brackets and local controllability: a sufficient condition for scalar-input systems. *SIAM J. Control Optim.*, 21(5):686–713, 1983.

[100] Héctor Sussmann. A product expansion for the Chen series. In *Theory and applications of nonlinear control systems (Stockholm, 1985)*, pages 323–335. North-Holland, Amsterdam, 1986.

[101] Hale Trotter. On the product of semi-groups of operators. *Proc. Am. Math. Soc.*, 10:545–551, 1959.

[102] Gérard Viennot. *Algèbres de Lie libres et monoïdes libres*, volume 691 of *Lecture Notes in Mathematics*. Springer, Berlin, 1978. Bases des algèbres de Lie libres et factorisations des monoïdes libres.

[103] James Wei. Note on the global validity of the Baker-Hausdorff and Magnus theorems. *J. Mathematical Phys.*, 4:1337–1341, 1963.

[104] Franz Wever. Operatoren in Lieschen Ringen. *J. Reine Angew. Math.*, 187:44–55, 1949.